Bayesian Network for Predicting Student Retention Based on Expert Elicitation

Gbolagade Kola Adegoke

A thesis submitted in partial fulfilment of the requirements of
Staffordshire University for the degree of
Doctor of Philosophy

Staffordshire University
School of Digital, Technologies and Arts
United Kingdom

November 2022

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification, except as specified.

I declare that this thesis conforms to the university research regulations and represents my own work.

The research findings are documented honestly and fairly, and work attributed to other researchers has been acknowledged and referenced.

Name: Gbolagade Kola Adegoke

Signature: GKA

Date: November 2022

Dedication

This Thesis is dedicated to my late grandparents and parents who had died before the commencement of my PhD studies. I wish you were alive to show you how proud I am being able to reach the point which you had aimed at seeing me.

Their desire for me to obtain this degree is beyond what my words can describe, and they so much wished me to reach this level of education, but that was not meant to be.

Your passion for quality education and decency of lifestyle are great values that I will always remember and appreciate. Thanks for being such great grandparents and parents.

Acknowledgments

This research would not have been completed without the help and guidance of Almighty God, all thanks and praises to Him.

I would like to express my sincere gratitude to my supervisors, Prof. Mohamed Sedky and Dr. Asma Patel, for their encouragement, support, and continuous proficient guidance they gave to me, and also for their inspiration, advice and friendship. Their marvellous knowledge made me completed my research successfully.

Also, I would like to extend my thanks to both Dr. David Trigg and Prof. Claude Chibelushi who started the supervision of the study but could not complete it, because both of them left the University at the same time before I completed this research, for their support and guidance.

Many thanks to those who serve as Domain Experts for the research for a job well done. Without them, there would be no data to use for the research. I really appreciate their efforts and contributions.

Of course, I will never forget to express my love and sincere thanks for the support, prayers, encouragement and patience of my wife, children, and sisters throughout my PhD research.

This PhD would be impossible without the funding of Tertiary Education Trust Fund (TETfund), Nigeria and Nasarawa State University, Keffi (NSUK) also in Nigeria. Many thanks go out to the Fund and the University for providing me with a financial support.

Special thanks to my late parents and grandparents, without their support all, I would not have been able to reach where I am now. I can only say I miss you all so much and hope the soul of every one of you is blessed. Many thanks for deeming it fit to educate me. I wish you were alive to show you how proud I am being able to reach the point which you had aimed at seeing me.

Last but not the least, thanks are also due to past and current colleagues at the Research Laboratory, School of Digital, Technologies and Arts, Staffordshire University, Stoke-on-Trent City, England, sUnited Kingdom, for their greatly appreciated support.

Abstract

This project aims to create a Bayesian network (BN) for predicting student retention based on expert elicitation. It aims to construct a Bayesian network (BN) that can be used to predict student retention among and between the variables in the domain.

The model is created primarily by eliciting the necessary information on the types and strength of relationships in the system from the student retention experts.

The participants involved in the elicitation session are experts in their field. It is envisaged that sessions involved interviews and interactive modelling. Session involved a group of domain experts.

The information required from the experts is a causal structure, detailing cause and effect relationships at work in the system being modelled and estimates of probabilities, to capture the strength and characteristic of these relationships.

The values elicited from the experts were used to create the Bayesian network (BN) model. The model was validated by the domain experts.

Academic progress, academic engagement, academic skills, attendance, and financial and moral support from: government, university, partner, family and friends are identified as key variables related to student retention.

Predicting student retention in order to detect students at risk in the early stages of education is essential so that higher education institutions (HEIs) can minimise students not graduating on time or drop out outrightly.

The study shows that in a situation where domain data is sparse or not available at all, the knowledge-driven approach is suitable to be used to elicit estimates of probabilities from domain experts.

Table of Contents

List of Figures

List of Tables

LIST OF APPENDICES

List of Acronyms

Acronym	Full form
A/L	Advanced Level
BN	Bayesian Network
CGPA	Cummulative Grade Point Average
DE	Direct Entry
GPA	Grade Point Average
HEI	Higher Education Institution
UTME	Unified Tertiary Matriculation Examination

Definition of Terms

Attrition – refers to students leaving school.

Dropout – refers to a student who has discontinued his/her study with no immediate plan to re-enrol.

Goal commitment – refers to the degree to which the student is committed or motivated to get a university degree in general.

GPA – refers to Grade Point Average.

CGPA – refers to Cummulative Grade Point Average.

Retained – refers to the students who graduated from the university.

Retention – refers to the ability of an institution to retain a student from admission to the university to graduation.

Retention rate – refers to the percentage of students who were enrolled at the university and stay there until they graduated.

Chapter 1

Introduction

1.1 Context

The PhD research, titled ‘Bayesian network for predicting student retention based on expert elicitation’, aims to construct a Bayesian network that can be used to predict student retention. The estimates of probabilities for the variables in the domain are elicited from student retention experts. The experts that are eligible to participate in this study involve student affiars officers, level coordinators, examination officers, lecturers, and faculty officers.

The domain experts were asked to provide information on the types and strength of relationships among and between the variables in the domain, such as “Change in Circumstances”, “Mode of Entry”, “Academic Skills”, “Academic Engagement”, “Academic Progress”, “Attendance” and a host of other variables as well as the interactions among and between the variables enumerated above.

Student retention is the measure of the rate at which students re-enrol from one academic session to the following one until they graduate. According to (Gragg, 2022), student retention is the measure of students that enroll, continue, and finish their academic studies in the same school. It is a fundamental mission of a university to retain students and advance them towards successful graduation (Kang & Wang, 2018). In efforts to attain this mission, a university must be able to recognise and understand all the factors impacting student retention and success (Bytheway & Venter, 2014).

Universities need to practise student retention as it ensures support systems are in place to enable students to remain at university and succeed. According to (Simpson, 2005), student retention improves graduate rates, decreases loss of tuition revenue from students that either drop out or transfer to another institution, and brings reputation to an institution as well. It helps students by improving their lives, the lives of their families as well as affording them the opportunity to make a positive contribution to their local community.

Generally, there are three main approaches for constructing a Bayesian network, namely: data-driven, a mixture of data- and knowledge-driven, and knowledge-driven. Typically, if domain data is available, then a data-driven approach can be used, but if domain data is sparse, elicitation can be used to supplement the sparse data (Meyer & Booker, 2001; O'Leary, 2015). However, in a situation where domain data is not available at all, the knowledge-driven approach is used to elicit estimates of probabilities from domain experts.

In this research, estimates of probabilities were elicited from the domain experts for the various variables and states in the Bayesian network. The estimates of probabilities provided by the domain experts served as input data that fed into the Bayesian network.
However, the knowledge-driven approach is associated with problems of heuristics and biases, such as anchoring and adjustment, availability, and representativeness (Tversky & Kahneman, 1973; Tversky & Kahneman, 1974; Kahneman, 2011). Since these problems could adversely affect the accuracy of the encoding process, the experts were checked against those problems previously mentioned, when they are providing probability distributions/summaries since these human judgemental (cognitive) errors might affect their assessments. In order to obtain accurate estimates of probabilities from the domain experts in the elicitation session, they need to be motivated by the facilitator by asking questions logically as well as avoiding multiple questions in one.

1.2 Aim of the Investigation

The aim of this project is to propose a Bayesian network (BN) for predicting student retention based on expert elicitation and to evaluate different methods of eliciting estimates of probabilities and identify the appropriate one for predicting student retention.

The following objectives will be developed to achieve this research aim:

1. To undertake a literature review on techniques to inform the development of a method of eliciting estimates of probabilities from domain experts, that is, student retention experts.

2. To undertake a literature review on techniques to inform the development of a method of eliciting estimates of probabilities from multiple experts.
3. To motivate domain experts when eliciting causal structure and estimates of probabilities from them in order to obtain better results.
4. To construct a Bayesian network for predicting student retention.
5. To test the Bayesian network in order to find out how well it performs.

1.3 Research Ouestion

The below research question was formulated for this research based on its objective.

To what extent does motivation impact on elicitation of estimates of probabilities from domain experts?

1.4 Scope of Investigation

The scope of the research described in this thesis covered elicitation of causal structure, and estimates of probabilities, how to obtain accurate estimates of probabilities from domain experts as well as construction and testing of Bayesian networks as shown in Figure 1-1 below. The thesis aims to obtain causal structure and estimates of probabilities from domain experts on student retention. The thesis investigates different techniques for obtaining accurate estimates of probabilities from domain experts in an elicitation session. The causal structure and estimates of probabilities obtained from the domain experts was used to construct a BN, which has been used to predict the probability of students to enrol the following academic year and continue to complete their study in the university or otherwise.

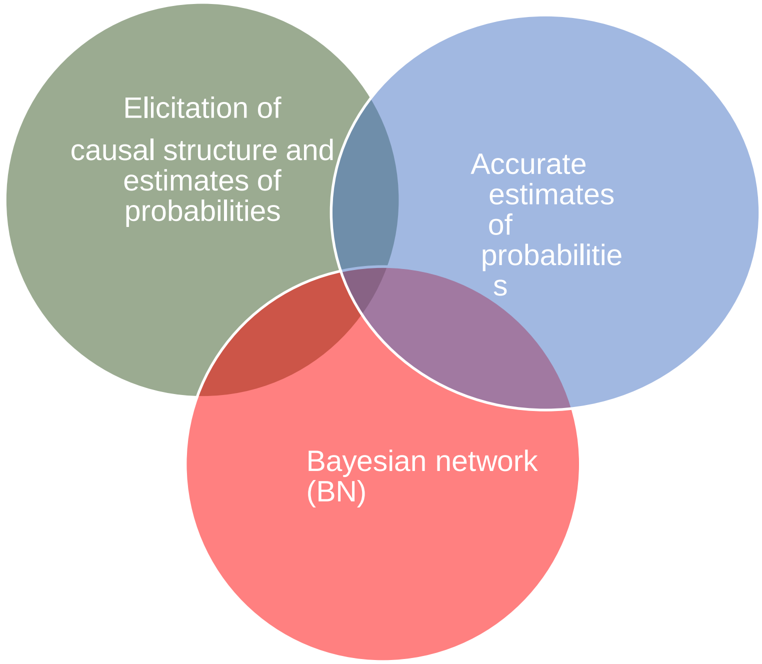


Figure 1-1 Scope of the work

1.5 Contributions to Knowledge

The primary contributions to knowledge include:

- Impact of motivation on elicitation
- A novel elicitation protocol
- Development of a software tool to predict student retention.

Secondary contributions:

- An evaluation on the effectiveness of different approaches to elicitation with a particular reference to trouble shooting of reachability and security issues in a computer network - Direct and Indirect elicitation methods.
- A study on the impact of motivation on elicitation of estimates of probabilities from domain experts.

1.6 Research Approach

The research method is involved with research approaches and philosophies. It is important the selected research method serves the research purposes (Saunders, et al., 2009).

The research approach can be classified into three types qualitative, quantitative, and hybrid methods. The two main categories are qualitative and quantitative. A qualitative approach is used to obtain non-numerical data to understand the reasons, opinions, and motivations. It is considered subjective since the researchers are involved in the experiment themselves (Creswell, 2014). The inductive approach and interpretivism philosophy are related to this method because they are commonly applied to understand facts and generate general principles which related to an interpretivism philosophy.

Conversely, a quantitative approach is objective, as it typically deals with numbers, statistical data, or statistical analysis (Creswell, 2014). A quantitative approach is applied to analyse the data involved with numerical data such as statistical data and graph (Saunders, et al., 2009). This method is commonly used with a deductive approach based on a positivism philosophy (scientific philosophy). The deductive approach is based on using knowledge and information to perceive or produce an opinion about something and a positivism philosophy believes in proved facts instead of ideas. Therefore, a research based on a positivism philosophy usually uses a deductive approach and quantitative method to find out the proved facts of something.

The hybrid method is a mixture of quantitative and qualitative approaches. It is applied to analyse data by using both qualitative and quantitative methods.

This research adopts the hybrid method (qualitative and quantitative approaches). The qualitative method based on the inductive approach and interpretivism philosophy as well as the quantitative method based on the deductive approach and positivism philosophy are used to elicit the causal structure of the variables in the domain (qualitative in nature) as well as the estimates of probabilities of their relationships (qualitative in nature) while analysis of estimates collected in the session is quantitative in nature. Furthermore, the research strategy adopted to address the topic was an experimental one.

The qualitative approach involves the interview method where estimates of probabilities are elicited from the domain experts in student retention.

The quantitative approach involves the use of a questionnaire which was completed and returned by the domain experts in student retention.

The research onion is a popular way to depict the research approach, philosophy, and strategy to be used for research. This research makes use of the positivism and interpretivism as indicated in the figure below.

To develop a robust algorithm to predict student retention, it is required to accurately quantify the performance over a range of data representative of operational conditions. The available tools to evaluate the effectiveness of such approach and visualize the results are the image quality performance metrics. The Figure 1-2 below shows the sequence diagram of the adopted research methodology.

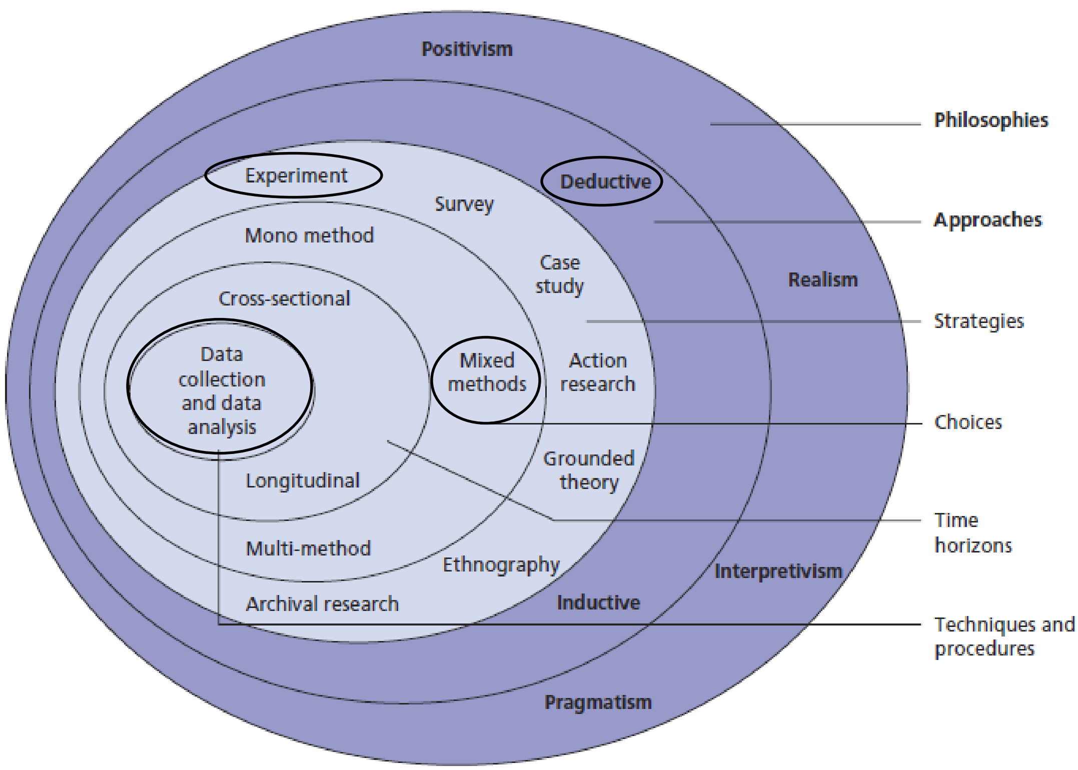


Figure 1-2 The research 'onion' (Saunders et al., 2009).

1.7 Ethical statement

The research complied with the Staffordshire University's and British Computer Society (BCS) ethical rules guiding the conduct of research. The reasons for the ethical compliance are to ensure transparency, confidentiality, and to protect the participants (Creswell, 2014). The research elicited a causal structure and estimates of probabilities from the participants who do not fall under any category of people considered vulnerable.

For the purposes of ethical approval, the Proportionate Review Ethics Form had been completed and the study had been approved by the School of Digital, Technologies and Arts Ethics Panel (see Appendix 9). The participants were not forced to answer any questions against their wish, and they had the right to withdraw at any stage of the study.

Special considerations were given to the ethical principle regarding academic honesty by acknowledging the authors whose works were used in this research. Furthermore, their names and the dates of publications of their works were cited, both in the text and reference sections of this PhD thesis.

The ethical considerations included the adherence to policies concerning computer usage such as privacy, confidentiality, anonymity, and security (Banegas & Villacañas de Castro, 2015).

1.8 Thesis organisation

This thesis consists of 6 chapters, in which this chapter has described the problem area, and the aims and objectives of the study. This thesis is organised in the following way:

Chapter 1 gives an overview, the objectives, as well as the contribution to knowledge of this research work are stated. Finally, the chapter concludes with a brief overview of the chapters to follow.

Chapter 2 investigates the existing literature on elicitation of estimates of probabilities student retention and Bayesian network.

Chapter 3 presents Primary Research 1, which focuses on Bayesian Networks and their applications with a particular reference to troubleshooting of reachability and security issues in a computer network based on elicitation of knowledge from a domain expert while Primary Research 2 focuses on impact of motivation on elicitation of estimates of probabilities from student participants who serve as experts in Internet usage in the UK due to their existing knowledge.

Chapter 4 provides a description of the experiment on elicitation of the causal structure and estimates of probabilities from domain experts.

Chapter 5 is devoted to data analysis and interpretation of results.

Chapter 6 summarises the findings of the study, makes recommendations and draws conclusions as well.

Chapter 2

Literature Review

This Chapter discusses the previous work carried out on elicitation of causal structure and estimates of probabilities. Furthermore, it discusses construction of Bayesian network (BN) and its application areas as well as that of student retention.

2.1 Elicitation of Estimates of Probabilities

Elicitation is the process of extracting expert knowledge about some unknown variable or variables and expressing that information as a probability distribution (O'Hagan, et al., 2006). In a situation where there is a little or no statistical data about a variable or variables of interest, probability values are typically elicited from domain experts (Huang & Darwiche, 1994). Usually, the use of expert opinion is adopted to provide probabilistic inputs where other sources of information are unavailable or are not cost effective.

Elicitation was used to estimate some potential causes of automobile failure in an experiment using the fault trees technique (Fischhoff, et al., 1978).

Elicitation was utilised to carry out a study of dry deposition of radioactivity, the experts were told that the deposition surface was northern European Grassland, but they were not told the length of the grass which is thought to be an important determinant of the rate of deposition (Harper, et al., 1994).

The use of expert opinion to provide probabilistic inputs was employed in the risk analysis of nuclear power generation in the particular aspects of power plant safety being investigated, that is, The Reactor Risk Reference Document (NUREG-1150) (Hora & Iman, 1989).

The elicitation and expert judgement were used to assess the performance of high-level radioactive waste repositories tagged NUREG/CR-5411 (Bonano, et al., 1989).

The expert judgement was also used in the safety assessment of the final disposal of radioactive waste. The method of formal expert elicitation was developed by the US

Nuclear Regulatory Commission during the safety studies of nuclear reactors and provided the basis for expert judgement methods used in the license application for the waste isolation pilot plant (WIPP) (Hora & Jensen, 2002).

The expert judgement was used between a manager (an expert) and an analyst (Kirkwood) to elicit the number of engines that a company would sell in a particular year (Shephard & Kirkwood, 1994). An elicitation exercise was undertaken where signs and symptoms are mapped out to obtain probabilities for Oesophageal cancer network (van der Gaag, et al., 2002).

In clinical trials, (Kadane, 1994) used elicitation of priors from experts to compare two alternative drug treatments, with the success measure being systolic blood pressure (SBP) after open-heart surgery.

Hurd & McGarry, (2002) explored survival analysis by eliciting people such as the following:

"Using any number from 0 to 10 where 0 equals *absolutely no chance* and 10 equals *absolutely certain*, what do you think are the chances you will live to be 75 or more?"

They concluded subjective survival analysis did have some predictive power.

In 2010, Johnson, et al., (2010) used "bins and chips" method to elicit the probability of being alive at 3 years for an average group of newly diagnosed SSc-PAH patients not treated with Warafin medication.

Zapata-Vazquez, et al.,(2012) elicited the efficacy of a new antibiotic in patients who are hospitalised in the Paediatric ICU (PICU) and who are severely infected by pneumococci in order to know the proportion of patients that would survive in good condition, to have a sequel or die.

A combination of data- and knowledge- driven approach was used to elicit causal structure and estimates of probabilities to construct a model on student retention (Dunn, 2016).

In clinical and pharmaceutical research, (Best, et al., 2020) and (Alhussain & Oakley, 2020a) used the prior elicitation approach to obtain estimates of probabilities about variance from domain experts.

In business research, (Ming, et al., 2016) utilised the prior elicitation approach for predicting Radar System based on new Dirichlet prior distribution.

In decision sciences, (Dias, et al., 2018) as well as (Hanea, et al., 2021) used expert knowledge elicitation to get estimates of probabilities from domain experts.

Prior elicitation approach was used to obtain estimates of probabilities from domain experts in Bayesian model (Stefan, et al., 2020).

In clinical research, (Azzolina, et al., 2021) utilised expert knowledge elicitation to get estimates of probabilities from domain experts for clinical trial design and analysis.

2.2 Student Retention

Student retention is the measure of the rate at which students re-enrol from one academic session to the following one (Caruth, 2018) and this assertion was also posited by (Haverila, 2020). It is the fundamental mission of a university to retain students and advance them towards successful graduation. In efforts to attain this mission, a university must be able to recognise and understand all the factors impacting student retention and success (Bytheway & Venter, 2014). Also, (Mahmoud & Zohair, 2019) noted that in most higher education, predicting students' retention has become one of the pressing challenges. This is critical to assist at-risk students and ensure their retention, as well as to provide excellent learning resources and experience and also to improve the university's ranking and reputation. Hence stakeholders in the education system are faced with difficult choices. Universities need to practise student retention as it ensures support systems are in place to enable students to remain at university and succeed (Caruth, 2018). Student retention improves graduate rates, decreases loss of tuition revenue from students that either drop out or transfer to another institution, and brings reputation to an institution as well. It helps students by improving their lives and that of their families, as well as affording them the opportunity to make a positive contribution to their local community (Sani, et al., 2020).

If the above outcomes were not obtained, it may result in a waste of government and private of funds as well as educational time. According to (Kapur, 2018) and (Oragwu, 2020), the entire number of years spent by students who are unable to complete their degrees programmes is referred to as educational wastage. Predicting students' retention may aid both educators and students in identifying and improving students deficiencies.

Students' inability to complete their degree programmes represents a serious waste in education because limited resources that could have been allocated to other profitable sectors are spent on education that does not yield the desired outcomes in terms of graduation (Silas, 2003). This assertion was also posited by (Ochuba, 2005) as well as (Odekunle, 2007).

Forecasting student retention would help educational institution administrators discover students who are at risk thereby lowering dropout rates and ensure on-time graduation.

In most educational institutes, predicting students' retention has become a dire need. This is critical to help at-risk students and ensure retention, to provide excellent learning resources and experience, as wqll as to improve the university's ranking and reputation (Mahmoud & Zohair, 2019).

A range of students' background characteristics as well as a range of characteristics associated with students' study: age; gender; socio-economic class/parental education level; mode of study (part-time/full-time); pre-Higher Educational country of domicile; Entry examination points obtained; distance travelled from journey time; attendance; health; religion and home sickness are some of the factors used in predicting student retention (Tinto, 2006); (Tinto, 1993).

According to (Tinto, 1993), a university is fundamentally responsible for student retention. In addition, failure to offer effective institutional environments, which promote student engagements in university activities as well as comprehensively understand student characteristics in addition to their cultural backgrounds, may bring about high dropout rates or transfer to another institution, which can actually work against the primary role of a university. Furthermore, (Tinto, 1993) states that for the objectives of a university to be realised and put into perspective, there is a high need for a university to understand the relationship between student characteristics including cultural settings, student engagements, and learning outcomes (Kapur, 2018).

Some of the factors that impact on student retention are academic failure, financial reasons, health reasons, personal problem, course choice, university choice, home sickness, peer relationships, staff-student relationship (Tinto, 1975) as well as (Tinto, 1993), and (Kapur, 2018).

Another factor that impacts student retention is the increase in tuition and living costs which have made it difficult for students and parents to pay for their university, hence some students have to abandon their studies (Kapur, 2018).

The method of retaining students being adopted by the University is the creation of a number of new programmes to keep students engaged in their classes and involved in campus. This includes welcome week, campus funded tutoring, freshman seminar courses, intramural sports (Astin, 1985)

When a student participates, he/she forms both social and emotional ties to the university that both encourage the student to do well academically and prevents the student from leaving the school before graduating (Tinto, 1975) and (Tinto, 2006). When a student likes his/her course as well as his university such a student is more likely to complete his/her study without having any cause to abandon his/her course due to home sickness.

Intelligent planning systems that can predict student performance using the Neuro-Fuzzy was created by (Saxena & Singh, 2012). The variables that were used are Cumulative Point Average (CPA) and Grade Point Average (GPA). The result showed that the Neuro-Fuzzy performed better than the Fuzzy Systems using the same instances. The drawback of this study is that other factors that affect student academic performance were not considered.

Student performance was predicted by (Singh & Kaur, 2016). The important finding of this study was that J48 algorithm gives more accurate result than REP Tree (Reduced Error Pruning Tree) algorithm for student performance prediction. The overall accuracy of J48 algorithm is 67.37% and for REP Tree algorithm is 56.78%.

The prediction of Malaysian student' academic performance was studied by (Ahmad, et al., 2015) using Grade Point Average (GPA), race, and gender as factors. The important finding of this study was that the Rule Based had the maximum accuracy of 71.3%.

Gender, first language, locality, and previous semester grades were used as variables by (Saa, 2016). This study concluded that the Naïve Bayes classifier has the highest accuracy value.

Student success employing student's grades (scores in each term) in previous academic year, in every topic, and academic period was predicted by (Rojas, et al., 2021). The accuracy was 90% when using a scale of 0 to 100. The research's shortcoming is that it has a high computational and time cost.

Student retention was predicted using the Bayesian network (BN) and machine language approach alongside a combination of data- and knowledge- driven approaches to elicit causal structure and estimates of probabilities to construct a model on student retention (Dunn, 2016).

(Riyadi, 2020) described how to forecast student graduation using Fuzzy Soft Set Theory where he used variables including 1st GPA, 2nd GPA, NE score and 3rd GPA. According to findings, the accuracy of this approach was up to 89.32%. The work's weakness is that no evaluation of existing methodologies was done to compare the results and choose the best to enhance accuracy. In addition, most elements that affect student performance were not taken into consideration.

Studies have shown the following: When staff-student relationship is good, a student is encouraged to complete his/her study and graduate. In the same vein, when peer relationship is good a student relates with his/her colleagues and succeed (Tinto, 1975) as well as (Tinto, 1993) ; (Tinto, 2006). In this case, it is more likely that a student will try to transfer to another school if he/she does not enjoy his/her present school or course as a result of having made the wrong choice of college or failing to settle at university and consequently have to move to another college which he or she prefers more (Dissanayake, 2016).

When a student enters a university with a minimum qualification of advanced level (A/L) such a student is more likely to perform academically better than a student that enters with GCE ordinary level (GCSE). Hence, a student with a previous academic qualification that is higher than the latter is more likely to complete his/her programme of study (Yorke, 1999) ; (Willmot & Lloyd, 2015).

Similarly, a student who studies on full-time basis is more likely to perform academically better than a student who studies on part-time basis. Since a full-time student would devote more time to his/her study and concentrate on it as well, he is more likely to complete his/her programme of study than a student who studies on part-time basis (Yorke, 1999).

When a student develops academic skills such as critical thinking, reflection, verbal, and written communications the student performs well academically and prevents the student from leaving the school before graduating (Kift *et al.* 2010a; Nelson *et al.* 2012; Loes *et al.*, 2012).

If a student is confronted with financial problems or concerns due to costly education or low education funding, he/she will find it difficult to pay tuition fees as well as to buy educational materials for his/her needs and this will affect his academic performance as he will not be able to concentrate with his/her studies. This situation can be arrested by receiving financial support from partner, family members and friends, as well as receiving scholarship, loan or bursary from government (Swaner & Brownell, 2008; Crosling *et al.*, 2009; (Dissanayake, 2016).

Journey time is another factor that hinders a student from performing well and consequently complete his/her study. It is advisable a student travels within one hour from his university in order to minimise cost (Dissanayake, 2016). If a student thinks that he or she is staying too far from school, in which case the student must rent an apartment closer to the university or seek university accommodation due to journey time, which could impose an additional financial stress on the student and could eventually cause the student to abandon his or her study and consequently drop out from university. This could later have an effect if the student decides to transfer to another university that is closer to his or her home. In this case, the student's present university has failed to retain him or her.

When a student is healthy, he/she will be able to attend classes, engage with his studies, participate, and consequently graduate. Health reasons could be either physical, mental or both (O'Keeffe, 2013).

When a student attends lectures regularly coupled with other factors, he/she is more likely to make substantial academic progress. Hence, it is a requirement that a student needs to attain at least 75% attendance before sitting for an examination at the end of a semester (NSUK, 2019) ; (FTU, 2020).

2.3 Bayesian Networks (BNs)

Bayesian networks (BN) are graphical structures that allow us to represent and reason about an uncertain domain" (Pearl, 1998) *;* (Korb & Nicholson, 2004) ; (Daly, et al., 2011). Bayesian network (BN) has a long history dating back to the 1970s.

A Bayesian network, as shown in Fig 2-1 below, consists of variables (nodes), directed arcs, which represent the relationships/dependencies among and between interconnected variables (nodes), and a conditional probability table (CPT), or the probability distributions that represent the estimates of the probability values provided by a domain expert(s). In a Bayesian network, the belief in each state of a node is normally updated whenever the belief in each state of any directly connected node changes (Wooldridge, 2003).

Bayesian methods provide a way to revise probabilities by incorporating new data. The Bayesian network approach allows the user to form a hypothesis (H) about the world based on the given evidence (e). The equation below depicts Bayes' theorem (Bayes & Price, 1763), in which the probability of H given e (the new data) is modelled as a function of the probability of e given H multiplied by the probability of H alone and divided by the probability of e, where e is the new data.

2.4 Bayes' Theorem

Principally, Bayesian methods provide a way to revise probabilities by incorporating new data. The Bayesian network approach allows the user to form a hypothesis H about the world based on the given evidence e. Equation 2.1 depicts Bayes' theorem (Bayes, 1763), in which the probability of H given e (the new data) is modelled as a function of the probability of e given H multiplied by the probability of H alone and divided by the probability of e.

$$p(H|e) = \frac{p(e|H)p(H)}{P(e)} \qquad \text{Equation 2-1}$$

where p(H|e) is called the posterior probability,

p(e|H) is called the likelihood of the evidence (the new data),

p(H) is called the prior probability, and

p(e) is a normalising constant (Heckerman, 1998).

For this reason, equation 2.1 is often referred to in terms of the prior, likelihood and posterior only:

$$posterior \propto likelihood \times prior \qquad \text{Equation 2-2}$$

Bayesian networks (BN) are graphical structures that allow us to represent and reason about an uncertain domain" (Pearl, 1998) *;* (Korb & Nicholson, 2004).

A Bayesian network for a set of variables X = {X1,,..., Xn} is the pair (S, P), where S is a directed acyclic graph, which we call the structure of the Bayesian network, and P is a set of local probability distributions (Chickering & Heckerman, 2000). A BN model consists of two parts namely: a qualitative component in the form of a directed acyclic graph (DAG), and a quantitative component in the form of prior and conditional probabilities of BN nodes.

A Bayesian network consists of variables (nodes), directed arcs, which represent the relationships/dependencies between a pair of variables (nodes), and a conditional probability table (CPT), or the probability distributions that represent the estimates of the probability values provided by domain expert(s). In a Bayesian network, the belief in each state of a node is normally updated whenever the belief in each state of any directly connected node changes (Wooldridge, 2003).

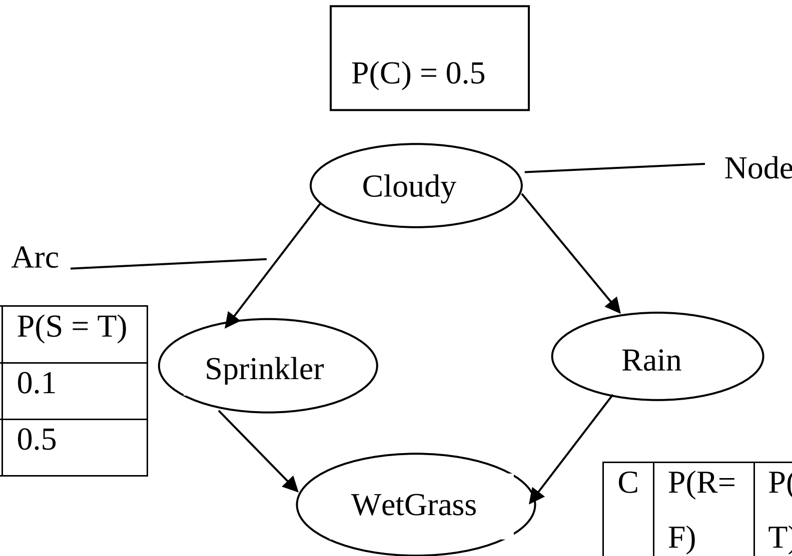
In a BN, a node X is said to be the parent of another node Y provided there is an arc from node X to node Y. Parent nodes have a direct influence on their child nodes and each child node X_i has a conditional probability distribution defined as $P(X_i \mid Parents\ (X_i))$, which quantifies the influence of the parents on the child node. If a node has no parent, then it is called a root node, and if a node has no children, then it is a leaf node. However, if a node is non-root or non-leaf, it is called an intermediate node.

Pearl (1998) noted that two random variables X and Y are conditionally independent given Z if and only if, given any value of Z, the probability distribution of X remains the same for all values of Y, and the probability distribution of Y remains the same for all values of X.

Using the conditional independence assumptions of BNs, the joint probability distribution of a set of random variables $\{X_1, X_2, X_3, \ldots, X_{n-1}, X_n\}$ can be determined using a chain rule as explained in (Pearl, 1998):

$$P(X_1, X_2, X_3, \ldots, X_{n-1}, X_n) = \prod_{i=1}^{n} P(X_i \mid Parent\,(X_i))$$

In the example below, all the nodes are binary denoted by F (false) and T (true). In a situation when rain (R = true) or the sprinkler is on (S = true), the event "grass is wet" (W = true). The conditional probability table (CPT) (Table 2-1) below depicts the strength of relationship between the two nodes. For instance, it can be seen that Pr(W = true | S = true, R = true) = 0.99, and hence, Pr(W = false | S = true, R = true) = 1 - 0.99 = 0.01 (i e., first row), thereby making each row sum to 1. The C node is a root node, its CPT depicts the prior probability, that it is cloudy, Pr(C = true) = 0.5. Assuming it is a cloudy season, it is less likely that the sprinkler is on and more likely that it is raining. The node "WetGrass" is independent of its ancestor "Cloudy" given its parents "sprinkler" and "rain". In Bayesian network, this phenomenon is referred to as conditional independence.



C	P(S = F)	P(S = T)
T	0.9	0.1
F	0.5	0.5

C	P(R= F)	P(R = T)
T	0.2	.80
F	0.80	0.20

S	R	P(W = F)	P(W = T)
T	T	0.01	0.99
T	F	0.1	0.9
F	T	0.1	0.9
F	F	1.0	0.0

Figure 2-1 Example showing a simple Bayesian network (Russel and Novig, 2010)

Bayesian networks, are considered to be appropriate for this study because of the following reasons (Geng et al., 2019 ; Wu et al., 2020):

- Bi-directional reasoning: A Bayesian network technology is capable of reasoning from cause to effect and vice versa thereby providing diagnostic and predictive reasoning. Hence the proposed system should be capable of predicting probability of students being retained the following academic session.
- A Bayesian network technology is capable of handling dynamic changes in dependency therefore the proposed student retention expert system should be capable of

permitting relationships among clinical variables to create patterns of dependence and independence which will change as evidence on the states of the variables are obtained.

- Holistic: A Bayesian network technology is holistic in nature hence the input variables of the proposed student retention expert system should produce evidence considered as a whole.
- Flexibility: A Bayesian network technology is flexible in nature therefore the proposed student retention expert system should be flexible such that evidence can be entered anywhere in the system and propagated in both cause-and-effect directions.
- The belief in each state of a node is normally updated whenever the belief in each state of any directly connected node changes (Wooldridge, 2003).
- Combination of domain experts’ knowledge and empirical data: In a situation where data is sparse, probability estimates can be elicited from the domain experts and same can be used to supplement the sparse data.

2.5 Bayesian Networks and their application areas

The ‘Expert systems’ is a branch of artificial intelligence that imitates the human mind in providing solution to a particular problem. The purpose of an expert system is to assist an individual in the performance of a task (Darlington, 2000). In this study, the Bayesian network technology is going to be used to create the student retention system, being a technique of artificial intelligence (AI).

Areas of application of Bayesian network (BN): Bayesian network technology (BN) has innumerable application areas hence it has been deployed successfully in many areas. These areas include and are not limited to student retention, enginneering, medicine and healthcare, military, ecology, site suitability, education, river pollution, and air pollution.

BN has been used for military applications (Malbasic & Duric, 2019) and, more specifically, in threat evaluation by Johansson & Falkman (2008) who developed a threat evaluation system that could handle data better than imperfect observations of humans and other systems.

Hossain, et al. (2019) applied a Bayesian network to quantify the resilience of the port infrastructure, which indicated that maintenance, alternate routing, and manpower restoration are the leading factors. Also, Hossain, et al. (2020) used a Bayesian-network based approach to assess and develop cyber resilience of a smart grid system. The research showed the efficacy of a BN to assess and enhance the overall cyber resilience of the smart grid system.

The use of Bayesian networks for ecological modelling has increased since the last two decades. It includes applications to a variety of problems including fisheries assessment (Lee, 1997; Kuikka, et al., 1999), habitat restoration (Rieman, et al., 2001). Applications to aquatic ecosystems include eutrophication in the Neuse River estuary, North Carolina (Borsuk, et al., 2004), an ecological assessment of the impacts of changed environmental conditions on native fish communities in a catchment in Victoria, Australia (Pollino, et al., 2007; Nicholson, et al., 2010). Ames, et al. (2005) used BNs to model phosphorous management in the East Canyon watershed in USA. BN was also used to assist in prioritizing river restoration options in response to changing flows and land use (Stewart-Koster, et al., 2005; Howitt, et al., 2007; Webb, et al., 2010).

Ticehurst, et al. (2007) have also used BNs in the assessment of the sustainability of eight coastal lake-catchment systems, which are located on the coast of New South Wales in Australia. They found that BNs factored all the variables in a less complicated and easier to understand way than other processes. Presenting a case study as evidence, they proposed that BNs should be used more often for ecological analysis, due to their simplicity of use and reliable output. BN was also used to assess and monitor the water quality of rivers based on macroinvertebrates see (Boets, et al., 2015; Forio, et al., 2015).

BN has been successfully used for natural language processing (NLP) of the English language, such as spell checking (Haug, et al., 2001), text categorisation and retrieval (Yang, 1994) and speech recognition (Bilmes, 2004).

Kwan et al. (2007) demonstrated the potentials of BN model in computer forensics analyses which is in accordance with the actual court verdict of guilty.

The BN has also been applied in the medical domain to diagnose various ailments. Azab et al. (1989), Radaszkiewicz et al (1992) and Valicenti et al. (1992) developed model about prognostic factors of primary gastrointestinal in non-Hodgkin Lymphoma. Lucas et al. (1998) developed a model which aimed to assist the clinician in exploring various clinical questions, among other questions concerning prognosis and optimal treatment of primary gastric non-Hodgkin Lymphoma. The model was incorporated into a computer-based system, that could be used as a decision-support system. Preliminary evaluation results indicated that the performance of the model matched the performance of experienced clinicians.

Nikovski (2000) found that Bayesian networks assisted in numerical probabilistic analysis when the information provided was incomplete or only partially correct. This occurs often when studies publish indirect statistics. Although nothing can replace the actual information, a BN is as close as it gets to filling in the blanks. The techniques were discussed in the practical contexts of designing diagnosis devices.

Zou & Conzen (2005) adapted a dynamic Bayesian network (DBN) for the prediction of the gene regulatory system from time course expression data. They presented the DBN-based approach and revealed that it had increased accuracy and reduced computational time compared to other DBN approaches, improving the process of prediction. Their approach limited potential regulators to genes with early and simultaneous expression changes, effectively allowing the BN to learn as it processed the information. This resulted in a limited number of potential regulators and a smaller search space. They also employed lag estimation to further increase the accuracy of predicting gene regulatory networks. Their results demonstrate that the approach can predict regulatory networks with greater accuracy and less computational time.

Velikova, et al. (2014) also applied BN in healthcare. They noted that although it is still difficult to bridge the gaps between Bayesian networks, they are still the best technology available for modelling medical problems, including personalisation of healthcare. Inputting knowledge of diseases based on the interpretation of patient data allows the BN to predict the progression of the disease. They used preeclampsia to illustrate the use of this model.

Student retention was predicted using the Bayesian network (BN) and machine language approach alongside a combination of data- and knowledge- driven approaches to elicit causal structure and estimates of probabilities to construct a model on student retention (Dunn, 2016).

Geng et al. (2019) constructed a BN model to predict the survival time for patients with advanced gallbladder carcinoma (GBC) after curative resection from the SEER database, with a high model accuracy. The prediction model supported the role of adjuvant therapy for advanced GBC patients. For patients with node-negative disease, the model estimated the survival benefit from the addition of adjuvant radiotherapy (XRT) and adjuvant chemoradiotherapy (cXRT). For patients with node-positive disease, adjuvant chemoradiotherapy is expected to improve the survival significantly.

Bradley et al. (2019) developed a prognostic BN network that makes personalised predictions of poor prognostic outcome post resection of pancreatic ductal adenocarcinoma. The model makes accurate predictions, even when data is missing. It was based on published survival analysis studies which is likely to carry a risk of bias.

Wu et al. (2020) developed a nomogram and a Bayesian network (BN) model for prediction of survival in gallbladder carcinoma (GBC) patients following surgery and compared the performance of the two models. The researchers discovered that the BN model was more accurate than a Cox regression-based nomogram for prediction of survival in GBC patients undergoing curative-intent resection.

Reijnen et al (2020) illustrated how BNs can be used for individualising clinical decision-making in oncology. The network also showed the complex interactions underlying the carcinogenetic process of endometrial cancer by its graphical representation.

2.6 Limitations of Bayesian Network

The demerits of the Bayesian network (BN) are enumerated below:

- Discretisation of continuous network: Bayesian networks can deal with continuous variables in only a limited manner (Jensen, 2001). The usual solution is to discretise the variables and build the model over the discrete domain.
- Collecting and structuring expert knowledge: It is difficult to reach an agreement on the BN structure and defining the parameters with expert opinion (Uusitalo, 2007; Pollino, et al., 2007).
- It has no support for feedback loops: Bayesian networks are acyclic, and thus do not support feedback loops that would sometimes be beneficial in environmental modelling (Jensen, 2001).
- Exponential growth: As the number of variables and states increases, the number of probability values to be derived grows exponentially (Jensen, 2001).

A BN's primary purpose is to predict the state of all the variables in the model based on whatever input information is available. This means that any number, type, and combination of variables can be used as input variables. The inputs can also be used for information purposes, permitting the comparison of actual and predicted values (Everaert, et al., 2011).

2.7 Summary and Conclusion

This chapter has described the previous work done in the areas related to this study, namely elicitation of estimates of probabilities from domain experts, student retention, and Bayesian network (BN).

The chapter has enumerated various areas where elicitation of estimates of probabilities had been used as a source of data in situations where there was little or no data.

Furthermore, it discusses student retention as a problem as well as its features. This study will use this problem as an application area for the combination of expert knowledge and Bayesian network (BN).

The application areas of Bayesian network (BN) are discussed in this chapter as an artificial intelligence technology which has many application areas that can be used to predict outcomes of events regarding uncertain quantities.

Conclusively, the combination of both elicitation of estimates of probabilities from domain experts and the use of Bayesian network (BN) technology have been and can still be deployed successfully in many applications.

Chapter 3

3.0 Primary Research 1: Elicitation of Knowledge from a Domain Expert

The Chapter aims to conduct a study on the effectiveness of different approaches to elicitation, and as well to demonstrate and reflect upon research skills in conducting face-to-face elicitation interviews for domain experts, to investigate whether there is a difference between direct elicitation approach and indirect elicitation approach with regards to troubleshooting computer network reachability and security issues. A network reachability and security model was created primarily by eliciting from a computer network expert the necessary information on the types and strength of relationships among and between the variables in computer networking domain.

This chapter also aims to conduct a study to develop techniques for motivating domain experts in order to obtain better estimates of probabilities from them in an elicitation session, that could be used later for the development of the main research. The study involved the development of impact of motivation on elicitation of estimates from domain experts prior to commencement of work on the Bayesian network (BN) for the prediction of student retention based on elicitation.

Computer network reachability and security: Diagnosing computer network reachability and security is paramount for network management such as troubleshooting. This diagnosis is based on network configuration called “Access Control Lists”. The model gives consideration to the troubleshooting of hardware such as the Network Interface Card and routers, and as well considers Internet Protocols (IPs) and mutual redistribution amongst the protocols. Due to complex hardware and software interactions in computer networking, quantifying network reachability often relies on computer network expert therefore Bayesian Networks (BNs) offer a robust and flexible method of encapsulating this expertise.

The study investigated the elicitation approach that is capable of producing better probability values, out of the direct and indirect approaches of elicitation. The participant involved in the elicitation session is an expert in the field of computer networking. The session which was one-to-one, involved interviews and interactive modelling.
The information required from the expert was a causal structure (graph), detailing cause and effect relationships at work in the system being modelled and estimates of probabilities to capture the strength of the relationships in the network and characteristic of these relationships. Lastly, the values elicited from the expert are utilised to create two BN models, one BN model for each elicitation approach.
Prior to this study, a skills audit was carried out in which the types of skills required for this work was explored, and as well compared to the existing skills. There was skills deficit and opportunities were identified, to rectify them by formulating an action plan and a timetable for doing so. The skills audit has led to the development of existing skills and acquisition of new ones such as: communication skills, face-to-face interview, critical thinking, self-reflection, taking responsibility for own research, planning, and time management. The skills enumerated above have helped in producing this Chapter as well as conducting this PhD research.

3.1 Background

Prior to commencement of work on the Bayesian network (BN) for predicting student retention based on expert elicitation, a pilot study was conducted. The main reason behind this pilot study was to develop techniques for elicitation of estimates of probabilities, creating and testing Bayesian networks (BNs) that could be used later for the development of the main research. The pilot study involved the development of a Bayesian Network for Troubleshooting Reachability and Security Issues in a Computer Network Based on Elicitation of Knowledge from an Expert. After identifying an expert, a face-to-face interview was conducted with the expert in order to obtain probability values from him. In the end, it was found out that the indirect elicitation approach is more suitable for this study because it produced better estimates of probabilities than the direct elicitation approach as well as better predictions from the created Bayesian **n**etwork (BN).

3.2 Context of investigation

The process of requesting for an unknown probability value or values from the domain expert(s) is known as elicitation of the said value(s) (O'Hagan, et al., 2006). (O'Leary, 2015) noted that probability distributions can be elicited from domain experts when limited or no observed data are available.

Values can be elicited from domain experts directly or indirectly (Kuhnert, et al., 2010). In the case of direct elicitation, an expert is asked to state his/her response. In line with this assertion, (O'Hagan, et al., 2006) noted that a typical example of a direct elicitation is "What is the probability of getting a 3 in the roll of a fair die?" or "What is the level of sales that corresponds to a 5% probability?". Kuhnert, et al. (2010) noted that an indirect elicitation requires probabilities to be inferred from preferences, such as elicitation of a frequency, category, weighting/ranking, relative measure, and verbal (linguistic) probabilities.

An elicitation session might involve using different approaches such as direct and indirect approaches in eliciting probability values from an expert in order, to identify the approach which is better than the other in terms of accuracy.

The following methods can be used to elicit probability values from domain expert or experts in an indirect elicitation session:

(i) Frequency: Frequency is a method of eliciting probabilities, and this can be converted to a proportion (Kuhnert, et al., 2010). A facilitator can elicit the frequency of an uncertain quantity of interest from the expert by asking question such as: Considering *n* students in a class, how many would you expect to graduate with first class? The facilitator will convert the frequency to a proportion, p_i , for each domain expert i, in order to form a prior.

(ii) Category: Categorical measures can be used to (for example) categorize species of animals into low, moderate, and high abundance at a particular site. Experts can be asked to consider a species of animal at a grazing land, and to specify whether they expect a low, moderate, or high number of that species, where low, moderate and high belongs to different bands. A facilitator can convert the categories provided by the expert into numbers by finding the mean and median of each category in order to form prior (Kuhnert, et al., 2010).

(iii) Weighting/Ranking: Ranking measures the correspondence between variables or criteria (Kendall, 1962), some rankings involve comparison of only two items at a time (David, 1988). A facilitator may ask the experts to express the rank for a chemical, using scale 1 to 3, and to express the weight for a criterion such as toxicity, using scale 1 to 3. The median from the experts' rankings and a corresponding interquartile range can be found in order to form a prior.

(iv) Relative measure: It can be used to ask experts whether (for example) the population of a town would increase, decrease or show no difference from the current level. The qualitative response provided by the experts will be converted to a quantitative response, e.g. increase (+1), decrease (-1), and no difference (0). In order to form a prior, the mean and standard deviation across experts can be found then ((Kuhnert, et al., 2010).

(v) Verbal expression of probabilities: This is an indirect and non-numerical method of eliciting probabilities which involves using the expressions such as "probable", "fifty-fifty", "improbable" etc. which can be interpreted as probability of about 0.85, 0.5 and 0.15, respectively (van der Gaag, et al., 1999).

The learning done in the directed study module was demonstrated to elicit probability values from an expert, while the probability values were used in creating BN networks that troubleshoot computer network reachability and security issues using the knowledge-driven approach for the direct and indirect elicitation methods.

Computer network reachability is the probability of data being transported from one point (source) to reach another point (destination) within a computer network. The Access Control Lists on routers are commonly used to limit reachability for security or privacy purposes (Liu & Khakpour, 2013). Teixeira & Rexford (2004) proposed a solution where an Omni server continuously maintains a view of routing changes in its own network, without requiring additional support from the underlying routers. Thereafter, they describe how to query the measurement servers along the forwarding path from the source to the destination to uncover the location and the reason for the routing change. Xie, et al. (2005) formulated the challenging problem of reachability of an Internet Protocol network, and the effects that packet filters, routing policy, and packet transformations have on the network's reachability. Feamster & Balakrishnan (2004) intended to explore how the Routing Control Platform (RCP) could improve routing efficiency such that RCP could make routing more efficient by aggregating prefixes for a particular router's forwarding table.

Liu & Khakpour (2013) verified the reachability between any two points in an interconnected and access-controlled network and formulated network reachability which considered the differences in connectionless and connection-oriented transport protocols, as well as the presence and absence of various packet transformers.

In a situation, where there are reachability issues, one or combination of the factors below must have been responsible for the issues. The factors that typically impact computer network reachability and security are: "Access Control Lists", "Network not announced", "Interface down", "Wrong configuration", and "Routing loop".

Access Control Lists or Filtering: An ACL controls the traffic into and out of a network. It controls whether a router forwards or drops packets (data) based on information found in the packet header. The Access Control Lists on routers are commonly used to limit reachability for security or privacy purposes (Liu & Khakpour, 2013).

Network not announced: This is a situation where wireless Access Points are configured not to announce their SSID (their wireless network name). This feature is enabled with the goal of preventing unauthorized users from being able to detect the wireless network from their wireless clients.

Interface down: An interface may fail to work properly due to a wrong configuration. Sometimes when an internet connection does not work properly, this may be due to the problems with the Ethernet interfaces.

Wrong configuration: Network configuration is the process of setting a network's controls, flow and operation to support the network communication of a network. This broad term incorporates multiple configuration and setup processes on network hardware, software and other supporting devices and components.

Routing loop: A routing loop is a type of network failure in which packets (data) continue to be routed in an endless circle rather than reaching their intended destination(s).

This study makes use of the principle and functions of the above features in choosing and using them as the variables for reachability and security issues domain**.**

3.3 Research question

Specifically, the focal research question for this study is:

"Effectiveness of different approaches to elicitation"

The research question tends to compare the effectiveness of direct elicitation approach to indirect elicitation approach.

3.4 Methods of investigation

The research approach can be classified into two main categories: qualitative or quantitative. A qualitative approach is considered subjective since the researchers are involved in the experiment themselves (Creswell, 2014). Conversely, a quantitative approach is objective, as it typically deals with numbers, statistical data, or statistical analysis (Creswell, 2014). This research used the hybrid method (qualitative and quantitative approaches) as it aimed to elicit the causal structure of the variables in the domain (qualitative in nature) as well the probability values of their relationships (quantitative in nature).

Prior to the elicitation session, an expert was identified and also asked to permit the researcher to extract knowledge from him on computer network reachability and security issues. Also, an arrangement was made for a venue where the sessions took place. Furthermore, all necessary software were installed for the elicitation session on a computer and the forms to be used for the elicitation session were designed. The research materials that were used for the study included textbooks, journals and conference articles, and flip papers.

The elicitation session employed face-to-face interviews to elicit information from an expert in computer networking. The technique involves the facilitator (interviewer) asking the expert (interviewee) to identify the variables in computer network reachability domain, and thereafter to provide some probability values, while the facilitator recorded the probability values that were provided by the expert, with the aid of a computer and a Bayesian network software. This technique consumed much time because of the iterative nature of the elicitation meetings which involves adjusting the estimates provided by the expert in order to improve the predictions from the network.

Construction of a causal structure: The variables in the domain were elicited, as well the states of each variable from the domain expert.by asking him to list the factors that typically cause network reachability issues, the states of each of the factors, and the relationship between the variables. The responses from the expert were used to create a causal structure for the Bayesian network by adding the variables (nodes) to the BN graphical interface. Thereafter, the researcher added arcs (links) from the parent nodes to the respective child nodes in order to establish causal relationships between the child node and its parents as shown in Figure 3-1 below.

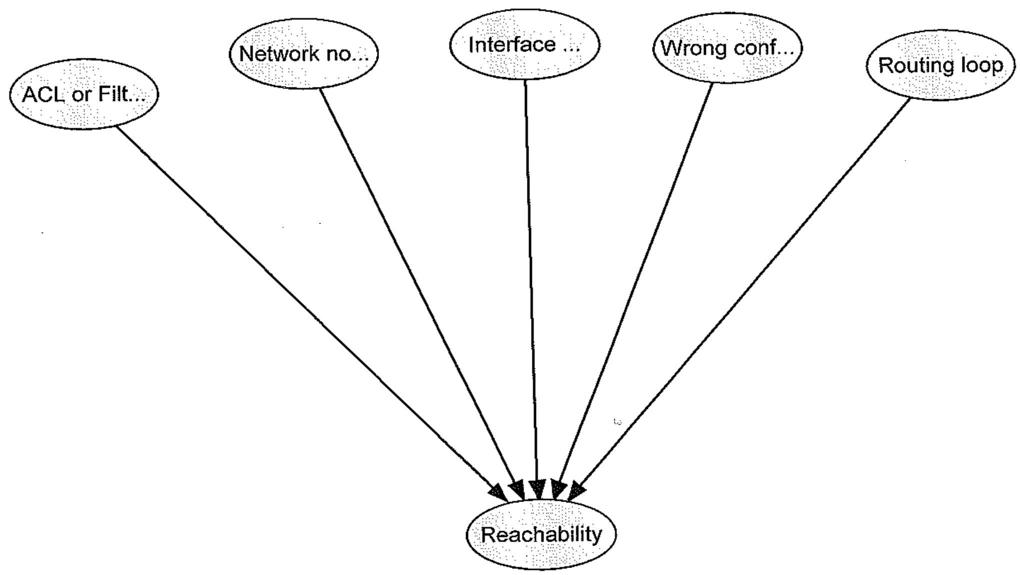


Figure 3-1 Causal structure for network reachability

3.4.1 Encoding process

The encoding process involved two experiments: one experiment on the direct elicitation method, and the other on the indirect method of eliciting probability values from a domain expert.

3.4.2 Direct elicitation method

In the direct elicitation method, probability values were elicited from a single expert by asking him to directly provide probability values for the states of each root node, by asking the expert questions, such as: What is the probability value for each state, namely; "on", "off", "down", "true", "false", etc. The probability values were recorded provided for each state with the aid of a computer.

Thereafter, the conditional probability of the child node (that is, *Reachability*) were elicited given the combinations of the states of its parents (namely; *ACL, Network not announced, Interface down, Wrong configuration, and Routing loop*) by asking the expert questions such as: What is the probability of *Reachability* being true, given the combination of the states of all of its' parents? What is the probability of *Reachability* being false, given the combination of the states of all of its' parents?

The examples of the questions asked from the expert are; What is the probability of reachability being true, given "*routing loop* "is *on*, "*Interface down*" is true, "*Network not announced*" is true, "*Wrong configuration"* is true, and "*Access Control List"* is on? What is the probability of reachability being true, given "*routing loop*" is *off*, "*Interface down*" is false, "*Network not announced*" is true, "*Wrong configuration"* is true, and "*Access Control List"* is on? etc. The researcher made sure that the values for the conditional probability provided for each of the questions were recorded, with the aid of a computer, Bayesian network (BN) software and elicitation record form.

On occasions when the values provided by the expert did not conform to the laws and theorems of probability calculus, namely the total probability, additivity and multiplicative laws of probability calculus (O'Hagan, et al., 2006), the researcher prompted the expert by asking him to adjust his probability values to conform with these laws and theorems.

Thereafter, the researcher constructed a Bayesian network from the values provided by the domain expert and compiled the network. The predictions from the Bayesian network were shown to the expert in order to get feedback from him. Initially, the predictions from the Bayesian network did not adequately represent his judgements hence he had to adjust his probability values until the predictions adequately represented his judgements.

As the facilitator, the researcher recorded the variables whose distributions were elicited, the time that the elicitation started, the estimates provided by the expert, feedback, and the time the elicitation ended. Furthermore, the researcher managed the interview with the expert, and checked the expert against cognitive biases such as; anchoring and adjustment, availability, and representativeness since these factors typically affect the accuracy of an encoding process. The researcher checked the expert against anchoring-and-adjustment heuristic, by asking him to provide the minimum (lower bound) and maximum (upper bound) values before he provides the median value. Also, the expert was asked to avoid basing his probability values on recent occurrence, and as well not to provide probability values that are based on situations that are wrongly perceived to be similar in order to avoid heuristics due to availability and representativeness, respectively (Tversky & Kahneman, 1973; Tversky & Kahneman, 1974; Kahneman, 2011).

3.4.3 Indirect elicitation method

The same causal structure was used for both the direct and the indirect elicitation methods. In the indirect elicitation method, the probability values for the root nodes and the child node were elicited from the expert by using the probability scale and verbal (linguistic) probability expressions method which served as an aid to the expert in providing his estimates (van der Gaag, et al., 1999). Furthermore, the verbal expressions were converted to probability values, and as well assisted the expert by asking him to adjust his probability value when the value provided by him was too small or too much (Martin, et al., 2011).

Thereafter, another Bayesian network was constructed from the assessments provided by the domain expert, under the indirect elicitation method. The researcher compiled and showed the predictions from the second Bayesian network to the expert in order to get a feedback from him. Initially, the predictions from the Bayesian network did not adequately represent his judgements hence he had to adjust his probability values until the predictions adequately represented his judgements.

As the facilitator, we recorded the variables whose distributions were elicited, the time that the elicitation started, the estimates provided by the expert, and the time the elicitation ended. Furthermore, we managed the interview with the expert, and we checked the expert against cognitive biases such as anchoring and adjustment, availability and representativeness since all of these factors typically affect the accuracy of the elicitation exercise (Tversky & Kahneman, 1973; Tversky & Kahneman, 1974; Kahneman, 2011)..

3.5 Results

The probability values provided by the domain expert during the elicitation sessions were entered into the table of a Bayesian network (BN) software before each network was compiled in order to get predictions from the two networks. The predictions from each network, for each approach can be found in Table 3-1. From the table, the results on a case-by-case basis, combining the states of all the parent nodes in the domain are as follows:

In Cases 6, 14, 18, 26, 28 and 32, the probability of reachability in each of these cases is much lower in the indirect elicitation approach than in the direct elicitation approach. The indirect elicitation approach performs better in terms of network security or privacy purposes because the lower the probability of reachability the more secured the network (Liu & Khakpour, 2013). However, Case 24 performs much better in the direct elicitation approach than in the indirect elicitation approach.

The two approaches perform equally in Cases 7, 8, 20, and 27. The probability of reachability for the direct elicitation and the indirect elicitation approach is the same in each case. The Cases also perform equally in terms of network security because the probability of reachability in each approach is the same.

In the remaining Cases, the probability of reachability for the direct elicitation and the indirect elicitation approach is almost the same, but the indirect elicitation approach performs a little bit better in most of the cases, in terms of network security because the lower the probability of reachability the more secured a network ((Liu & Khakpour, 2013).

On the overall, the indirect elicitation approach performs better than the direct elicitation approach. The network for the indirect elicitation approach performs better than the network for the direct elicitation approach in 23 out of 32 cases. While the network for the direct elicitation approach performs better than the network for the indirect elicitation approach in 5 out of 32 cases, but both perform equally in 4 out of 32 cases.

Table 3-1 Elicited probability values

Case	Routing loop	Interface Down	Network not announced	Wrong configuration	ACL	Reachability (%)	
						Direct Elicitation	Indirect Elicitation
1	On	Software error	True	True	On	9.00	4.00
2	On	Software error	True	True	Off	20.00	10.00
3	On	Software error	True	False	On	3.00	6.00
4	On	Software error	True	False	Off	5.00	3.00
5	On	Software error	False	True	On	2.00	6.00
6	On	Software error	False	True	Off	30.00	4.40
7	On	Software error	False	False	On	10.00	10.00
8	On	Software error	False	False	Off	80.00	80.00
9	On	Software error	True	True	On	9.00	8.50
10	On	Software error	True	True	Off	18.18	4.00
11	On	Software error	True	False	On	3.41	6.00
12	On	Software error	True	False	Off	5.10	3.50
13	On	Software error	False	True	On	2.00	5.80
14	On	Software error	False	True	Off	24.19	4.70
15	On	Software error	False	False	On	11.11	18.00
16	On	Software error	False	False	Off	80.00	75.00
17	Off	Hardware error	True	True	On	5.00	9.00
18	Off	Hardware error	True	True	Off	30.00	10.00

Case	Routing loop	Interface Down	Network not announced	Wrong configuration	A C L	Reachability (%)	
						Direct Elicitation	**Indirect Elicitation**
19	Off	Hardware error	True	False	On	20.00	23.00
20	Off	Hardware error	True	False	Off	40.00	41.00
21	Off	Hardware error	False	True	On	5.00	3.80
22	Off	Hardware error	False	True	Off	9.00	4.90
23	Off	Hardware error	False	False	On	20.00	18.00
24	Off	Hardware error	False	False	Off	72.00	99.50
25	Off	Hardware error	True	True	On	6.25	8.00
26	Off	Hardware error	True	True	Off	46.15	15.00
27	Off	Hardware error	True	False	On	19.05	19.00
28	Off	Hardware error	True	False	Off	66.07	45.00
29	Off	Hardware error	False	True	On	6.10	10.00
30	Off	Hardware error	False	True	Off	10.71	7.50
31	Off	Hardware error	False	False	On	16.95	5.00
32	Off	Hardware error	False	False	Off	84.71	99.50

We conducted one session for each elicitation approach and these accounted sufficiently for the variability of the domain expert judgements hence one session for each elicitation approach did not have any impacts on the probability values provided by the experts.

We changed the order of the elicitation sessions in order to detect the impact that sequence might have on the elicitation sessions, and we did not give a preferential treatment to the elicitation approach used in the second place.

We checked the expert against heuristics and biases that can mar the encoding process by checking him when his probability values are either too small or large. On such occasions, we asked him to revise his judgements or opinions, and this led to provision of better probability values. Consequently, this improves the predictions from the Bayesian network.

The following analyses were done:

(i) The duration of the elicitation, that is, the time taken to complete the elicitation in each session:

Under the direct elicitation method, it took 30 minutes to elicit the probability values for the states of the root nodes and that of the child node, that is an average of one question per minute, while it took 50 minutes under the indirect elicitation method, that is an average of one question for one and a-half minute.

(ii) Time taken to elicit a group of five states:

Under the direct elicitation method, it took 12 minutes to elicit the probability values for the states in the group, while it took 35 minutes to elicit the probability values for the states in the same group, under the indirect elicitation method.

(iii) Predictions from each network in terms of accuracy, which is how well the Bayesian network predicts.

We evaluated the Bayesian network, after developing the models' structures and obtaining the marginal and conditional probability values from the expert. We did the qualitative evaluation by showing the predictions from the networks to the expert as shown in Figure 3-2 to 3-5 on page 40 to 43 below, for some of the predictions from the Bayesian network (BN).

We showed the predictions to the domain expert who confirmed in his feedback that the predictions adequately represented his judgements.

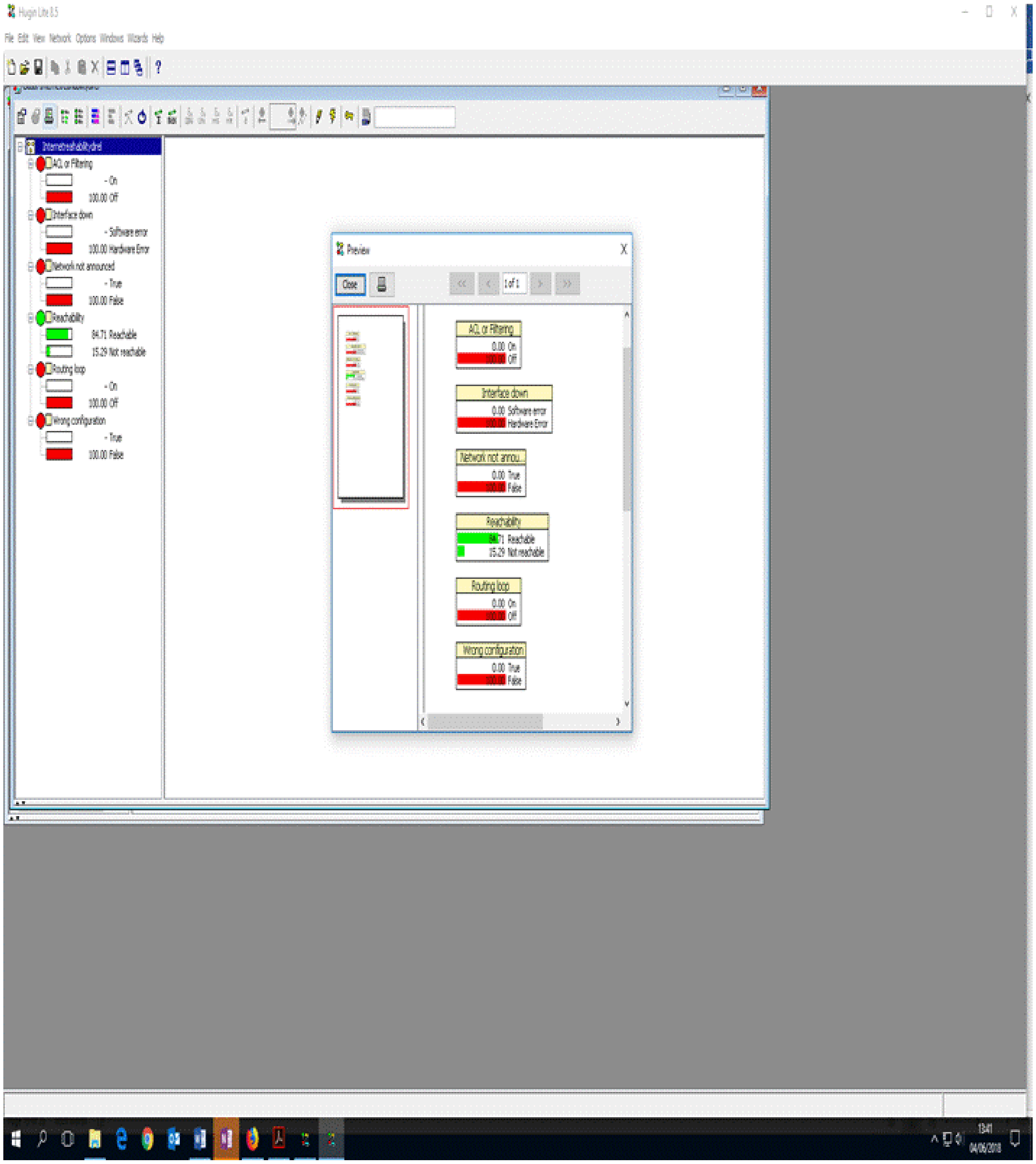


Figure 3-2 A Screenshot of prediction from the Bayesian network

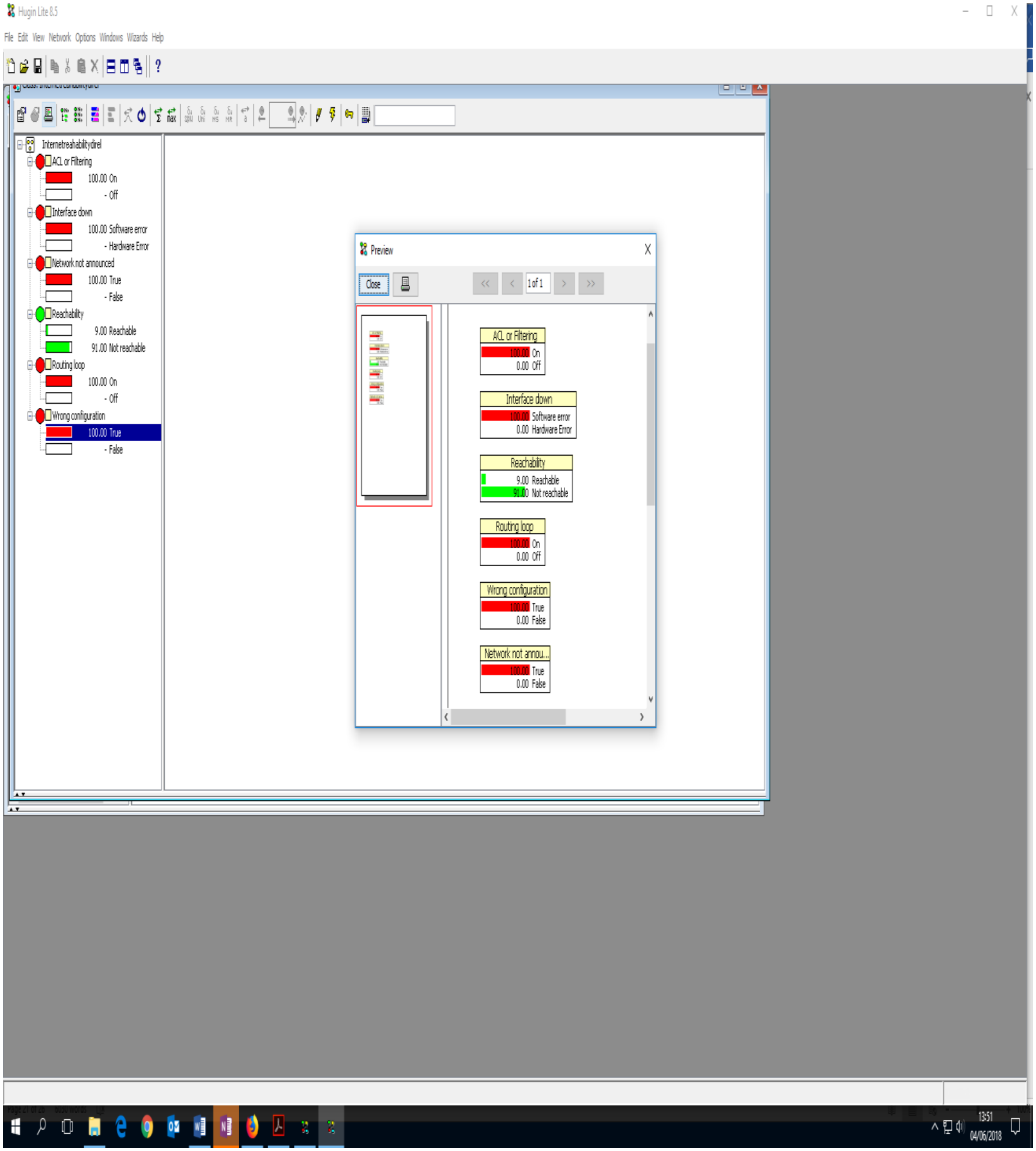


Figure 3-3 A Screenshot of prediction from the Bayesian network

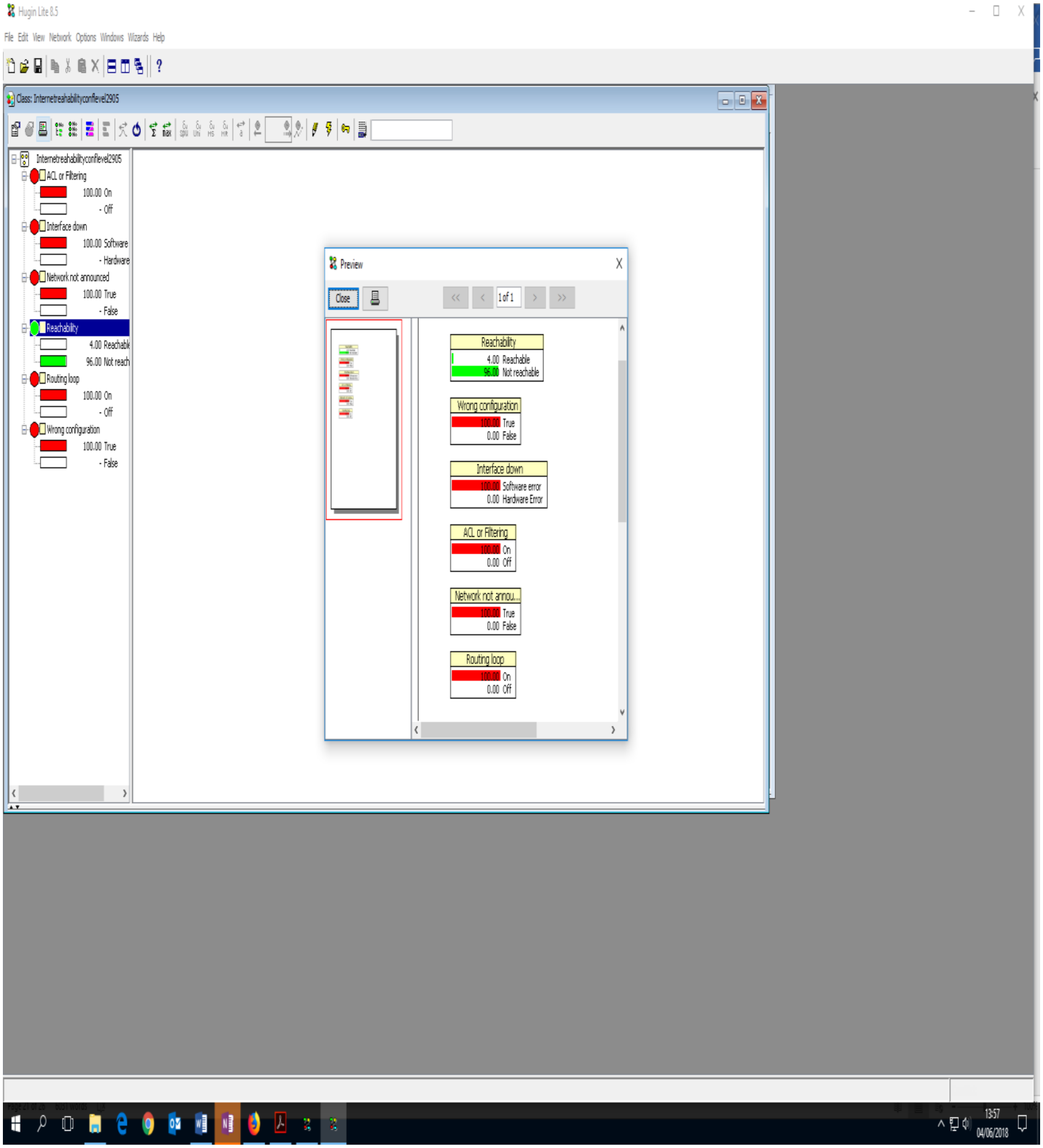


Figure 3-4 A Screenshot of prediction from the Bayesian network

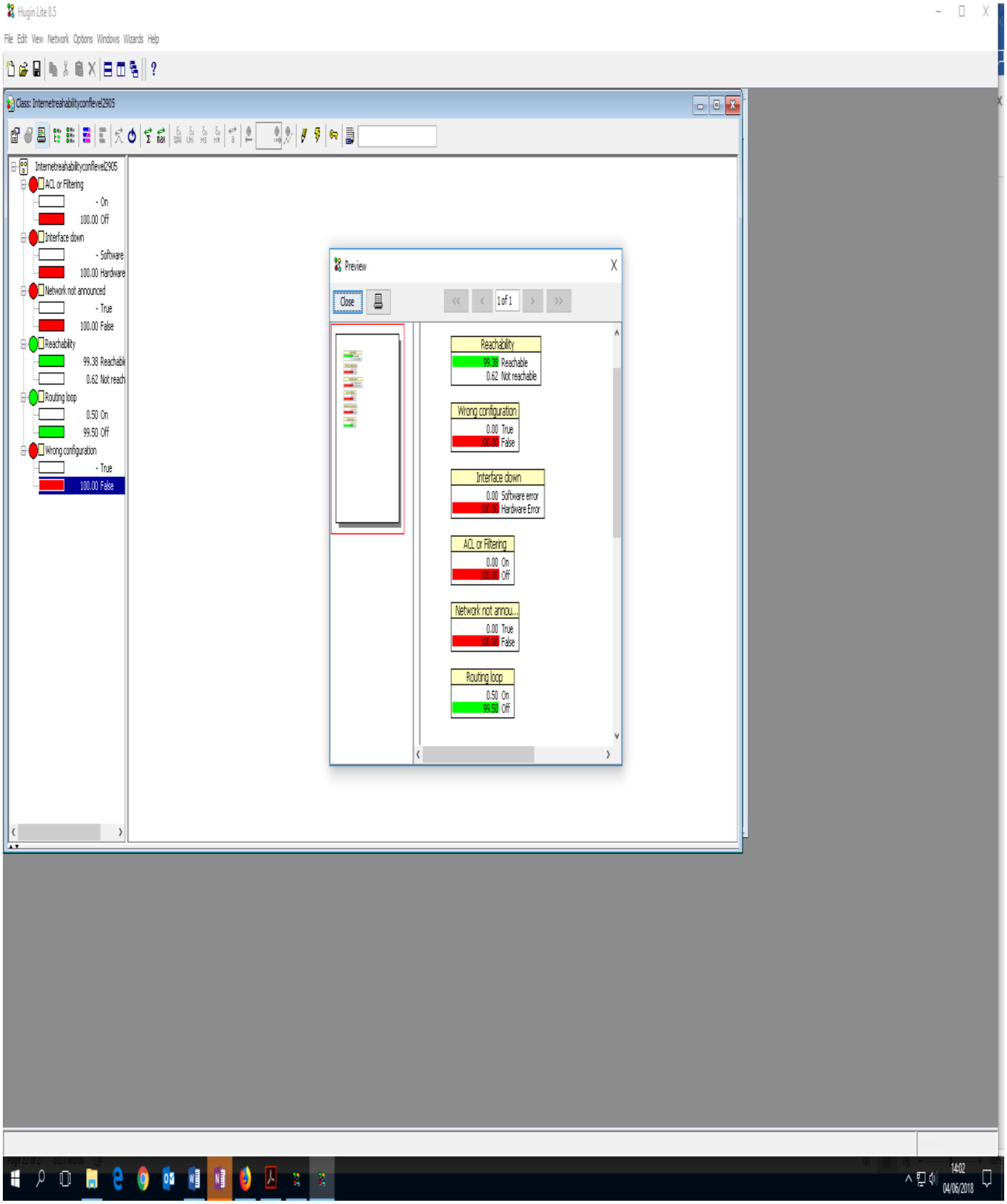


Figure 3-5 A Screenshot of prediction from the Bayesian network

3.6 Primary Research 2: Impact of motivation on elicitation of estimates of probabilities

This chapter discusses methodological issues. Then the research instrument, that is, two sets of questionnaires (see Appendix 1 and 2) that were used for the experiments. The study adopted the within-subjects experimental approach. The study is then described in terms of unmotivated and motivated conditions.

3.7 Motivation

Motivation is the reason for one's actions, desires, and needs. Motivation is also one's direction to behaviour, or what causes a person to repeat a behaviour. In the context of motivated reasoning, (Kunda, 1990; Knol et al, 2010; Montibeller & Winterfeldt, 2015) defined motivation as any wish, desire, or preference that concerns the outcome of judgements and decision making. Motivation is applicable to reasoning tasks in which people are to be accurate in their judgements and decisions as well as when they are to arrive at particular directed judgements and decisions.

Kruglanski and Klar (1987) noted that motivated reasoning phenomena can be divided into two major categories, namely; those in which the motive is to arrive at an accurate conclusion, and those in which the motive is to arrive at a particular directional conclusion. In a situation where people are to be accurate, they should utilize more cognitive effort, process information thoroughly and apply appropriate rules, however directional goals makes use of the thinking strategy that will produce the particular directional result (Kunda, 1990). This study falls in the category in which the motive is to arrive at accurate conclusions by obtaining accurate estimates of probabilities from the domain experts in an elicitation session.

Basically, motivation can be classified into two types: intrinsic or internal motivation, and extrinsic or external motivation. Ryan and Deci (2000) noted that whenever an individual performs a certain reasoning task without being affected by some external inducement such as rewards, such form of motivation is called intrinsic or internal motivation. Conversely, if an individual performs a certain reasoning task because the individual is affected by some external inducement such as rewards or punishments, such form of motivation is called extrinsic or external motivation (Ryan and Deci, 2000).

This study investigated the impact of motivation on elicitation of estimates of probabilities from participants in an elicitation session, using “Adults’ Internet Usage in the UK” as a domain. The participants were asked to state estimates of probabilities for variables such as; age, gender, social-economic class, online activities, and devices used by adults to access the Internet. In order to obtain accurate estimates of probabilities from participants in an elicitation session, a facilitator needs to motivate the participants. In this study, estimates of probabilities were elicited from participants in two experiments in order to investigate the impact of motivation on elicitation of estimates of probabilities. The participants were not motivated in the first experiment, while they were in the second. In the second experiment, the motivation made the experts to think slowly and hard, and thereby avoided the effects of heuristics and biases that typically lead to underestimation or overestimation of estimates of probabilities. It was found out that the experts performed better in the second experiment than in the first. The improvement in performance was as a result of incentive that was attached to the second elicitation session.

In the literature, some researchers have examined the impact of motivation and found out that it makes people to accomplish a task in order to get a financial or non-financial reward attached to such task. For instance:

Buehler et al. (1997) examined the impact of motivation on predicted and actual times to complete a task and found out that the difference between predicted and actual completion times was greater for participants expecting a tax refund than those not expecting a refund. The authors found out that motivation could increase the speed at which people expected to complete a task.

In another study, Buehler et al. (1997) drew a more direct link between financial incentives and the speed to accomplish a task such as solving word puzzles quickly, which was done as a result of manipulation using monetary incentives for speed. The authors found out that the financial incentive that was involved, caused the participants to solve the word puzzles quickly.

Bower et al. (2002) noted that here are three main types of financial incentives used on construction contracts, namely; share of savings incentives, where cost savings are shared between the client and the contractor based on an agreed formula; schedule incentives, where a premium is offered to the contractor for the early completion of the project; and technical performance bonuses for meeting performance targets, other than cost and schedule. A performance bonus arrangement can be applied to a wide range of performance areas such as quality and functionality. The authors found out that financial incentives can enhance the positive impact of performance-enhancing initiatives.

Timothy and Manley (2011) noted that financial incentives aim to increase the efficiency and effectiveness of projects by stimulating the motivation to work harder and smarter in pursuit of such goals.

Maheady et al. (2006) set up an experiment in which they gave sixth grade science students a science test without giving them any incentives. Thereafter, they gave them another science test in which a financial incentive was attached. In the end, the experimenters found out that the students performed better in the second test because of the attachment of a financial incentive which caused the class means to rise. In the above studies, the authors found out that the financial incentive that was involved, enhanced the performance of the participants.

Chen (2009) conducted a study wherein partial course credit was awarded to undergraduate students in exchange for taking part in the study “Cognition and Social Judgement”. The experimenter discovered that the motivation increased the students’ desire for accuracy in the experiment.

Gardner (2008) set up an experiment in which groups of undergraduate students were asked to look at a list of courses and choose the ones that they might like to take in future. One group did this with no further motivational information. A second group got further motivational information from students who had taken the courses previously. The researcher found out that the second group made the right choices as a result of the motivational information they received from the students who had taken the courses previously.

Liu et al. (2012) used motivational instruction in an experiment in which he reminded one group of students in a college that their performances in a test may affect the ranking of their institution and therefore affects the value of their diploma while he did not give the other group any instructions at all. The experimenters found out that the group of participants that received the motivational instruction/information performed better than the other group.

Aczel et al. (2015) set up an experiment in which one group of students were given an elicitation training before an elicitation session while the other group was not given. In the end, the experimenters discovered that the group that received the elicitation training provided better estimates of probabilities as a result of the knowledge they gained from the training they had previously received on estimates of probabilities.

Shirley and Smidts (2018) performed an experiment on the factors that motivate an operator's actions in a nuclear power plant environment and evaluated unbiased elicitation, perfect bias and operator expectation.

Boiney et al (1997) performed an experiment on a sales forecasting exercise in which they asked participants to play the role of a marketing manager of a company. The participants were motivated by telling them that their forecasts could critically influence the company's decisions to introduce its products, hence the participants should take the tasks very seriously. In the end the forecasts provided by three experts were clustered, while the fourth expert's forecast was the highest of all the four forecasts. The experimenter found out that the other three participants were more accurate in their forecasts, while the fourth participant's forecast was done in order to influence the decision of the company to introduce its products.

Method: The aim of this study is to obtain estimates of probabilities from participants on Internet usage in the UK. In preparation for this study, the literature on elicitation of estimates of probabilities and also on motivation had been searched. Furthermore, twenty questions were drawn and randomly split into two equal parts, resulting in two fill-in questionnaires. The participants need to provide their estimates of probabilities for each question on the questionnaire, during each elicitation session. Furthermore, the experimenter had completed and submitted a proportionate ethics form which has been approved by the Ethics Committee of the University.

The students in the School of Computing and Digital Technologies, Staffordshire University, UK have been identified as participants for the elicitation sessions. The students were chosen to participate in the session because of their existing knowledge in Internet usage in the UK, the various activities on the internet, and the devices that can be used to access the internet.

The estimates of probabilities on Internet usage in the UK were elicited from the participants, based on the variables such as gender, age group, device, activities, and socio-economic class. The participants were asked to provide their estimates in form of percentages in order to make the estimations easier for them. Below are the two experiments to assess the impact of motivation on elicitation of estimates of probabilities.

The aim of the two experiments is to assess the impact of motivation on elicitation of estimates of probabilities from the students who served as participants.

This study consists of two experiments, namely experiment 1 and experiment 2, respectively. The first experiment aimed to assess the individual performances of the participants without attaching any incentives to their participation in the first elicitation session, while the second experiment aimed to assess the individual performances of the participants after attaching an incentive to their participation in the second elicitation session. The same set of participants were used for the two experiments and as well exposed to both unmotivated and motivated conditions hence this study adopted the within-subjects experimental design.

Finally, the outcomes of both experiments were compared in order to assess the impact of motivation on the estimates of probabilities provided by the participants in the first and the second elicitation session.

3.7.1 Experiment 1

Ninety-one undergraduate students at Staffordshire University were recruited on voluntary basis as participants for the experiment, but thirty participants eventually took part in this experiment (Ritchie et al, 2003).

Design and Procedures

The experiment took place in a lecture theatre in the university where the experimenter addressed the participants before the commencement of the experiment. The researcher introduced himself to the participants and also informed them that he wanted to use the study for transfer of his registration from that of a MPhil to a PhD status. Furthermore, the researcher informed the participants that the study would consist two experiments which would involve elicitation of estimates of probabilities on Internet usage in the UK, in two elicitation sessions.

In order to control the experiment, it took an in-class format whereby the participants provided their estimates at the same venue, duration of time, atmospheric condition and in the presence of the experimenter. Every participant was given an Information Form (See Appendix 3) and a Consent Form (Appendix 4) to read, understand, and endorse before the commencement of the experiment. Thereafter, the first questionnaire (Questionnaire 1, see Appendix 1), that was designed for the Internet usage in the UK, was distributed to the participants in order for them to provide their estimates for the variables, in the spaces provided in the questionnaire. Being a predominantly survey research method, a questionnaire is a set of standardised questions for obtaining responses from a large group of individuals and is very often used to provide quantitative data (Burton and Bartlett, 2016). The completed questionnaires were collected from the participants after a duration of fifteen minutes which was allocated to the elicitation session.

At the end of the experiment, the experimenter showed his appreciation to the participants by thanking them for participating in the experiment as well for sparing their invaluable time. The experimenter also informed them that the second experiment would take place the following week, at the same venue and at the same time. He also instructed them that only those who participated in the first experiment would be eligible to participate in the second experiment.

Thereafter, the experimenter compared the estimates of probabilities provided by the participants to a survey data on Internet usage in the UK, in order to calculate the error rates in the estimates of probabilities provided by the participants in the elicitation session. An error rate in estimation of probability is the absolute difference between the estimated and the actual value. The error rates in estimates of probabilities provided by each participant were calculated using the formula below:

ABS(estimated value – actual value)

Where ABS = Absolute value of the difference between the estimated value and the actual value.

The error rates in estimates of probabilities were used to determine the accuracy of the estimates of probabilities provided by the participants in the elicitation session. This was based on the fact that the smaller the error rate is, the closer the estimated value is to the actual value. Thereafter, the average error rate of each participant in the study was calculated in order to find the performance of each participant in the study. The error rates and the average error rates in the estimates of probabilities provided by the participants served as input data that was analysed later on.

3.7.2 Experiment 2

The thirty undergraduate students at Staffordshire University who eventually participated in the first experiment are also the set of participants for the second experiment.

The experiment took place at the same venue where Experiment 1 was conducted, also the format for this experiment is not different from that of Experiment 1. But the procedure for this experiment differs because of the attachment of an incentive to the elicitation of estimates of probabilities from the participants. This is in order to motivate them to provide better estimates of probabilities.

Before this elicitation session, the experimenter informed the participants that each of the best three performers will be presented with a Love2Shop motivating voucher, with which they can buy items from high street shops such as Wilko, Shoezone, Iceland, Argos, etc. Furthermore, in order to win a Love2Shop voucher, the experimenter advised the participants to read and understand every question very well, and to think slowly and hard, before they provide an estimate for each of the variables in the questionnaire.

Thereafter, the second questionnaire (Questionnaire 2, see Appendix 2) that was designed for the Internet usage in the UK, was distributed to the participants in order to provide their estimates for the variables, in the spaces provided in the questionnaire. The completed questionnaires were collected from the participants after the expiration of fifteen minutes which was allocated to the elicitation session.

At the end of the experiment, the experimenter showed his appreciation to the participants by thanking them for participating in the first experiment and in the second experiment as well for sparing their invaluable time on each occasion. The experimenter also informed them that the presentation of Love2Shop motivating vouchers to the best three performers in the second elicitation session would take place shortly after the second experiment.

Thereafter, the experimenter compared the estimates of probabilities provided by the participants to a survey data on Internet usage in the UK. The error rates in each variable and the average error rate were calculated as previously done at the end of first experiment.

The top three performers in the second elicitation session were presented with a Love2Shop voucher each, which is worth of £25.00, £20.00 and £15.00, respectively. The participants were rewarded according to individual performances in accomplishing the task.

3.8 Ethics

The research complied with the Staffordshire University's and British Computer Society

(BCS) ethical rules guiding the conduct of research (see Appendix 5). The reasons for the ethical compliance are to ensure transparency, confidentiality, and to protect the participants (Creswell, 2014). The research adopts completion of two fill-in questionnaires designed for this study (See Appendix 1 and 2), where the estimates of probabilities are entered by the participants who do not fall under any category of people considered vulnerable.

3.9 Data Analysis and Results

This section aims to analyse the estimates of probabilities provided by the participants in the first and second elicitation session. The error rate serves as input data that was analysed. The data were entered into MS-Excel and analysed in Statistical Package for Social Sciences (SPSS) software. The following analyses were performed on the data: descriptive, paired samples t-test and Spearman ranking to analyse the quantitative data from the questionnaires. The descriptive analysis describes or summarizes the characteristics of the error rates under unmotivated and motivated conditions, the paired samples t-test was used to test whether or not the participants' performances are different under unmotivated and motivated conditions, while Spearman ranking test was used to rank the performance of the participants in the two experiments.

It was mentioned previously under method, that twenty questions were drawn and randomly split into two equal parts, the experimenter had to remove some questions and participants in order to ensure that the analysis is fairer.

3.9.1 Thirty participants, ten questions:

The number of participants and questions in this data analysis is thirty and ten, respectively. The error rates under unmotivated and motivated conditions were analysed in order to find its summary, to compare the performances of the participants in the two tests, and to rank the participants in the two experiments in which all the participants provided estimates of probabilities for ten questions in each experiment.

Below is the descriptive statistics for the thirty participants that provided estimates for ten questions, see Table 3-2 below.

Table 3-2 Descriptive statistics for comparison of error rates (N = 30, Questions = 10)

		Unmotivated	Motivated
N	Valid	30	30
	Missing	0	0
Mean		10.8583	8.3167
Std. Error of Mean		1.01585	.55861
Median		8.9500	8.0000
Mode		6.60	5.20[a]
Std. Deviation		5.56406	3.05965
Variance		30.959	9.361
Range		22.40	13.00
Minimum		4.20	5.10
Maximum		26.60	18.10
Sum		325.75	249.50

a. Multiple modes exist. The smallest value is shown.

The descriptive statistics for the comparison of error rates in the estimates of probabilities provided by the participants in the two experiments is displayed in Table 3-2 above. Results in Table 3-2 show that a higher average error rate (Mean = 10.86, SD = 5.56) was obtained when the participants were not motivated by the experimenter. Conversely, a lower average error rate (Mean = 8.32, SD = 3.06) was obtained when they were motivated. Since the average error rate fell, this is an indication that better estimates of probabilities were provided by the participants when an incentive was attached to the elicitation exercise.

Below is the Paired samples t-test for the thirty participants that provided estimates for ten questions:

Table 3-3 Paired Samples Statistics (N = 30, Questions = 10)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Unmotivated	10.8583	30	5.56406	1.01585
	Motivated	8.3167	30	3.05965	.55861

Table 3-3 above displays the paired samples statistics of the unmotivated and motivated participants in terms of mean, population of participants in each group (N), standard deviation and standard error mean.

Table 3-4 Paired Samples t-Test (N = 30, Questions = 10)

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	T	Df	Sig. (2-tailed)
Pair 1	Unmotivated – Motivated	2.54167	5.41596	.98881	.51931	4.56402	2.570	29	.016

A paired-samples t-test statistic was conducted to compare the performances of the participants before and after attaching the incentive to the elicitation sessions. Results in Table 3-4 shows that Sig. (2-tailed) value is .016. Since this value is less than .05, it can be concluded that there is a statistically significant difference in performance when an incentive was attached to the second elicitation session. The table also shows that the average error rate for incentive condition was less than the average error rate for the no incentive condition, therefore we can conclude that estimates of probabilities provided under the incentive condition are significantly better than those in the no incentive. The test shows that motivation reduces heuristics and biases and thereby improves the performance of the participants in providing better estimates of probabilities (Hasannejad et al., 2017).

Below is the Spearman's rho ranking test for the thirty participants that provided estimates for ten questions: Rho is the Greek pronounciation for the Spearman rank-order correlation coefficient, which is usually denoted by the symbol r_s or the Greek letter ρ.

Table 3-5 Spearman's rho ranking test (N = 30, Questions = 10)

Correlations

			Unmotivated	Motivated
Spearman's rho	Unmotivated	Correlation Coefficient	1.000	.409*
		Sig. (2-tailed)	.	.025
		N	30	30
	Motivated	Correlation Coefficient	.409*	1.000
		Sig. (2-tailed)	.025	.
		N	30	30

*. Correlation is significant at the 0.05 level (2-tailed).

The individual performances before and after attaching the incentive to the elicitation sessions were obtained by conducting a Spearman rho ranking statistic test. Result in Table 3-5 shows that the Sig. (2-tailed) value is .025. Since this value is less than .05, it can be concluded that there is a statistically significant difference in individual performance when an incentive was attached to the second elicitation session. A value of 0.05 is important in this test because it implies that this finding has 5% chance of not being true. This test shows that motivation improves the performance of the participants. Also, one can conclude that the correlation in performance of the participants between the two elicitation sessions (0.41), is a moderately positive relationship. This implies that some participants performed well in the first test and as well in the second. Also, motivation did not work equally in the sense that it worked for some while it did not work for others.

3.9.2 Thirty participants, seven questions:

Originally, this study consisted of twenty questions which was randomly split into two equal parts, resulting in two questionnaires. In order to find out whether or not the questions in one questionnaire have the same level of difficulty with the questions in the other, the three best answered questions in each questionnaire were removed from the study. The three questions are the ones that have the lowest average errors than the remaining seven questions. The three well-answered questions were not included in the data analysis displayed in Table 3-6 below. The three-well-answered questions were removed because simple questions do not require much motivation unlike difficult ones. This step was taken in order to find out whether the variation in performance, was as result of the participants finding one test harder than the other.

Table 3-6 Descriptive statistics for comparison of error rates (N = 30, Questions = 7)

		Unmotivated	Motivated
.N	Valid	30	30
	Missing	0	0
Mean		13.5500	9.6573
Std. Error of Mean		1.27067	.65537
Median		12.1450	9.0700
Mode		13.29	6.86
Std. Deviation		6.95972	3.58962
Variance		48.438	12.885
Range		26.57	16.57
Minimum		4.57	5.29
Maximum		31.14	21.86
Sum		406.50	289.72

Results in Table 3-6 show the descriptive statistics for the comparison of error rates in the estimates of probabilities provided by the thirty participants for the remaining seven questions. A higher average error rate (Mean = 13.55, SD = 6.96) was obtained when the participants were not motivated by the experimenter. Conversely, a lower average error rate (Mean = 9.66, SD = 3.59) was noted when the participants were. Since the mean error rate fell, this is an indication that better estimates of probabilities were provided by the participants when an incentive was attached to the second elicitation session.

Table 3-7 Paired samples statistics (N = 30, Questions = 7)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Unmotivated	13.5500	30	6.95972	1.27067
	Motivated	9.6573	30	3.58962	.65537

Table 3-7 above displays the paired samples statistics of the unmotivated and motivated participants in terms of mean, population of participants in each group (N), standard deviation and standard error mean.

Table 3-8 Paired samples t-test (N = 30, Questions = 7)

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	T	Df	Sig. (2-tailed)
Pair 1	Unmotivated – Motivated	3.89267	6.94126	1.26729	1.30076	6.48458	3.072	29	.005

A paired-samples T-Test statistic was conducted to compare the performances of the participants after removing the three variables. Result in Table 3.8 shows that the Sig. (2-tailed) value is .005. Since this value is less than .05, it can be concluded that there is no statistically significant difference in the level of difficulty of the questions in the two questionnaires. This is an indication that the participants did not find one elicitation session harder than the other, but the difference in performance is due to motivation. Therefore, the questions in one questionnaire have the same level of difficulty with the questions in the other.

3.9.3 Twenty-eight participants, ten questions:

Originally, this study consisted of thirty participants, but two participants were removed from the study as a result of providing highly incorrect estimates of probabilities. The participants must have provided the estimates of probability which readily came to their mind, this could simply be because that they do not have an existing knowledge in probability estimation and thereby leading to overestimations or underestimations of estimates of probabilities. Although, we attempted to motivate all the participants, but the study reveals that some of the participants do not have an existing knowledge in assessing probabilities. Therefore, they were unable to provide better estimates of probability, which led to their exclusion from the study.

Table 3-9 Descriptive statistics for comparison of error rates (N = 28, Questions = 10)

		Unmotivated	Motivated
N	Valid	28	28
	Missing	2	2
Mean		10.3411	7.9750
Std. Error of Mean		1.01708	.52437
Median		8.5500	7.8000
Mode		6.60	5.20[a]
Std. Deviation		5.38190	2.77470
Variance		28.965	7.699
Range		22.40	13.00
Minimum		4.20	5.10
Maximum		26.60	18.10
Sum		289.55	223.30

a. Multiple modes exist. The smallest value is shown

Results in Table 3-9 show that a higher average error rate (Mean = 10.34, SD = 5.38) was obtained when the participants were not motivated by the experimenter. Conversely, a lower average error rate (Mean = 7.98, SD = 2.77) was noted when the participants were. Since the average error rate fell, this is an indication that better estimates of probabilities were provided by the participants when an incentive was attached to the second elicitation session. This is an indication that motivation has improved the performance of the participants.

Table 3-10 Paired samples statistics (N = 28, Questions = 10)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Unmotivated	10.3411	28	5.38190	1.01708
	Motivated	7.9750	28	2.77470	.52437

Table 3-10 above displays the paired samples statistics of the unmotivated and motivated participants in terms of mean, population of participants in each group (N), standard deviation and standard error mean.

Table 3-11 Paired samples t-test (N = 28, Questions = 10)

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	T	Df	Sig. (2-tailed)
Pair 1	Unmotivated – Motivated	2.36607	5.46265	1.03234	.24788	4.48427	2.292	27	.030

A paired-samples t-test statistic was conducted to compare the performance of the remaining twenty-eight participants. Result in Table 3-11 shows that the Sig. (2-tailed) value is .03. Since this value is less than .05, it can be concluded that there is statistically significant difference in performance after excluding the two participants. Although, we attempted to motivate all the participants, but the study reveals that some of the participants do not have an existing knowledge in assessing probabilities. Therefore, they were unable to provide better estimates of probability, which led to their exclusion from the study.

3.9.4 Twenty-eight participants, seven questions:

Initially, this study consisted of thirty participants and ten variables, but two participants and three variables were simultaneously removed from the study due to provision of highly incorrect estimates of probabilities. Although, we attempted to motivate all the participants, but the study reveals that some of the participants do not have an existing knowledge in assessing probabilities. Therefore, they were unable to provide better estimates of probability, which led to their exclusion from the study.

Table 3-12 Descriptive statistics for comparison of error rates (N = 28, Questions = 7)

		Unmotivated	Motivated
N	Valid	28	28
	Missing	2	2
Mean		12.8546	9.5104
Std. Error of Mean		1.24936	.68884
Median		11.0000	8.7850
Mode		13.29	6.86
Std. Deviation		6.61097	3.64499
Variance		43.705	13.286
Range		26.57	16.57
Minimum		4.57	5.29
Maximum		31.14	21.86
Sum		359.93	266.29

Results in Table 3-12 show that a higher average error rate (Mean = 8.99, SD = 4.63) was obtained when the participants were not motivated by the experimenter. Conversely, a lower average error rate (Mean = 6.66, SD = 2.55) was obtained when the participants were motivated. Since the average error rate fell, this is an indication that better estimates of probabilities were provided by the participants when an incentive was attached to the elicitation exercise. The participants and the questions were removed to make the analysis fairer or even.

Table 3-13 Paired samples Statistics (N = 28, Questions = 7)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Unmotivated	12.8546	28	6.61097	1.24936
	Motivated	9.5104	28	3.64499	.68884

Table 3-13 above displays the paired samples statistics of the unmotivated and motivated participants in terms of mean, population of participants in each group (N), standard deviation and standard error mean.

Table 3-14 Paired samples t-test (N = 28, Questions = 10)

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	T	Df	Sig. (2-tailed)
Pair 1	Unmotivated – Motivated	3.34429	6.84760	1.29407	.68907	5.99951	2.584	27	.015

A paired-samples t-test statistic was conducted to compare the performance of the remaining twenty-eight participants. Result in Table 3.14 shows that the Sig. (2-tailed) value is .015. Since this value is less than .05, it can be concluded that there is statistically significant difference in performance when the two participants and the three variables were excluded. The participants and the questions were removed to make the analysis fairer or even.

3.9.5 Summary and Conclusion

The study aimed at auditing research skills required for this study, and as well to compare the required skills with the existing ones. Skills deficits were identified and rectified by formulating an action plan and a timetable for doing so. The skills audit led to the development of existing skills and acquisition of new ones such as: face-to-face interview, communication skills, critical thinking, self-reflection, taking responsibility for own research, planning, and time management. The skills above have helped me in conducting this study, and as well in producing this report.

The research skills that were developed and acquired were used in extracting knowledge from a computer network expert about the causal structure and estimates of probabilities for the nodes (variables) in the domain.

The issues with the study are:

(i) Problem of getting domain expert(s): We spent about three months looking for domain experts in any field, within and outside the University without success, hence the only option was to use the person that agreed and was ready to serve as an expert for the study.

(ii) Domain consisting of discrete variables only and each having two states: This issue can be addressed by waiting until we could get expert(s) that can provide probability values for a domain that consists of both the discrete and continuous variables. The only option we had, was to make use of the expert that was available to us in order to complete the study as scheduled.

(iii) Use of single expert: Using more than one expert could have improved the outcome of the study but all efforts to get and use more than one expert, proved abortive since we could not get an additional expert.

(iv) Confidence level: The expert expressed 50% confidence level in all the estimates he provided, which translates to underconfidence in his estimates.

(v) Much time was spent on the study because of the iterative nature of the elicitation meetings which involves adjusting the estimates provided by the domain expert in order to improve the predictions from the network.

The more accurate approach of elicitation for the study is the indirect approach of elicitation. This was evaluated in terms of quality and determined from the predictions of the network obtained in using the indirect approach. Although, the indirect elicitation approach produced the network that predicts better, it took a longer duration than the direct elicitation approach to accomplish because it involves offering assistance to the expert in providing the probability values.

In Primary Research 1, we were able to construct Bayesian networks to troubleshoot computer network reachability and security issues, using the knowledge-driven approach, and as well validate the Bayesian networks in order to assess their performances.

Finally, we were able to identify indirect elicitation as the more effective approach to use when eliciting the causal structure of the Bayesian network and the probability assessments from domain expert(s) since indirect elicitation approach produced better estimates of probabilities as well as better predictions from its Bayesian network (BN).

Primary Research 2

In the first experiment, since the participants were not expecting any rewards from the elicitation session, they did not spend enough time to understand and answer the questions, therefore they provided highly incorrect estimates of probabilities which is as a result of fast thinking.

The participants performed better in the second experiment than in the first because they were expecting to get Love2Shop motivating vouchers in the second elicitation session. The reward made them to think slowly and hard, thereby providing better estimates of probabilities in the second elicitation session (Kahneman, 2011).

Some participants were excluded from the study because their estimates were highly incorrect thereby having a negative effect on results. This led to overestimation or underestimation of estimates of probabilities by the participants (O'Hagan, et al., 2006).

Although, we attempted to motivate all the participants, but the study reveals that some of the participants do not have an existing knowledge in assessing probabilities. Therefore, they were unable to provide better estimates of probability, which led to their exclusion from the study.

Although, all the questions were difficult, but some are relatively easy. Therefore, we removed the relatively easy questions from the study. The reason behind this, is that relatively simple questions do not require slow and deep thinking to answer (Kahneman, 2011). Motivation caused the participants to provide better estimates of probabilities to the difficult questions they were asked in the elicitation session.

In conclusion, motivation made the participants to expend effort to minimize heuristics and biases and thereby provide better estimates of probabilities. See Section 3.10.2 of this chapter.

In Primary Research 2, we were able to show that domain experts would perform better in an elicitation session if they are motivated. The motivator would make the domain experts to think slowly and hard, thereby enabling them to provide better estimates of probabilities in an elicitation session.

Chapter Four

Elicitation of Causal Structure and Estimates of Probabilities

4.1 Introduction

This chapter is both on elicitation of causal structure and estimates of probabilities. Generally, there are three main approaches for constructing a Bayesian network, namely data-driven, knowledge-driven and combination of data- and knowledge-driven. Typically, if domain data is available, then a data-driven approach can be used but if domain data is not available then the knowledge-driven approach is used to elicit estimates of probabilities from the domain experts.

4.2 Elicitation of Causal Structure and Estimates of Probabilities from Domain Experts

The process of requesting for an unknown quantity or quantities of interest from domain expert(s) is known as elicitation (O'Hagan, et al., 2006) ; Martin *et al*. (2011). Note that causal structure and probability distribution can be elicited from domain experts in a situation where there is a little or no data about a quantity or quantities of interest.

The first stage in construction of a BN is to elicit the causal structure from a domain expert in an elicitation session. Pearl (1988) and Jensen (1996) suggest that construction process of Bayesian model has four main stages, namely:

(i) Define the domain variables and order them topologically.

(ii) Define limits on the number of states and causal relationships for each node.

(iii) Build the causal structure, and

(iv) Define the conditional probability matrices or elicit the estimates of probabilities.

A Bayesian network consists of two parts namely the causal structure (qualitative part) and the probability estimates (quantitative part), the two parts can be elicited from student retention experts.

It is expected that the student retention experts will be able to provide an accurate causal structure and estimates of probabilities based on their expertise in student retention domain provided they can guide against the types of human judgement errors mentioned previously.

Elicitation of causal structure typically takes place before the elicitation of estimates of probabilities. For a given model, a set of variables of interest are to be identified and ordered. Thereafter, the identification of the causal links of the model would take place ensuring that they adequately represent the dependence and independence relations.
In this research, the causal structure is elicited from student retention experts by using the face-to-face elicitation technique. An elicitation session for a causal structure usually concerns the construction of the probabilistic graph (causal structure) by a facilitator and domain expert(s) using face-to-face interview technique. The facilitator asked each expert to identify the variables of interest and as well as identify the causal links between the variables of the proposed model. In a situation where the domain experts provided different probabilistic graphs (causal structures), they are asked to jointly review the structure in order to produce a consensus probabilistic graph (causal structure). After the reviews, the probabilistic graph (causal structure) is refined to include variables of cause (parent nodes) and effect (child nodes) (O'Hagan, et al., 2006).

4.2.1 Eliciting Univariate Probability Distribution

The student retention experts might be asked by a facilitator for a fixed interval or variable interval judgement. In the fixed interval method, the experts are asked for probabilities of the form $P(a < x < b)$ while in the variable interval method, they are asked for quartiles (lower quartile, median and upper quartile), for instance, for the value of x such that $P(a < x) = 0.25$

where,

P = Probability

a = lower bound or lower estimate

x = most likely or best estimate, and

b = upper bound or upper estimate

4.2.2 Fixed-interval method of probability judgement

In the method, a facilitator presents intervals to the expert, and thereafter asks the domain expert the probabilities of X lying in each interval.

- The roulette method: It is a fixed-interval method used to elicit graphically from experts the probability of X lying in a particular bin within a grid of *m* equally-sized bins (Morris *et al.*, 2014). "*An expert can see the shape of his or her distribution visually as he or she allocates the chips hence he or she does not have to specify probability distributions while a facilitator needs to fit distributions to the elicited estimates"* (Jeremy *et al*., 2014).

A facilitator has to fit distributions in order to get the parameters to represent the distribution in a Bayesian network (BN) package such as Netica, Hugin Expert or Bayes Server BN software.

4.2.3 Variable-interval methods of probability judgement

In the method, the facilitator would request the expert to choose a value *m (median),* such that the expert judges two intervals [0 , *m*] and [m ,1] to have the same probability of containing X

- The quantile method: The method is also referred to as the bisection method. It is based on the elicitation of the 25th, 50th and 75th percentiles from experts (O'Hagan, et al., 2006)

In the variable-interval method, experts need to specify their probability distributions while a facilitator has to fit distributions to the elicited estimates of probability values.
A statistical distribution could be discrete or continuous in nature. A discrete distribution is the one in which variables can take on a finite value within a specified interval, while a

continuous distribution is the one in which variables can take on any value within a specified interval. A continuous probability distribution function, $f(x)$, is a function which satisfies the following properties:

(i) The probability that x is between two points a and b is,

$$p[a \leq x \leq b] = \int_a^b f(x)dx$$

(ii) $f(x) \geq 0$. That is, it is non-negative for all real x

(iii) The integral of the probability function is one, that is,

$$\int_{-\infty}^{\infty} f(x)dx = 1 \qquad \text{(Dekking } et\ al.\text{, 2005).}$$

A probability distribution could involve one or more random variables. "*When the value of a variable is determined by a chance event, that variable is called a random variable"* (Stat Trek, 2015). A univariate probability distribution is a probability distribution involving only one random variable, denoted by θ while a multivariate probability distribution is a probability distribution involving two or more random variables, denoted by $\theta = (\theta_1, \theta_2, ..., \theta_n)$ which is a representation of the experts' joint probability distribution for those variables. The beliefs of experts can be captured about the parameters (that is, means, standard deviation, variance, covariance and correlation) of a quantity or quantities of interest representing them with a univariate or multivariate probability distribution. In this research, we are interested both in univariate and multivariate probability distributions because the study will involve elicitation of univariate and multivariate probability distributions from student retention experts.

4.2.4 The Tertile method

In the method, a facilitator presents intervals to the expert, and thereafter asks the expert the probabilities of x lying in each interval. *An expert does not have to fit distributions to the elicited estimates"* (Jeremy *et al*., 2014).

The tertile method: The method is used to elicit a median and three equally likely intervals from experts. It is based on the elicitation of the 33rd, 50th and 66th percentiles from experts (O'Hagan, 2012).

A facilitator might decide to fit a distribution in order to get the parameters to represent the distribution in a BN package such as Netica, Hugin Expert or Bayes Server software.
In this method, experts have to specify their probability distributions while a facilitator has to fit distributions to the elicited estimates.

4.2.5 Correlation and Regression: These methods can be used to encode the judgements of the domain experts by eliciting estimates of probabilities from them during a group elicitation session where experts would be asked to provide their lower and upper bounds and the most likely values, and they will be asked to provide their lower and upper quartiles (25th percentile and 75th percentile) and median (50th percentile) (James et al., 2010). The values provided by the domain experts will serve as inputs into an elicitation software in order to obtain the mean and variance of the input and output variables.
Thereafter, we elicited the relationship between the dependent and independent variables. The questions were asked clearly in such a way that the experts provided accurate values. On each occasion we asked the experts to assure us how confident they are in the values they have provided, and as well allow them to revise their estimates, if need be. We checked for coherence by making sure that the values provided by the experts conform to the laws and theorems of probability, for example, multiplicative and addition laws of probability. The predictions from the BN were shown to the domain experts in order to serve as a feedback for the domain experts (O'Hagan, et al., 2006). We also debiased the experts when values are elicited from them during the elicitation session (Tversky and Kahneman, 1973 & 1974; Kahneman, 2011).

3.3 Aggregation of results

Elicitation usually involves multiple experts in order to allow experts to share knowledge and to discuss the quantity or quantities of interest in question, and it can be performed on the individuals separately or as part of a group. In case of a group of experts, each domain expert provides his/her own result while a facilitator should combine the individual results by applying the behavioural or mathematical aggregation approach (Garthwaite *et al.*, 2005; (O'Hagan, et al., 2006). In the case of an elicitation session which involves a single domain expert, there is no need for aggregation of results. In this research, the elicitation involved multiple experts in order to allow experts to share knowledge and to discuss the quantity or quantities of interest in question. Whenever there is a divergent opinion, the domain experts applied the behavioural approach in order to reach consensus.

3.3.1 Behavioural aggregation approach

In this approach, the experts are asked to agree on a single result if there are conflicting results (Morris *et al.*, 2014). A consensus distribution involves negotiation and compromise amongst the experts which is a behavioural approach of combining information from multiple experts.

3.3.2 Mathematical aggregation approach

This is made up of linear and logarithmic opinion pools.

The linear opinion pool is a weighted average of the individual probability distributions comprising it (Garthwaite *et al.*., 2005).

It is expressed mathematically as:

$$p(\theta) = \sum_{i=1}^{n} w_i \, p_i(\theta) \,.$$

where:

w_i is a non-negative weight given to expert i of n.

The logarithmic opinion pool is a normalized weighted geometric mean which is equivalent to applying a linear pool to the logarithms of the individual probability densities and then normalizing the result (Garthwaite *et al.*., 2005)

It is expressed mathematically as:

$$p(\theta) = k \prod_{i=1}^{n} p_i(\theta)^{w_i} \,.$$

where:

w_i is a non-negative weight given to expert i of n and k is a normalizing constant which should integrate to 1.

4.4 Structured elicitation of expert knowledge

Elicitation is the process of extracting expert knowledge about some unknown variable or variables and expressing that information as a probability distribution (O'Hagan, et al., 2006). In a situation where there is a little or no statistical data about a variable or variables, estimates of probabilities are typically elicited from domain experts (Huang & Darwiche, 1994).

For an elicitation session to be successful, the process should be structured well by devising a new or adopting an existing protocol to use. The aim of a structured expert elicitation is to increase the accuracy of estimates of probabilities (O'Hagan, et al., 2006; Knol, et al., 2010).

Elicitation protocol has been covered by several authors and each of them came up with different suggestions. (Shephard & Kirkwood, 1994) suggested a five-stage method: motivating, structuring, conditioning, encoding and verifying. (Philips, 1999) suggested a four-stage process: introduction and training, motivation, conditioning and encoding once the experts have been identified and recruited. Clemen and Reilly (2001) suggested a seven-stage method; background; identification and assessment of experts, motivating experts, structuring and decomposition; probability and assessment training, probability elicitation and verification, and aggregation of experts' probability distribution. Garthwaite et al. (2005) suggested a four-stage method for the actual elicitation stage: setup, elicit summaries of the experts' distributions, assess the adequacy of the distributions. A five-stage method was suggested by (O'Hagan, et al., 2006): background and preparation, Identify and recruit experts, motivating and training the experts, structuring and decomposition, the elicitation itself (elicitation session).

The number of stages involved slightly differ from one protocol to another, but they have some components in common. Furthermore, there are differences in the order in which the components are to be carried out.

The Sheffield Elicitation Framework protocol is adopted for this work because it's components are complete, and the order in which the components are carried out is logical (O'Hagan, et al., 2006).

In this study, an actual elicitation (elicitation session) is going to be experimented with. This will involve elicitation of estimates of probabilities on internet usage in the UK from the participants.

A model for the elicitation process consisting of the following stages was given by (O'Hagan, et al., 2006):

(i) background and preparation
(ii) identifying and recruiting the expert(s)
(iii) motivation and training the expert(s)
(iv) structuring and decomposition
(v) the elicitation itself (elicitation session).

4.4.1 Background and preparation

After realising that there is a need for obtaining estimates of probabilities on student retention, the researcher first planned the elicitation session. The preparation of background information and the documents for recording the elicitation session took place at this step.

4.4.2 Identifying and recruiting the experts

The role of expert(s) in an elicitation exercise cannot be over-emphasized hence the recruitment of expert(s) is to be conducted very well because the outcome and acceptability of elicitation exercise greatly depends on the competency of the expert(s). The expert(s) selected for elicitation exercise are subject-matter expert(s) in the problem domain. Hora and von Winterfeldt (1997) list the following criteria for the experts: tangible evidence of expertise; reputation; availability and willingness to participate; understanding of the general problem area; impartiality and lack of an economic or personal stake in the personal findings.

4.4.3 Motivating and training the experts

At the beginning of the elicitation session, the facilitator clearly explained the purpose of the elicitation session to the expert(s), how the distributions elicited from them will be utilised and as well enlightens them about the difference between representing subjective uncertainty with probability distributions and estimating a probability. The facilitator also advised the experts to guide against vices associated with human judgement such as conflict of interest; biases; heuristic errors; anchor (initial value) and adjustment effects; and overconfidence. Lastly, the facilitator should conduct a mock elicitation exercise for the experts.

4.4.4 Structuring and decomposition

Here, a precise decision about the variables to be elicited, and how to elicit distribution for quantities of interest are made. The structure, dependencies and relationships as well are elicited from the experts at this stage.

4.4.5 The elicitation itself (elicitation session)

The facilitator conducted the elicitation session after the structuring and decomposition stage. Eliciting a distribution for quantities of interest involves the following steps: elicitation of estimates, debiasing, recording of estimates, feedback, aggregation and record of the session.

4.5 Question Framing

The way a question is framed often has an influence on how people answer that question. A slight paraphrasing of a particular question could lead to different estimates of probabilities.

In the literature, facilitators have used different ways in framing questions meant to be answered by the experts when eliciting estimates of probabilities from the domain experts

(O'Hagan, et al., 2006) ; Knol et al., 2010). For the domain experts to provide correct estimates of probabilities they must understand the questions a facilitator asked from them (Kuhnert, 2010).

The errors in question framing is typically caused by linguistic uncertainty. A facilitator should avoid linguistic uncertainty when eliciting estimates of probabilities from an expert. Regan *et al.* (2002) noted that linguistic uncertainty can be classified into four main types, namely: vagueness, insufficient background information, linguistic ambiguity and underspecificity.

Vagueness: A facilitator should avoid using a vague word by using clear definitions and at the same time ask the questions logically (Kuhnert, 2010)..

Insufficient background information about the variables that are conditionally dependent.

Linguistic ambiguity: Linguistic ambiguity occurs when words can have two meanings and, the expert is unclear which word a facilitator means (Kuhnert, 2010).

Underspecificity: This linguistic uncertainty occurs when too much room for interpretation is left as a result of not providing an expert with enough details needed to provide the required estimates of probabilities.

4.6 Heuristics and Biases

Heuristics and biases are the short-cut that experts typically use when asked to provide estimates of probabilities. The short-cuts typically affect the accuracy of the estimates provided by the experts. Tversky and Kahneman (1974); Kahneman (2011), Knol et al. (2010); and (O'Hagan, et al., 2006) list errors in human judgement as the drawbacks of the knowledge-driven approach. They cited issues such as conflicts of interest; biases; heuristic errors such as; availability, representativeness, anchor (initial value) and adjustment effects. Thinking is the action of using one's mind to produce thoughts. Kahneman (2011); Soll *et al.,* (2013) noted that there are two modes of thinking, namely fast thinking or System 1, and slow thinking or System 2. System 1 is an automatic, fast, effortless, and often unconscious way of thinking. Conversely, System 2 is slow, effortful, conscious, and can only process information in an orderly manner. System 1 is prone to biases and systematic errors hence System 2 typically monitors the System 1 mode of thinking. Motivation could cause the domain experts to think slow and hard while providing estimates of probabilities by adopting System 2 mode of thinking, which typically suppresses heuristics and biases, overestimation and underestimation of estimates of probabilities. Typically, accuracy goal leads to more intense processing while the directional goal creates bias.

Also, the participants are motivated by the experimenter by explaining the purpose of the elicitation session to them, and how the estimates of probabilities elicited from the participants will be used. Furthermore, the experimenter elicited the estimates of probabilities from them in such a way that they understood his questions, thereby, the participants would provide better estimates of probabilities for the variables being elicited for from the domain experts.

4.6.1 Kinds of biases and heuristics that can occur during an elicitation session.

Overconfidence: Overestimating the accuracy of their beliefs or alternatively underestimating the uncertainty in an elicitation exercise.

Conservatism: It relates to the process of experts understating their belief.

Availability: Basing a response on more recent available information and not considering past events.

Representativeness: providing opinions that are based on situations that are wrongly perceived to be similar.

Anchoring and adjustment**:** The tendency for experts to anchor around initial estimates (base value) and adjust their final estimates from the base value irrespective of the initial estimates' accuracy.

Misunderstanding of conditional probabilities: Confusion regarding the definition of conditional probability and failure to adhere to the axioms of conditional probability.

Translation: Confusion regarding the translation of a response to another scale.

Affect: Expert's emotions entering into the judgement making.

Hindsight bias: Expert places too much emphasis on past events and outcomes.

Law of small numbers: Expert bases their opinion on small pieces of information and assumes that this extrapolates to the population.

Linguistic uncertainty: Misunderstanding the question and/or applying different interpretations to the same term.

4.6.2 Heuristics and biases debiasing strategies

Below is a summary of some of the key heuristics, judgements or mental operations that can result in bias when eliciting estimates of probabilities from domain expert(s) and debiasing strategies (see Table 4-1).

Table 4-1: Heuristics and biases debiasing strategies

S/N	Heuristic	Bias	Debiasing strategy	Source
1	Representativeness: (Similarity in exactly the same way/relationship/ conditionality/ correlation)	(a) Insensitivity to prior probability of outcomes	Giving specific description or evidence	Kuhnet et al. (2010).
		(b) Insensitivity to sample size	Appreciation of the role of sample size in the formulation of a problem	(O'Hagan, et al., 2006).
		(c) Misconceptions of chance	Avoidance of selection of samples of inadequate size and over interpretation of findings	Garthwaite et al. (2005). Tversky and Kahneman, (1974).
		(d) Insensitivity to predictability	Avoidance of sole prediction in terms of the favourableness of the description of quantity of interest	Kahneman (2011)
		(e) Illusion of validity	Avoiding having great confidence in	

			predictions based on redundant input variables	
		(f) Misconception of regression	By ensuring that a predicted outcome is maximally a representative of the output.	
2	Availability	(a) Biases due to the retrievability of instances	To avoid assessing probability by the availability	Kuhnet et al. (2010).
		(b) Biases due to the effectiveness of a search set	of the contexts in which they occur or appear	O'Hagan et al. (2006).
		(c) Biases of imaginability	By generating and evaluating instances according to a given rule	Garthwaite et al. (2005). Tversky and Kahneman (1974).
		(d) Illusory correlation	By assessing probability of events that co-occur by the strength of the associative bond between them.	Kahneman (2011)
3	Adjustment and anchoring: (Starting from an Initial value that is revised in order to produce the final answer)	(a) Insufficient adjustment	Correct formulation of problem or/and accurate computation	Kuhnet et al. (2010). (O'Hagan, et al., 2006)
		(b) Biases in the evaluation of conjunctive and disjunctive events	By avoiding overestimation due to the chain-like	Garthwaite et al. (2005).

				structure of conjunctive events and underestimation due to funnel-like structure of conjunctive events	Tversky and Kahneman, (1974). Kahneman (2011)
		(c)	Anchoring in the assessment of subjective probability distributions	By starting with one's best estimate and adjusting upwards or downwards, as the case may be, in order to avoid a probability distribution that is too tight.	Kuhnet et al. (2010). (O'Hagan, et al., 2006) Garthwaite et al. (2005). Tversky and Kahneman (1974). Kahneman (2011)
4	Overconfidence		Overestimating the accuracy of his/her beliefs	By avoiding underestimation due to the chain-like structure of conjunctive events.	
5	Underconfidence		Underestimating the accuracy of his/her beliefs	By avoiding underestimation due to the chain-like structure of conjunctive events.	Kuhnet et al. (2010). (O'Hagan, et al., 2006)). Garthwaite et al. (2005). Tversky and Kahneman (1974). Kahneman (2011)

4.7 Gaps

In this research, the following gaps are filled:

- Motivation of domain experts (see Experiment 2 in Chapter 3)
- Protocol
- Software tool

The newly developed protocol is built upon the the Sheffield elicitation framework while the novel aspect that has been added is keeping of the record of an elicitation process. This is the sixth stage as well as the concluding step of the elicitation protocol.

When an elicitation session is completed, the estimates of probabilities provided by the domain experts and their aggregation judgement should be recorded. Every step taken and results obtained during the elicitation should be documented to provide a record that is clear and easy to understand of the process and results. In addition, the outputs and the aggregation estimates should be among the record to be kept. Every domain expert that took part in the elicitation process is entitled to have a copy of the elicitation process record. After the elicitation, the record should be used to duplicate the elicitation procedure with the same domain experts at a later date to check the self-consistency of the domain experts. In order to maintain the validity of the result of this study, this could be repeated at an interval of 3 years (see Figure 4-1).

4.7.1 Development of Softwaare Tool

The aim of this tool is to provide assistance for the domain experts while providing estimates of probabilities. In this regard, this research leads to the development of a tool that will assist domain experts in providing estimates of probabilities during the elicitation ession.

The prediction tool was developed in MATLAB. Matlab is a product of The MATHS Works, Inc. which is an advanced interactive software package specially designed for scientific and engineering computation. The Matlab environment integrates graphic illustrations with precise numerical calculations, and is a powerful and comprehensive tool for performing all kinds of computations.

The tool was implemented on a computer system with Intel (R) (TM) i5 - 8250U X64 based Processor 8.00 GB RAM, running on Microsoft Windows 11 Professional Version 23H2 while MATLAB 9.7 was used to implement the system design.

Instantiation of the variables in the student retention domain were derived from the elicited variables (nodes) obtained from the domain experts and were loaded on MATLAB 9.7 64-bit version. Graphical User Interphase (GUI) was dessigned to present the inputs and the results. The results produced by the tool could be used by a univesity to make prediction for probability of a student being retained or otherwise.

The tool was developed to reason with a mechanism for six input variables namely, “Change in circumstances”, “Mode of entry” , Academic Skills”, Academic Engagement”, “Academic progress”, and “Attendance” while the only output variable is “Retention”. An interactive graphical user interface was designed as front end to accept values of the input variables and to display result (output) as well. The flowchart for predicting the results is shown in Figure 4-68.

The tool has main window to key in student identification and the six input variables. On pressing “Predict” command button, the tool will display student’s probability of being retained (that is, retention result). The “Print” button permits a user to print the front end of the tool which is capable of presenting the input variables and the output while the “Reset” button permits a user to clear the contents of the front end of the tool in order to make a fresh entry. Furthermore, the “Close” button permits a user to exit the tool by pressimg on it.

The Bayesian network (BN) was integrated into the tool in order to calculate the prediction based on the elicitation inputs by connecting it to MATLAB via application programmer interfaces (API). This is made possible by Netica software which has the capability of integrating with MATLAB via API. In addition to language integration, Netica can also connect to many different data sources, including spreadsheets, databases, and even custom data formats.

The outputs as shown in Figure 4-69 to 4-132 were produced by the software tool developed in MATLAB. The outputs from the software tool were produced by different Bayesian networks (BNs) (see Figure 4-3 to 4-66).
After creating the protocol and the development of the software tool that was used for this work, the researcher carried out the elicitation session described below with the domain experts.

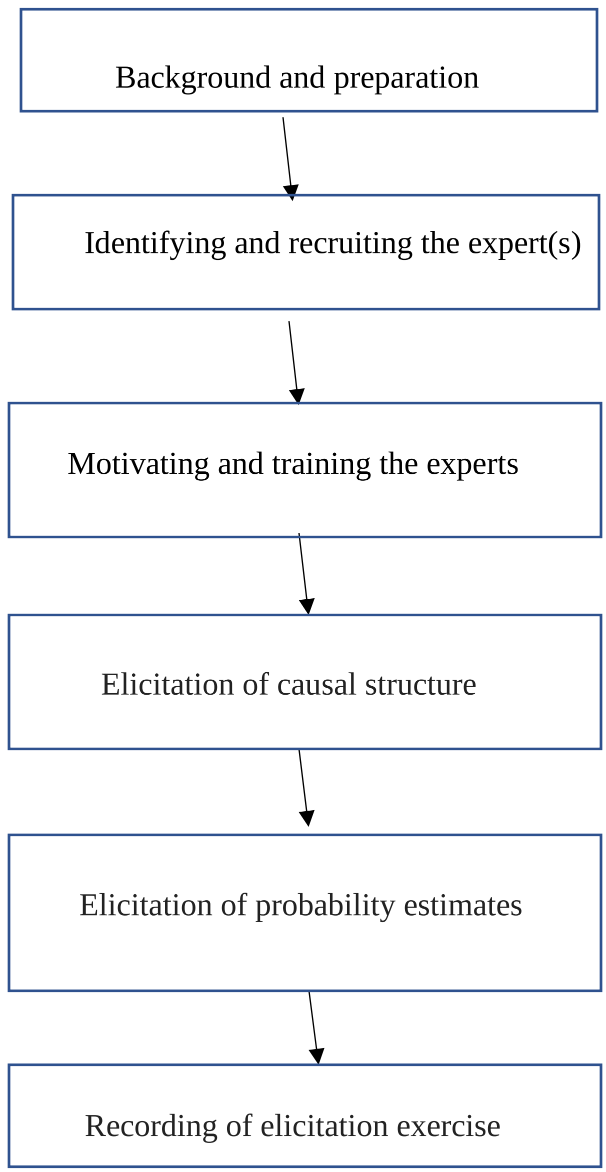


Figure 4-1: Flowchart for the protocol

For an elicitation session to be successful, the process should be structured well by devising a new or adopting an existing protocol to use. The aim of structured expert elicitation is to increase the accuracy of estimates of probabilities.

A protocol that has six stages is adopted for the elicitation session as shown in Figure 4-1. The protocol regarded identification of the variables in student retention domain which probability estimates are elicited from student retention experts and planning of the elicitation as the first step. This step was followed by identification and assessment of student retention experts who participated in the elicitation session. Furthermore, the reason for obtaining estimates of probabilities from the domain experts and how their estimates are going to be used, and provision of training for the domain experts. Thereafter, the structure, dependencies, and functional relationships of the variables in the domain were elicited, and the quantities to be elicited were defined precisely, including a specification of their units of measurement. Probability estimates are obtained from the domain experts during an elicitation session and the estimates of probabilities were verified. The estimates of probabilities provided by the experts are aggregated. Furthermore, an elicitation tool is developed to predict the probability of retention. Lastly, a report was produced on the elicitation session for record purposes.

3.8 Background and preparation

The researcher designed a structured elicitation protocol for the process. Also, the necessary software are installed on a computer prior to the elicitation session. Furthermore, a questionnaire is designed as an instrument of the research (see Appendix 6) where estimates of probabilities for the variables in the domain are filled in by the experts. The aim of this stage is to make adequate preparation for the elicitation session in order to achieve better estimates of probabilities from the domain experts as shown in Figure 4-1.

4.9 Identifying and recruiting the experts

The role of domain experts in an elicitation exercise cannot be over-emphasized hence the recruitment of domain experts was conducted well because the outcome and acceptability of elicitation exercise greatly depends on the competency of the domain experts. The experts selected for the elicitation exercise are subject-matter experts in student retention domain. Hora & Winterfeldt (1977) listed the following criteria for the expert(s): tangible evidence of expertise; reputation; availability and willingness to participate; understanding of the general problem area; impartiality and lack of an economic or personal stake in the personal findings. The above criteria are considered while identifying and recruiting domain experts for this research as shown in Figure 4-1.

4.10 Motivating and training the experts

Prior to the elicitation session, the facilitator clearly explained the purpose of the elicitation session to the experts, how the distributions elicited from them are to be utilised and as well enlightened them about the difference between representing subjective uncertainty with probability distributions and estimating a probability. The facilitator also advised the experts to guide against vices associated with human judgement such as conflict of interest; biases; heuristic errors; anchor (initial value) and adjustment effects; and overconfidence. Lastly, the facilitator conducted a mock elicitation exercise for the experts as shown in Figure 4-1.

4.11 Elicitation of causal structure

Here, a precise decision about the variables elicited and how to elicit distribution for a quantities of interest was made. The structure, dependencies and relationships are elicited from the domain experts at this stage as well as shown in Figure 4-1.

4.12 Elicitation of probability estimates

A facilitator can conduct an elicitation session after the structuring and decomposition stage. The elicitation of estimates of probabilities for quantities of interest involves the following steps: elicitation of estimates, debiasing, recording of estimates, feedback, and aggregation of estimates as shown in Figure 4-1.

4.13 Record of the session

For record purposes, a report about the elicitation session of the study should be produced as shown in Figure 4-1.

4.14 Elicitation Session

The elicitation session adopted the face-to-face approach in which the domain experts were asked to list the variables in the domain and to list the states of each of the variables as well. Thereafter, the variables were combined by establishing relationships between the states of each of the input variables to become a case (row). Each case has its own probability estimates which would represent the estimates of probability whether or not a student would be retained.

The indirect elicitation technique was used to elicit estimates of probabilities from the domain experts. The domain experts were asked to provide their estimates for each scenario in the questionnaire designed for this elicitation (see Appendix 6).

4.15 Heuristic and biases

Heuristics and biases are the short-cut that experts typically use when asked to provide estimates of probabilities. The short-cuts typically affect the accuracy of the estimates provided by the experts. The errors in human judgement are the drawbacks of the knowledge-driven approach. Issues such as conflicts of interest; biases; heuristic errors such as availability, representativeness, anchor (initial value) and adjustment effects as displayed in Table 4-1.

4.16 Elicitation of Causal Structure for the Student Retention Domain

Prior to the session for causal structure for student retention from the domain experts, the researcher have completed the Proportionate Review form and prepared accompanying documentation (letter of invitation, consent form (see Appendix 8), information sheet (see Appendix 7), and interview schedule for the session) for the purpose of ethical approval since the research raised only minimal ethical risk and it directly engaged human participants,that is, the student retention experts).

After identifying the domain experts, the researcher sent them letter of invitation to participate in the study. Thereafter he sent them a consent form, information sheet, and interview schedule for the elicitation session in advance for them to read, understand, and endorse before the commencement of the experiment.

The researcher obtained informed consent from the experts by issuing them consent forms to complete before they attend for the interview. The experts were given two weeks to withdraw their consent should they so choose. The consent form was accompanied with an information sheet which contains summary of the research as well as the rights they do have. The experts had the opportunity to ask questions before they consent to take part in this study. Completion and return of the questionnaire was taken as consent.

Furthermore, the researcher had arranged for an office space equipped with furniture (chairs and tables), which served as a venue for the session as well as for resources and items such as computers, computer software, projector, printer, flipboard, papers, whiteboard, and markers.

The elicitation of causal structure for student retention from the domain experts, which was the first part of the session lasted an hour. Thereafter, a probability training, and guidance against heuristics and biases, which was the second part of the session was conducted for the domain experts. This part also lasted an hour as well.

Twenty student retention domain experts were recruited on voluntary basis as participants for the experiment, but ten participants eventually took part in this experiment (Ritchie, et al., 2003). The student retention domain experts for this study involved students' affiars managers, level coordinators, student support officers, examination officers, lecturers, and faculty officers.

The sessions involved a group of student retention experts. We were able to recruit and make use of 10 domain experts. The justification for this population is that it captured at least one expert from the stated groups of student retention experts previously enumerated.

The criteria for selecting the experts is based on understanding of the general problem area on their part, tangible evidence of expertise, availability, and willingness to participate (Hora and von Winterfeldt,1997).

The experts that are eligible to participate in this study involved students' affairs managers, level coordinators, student support officers, examination officers, lecturers, and faculty officers. It only involved student retention experts who speak and understand written English in order to avoid translation.

A staff of the university who is not a students' affairs manager, level coordinator, student support officer, examination officer, lecturer, or faculty officer was excluded from this study. This is due to lack of understanding of the general problem area and tangible evidence of expertise in student retention domain. A student retention expert who does not speak and understand written English was also excluded from this study because we have no funds for translation.

The researcher started the session by welcoming and introducing himself to the domain experts (participants) and he asked them to introduce themselves to him in turn.

After the introduction, the researcher informed them that he will use the output of the elicitation session for his PhD project. Furthermore, the researcher informed the participants that the study would consist of two experiments which would involve elicitation of causal structure and elicitation of estimates of probabilities on student retention domain. Also, the researcher told them that he will confidentially keep the information given to him by the domain experts and that the elicitation session is only for academic purposes.

The elicitation session was the face-to-face type and it took the structured interview format. The researcher made use of open-ended questions to gather relevant information from the experts. Open-ended question is a type of question that does not limit the respondents to a set of answers. In other words, open-ended questions are free-form questions that gave the experts the freedom to express their knowledge, experiences, and thoughts about the domain (Wilson and Maclean, 2011).

Open-ended questions are typically used for qualitative observation where attention is paid to an in-depth description of the research subjects. Hence, the researcher used open-ended questions to elicit full and detailed responses from the domain experts, rather than the close-ended questions type that require brief responses.

There were follow-up questions depending on the domain experts' (participants') responses. The domain experts (participants) were prompted when they didn't mention the variables that impact on student retention. When the experts (participants) state different variables and didn't all agree on what the variables are, they were asked to reach a concensus by agreeing on a common causal structure (diagram). When the student retention experts had divergent opinions about the causal structure provided, they were asked to reach a concensus by agreeing on a common causal structure (diagram).

The researcher built a causal structure for the BN based on the information he obtained from the student retention experts, with the aid of Netica BN software. Thereafter, the researcher showed the causal structure (graph) to the student retention experts in order to obtain feedback from them whether or not the causal structure adequately represents the information obtained from them about the variables in the domain. The causal structure elicited from the student retention experts is used for constructing a Bayesian network (BN).

The researcher asked the participants, that is, student retention domain experts the following questions:

(i) Can you list all of the nodes (variables) in the student retention domain?

(ii) Can you topologically order the nodes (variables) that you have listed above in the student retention domain?

(iii) Can you specify the kinds of nodes (variables) in the student retention domain?

(iv) Can you define limits on the number of states and causal relationships for each node?

(v) Can you list all of the states (categories) of every node (variable) in the student retention domain?

(vi) Can you specify the scale of measurment, data type and unit for each variable and state (category) in the domain, e. g. nominal (label), binary, categorical, logical, discrete, or continuous.

In response to the above questions, the participants, that is, the student retention domain experts provided the following answers:

(i) They listed nodes (variables) ranging from 18 to 22 in number, but they finally agreed on 20, after reaching a concensus. This is in order to only use the important variables and ignore the less important ones as well as to group identical variables into one. The nodes listed are: gender, age, classification of students, staff/student relationship, peer relationship, mode of entry, academic skills, change in circumstances, family and partner moral support, visa issues, financial problem, journey time, health status, mode of study, enjoy course and university, attendance, home sickness, academic progress, academic engagement, and retention. (see Table 4.2).

(ii) They topologically ordered the nodes (variables) starting from the root nodes (the nodes at the top)

(iii) the intermediate nodes (the nodes in the middle), and

(iv) the target node (leaf node or the output node).

The variables of this study are formed based on elicitation from domain experts, it demonstrates the ideas among relevant variables under this study and how the variables are related to one another. The variables used for the study are defined as follow:

Table 4-2: Variables and States in the Domain

SN	Node (Variable) Name	Node (Variable) Type	Node's (Variable's) States (Categories)	Scale of Measurement	Data Type	Unit
i	Age	Root node (Independent node)	1) 16 - 21 2) 21+	Ratio	Continuous Discretised	Year
ii	Gender	Root node (Independent node)	1) Female 2) Male	Nominal	Categorical Discrete	None
iii	Classification of students	Root node (Independent node)	1) Domestic 2) International	Nominal	Categorical Discrete	None
iv	Journey time	Root node (Independent node)	1) 0 – 1 2) 1 – 2 3) 2 – 3	Ratio	Continuous Discretised	Hour
v	Staff/Students Relationships	Root node (Independent node)	1) Good 2) Poor	Nominal	Categorical Discrete	None
vi	Peer relationships	Root node (Independent node)	1) Good 2) Poor	Nominal	Categorical Discrete	None
vii	Financial problems	Root node (Independent node)	1) True 2) False	Nominal	Logical	None
viii	Mode of Entry	Root node (Independent node)	1) UME 2) Direct Entry	Ordered	Discrete	None
ix	Academic skills	Root node (Independent) node)	1) Good 2) Poor	Nominal	Discrete	None
x	Visa issues	Root node (Independent node)	1) True 2) False	Nominal	Logical	None
xi	Family and partner moral support	Root node (Independent node)	1) Great 2) Little	Nominal	Discrete	None

Table 4-2: Variables and States in the Domain

SN	Node (Variable) Name	Node (Variable) Type	Node's (Variable's) States (Categories)	Scale of Measurement	Data Type	Unit
xii	Change in circumstances e. g. family ties, career, religion, dietary, etc.	Root node (Independent node)	1) True 2) 21+	Nominal	Logical	None
xiii	Health status	Child node (Dependent node)	1) Physical health 2) Mental health	Nominal	Discrete	None
xiv	Mode of Study	Child node (Dependent node)	1) Full-time 2) Part-time	Nominal	Discrete	None
xv	Enjoy Course and University	Child node (Dependent node)	1) False 2) True	Nominal	Logical	None
xvi	Attendance	Child node (Dependent node)	1) At least 70% 2) Less than 70%	Nominal	Discrete	None
xvii	Home Sickness	Child node (Dependent node)	1) True 2) False	Nominal	Logical	None
xviii	Academic engagement	Child node (Independent node)	1) Engaged 2) Not engaged	Nominal	Discrete	None
xix	Academic progress	Child node (Dependent node)	1) Normal 2) Slow	Nominal	Categorical	None
xx	Retention	Target (Leaf) node	1) Retained 2) Not retained	Nominal	Categorical	None

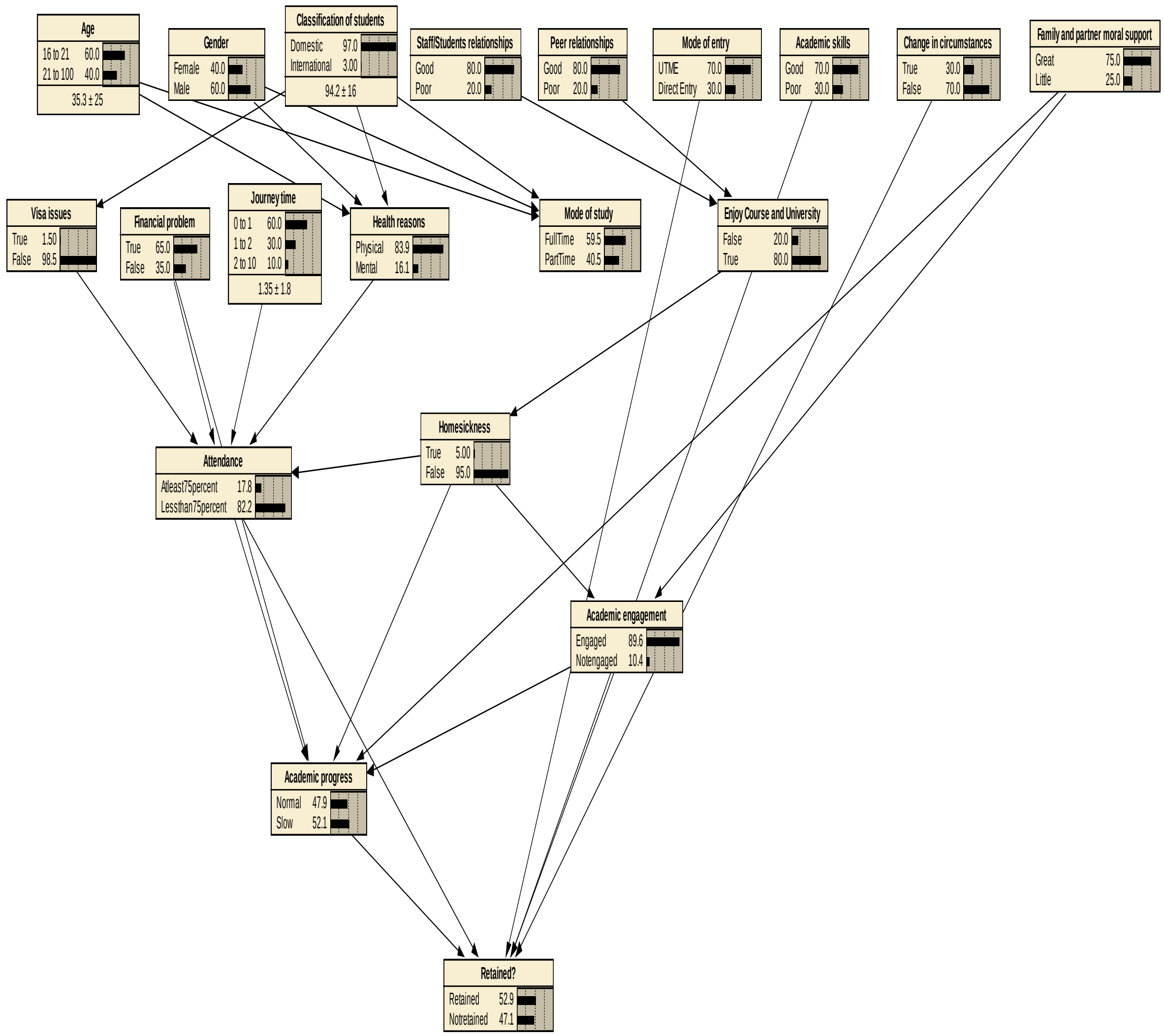


Figure 4-2: Bayesian network (BN) for Student Retention

(iv) They defined limits on the number of states and causal relationships for each node in the student retention domain (see Table 4-2 and Figure 4-2).

(v) They listed all of the states (categories) for each node (variable) in the student retention domain (see Table 4-2 and Figure 4-2).

(vi) They specified the unit and data type for each variable (node), and the states (categories) in the domain to be nominal (label), categorical, ordered, ratio, logical, discrete, or continuous (see Table 4-2).

The researcher built a causal structure for the BN based on the information obtained from the student retention experts, with the aid of a BN software. Thereafter, the researcher showed the structure to the student retention experts in order to obtain feedback from them whether or not the causal structure adequately represents their information about the variables in the domain. The causal structure elicited from the student retention experts was used for constructing a Bayesian network (BN). The causal structure is the output of this part of the session.

After checking the constructed causal structure, the domain experts confirmed that it is an adequate representation of their knowledge (opinion or belief).

The researcher was actively involved in the elicitation session by asking the domain experts various questions, taking down responses provided by the domain experts, asking the experts to reach a consensus whenever they had a divergent opinion about the variables in the domain, as well as constucting a causal structure for a Bayesian Network (BN).

The elicitation session for causal structure interview lasted for an hour. The output of this session is a causal structure (graph) obtained from the domain experts

At the end of the elicitation, the researcher showed his appreciation to the participants by thanking them for participating in the first part of the session and for sparing their invaluable time as well. The experimenter also informed them that the second part of the session would involve probability training and guidance against heuristics and biases during elicitation session.

4.16.1 Probability training for Domain Experts

The aim of the probability training is to guide the domain experts about how to express their opinion probabilistically. Although, the domain experts are experts in their field, but most experts do struggle to express their estimates in probabilistic form.

Prior to elicitation of estimates of probabilities from the experts, they were guided about how to provide probability estimates in terms of percentages, verbal expressions and the range (interval) method.

In order to guide the experts, the following methods are used to elicit probability estimates from the domain experts during the training:

(i) Frequency: Frequency is a method of eliciting probabilities, and this can be converted to a proportion (Kuhnert et al, 2010). A facilitator can elicit frequency from student retention experts by asking them a question such as: Considering 100 students who do not enjoy their course and were home sick as well, how many of them would you expect to remain and succeed? The facilitator converted the frequency to a proportion, for each domain expert *i*, to form a prior. This wass done by framing each question such that the experts would understand the question that was put across to them by the researcher (that is, question framing).

(ii) Verbal expression of probabilities: This is a method of eliciting probabilities which involves using the probabilistic expressions such as “probable”, “fifty-fifty”, “improbable”, “impossible”, etc. which can be interpreted as probability of about 85% (0.85), 50% (0.5), 15% (0.15), and 0% (0.0), etc., respectively (Van De Gaag et al.,1999 ; 2002). In this method, the experts expressed their estimates of probabilities using verbal expressions, and the researcher interpreted them as probabilities.

(iii) Elicitation of estimates of probabilities using the range or interval method (lower bound, upper bound, and the most likely value). In this method, the domain experts chose their extreme estimates (lower bound and upper bound) in the first instance, and thereafter adjusted their estimate to a value which is greater than their lower bound (lb) but smaller than their upper bound (ub).

Whenever the experts provided estimates of probabilities that did not conform to the fundamental laws and theorems of probability, namely, the total probability law, addition law, and multiplicative law of probability calculus (O'Hagan, et al., 2006), the researcher guided the experts by asking them to adjust their probability estimates to conform with these laws.

The researcher was actively involved in this part of the session by explaning how to assess probability estimates to the domain experts, entertaining and answering questions posed by the domain experts as well.

The probability training and guidance against heuristics and biases lasted for an hour. The output of this part of the session is to train the domain experts how to avoid cognitive errors while providing estimates of probabilities as this will affect their estimates.

4.16.2 Heuristics and biases training for Domain Experts

The aim of the heuristics and biases training is to let the domain experts be aware of the various types of heuristics and biases that can crop up in an elicitation session as well as how to guide against them since these vices can affect the accuracy of their estimates of probabilities.

Prior to elicitation of estimates of probabilities from the experts, they were guided guided against, or at least have awareness of human factors in terms of potential heuristics and cognitive biases.

The researcher enumerated and explained types of heturistic and biases that can occur during elicitation session to the domain experts, such as:

Availability bias: The availability bias is the tendency to overestimate the likelihood of events as they readily come to mind (O'Hagan *et al.*, 2006). The debiasing strategy for this type of heuristic and bias is that student retention experts should avoid assessing probability based on recent occurrence, familiarity and retrievability.

Anchoring bias: Anchoring bias is overreliance on a single piece of information (O'Hagan, et al., 2006). The debiasing strategy for this type of heuristic and bias is that student retention experts should start with their best estimate and adjust upwards or downwards, as the case may be, in order to avoid a probability distribution that is too tight to adjust (narrow).

Overconfidence bias: Overconfidence bias is an inflated opinion of experts ability leading to subsequent error (O'Hagan, et al., 2006). Their confidence in their judgements does not align with the accuracy of these judgements. The debiasing strategy for this type of heuristic and bias is that student retention experts should avoid overestimating the accuracy of their beliefs.

Representativeness bias: Misinterpreting the likelihood of an event considering both the key similarities to its parent population, and the individual characteristics that define that event (O'Hagan, et al., 2006) . The debiasing strategy for this type of heuristic and bias is that student retention experts should ensure that their predicted outcome is a true representative of their beliefs.

The researcher also trained the domain experts to adopt System 2 mode of thinking (Kahneman, 2011), which is a slow, effortful, conscious, and can only process information in an orderly manner. The System 2 mode of thinking typically suppresses heuristics and biases, overestimation, and underestimation of estimates of probabilities unlike System 1.

The researcher was actively involved in this part of the session by listing and explaning various types of heuristics and biases that exist to the domain experts, and how to supress them. He entertained and answered questions posed by the domain experts.

The probability training and guidance against heuristics and biases lasted for an hour. The output of this part of the session is to train the domain experts how to assess probabilities as well as how to mitigate cognitive biases while providing estimates of probabilities during elicitation session.

Prior to provision of estimates of probabilities by the student retention experts, they were given a questionnaire into which they were asked to provide their estimates of probabilities for the variables based on the causal structure that they had provided. The questionnaire is MS-Word file and there are rows into which the experts need to fill in their estimates of probabilities for each row or scenario. The questionnaire is easy to read, understand as well as less bulky as each row represents a scenario. Thereafter, the experts were asked to complete and return their completed questionnaire to the researcher for analysis, three weeks after collecting the research instrument (questionnaire), see Appendix 6. The experts were not able to meet up, but the researcher had to communicate and remind them to complete and forward the questionnaire to him for analysis.

4.16.3 Missing data

In situations where some of the experts could not produce estimates of probabilities for a particular variable, the researcher managed any missing data by calculating the mean or finding the mode of the estimates provided by other experts and use this to represent the estimate of probability for that row, this is called the imputation method (Little et al., 2012). In a situation where none of the experts could produce estimates of probabilities for a particular node (variable), the researcher managed any missing data by dividing the node s(variable) by the number of states it has. For instance, if a node (variable) has five states, 20% or 0.2 (that is, 1 / n) was used to represent the probability of each state of such node (variable) using the fact that the states of the node (variable) are equiprobable, that is, they have equal probability. Hence, they are equally likely to occur therefore each child state is given the same probability swhich is equal to 1 / n, where n is the number of states of such node (variable) (Cain, 2001). This conforms with the rule of probability which states that the sum of all of the possible outcomes of an event must equal to 100% or 1.0.

4.16.4 Quality checks

Quality checks such as reviewing the causal structure during the interview was done by the researcher. The researcher also reviewed the estimates of probabilities after receiving the questionnaires from the experts by mathematically aggregating the estimates of probabilities provided by the experts. In the case the student retention experts have divergent opinions about the estimates of probabilities they had provided, their estimates were aggregated mathematically by using the linear opinion pool (Garthwaite et al , 2005), which is a weighted average of the individual probability distributions. Furthermore, each expert was asked to choose and use any desired pseudonym and state the desired pseudonym in the space provided on a consent form for anonymisation purposes. Data were anonymised by identifying the experts by their pseudonym. This would also help with tracking the anonymous response for removal if a participant decided to withdraw their data after two weeks.

The researcher passed the causal structure and estimates of probabilities to his supervisors to check them for consistency and accuracy. Thereafter, the researcher showed the results to other experts in the field for validation purposes.

The estimates of probabilities that were elicited from the domain experts were provided based on their existing knowledge in the domain. Their estimates of probabilities serve as input data for the BN.

Thereafter, the researcher constructed a Bayesian network from the probability estimates provided by the domain experts and compiled the network as well. The researcher showed the predictions from the Bayesian network to the experts in order to get feedback from them. Initially, the predictions from the Bayesian network did not adequately represent the experts' probability estimates. Consequently, the researcher asked the experts to adjust their probability estimates until the predictions from the BN adequately represent the experts' judgements (O'Hagan, et al., 2006) in order to validate the BN. At last, the constucted Bayesian network was capable of predicting the probability of students' being retained (enroled) the following academic session and these were expressed in form of percentages.

Twenty domain experts in student retention were recruited on voluntary basis as participants for the experiment, but ten participants eventually took part in this experiment (Ritchie, et al., 2003).

Thereafter, the questionnaire (see Appendix 6) that was designed for the study was distributed to the participants in order to provide their estimates for the scenarios, in the spaces provided in the questionnaire. A questionnaire is a set of standardised questions for obtaining responses from a large group of individuals and is very often used to provide quantitative data (Burton & Bartlett, 2016). The completed questionnaires were collected from the participants 3 weeks thereafter.

At the end of the elicitation, the researcher showed his appreciation to the participants by thanking them for participating in the experiment and for sparing their invaluable time as well.

4.17 Bayesian Networks (BNs) Predictions

The estimates shown in Figure 4-3 to 4-66 were elicited from the domain experts and these were entered into Netica software (a Bayesian network software), where the estimates were compiled and ran. The different Bayesian networks (BNs) below (Figure 4-3 to Figure 4-66) produced different results from the software tool (see Figure 4-69 to 4-132).

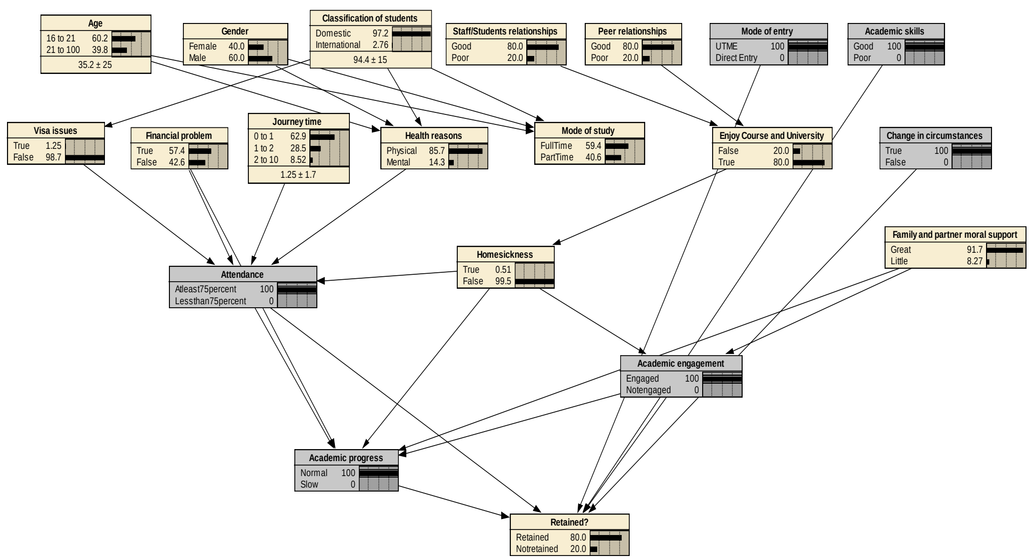


Figure 4-3: Bayesian network (BN) prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or more. This student represents a very low risk of attrition as his/her probability of being retained is 80%.

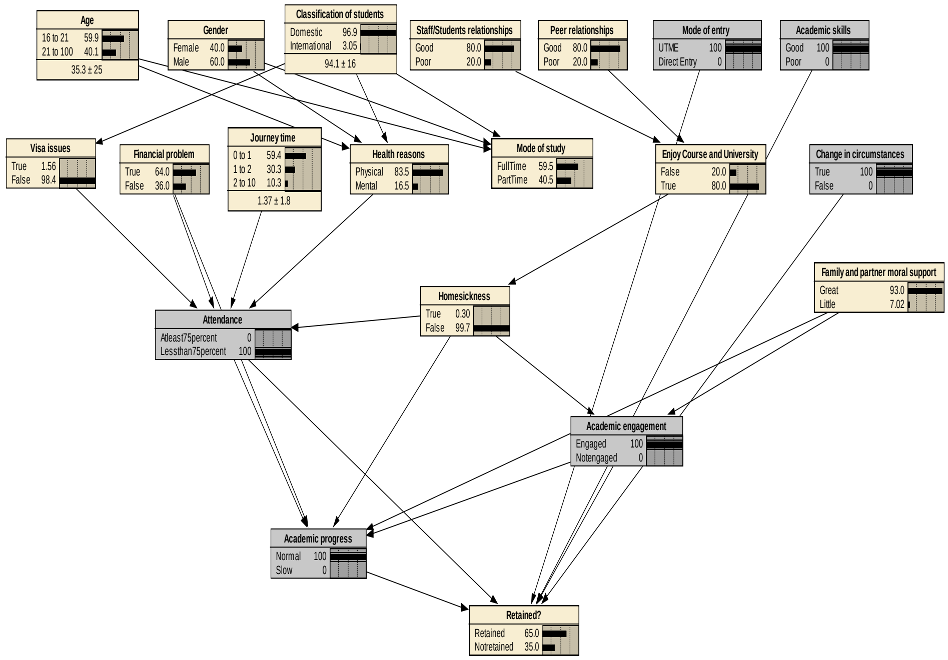


Figure 4-4: Bayesian network prediction for a student whose Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%. This student represents a risk of attrition as his/her probability of being retained is 65%.

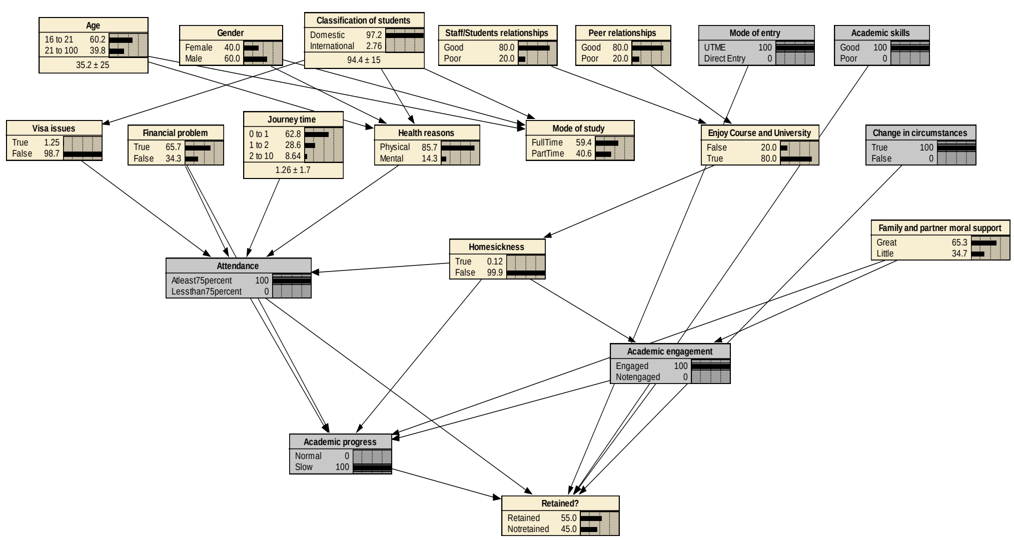


Figure 4-5: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a fairly high risk of attrition as his/her probability of being retained is 55%.

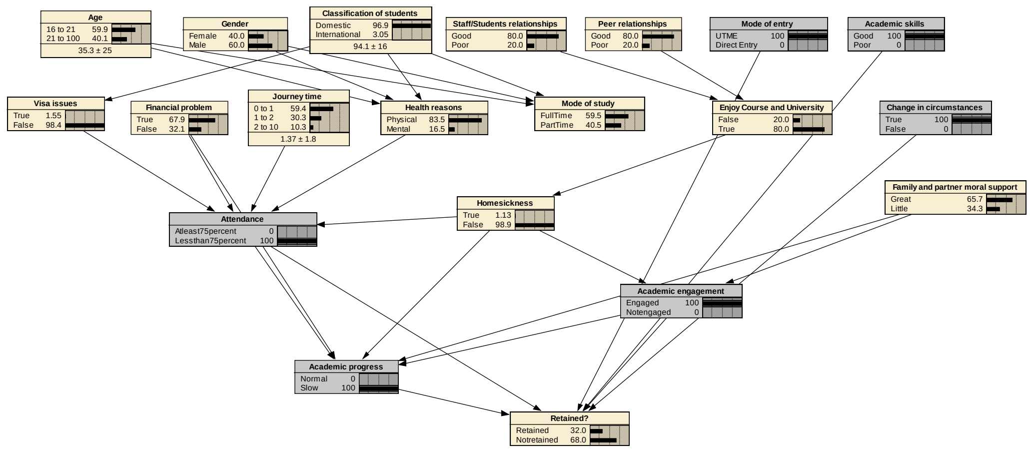


Figure 4-6: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 32%.

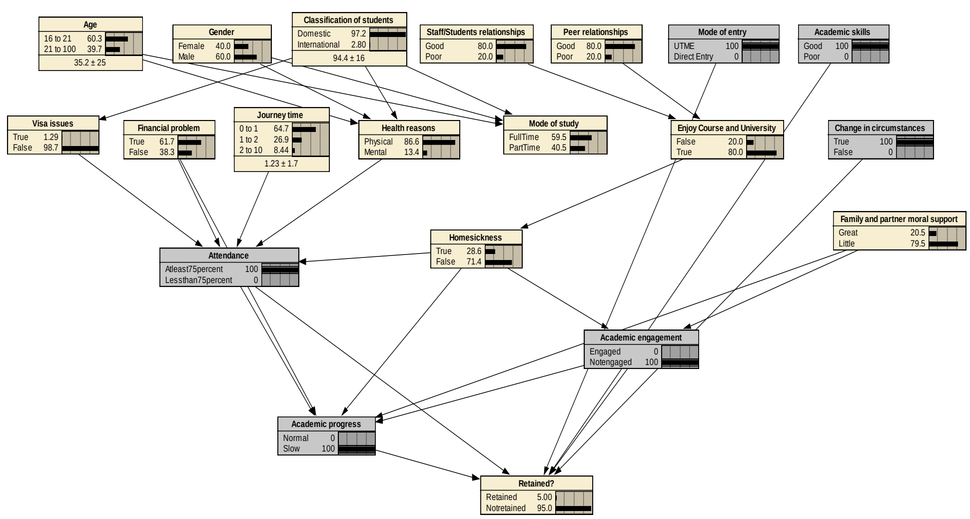


Figure 4-7: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic

engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

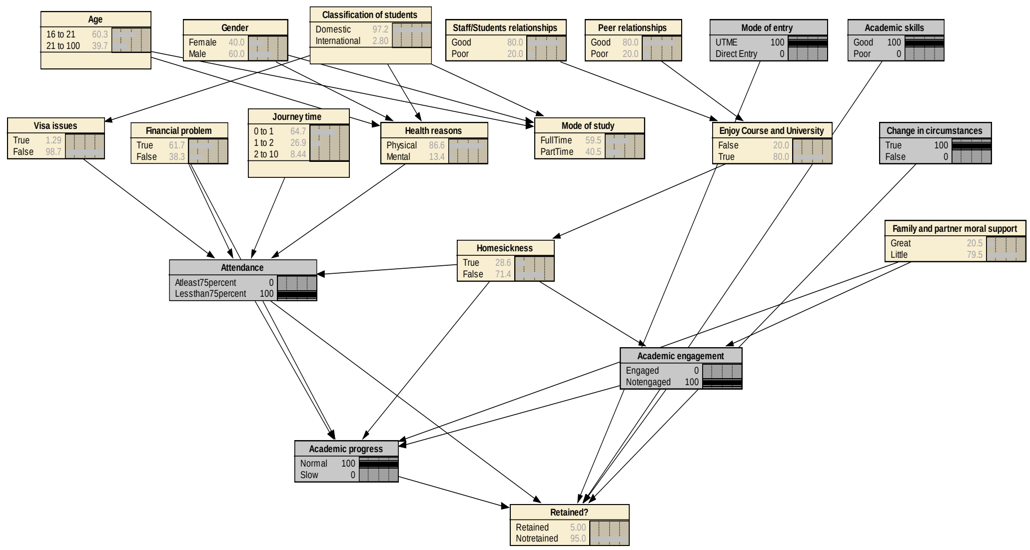


Figure 4-8: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 10%.

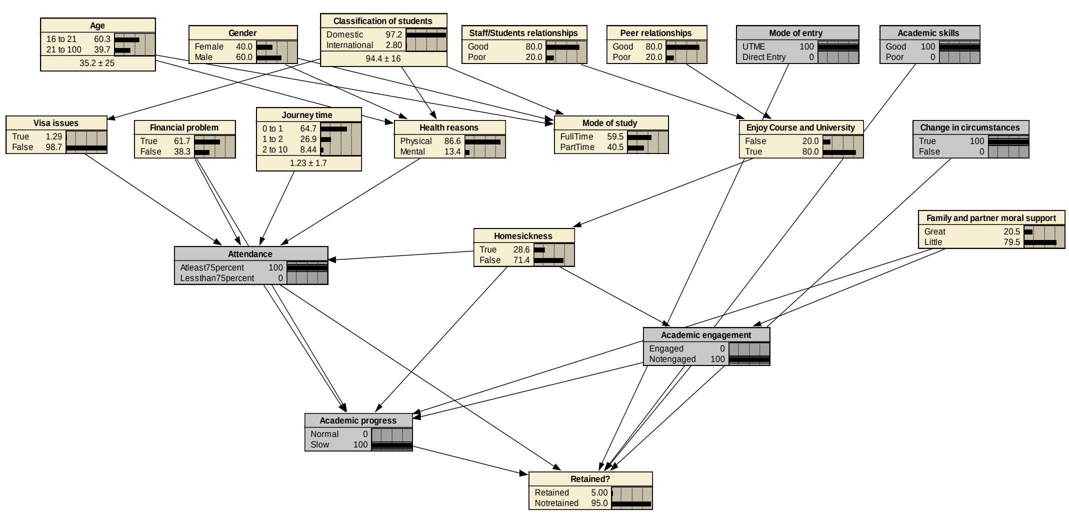


Figure 4-9: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic

engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

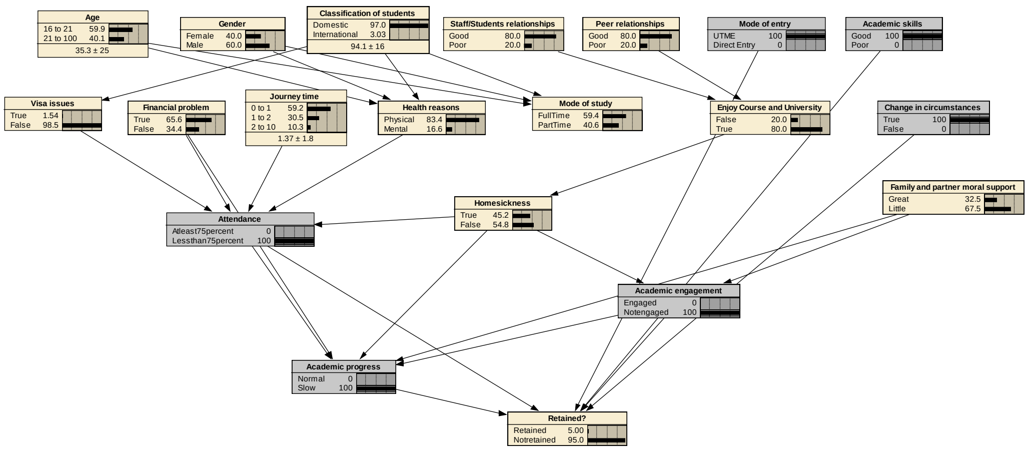


Figure 4-10: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

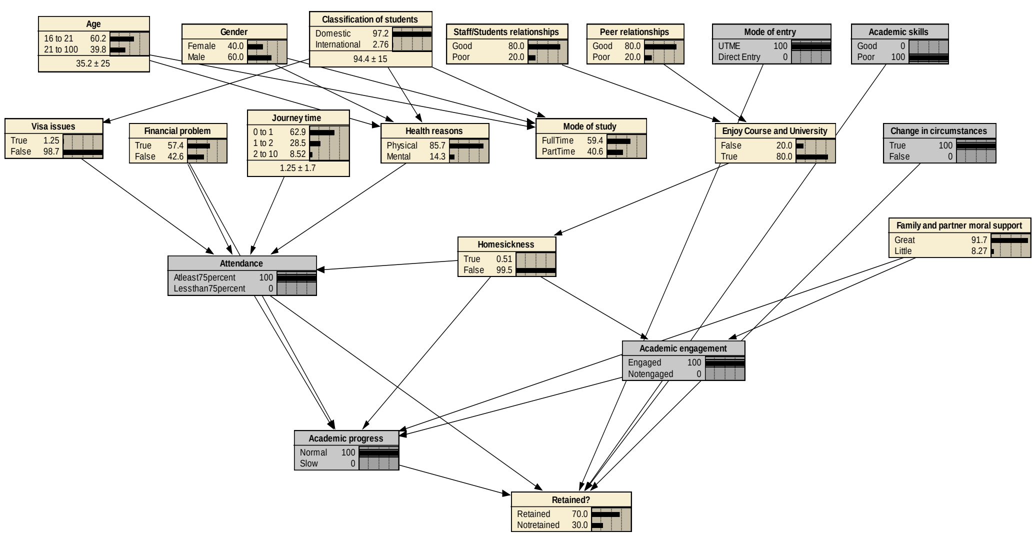


Figure 4-11: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a low risk of attrition as his/her probability of being retained is 70%.

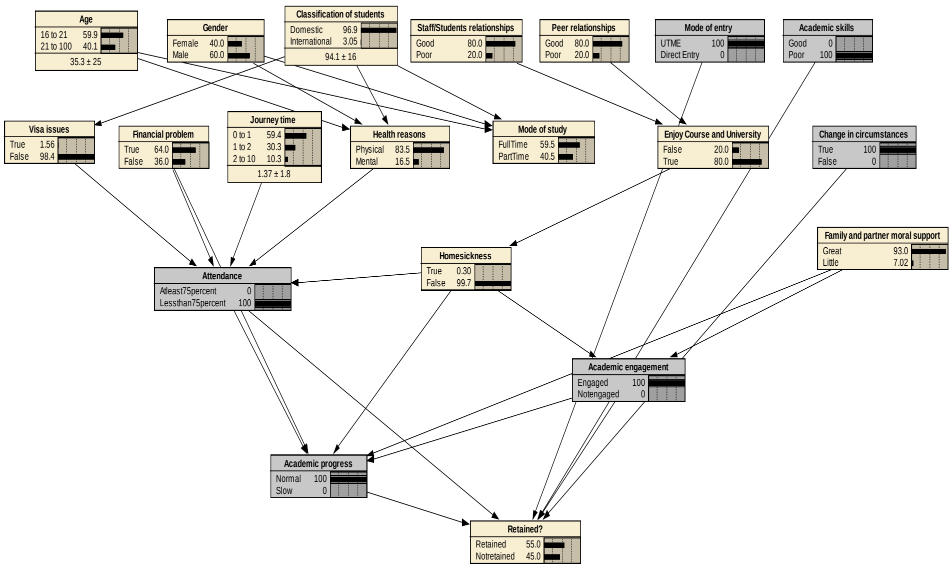


Figure 4-12: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a fairly high risk of attrition as his/her probability of being retained is 55%.

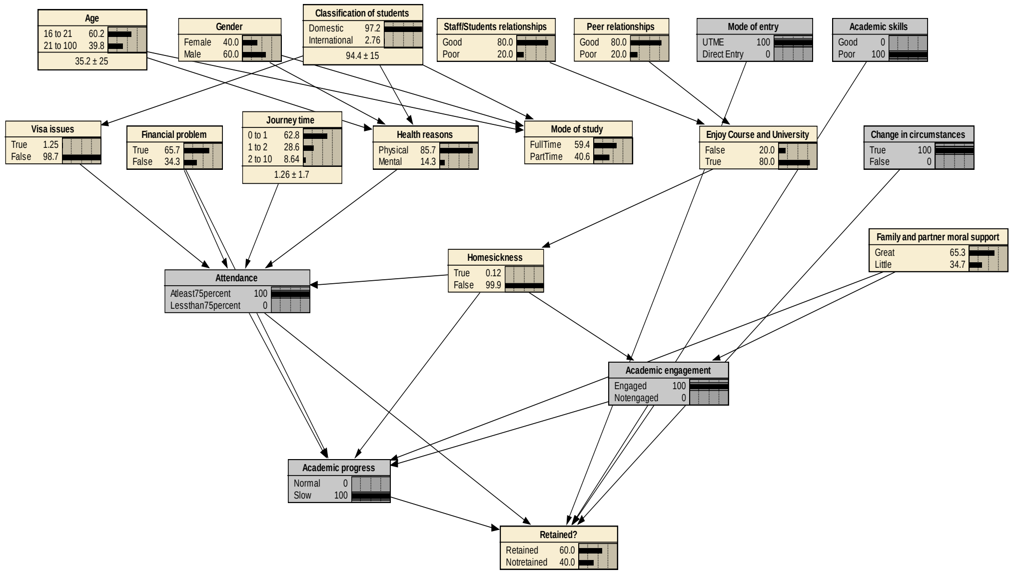


Figure 4-13: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a fairly high risk of attrition as his/her probability of being retained is 60%.

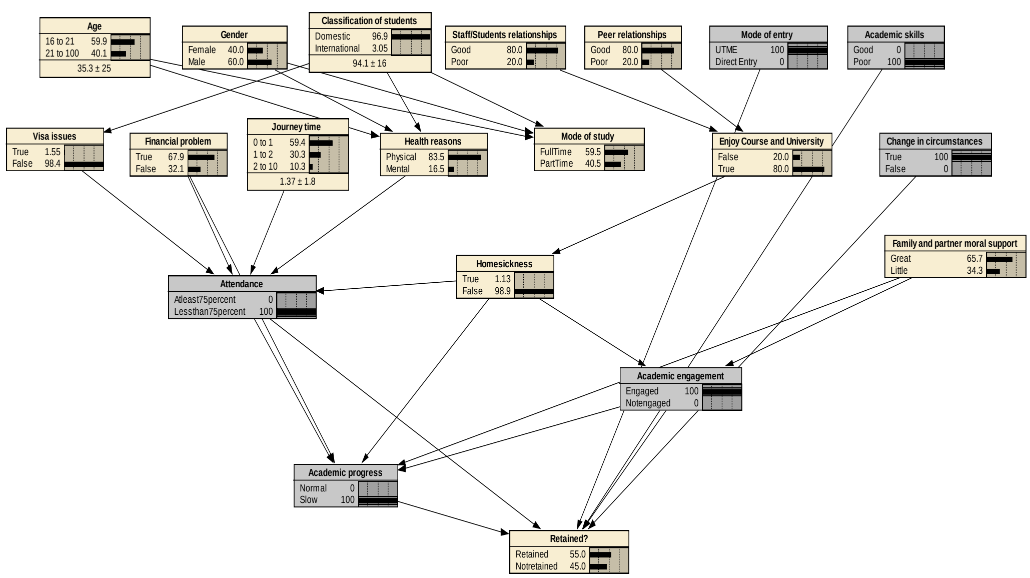


Figure 4-14: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a fairly high risk of attrition. as his/her probability of being retained is 55%.

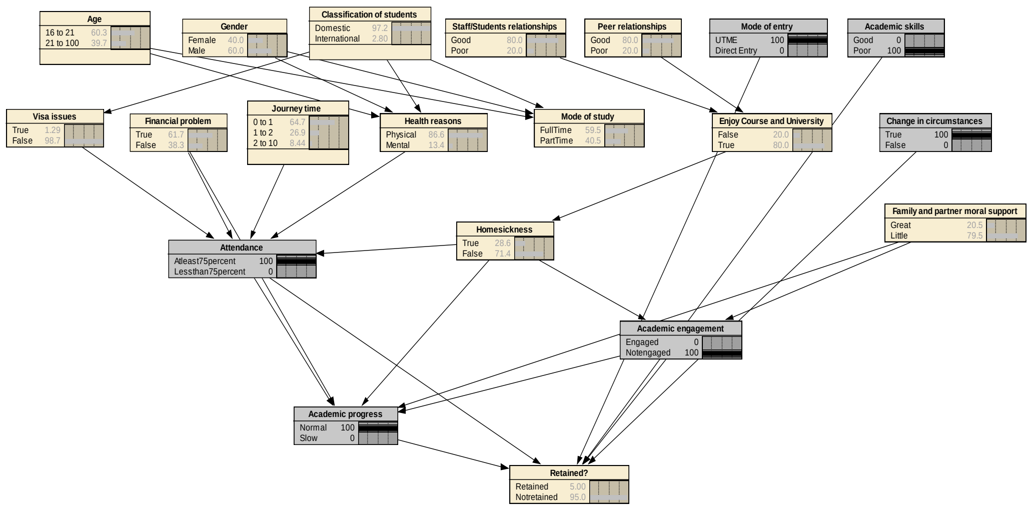


Figure 4-15: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 20%.

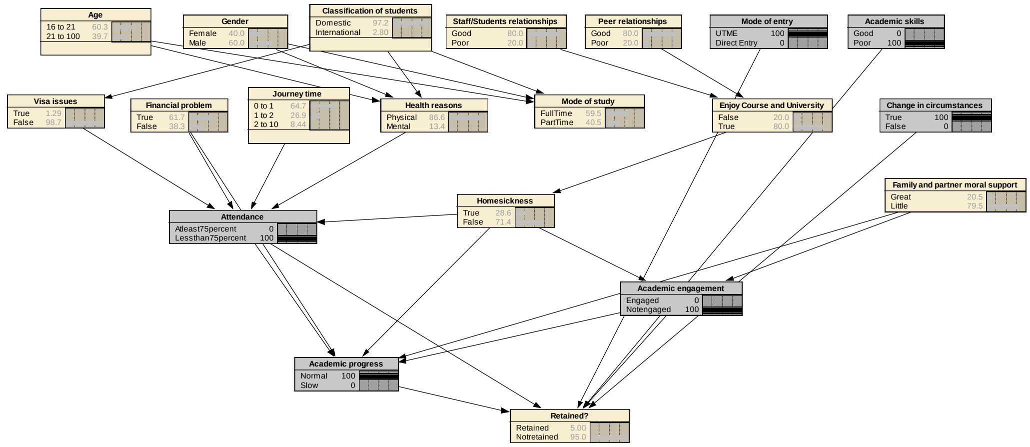


Figure 4-16: Bayesian network prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

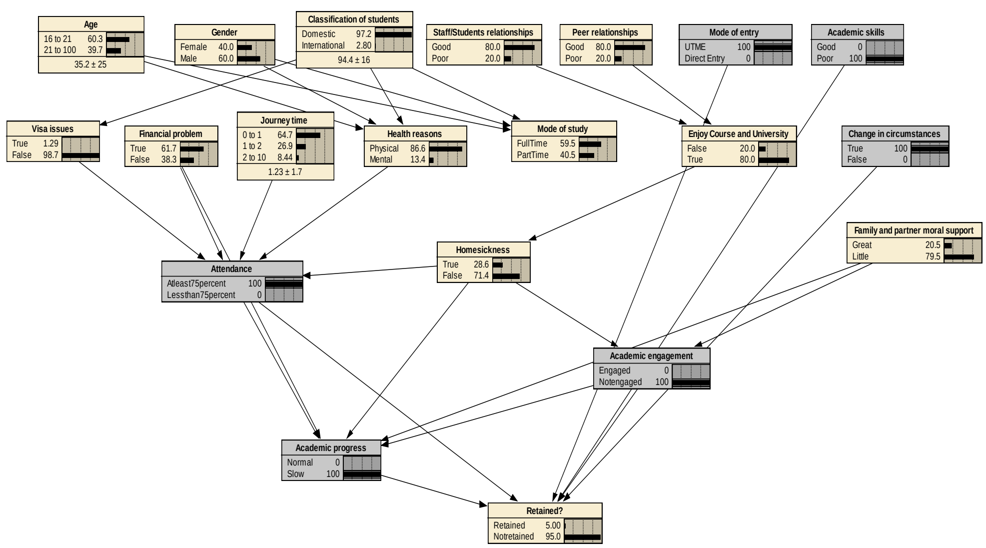


Figure 4-17: Bayesian network prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. This student

represents a very highs risk of attrition as his/her probability of being retained is 5%.

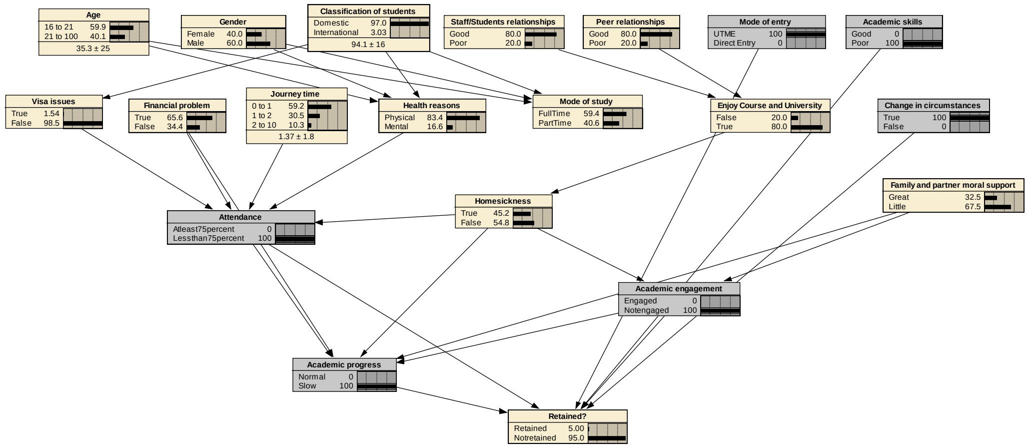


Figure 4-18: Bayesian network prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75% . This student represents a very high risk of attrition as his/her probability of being retained is 5%.

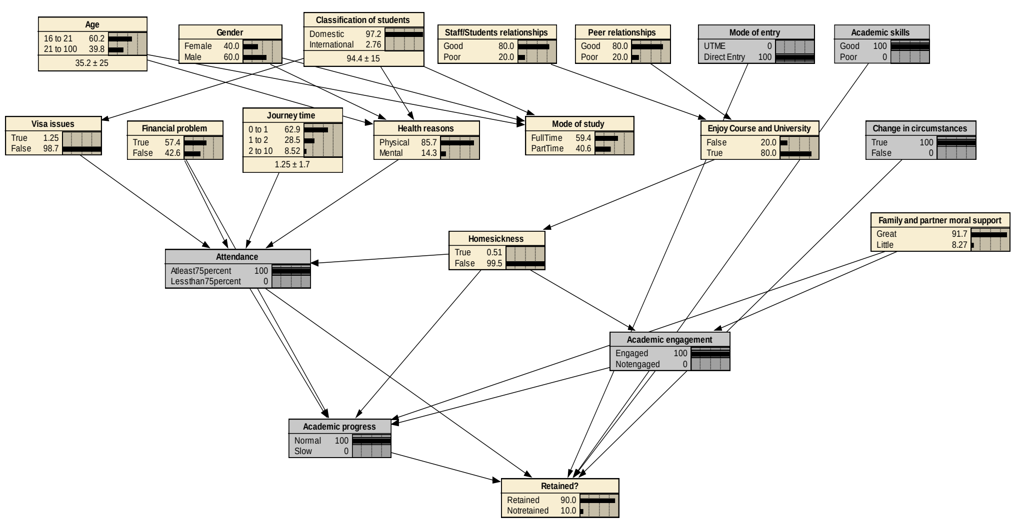


Figure 4-19: Bayesian network prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic

engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 90%.

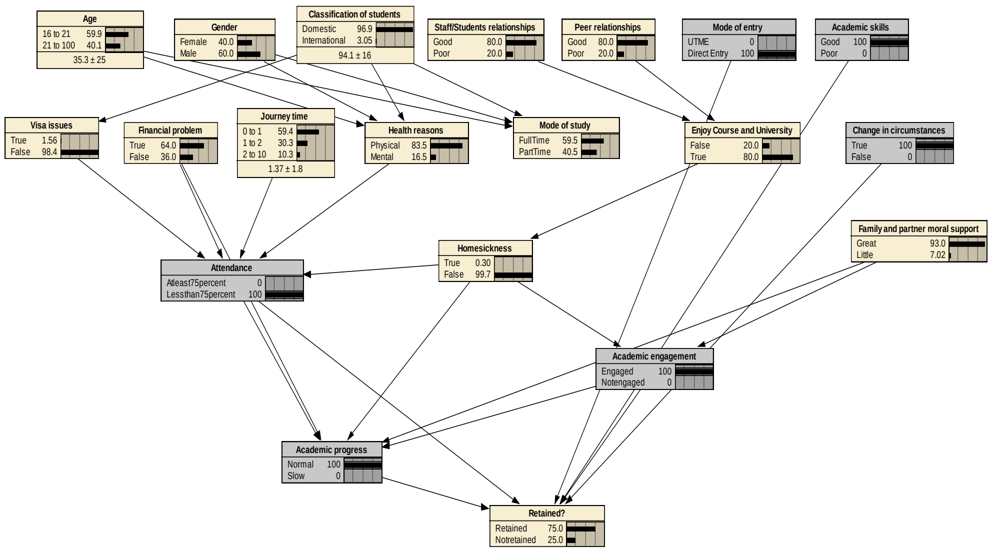


Figure 4-20: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a low risk of attrition as his/her probability of being retained is 75%.

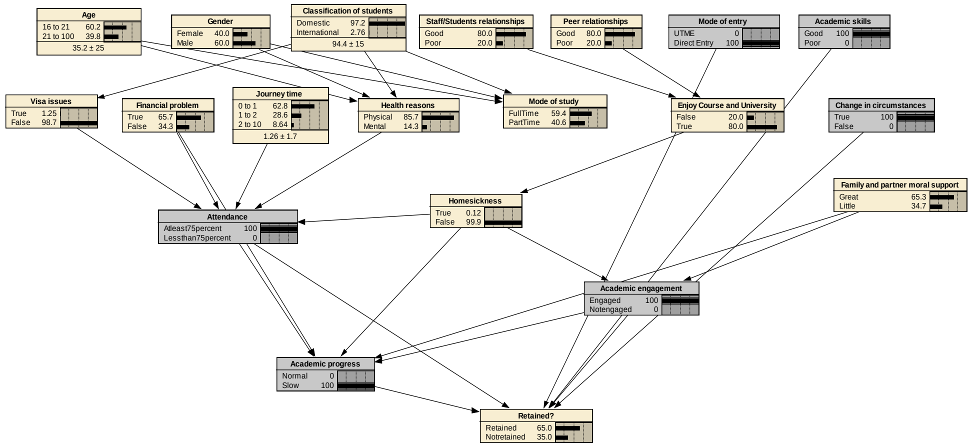


Figure 4-21: Bayesian network prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a low risk of attrition as his/her probability of being retained is 65%.

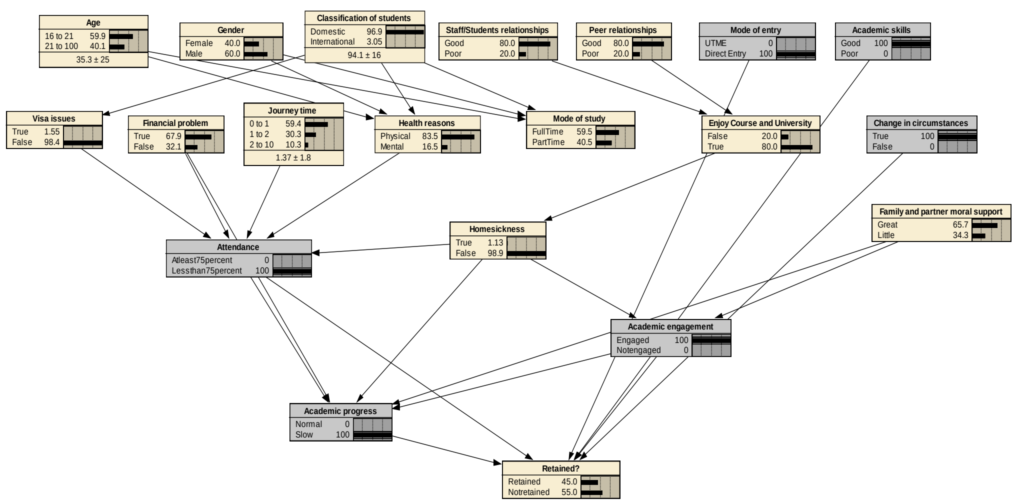


Figure 4-22: Bayesian network prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a fairly high risk of attrition as his/her probability of being retained is

45%.

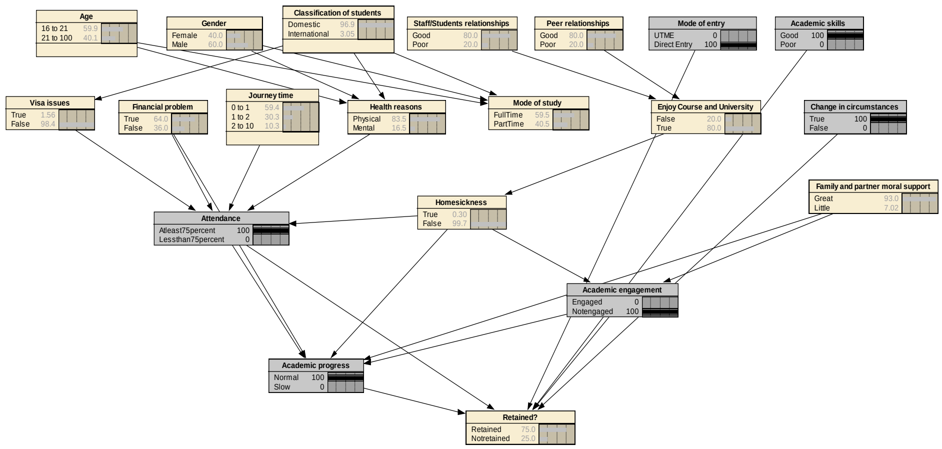


Figure 4-23: Bayesian network prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 20%.

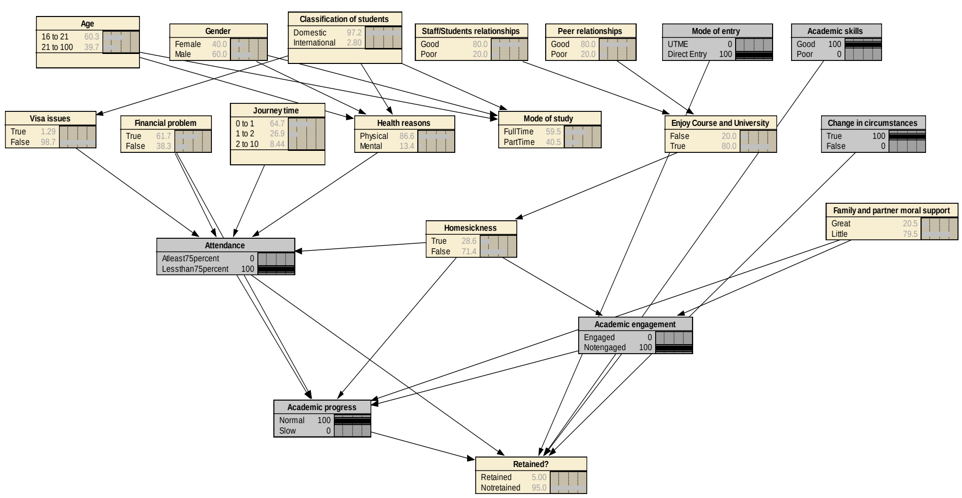


Figure 4-24: Bayesian network prediction for a student whose Change in circumstances is

‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

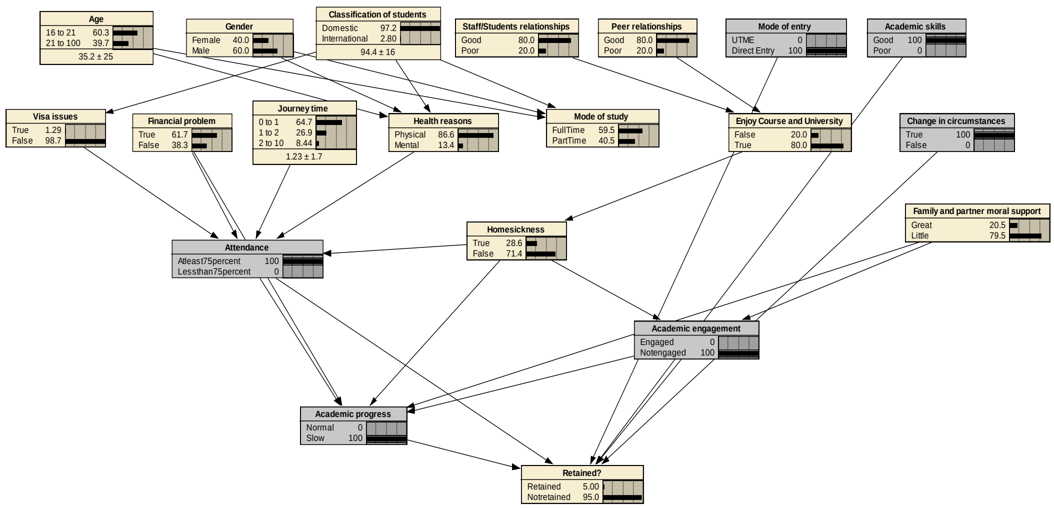


Figure 4-25: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

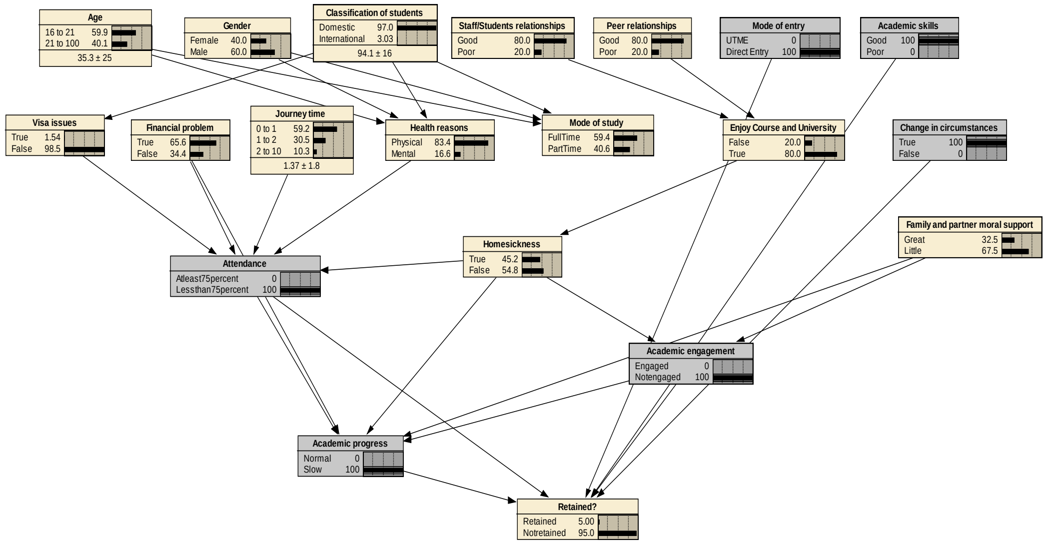


Figure 4-26: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

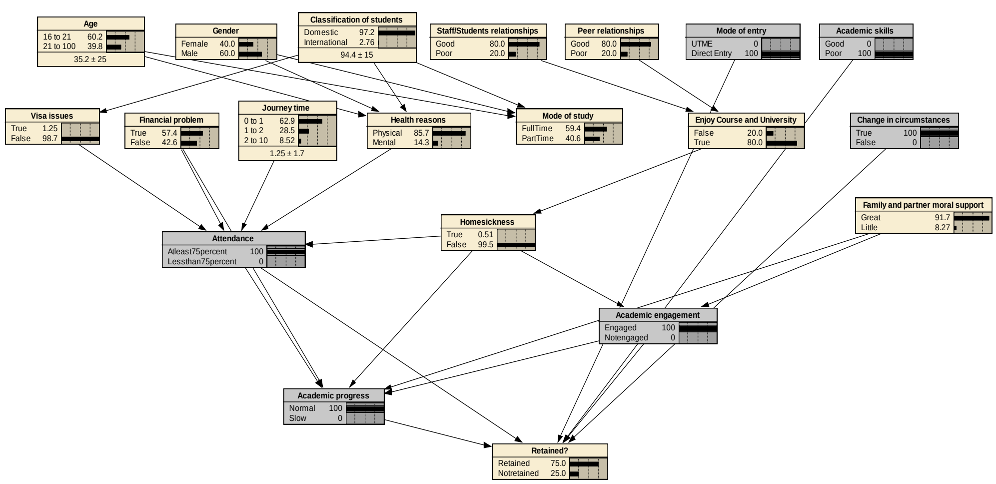


Figure 4-27: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic

engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 75%.

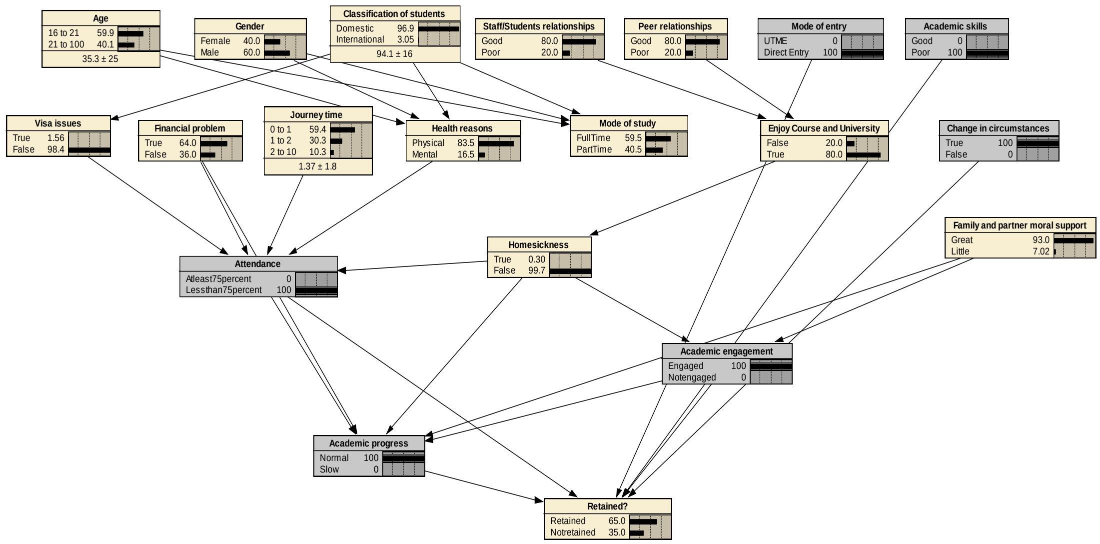


Figure 4-28: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a low risk of attrition as his/her probability of being retained is 65%

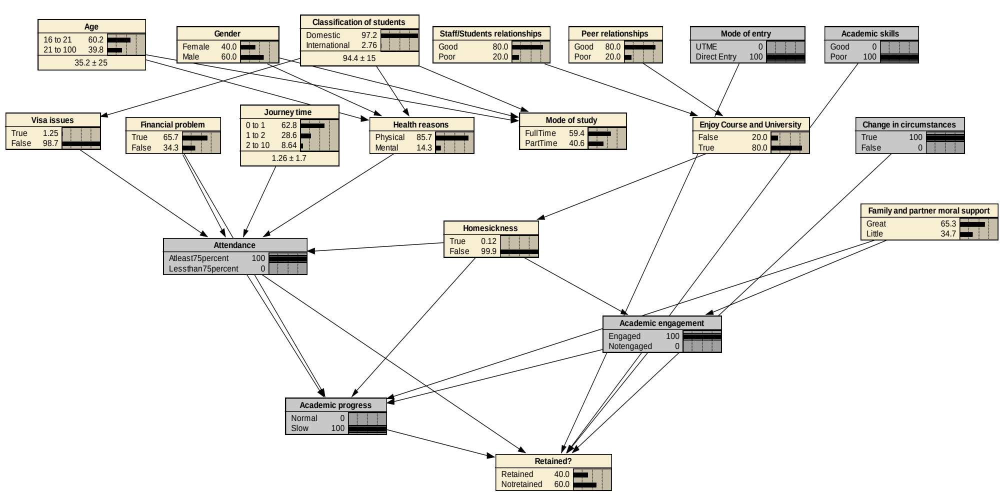


Figure 4-29: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 40%.

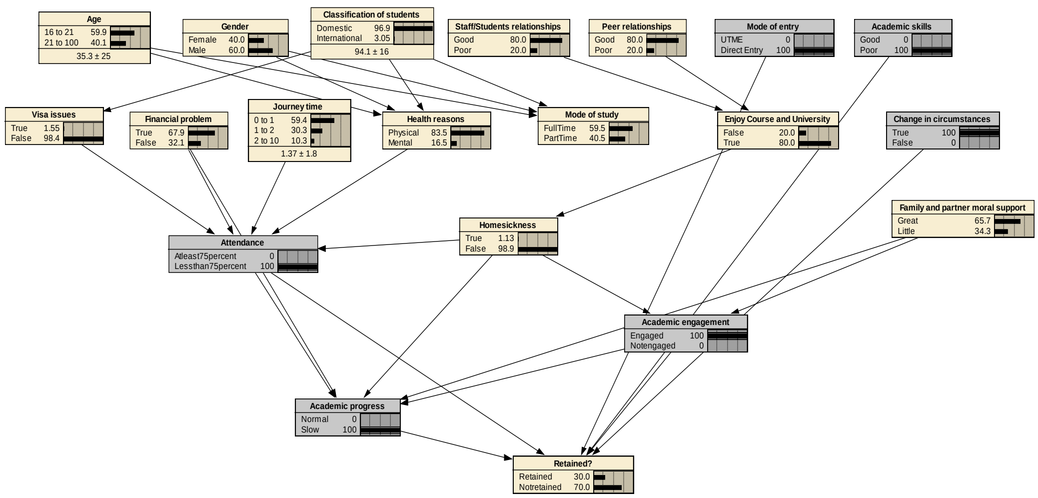


Figure 4-30: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 30%.

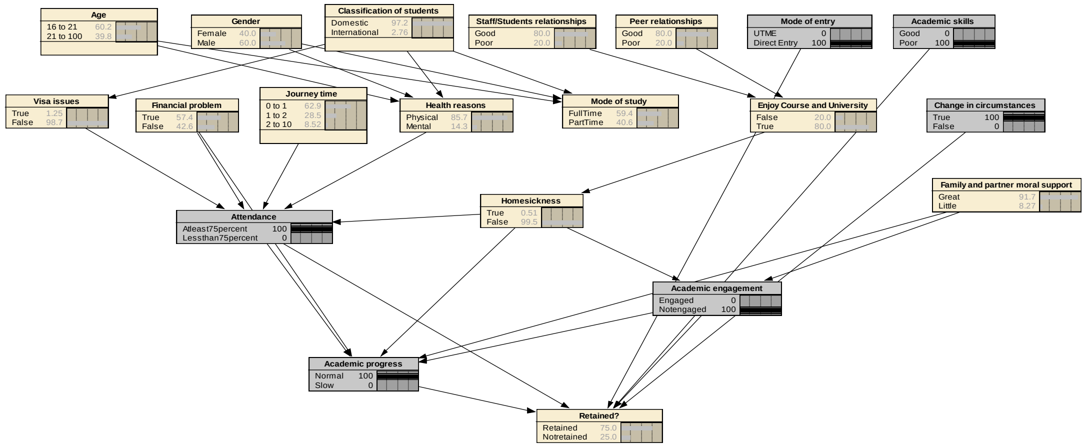


Figure 4-31: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 20%.

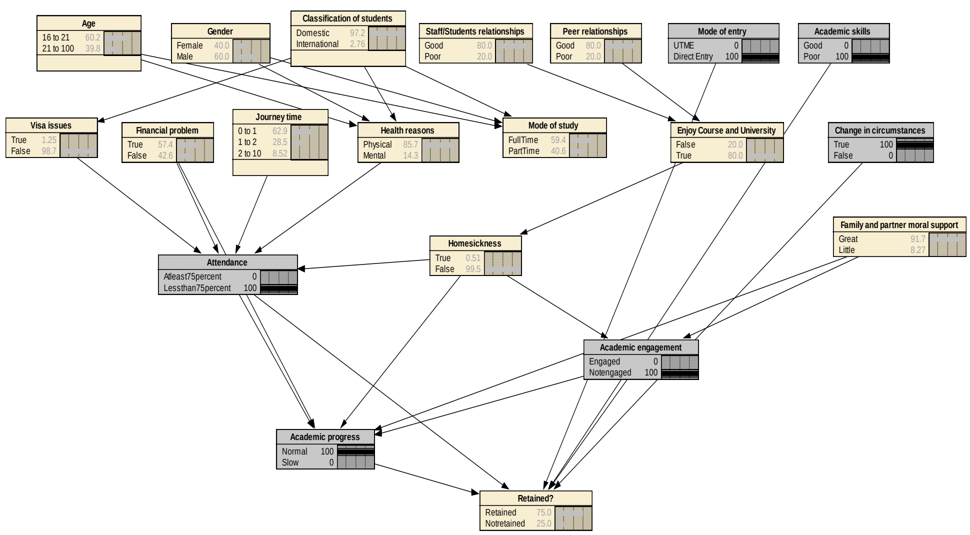


Figure 4-32: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic

engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

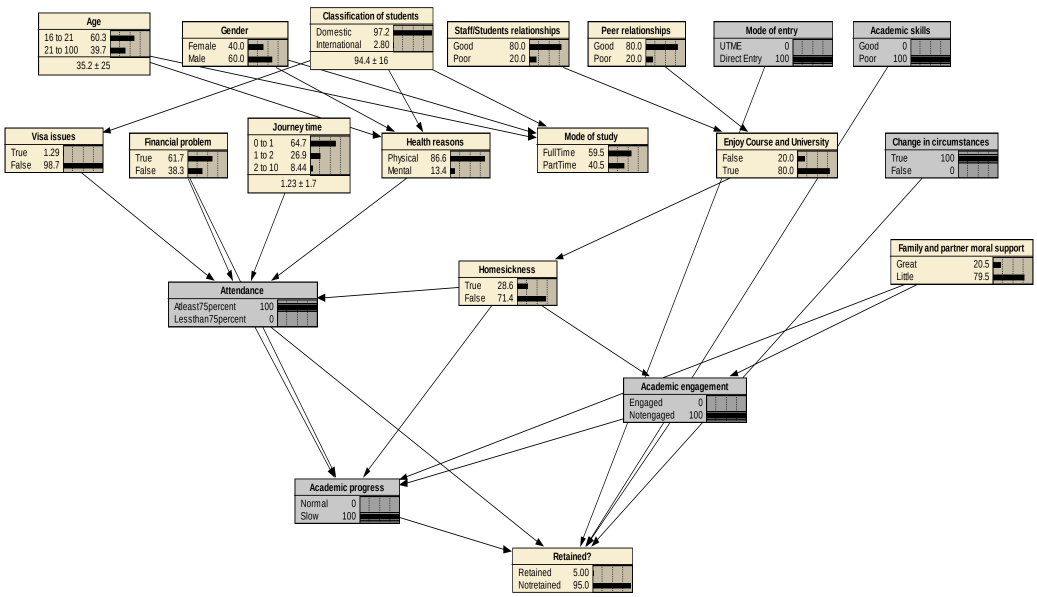


Figure 4-33: Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

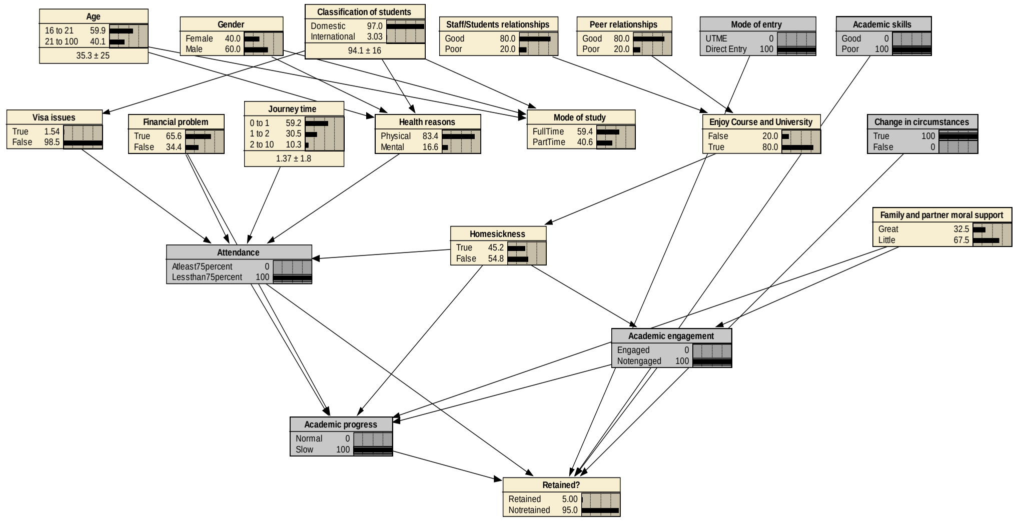


Figure 4-34: Bayesian network prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

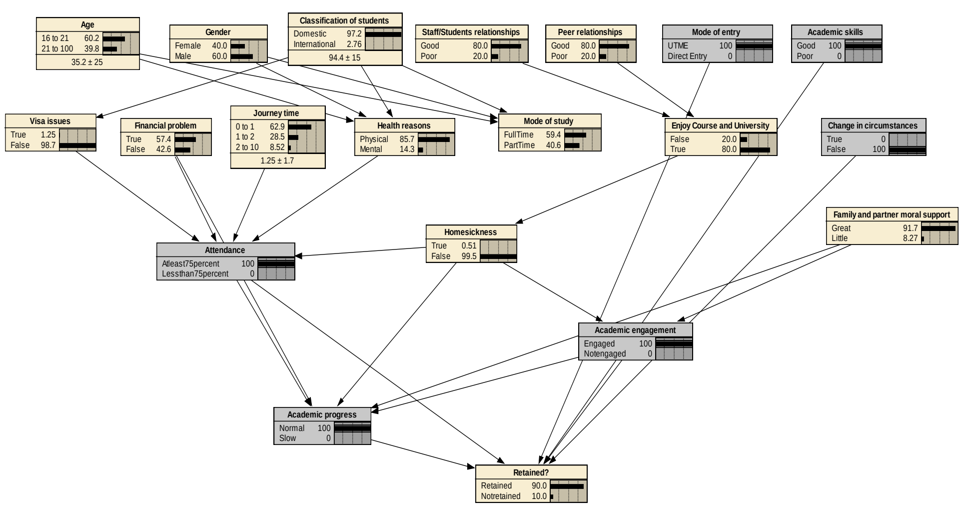


Figure 4-35: Bayesian network prediction for a student whose Change in circumstances is

‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 90%.

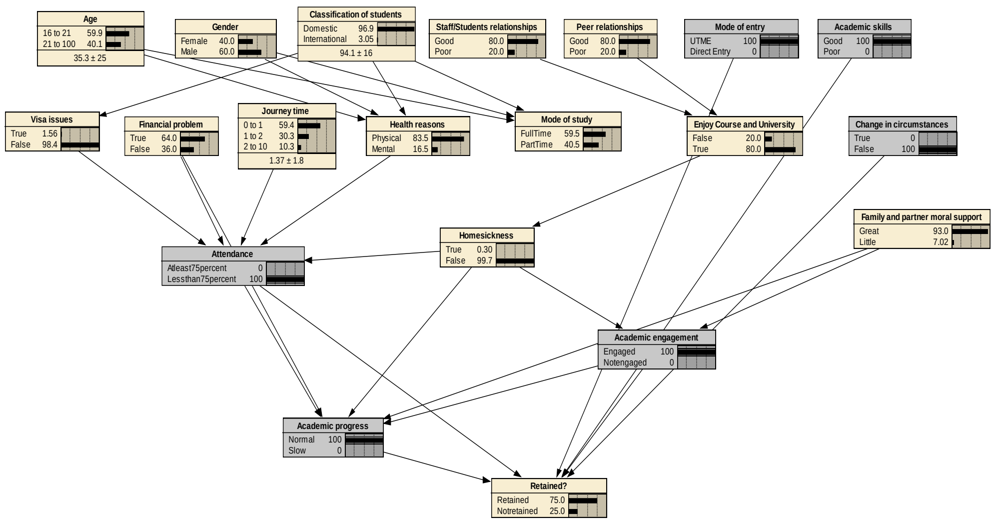


Figure 4-36: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very low risk of attrition as his/her probability of being retained is 75%.

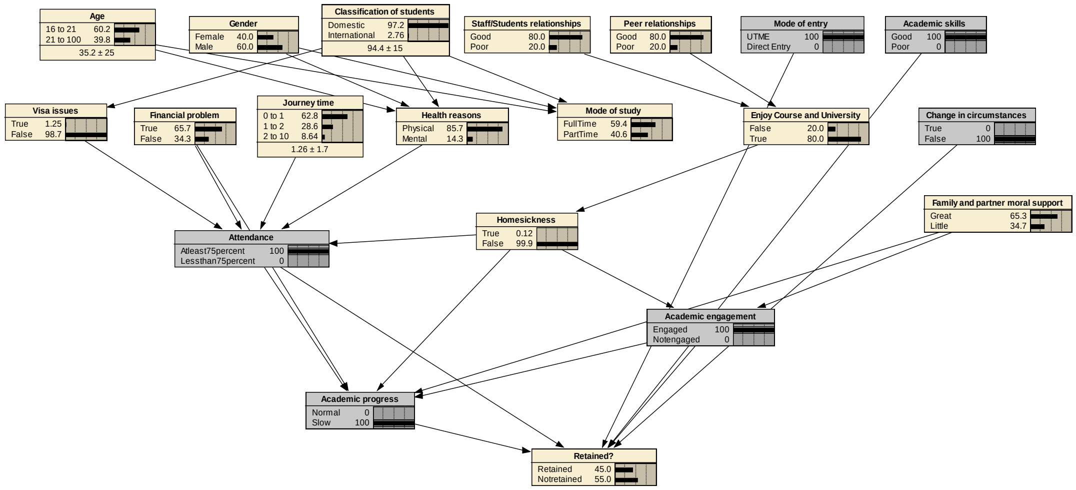


Figure 4-37: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a fairly high risk of attrition as his/her probability of being retained is 45%.

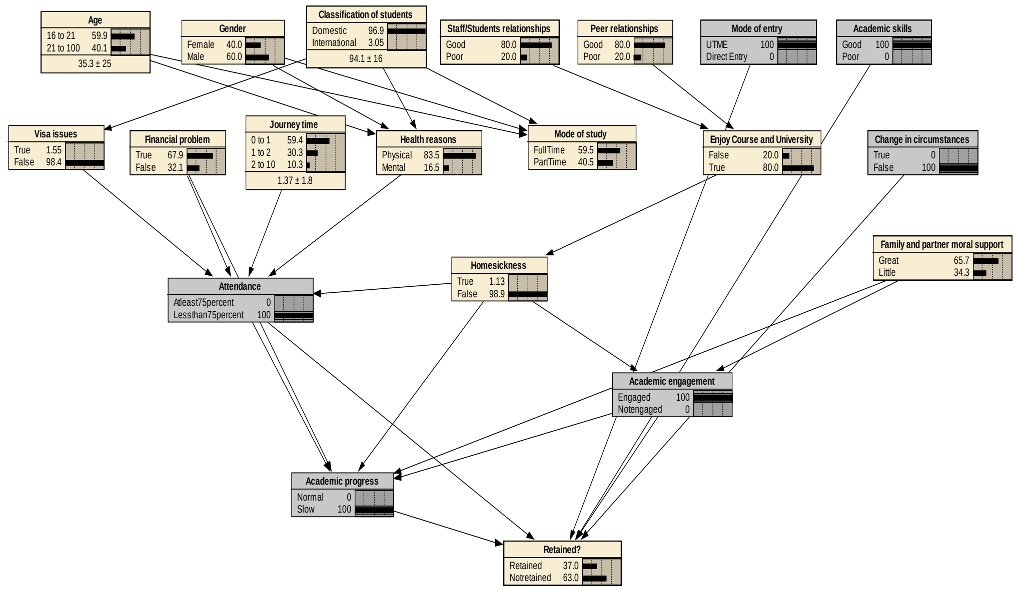


Figure 4-38: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 37%.

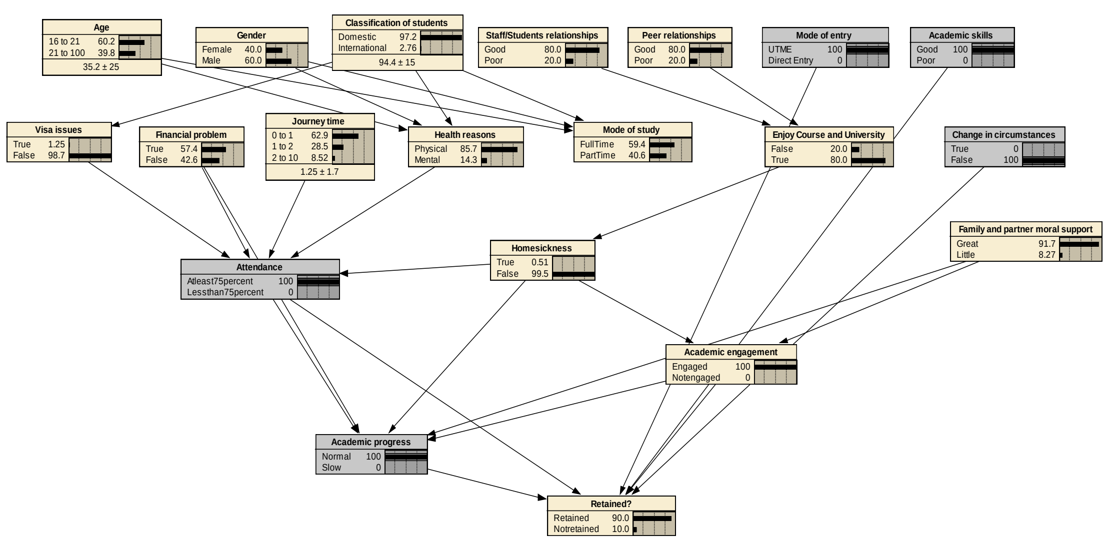


Figure 4-39: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 20%.

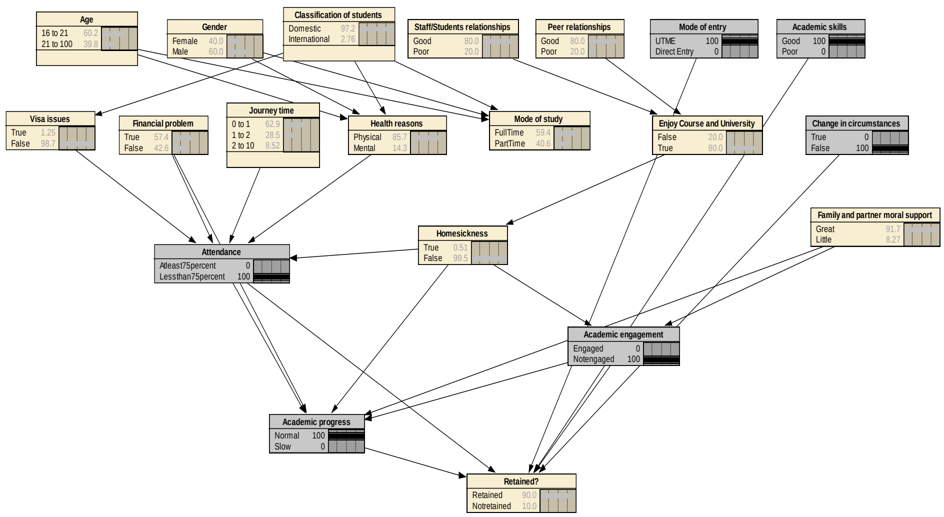


Figure 4-40: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘good’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

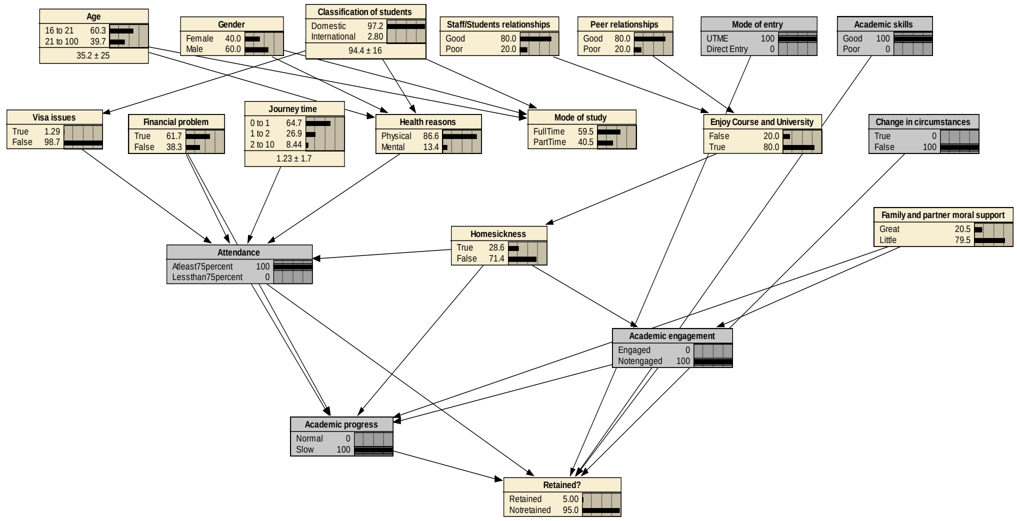


Figure 4-41: Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

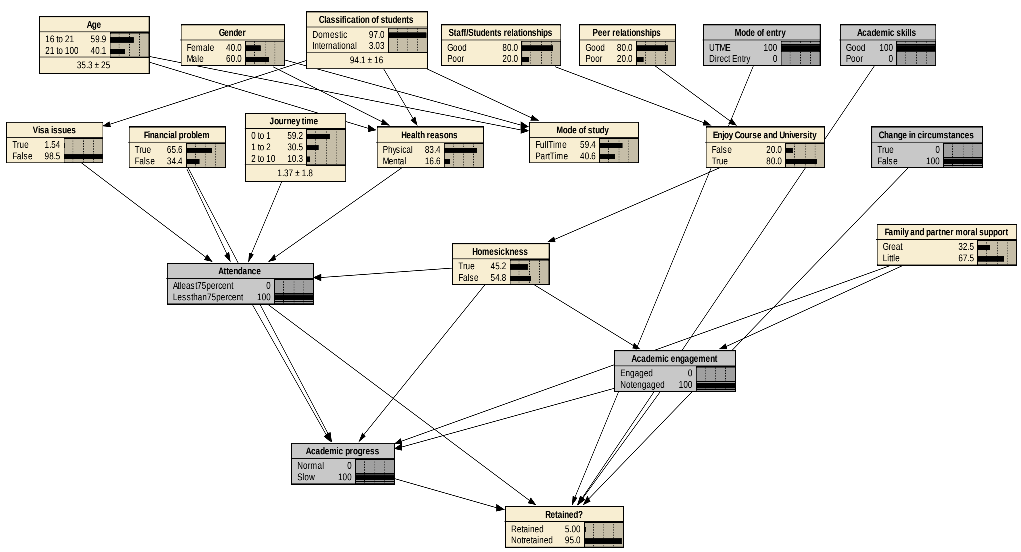


Figure 4-42: Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

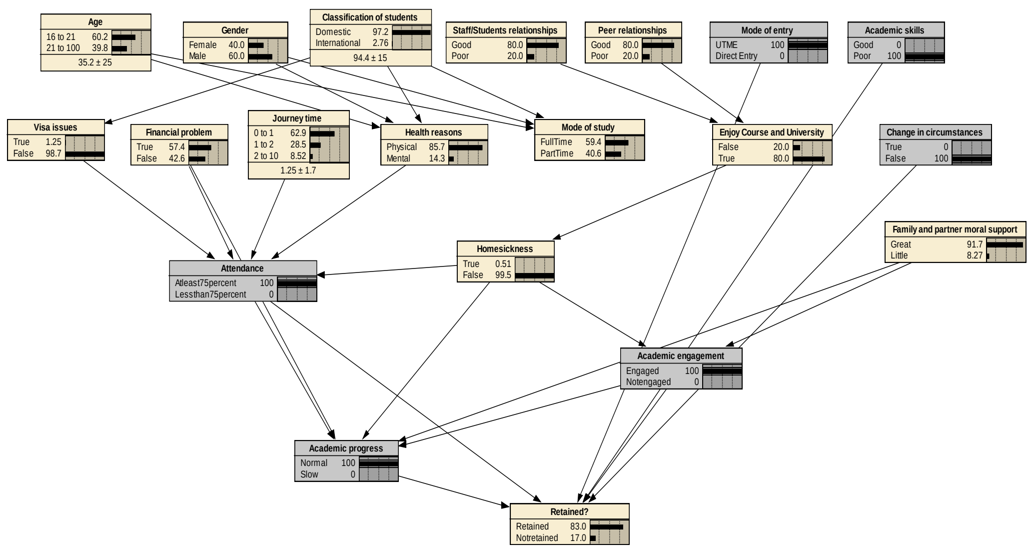


Figure 4-43: Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 83%.

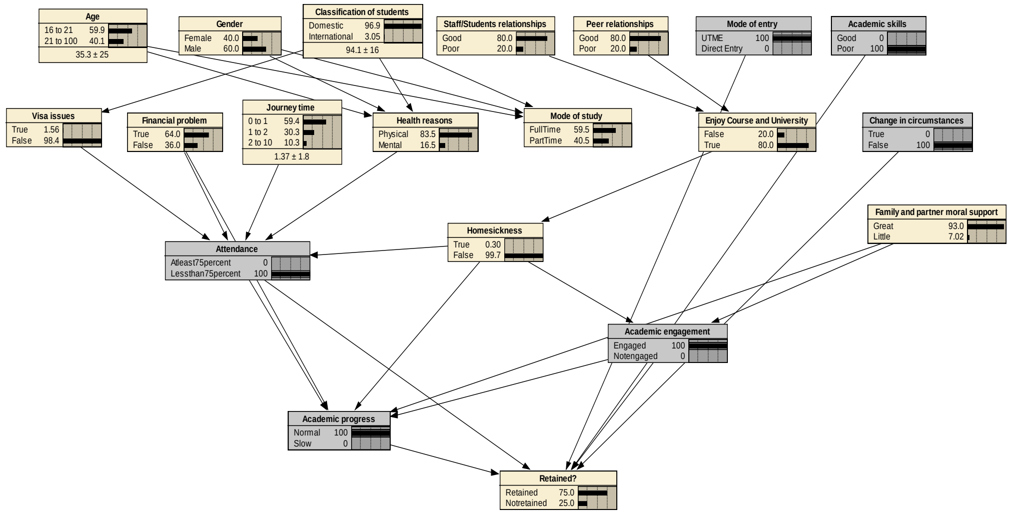


Figure 4-44: Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a low risk of attrition as his/her probability of being retained is 75%.

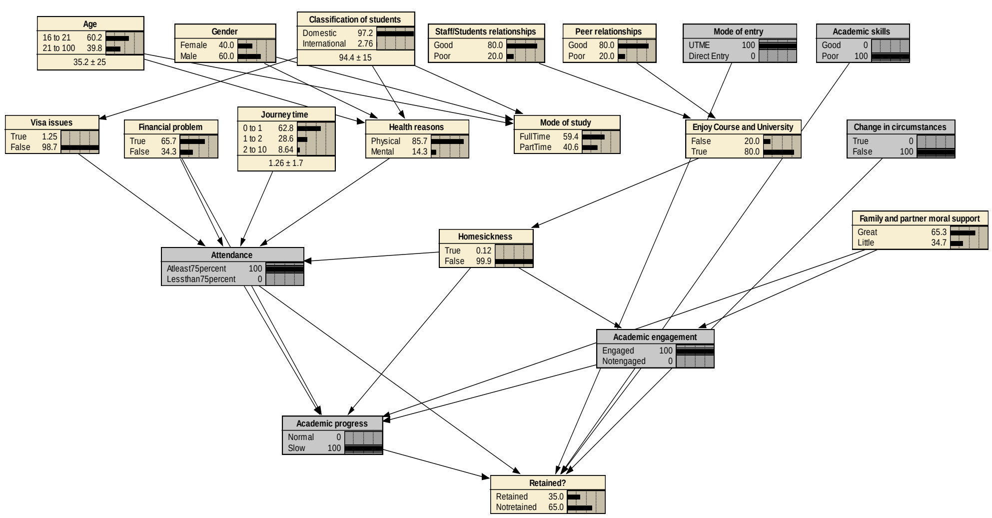


Figure 4-45: Bayesian network prediction for a student whose Change in circumstances is

‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 35%.

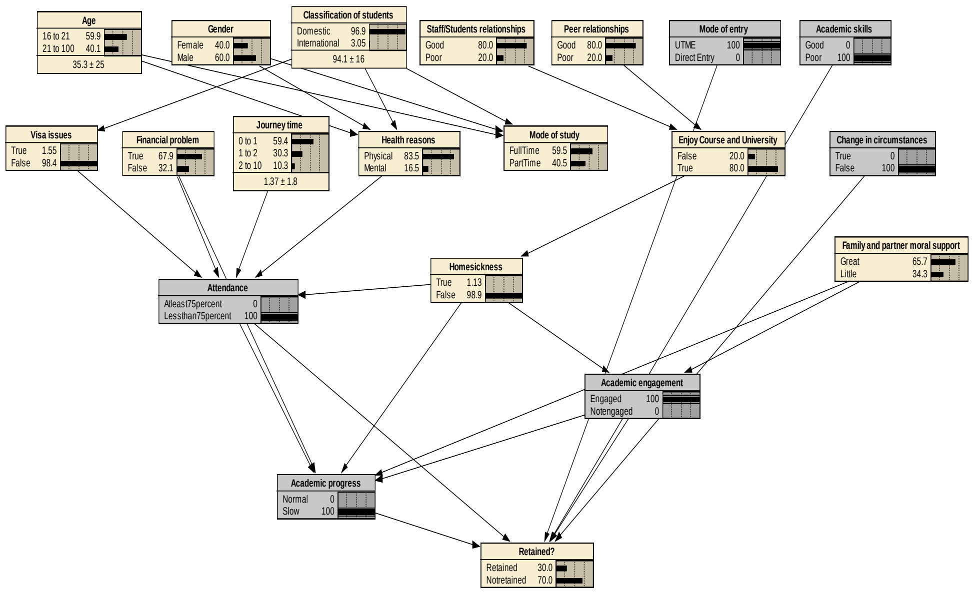


Figure 4-46: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 30%.

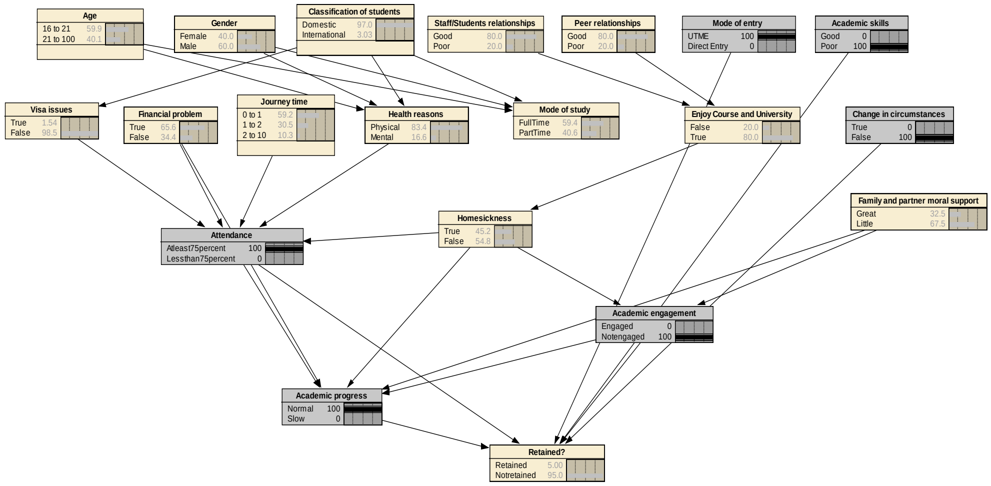


Figure 4-47: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

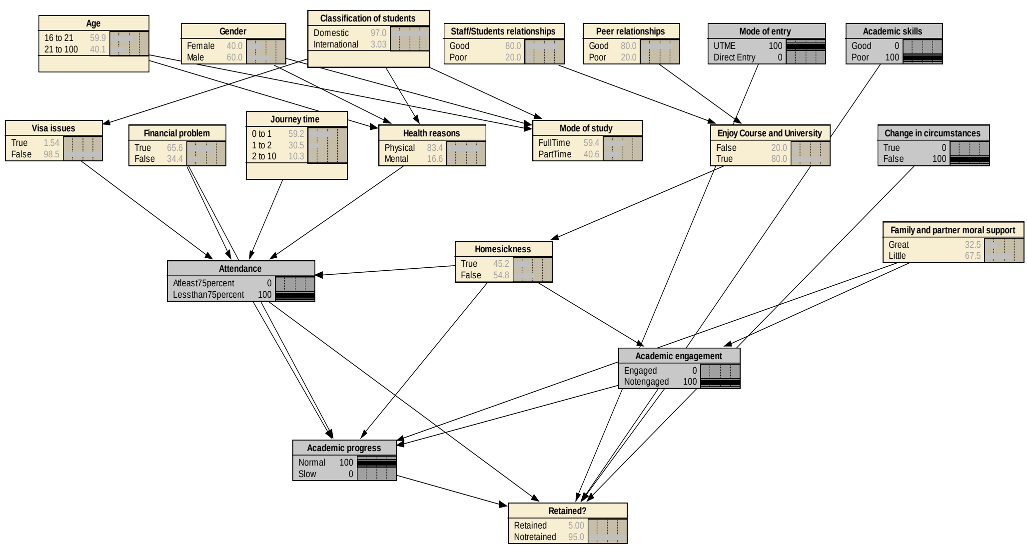


Figure 4-48: Bayesian network prediction for a student whose Change in circumstances is

‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 10%.

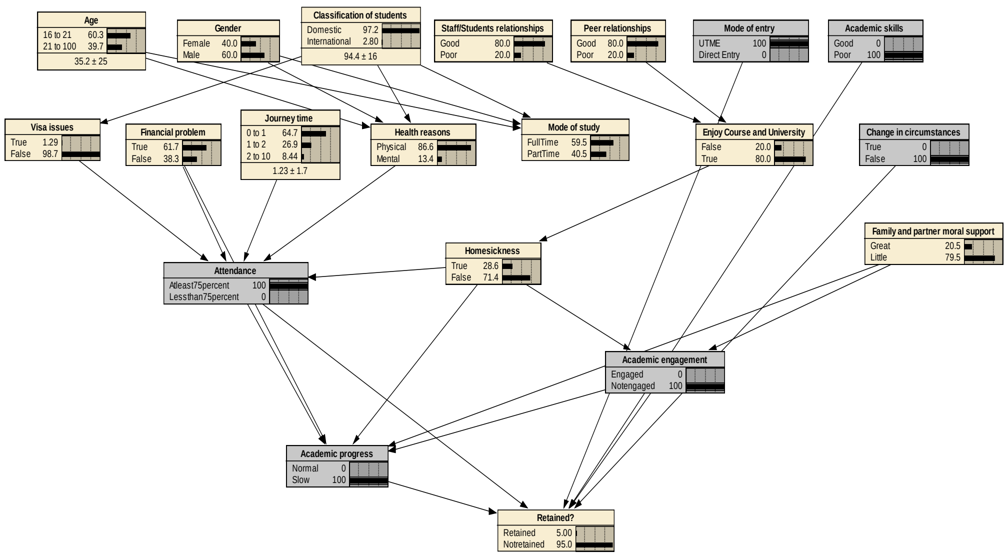


Figure 4-49: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

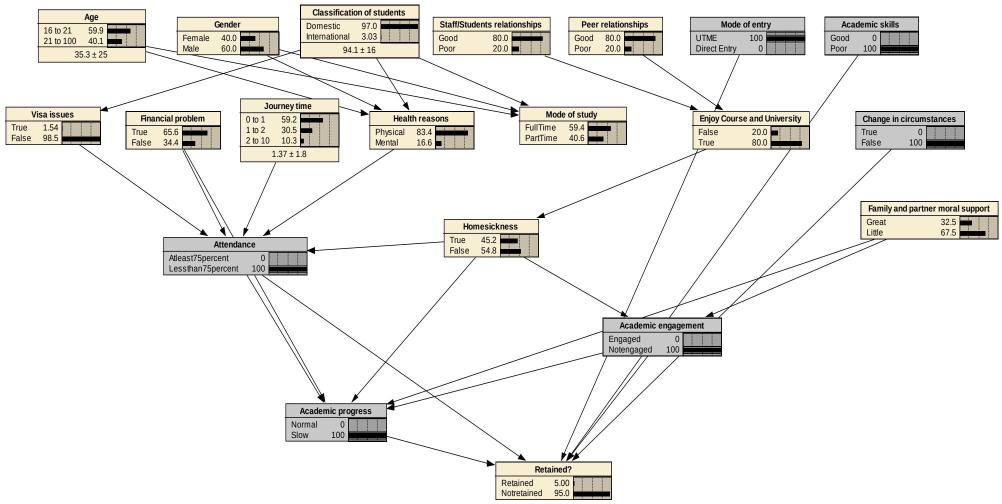


Figure 4-50: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

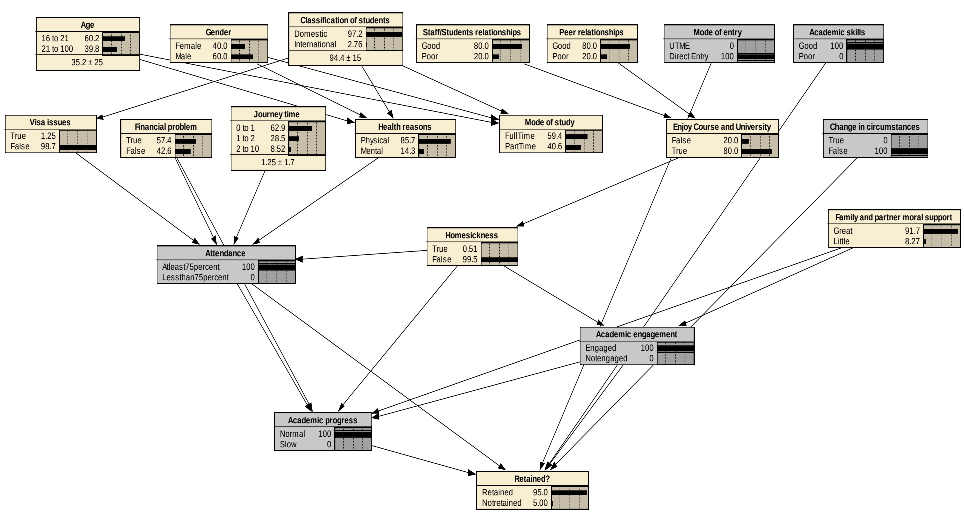


Figure 4-51: Bayesian network prediction for a student whose Change in circumstances is

‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 95%.

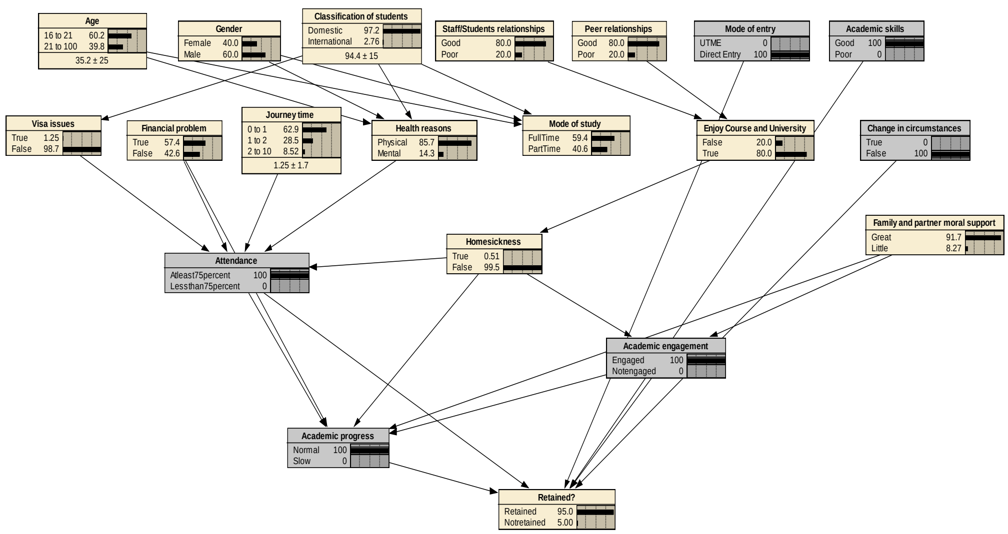


Figure 4-52: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very low risk of attrition as his/her probability of being retained is 90%.

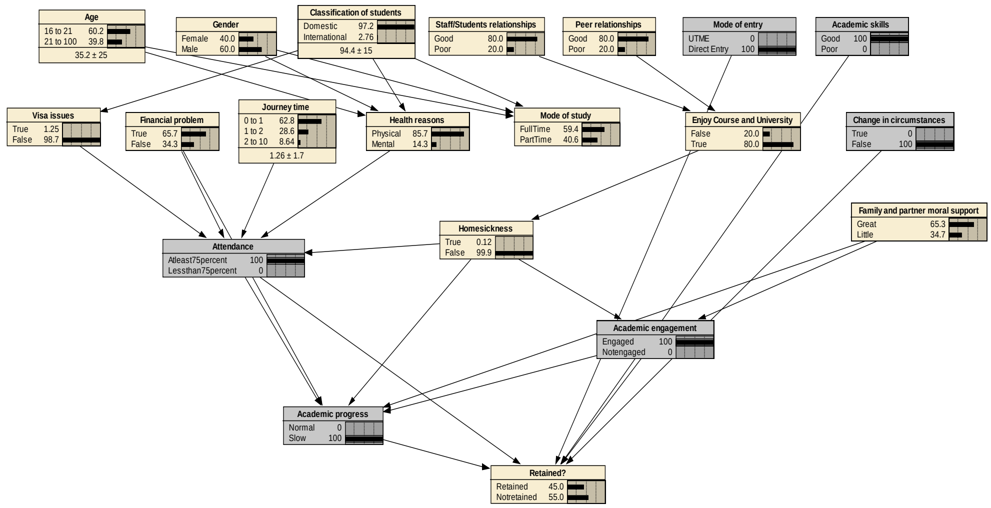


Figure 4-53: Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a fairly high risk of attrition as his/her probability of being retained is 45%.

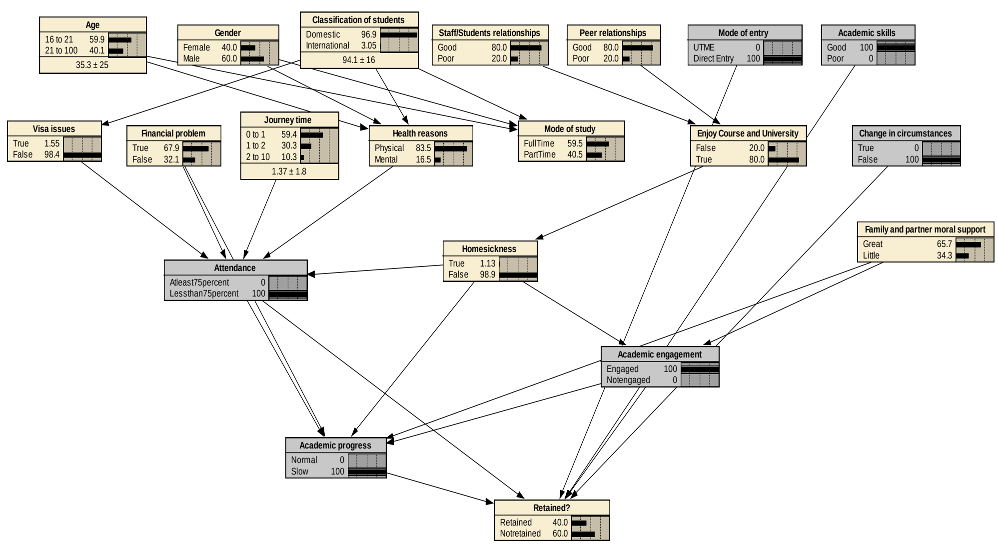


Figure 4-54: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 40%.

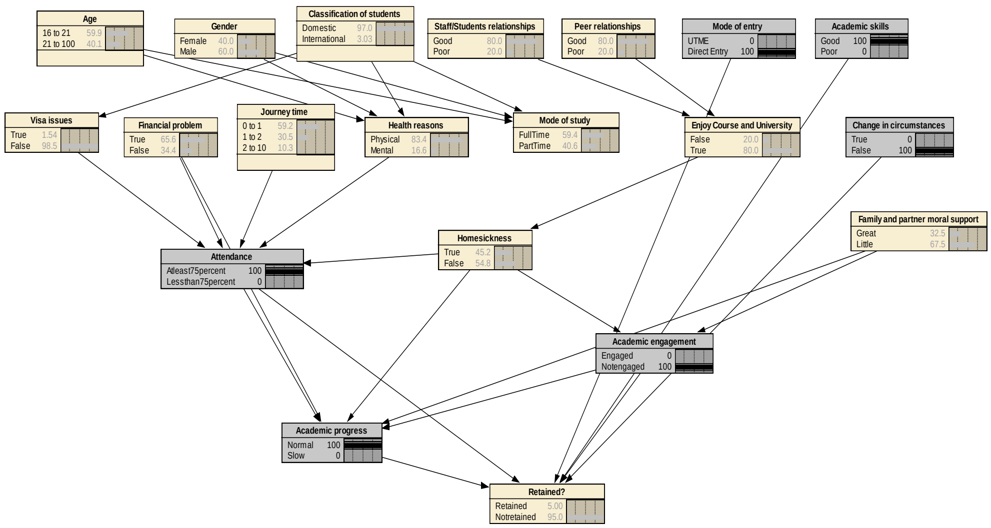


Figure 4-55: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 25%.

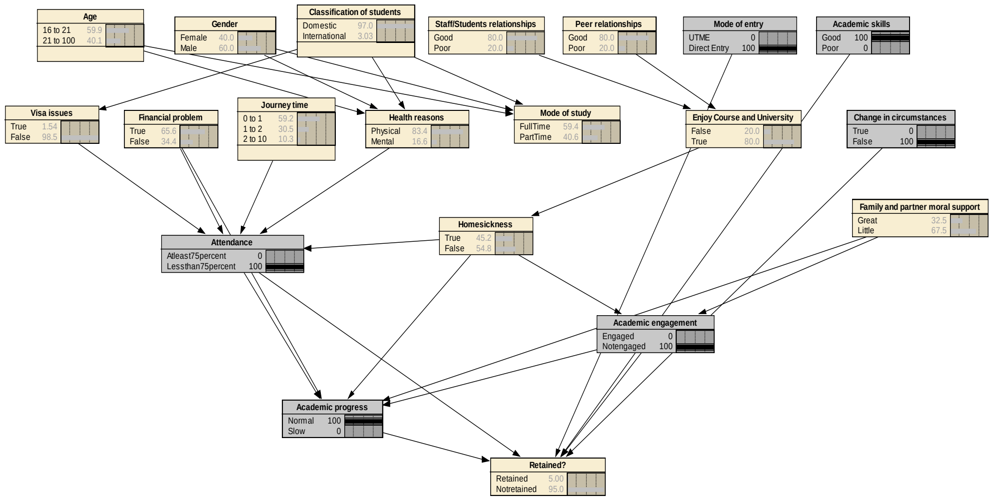


Figure 4-56: Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition. as his/her probability of being retained is 20%.

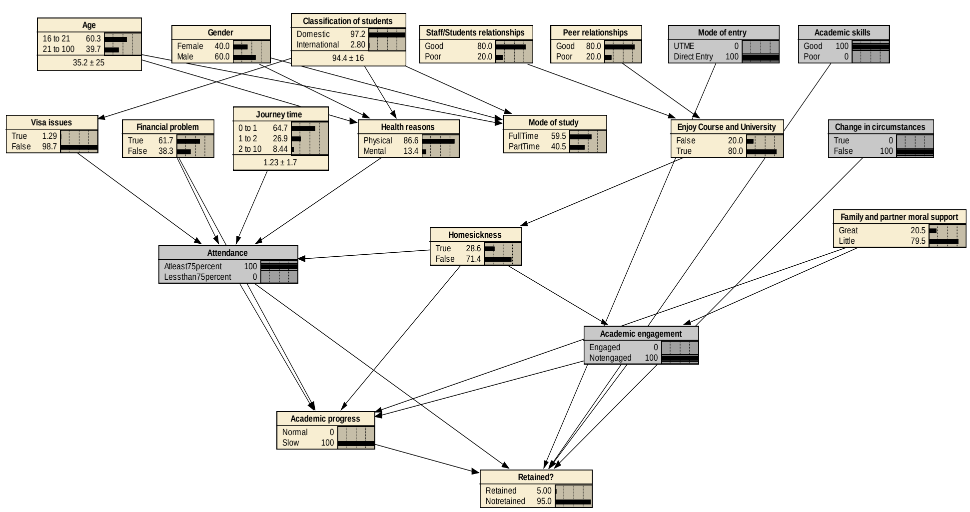


Figure 4-57: Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic

engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

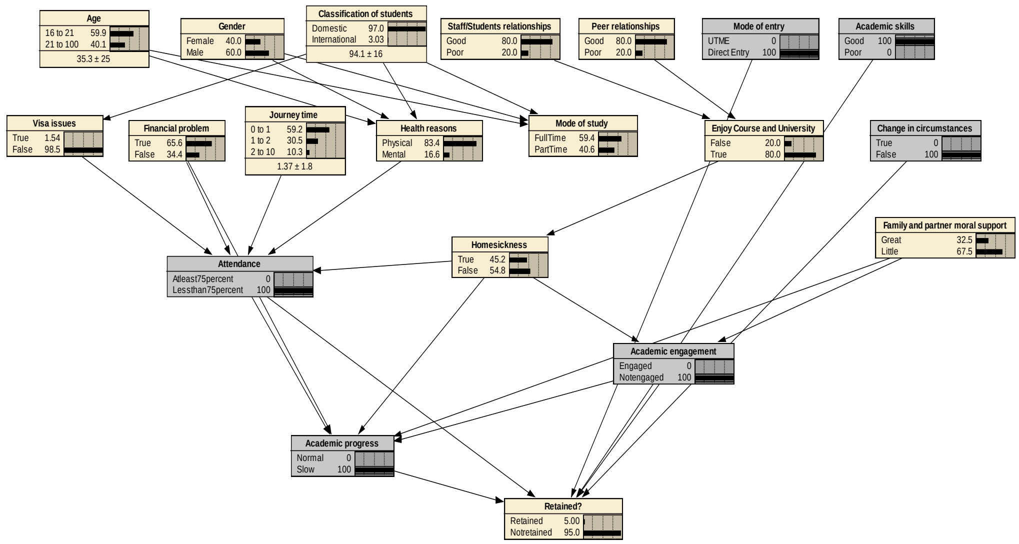


Figure 4-58: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

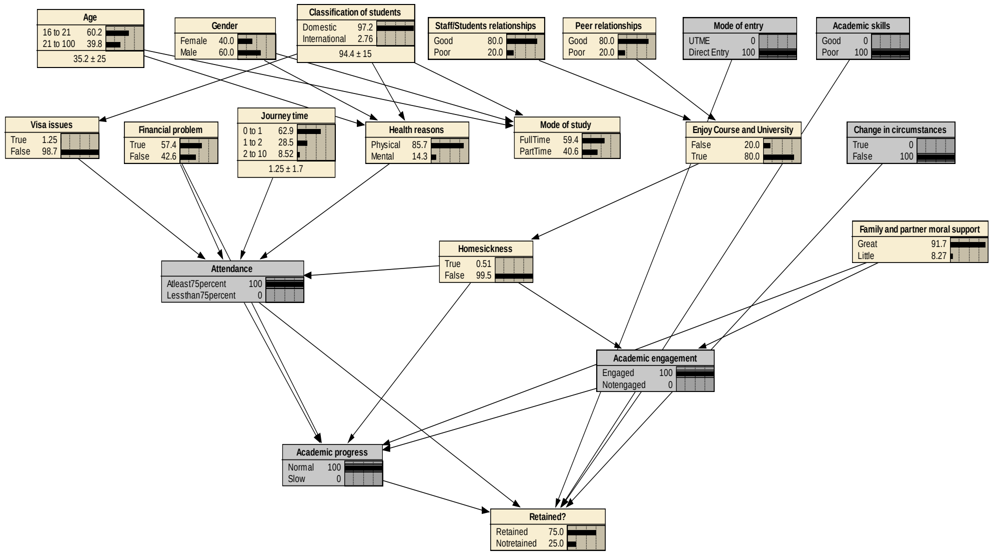


Figure 4-59: Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 75%.

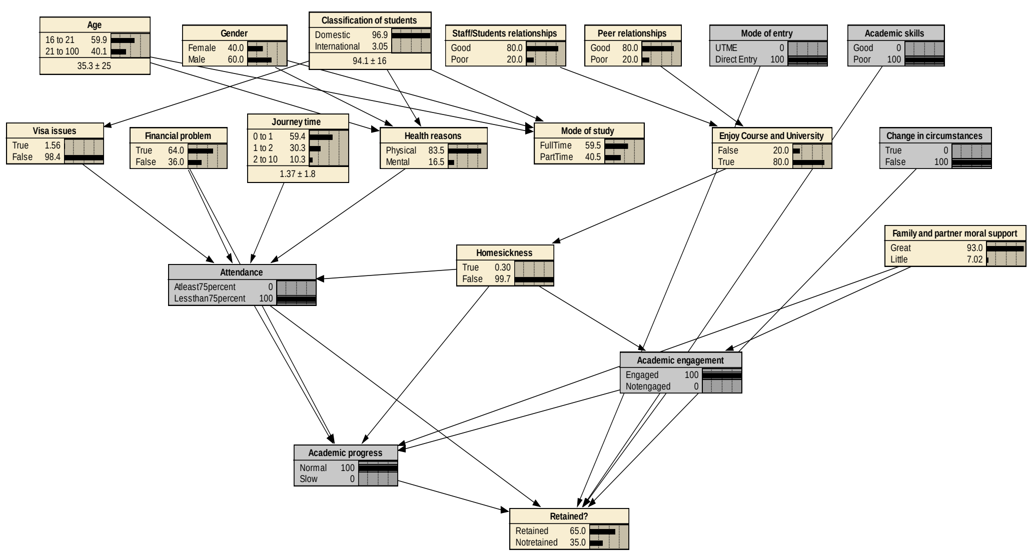


Figure 4-60: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a low risk of attrition as his/her probability of being retained is 65%.

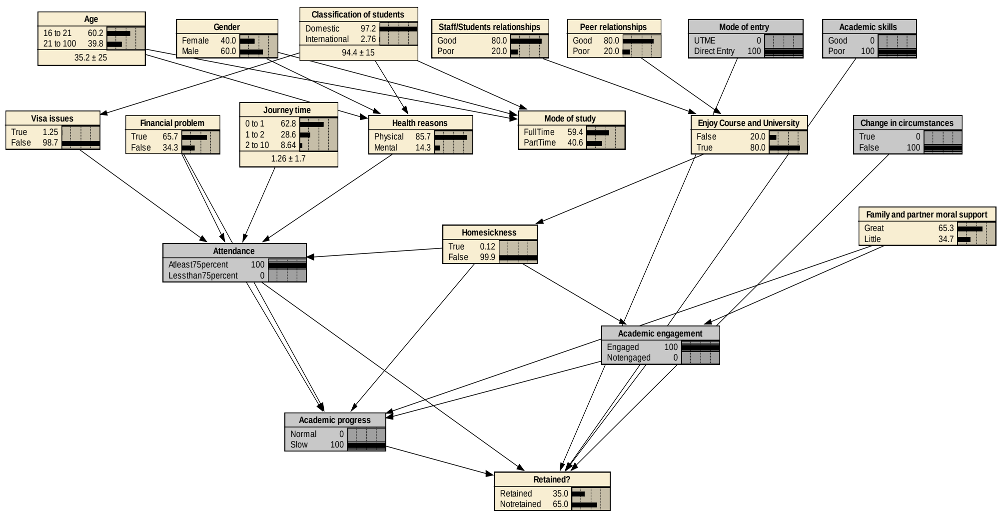


Figure 4-61: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 35%.

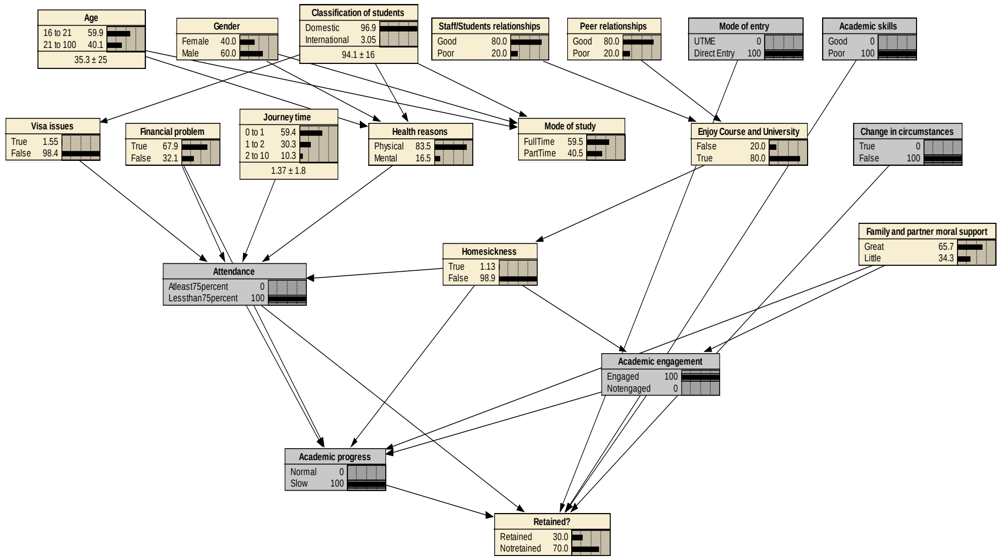


Figure 4-62: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 30%.

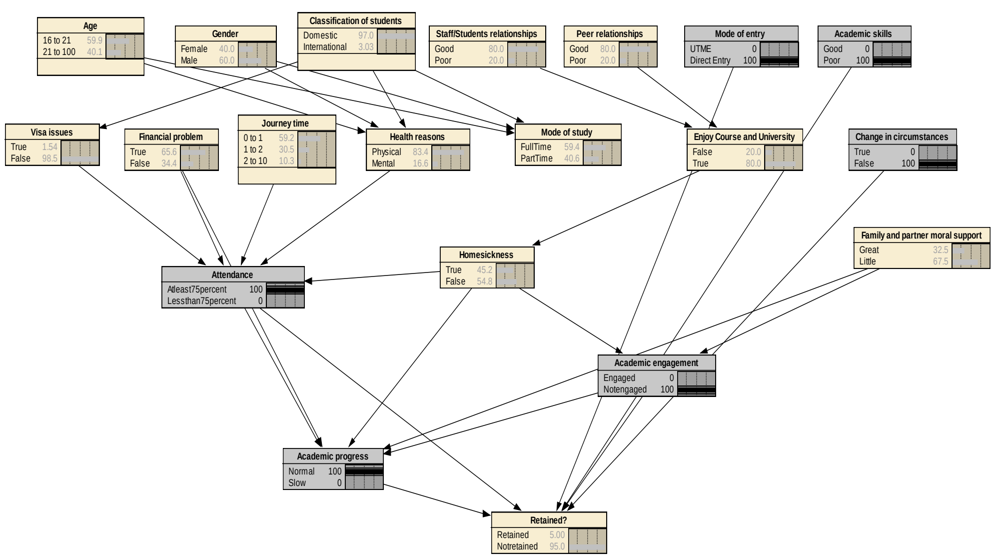


Figure 4-63: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 20%.as his/heras his/her

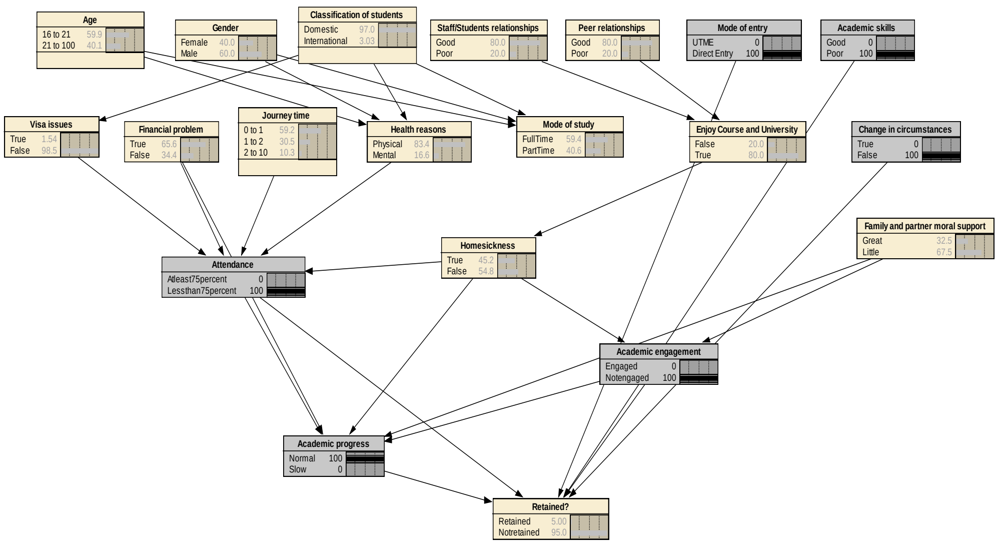


Figure 4-64: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

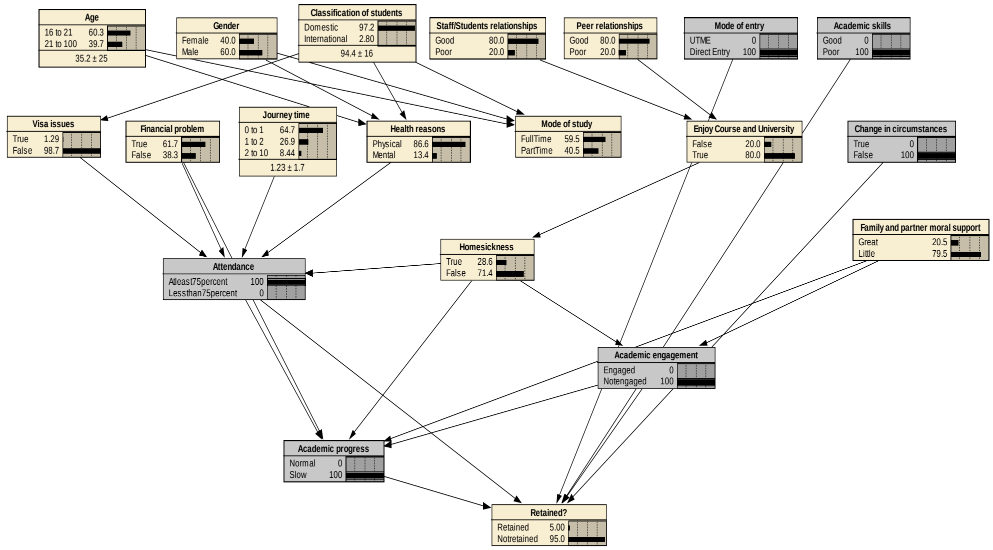


Figure 4-65: Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

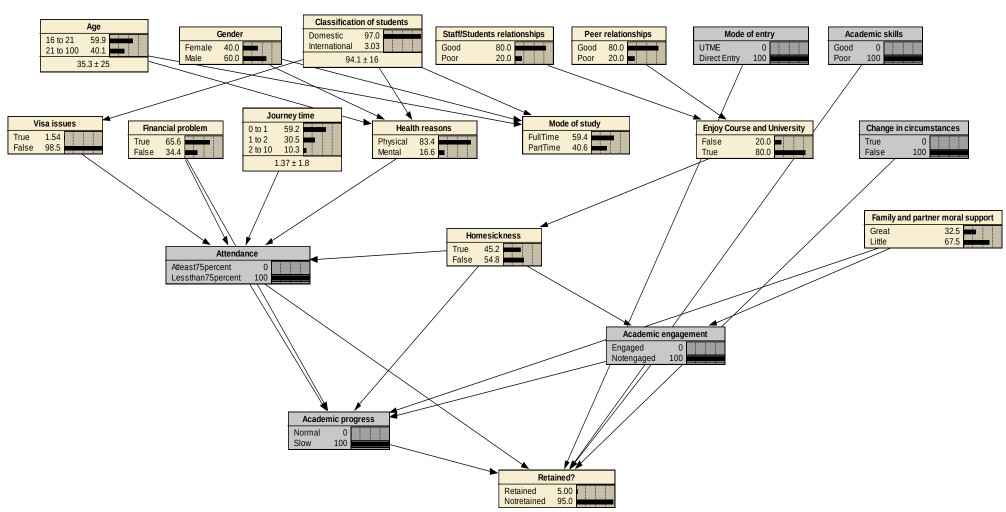


Figure 4-66: Bayesian network prediction for a student whose Change in circumstances is

‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

4.18 Software tool

Prior to the elicitation session, an interactive tool had been developed to predict student retention. The software tool which was developed with the aid of Matlab is capable of displaying the probability that a student would be retained based on combination of the variables in the domain. An end-user of the tool will be able to input some variables and get an output as a result of combining the variables. Thereafter, the probability of a student being retained will be displayed by the tool.

The software tool is menu-driven, which makes it easier for the end-users to utilise. Unlike some tools, an end-user does not need to have knowledge of statistics and probability, to use this tool. The tool has push buttons (command buttons) that end-users need to interact with in order to perform certain tasks. A screenshot of the software tool is shown in Figure 4-67.

Unlike the oldest tools such as probability scale and probability wheel that require the experts to assign the estimates of probabilities by mere looking at a linear scale whose gauge is from 0 to 1, or a circular scale whose gauge is from 0 to 360 degrees, respectively in order to provide their’ judgements.

Also, unlike verbal expression of probabilities which is a non-numerical method of eliciting probabilities that involves using the expressions such as “probable”, “fifty-fifty”, “improbable” etc. which can be interpreted as probability of about 0.85, 0.5 and 0.15, respectively (van der Gaag, et al., 1999). The disadvantage of this method is that a verbal expression has different interpretations to different people.

The experts were provided with guidelines that assisted them in providing probability estimates for the relationships among and between the variables in the domain. This is in order to avoid linguistic uncertainty such as vagueness, insufficient background information about the variables, linguistic ambiguity, and under-specificity.

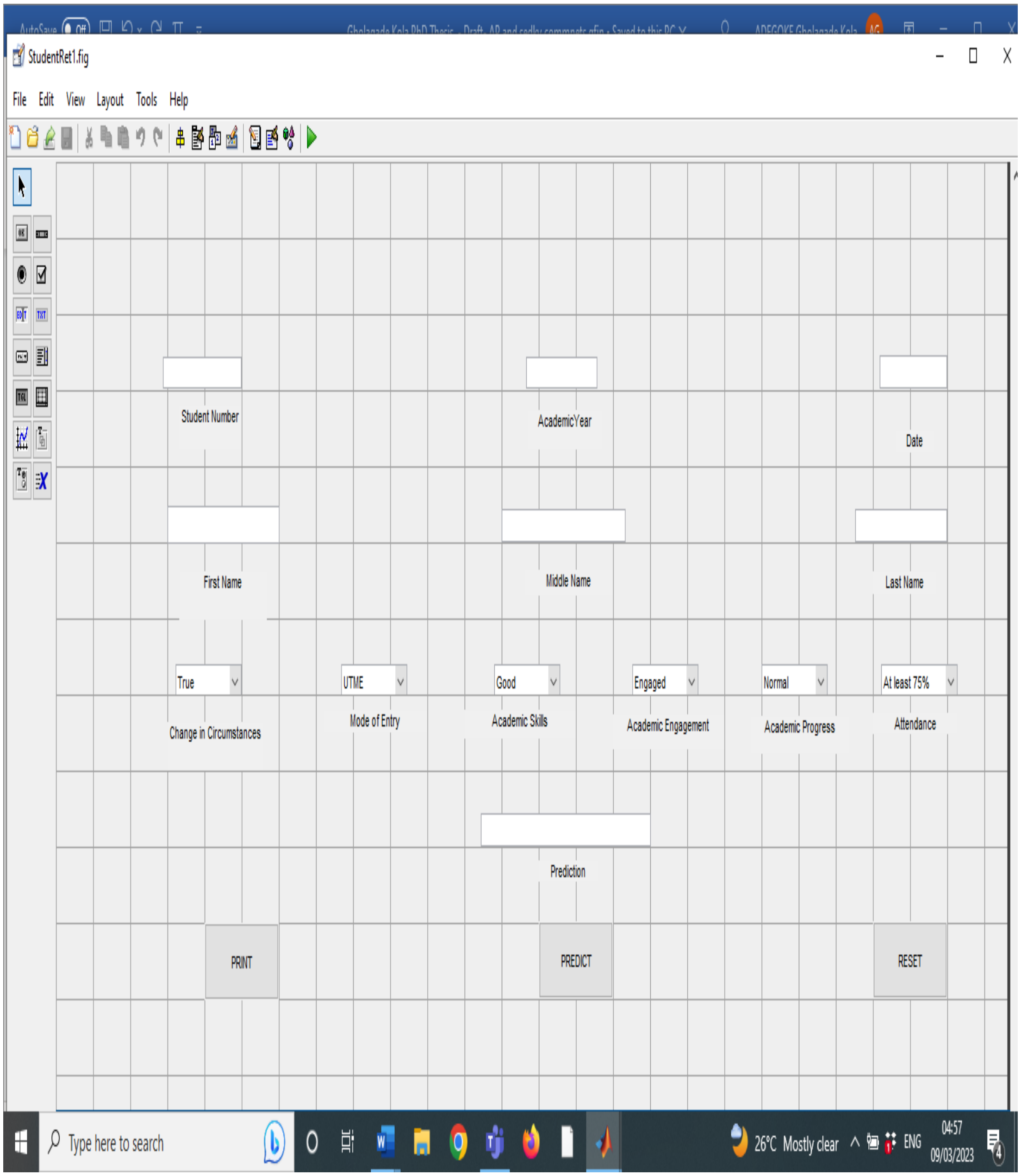
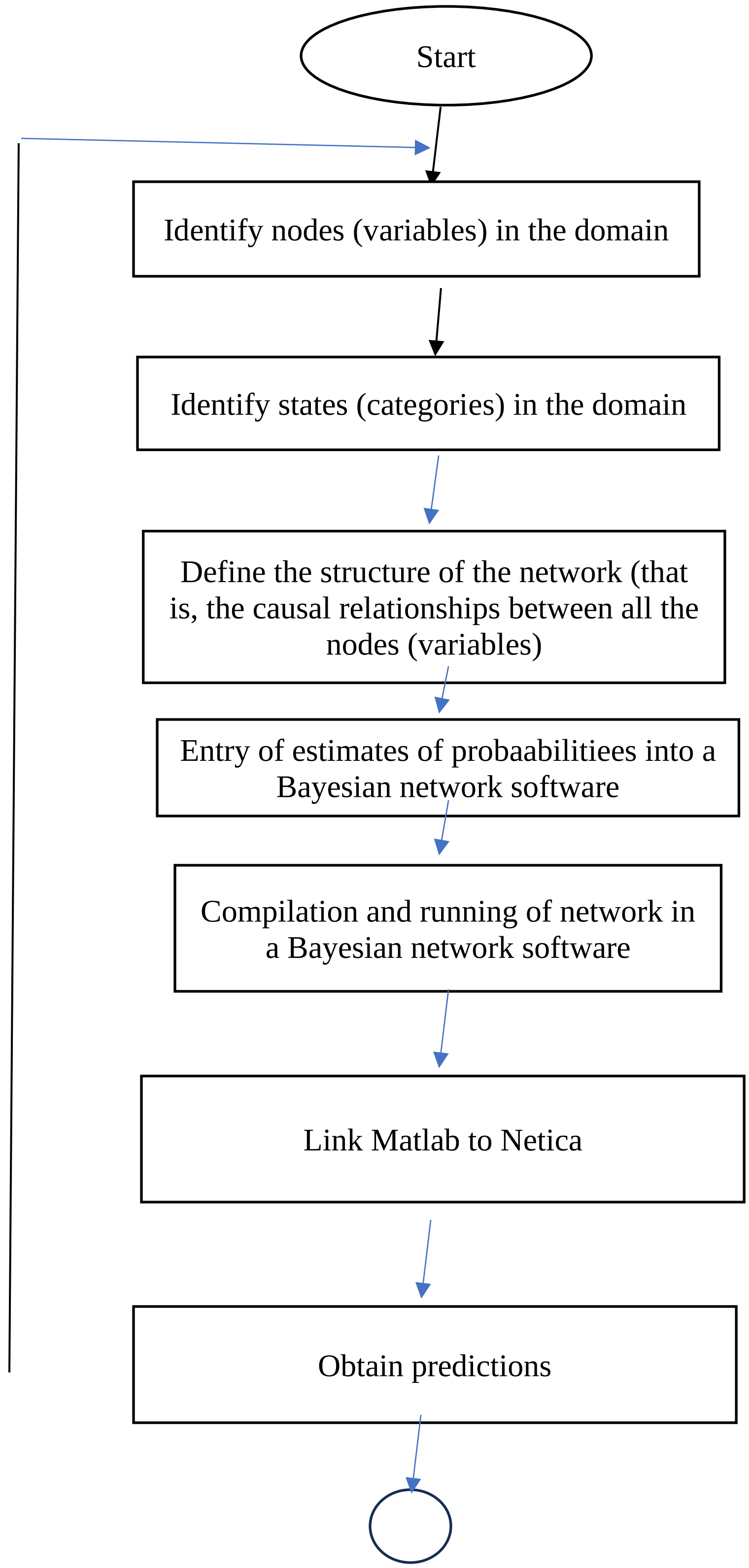


Figure 4-67: A Screenshot of the Software tool


Start
Identify nodes (variables) in the domain
Identify states (categories) in the domain
Define the structure of the network (that is, the causal relationships between all the nodes (variables)
Entry of estimates of probaabilitiees into a Bayesian network software
Compilation and running of network in a Bayesian network software
Link Matlab to Netica
Obtain predictions

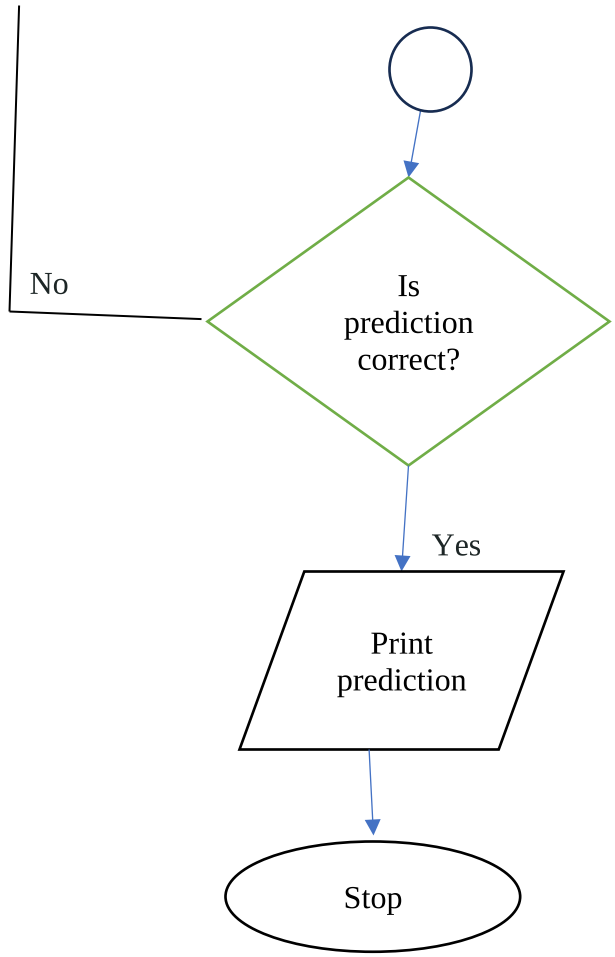


Figure 4-68: Flowchart for results predictions

4.19 Results from the tool

The results produced by the tool are shown in Figure 4-69 to 4-132 below. The different Bayesian networks (BNs) shown in Figure 4-3 to 4-66 produced different results from the software tool.

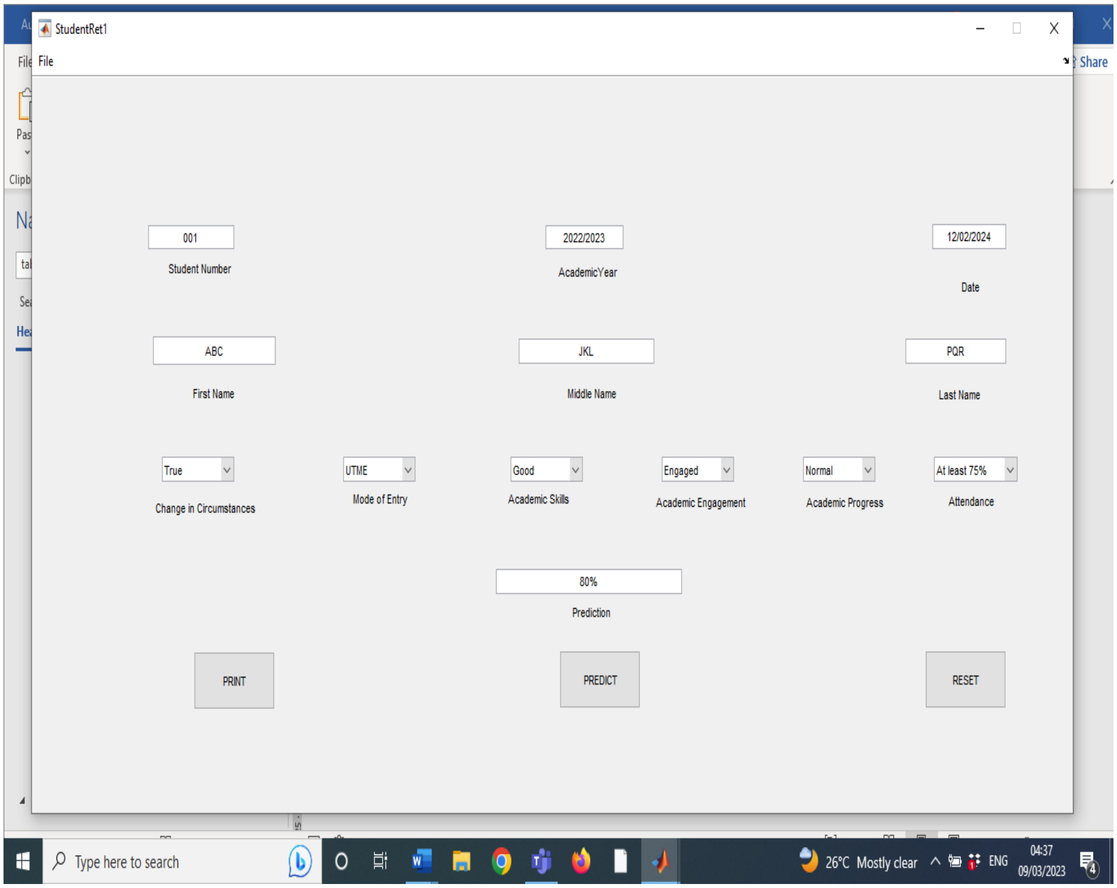


Figure 4-69 (Output for Figure 4-3): Tool's prediction for a student whose Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 80%.

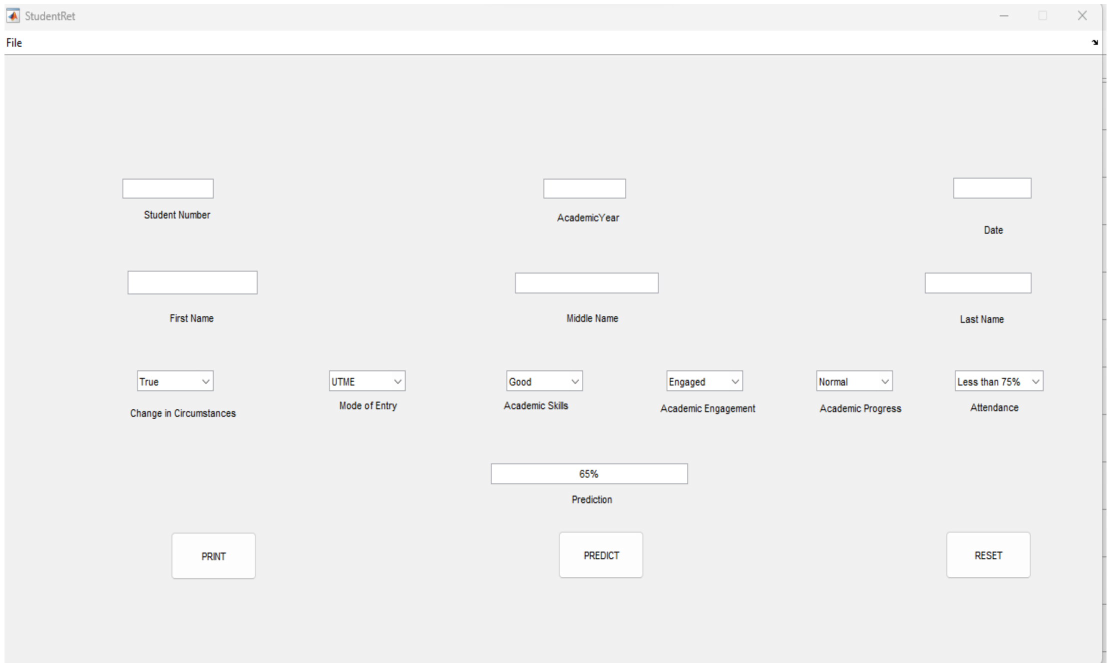


Figure 4-70 (Output for Figure 4-4): Tool's prediction for a student whose Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a low risk of attrition as his/her probability of being retained is 65%.

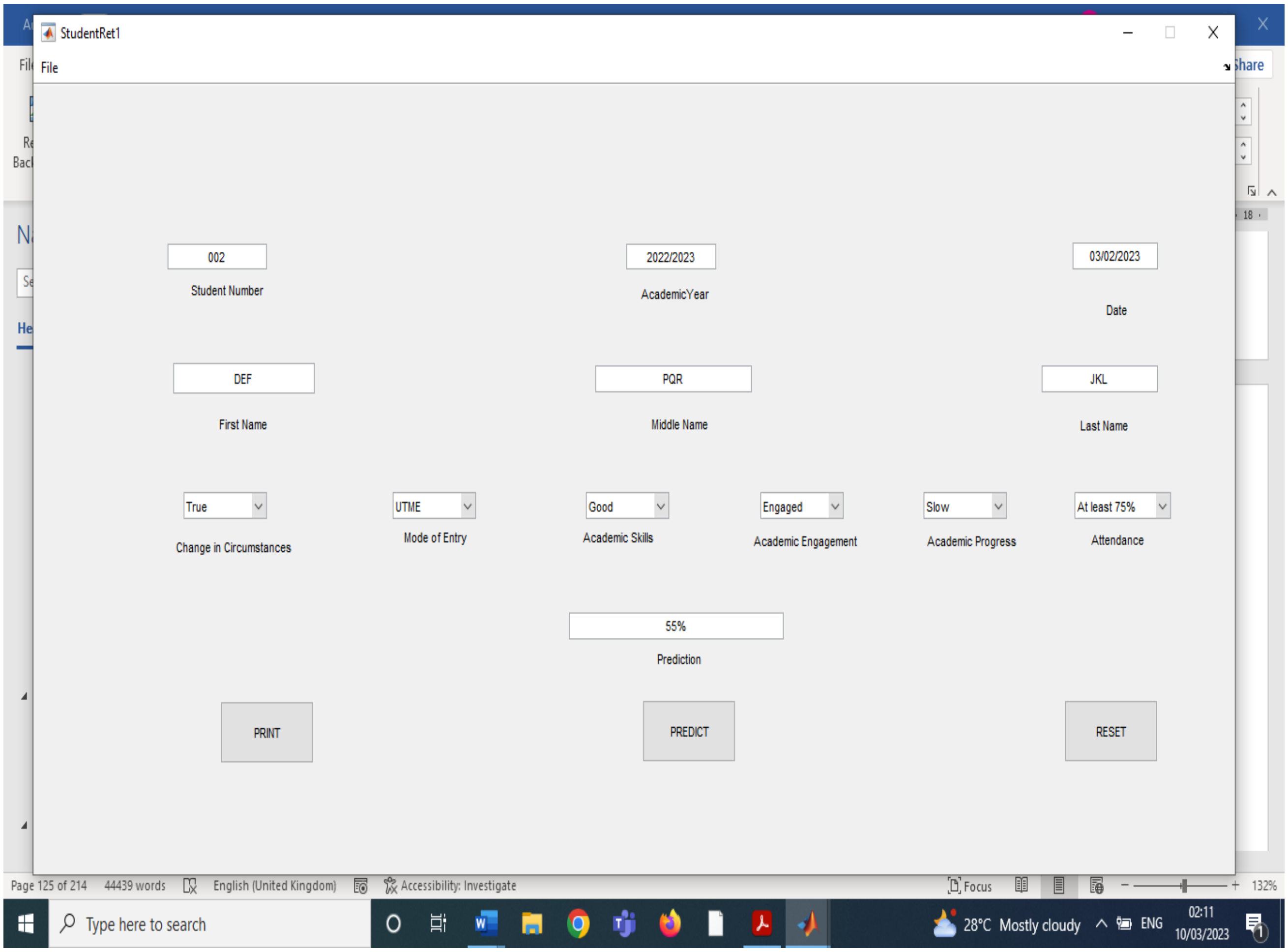


Figure 4-71 (Output for Figure 4-5): Tool's prediction for a student whose Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a fairly high risk of attrition as his/her probability of being retained is 55%.

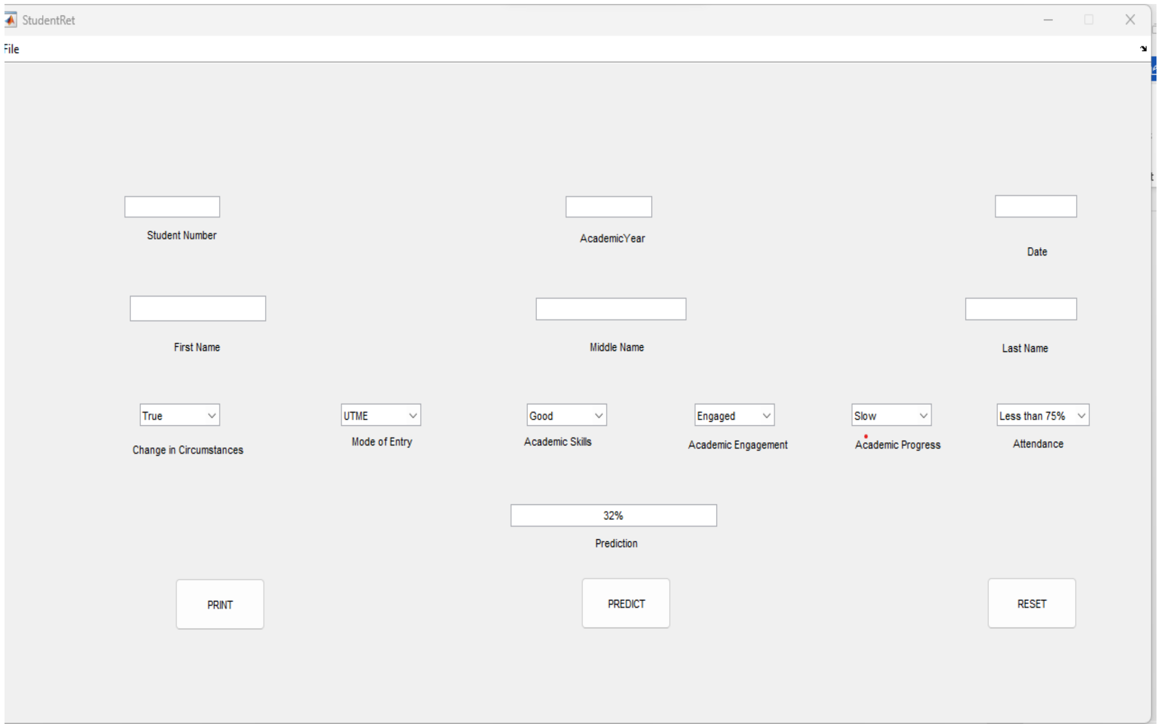


Figure 4-72 (Output for Figure 4-6): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 32%.

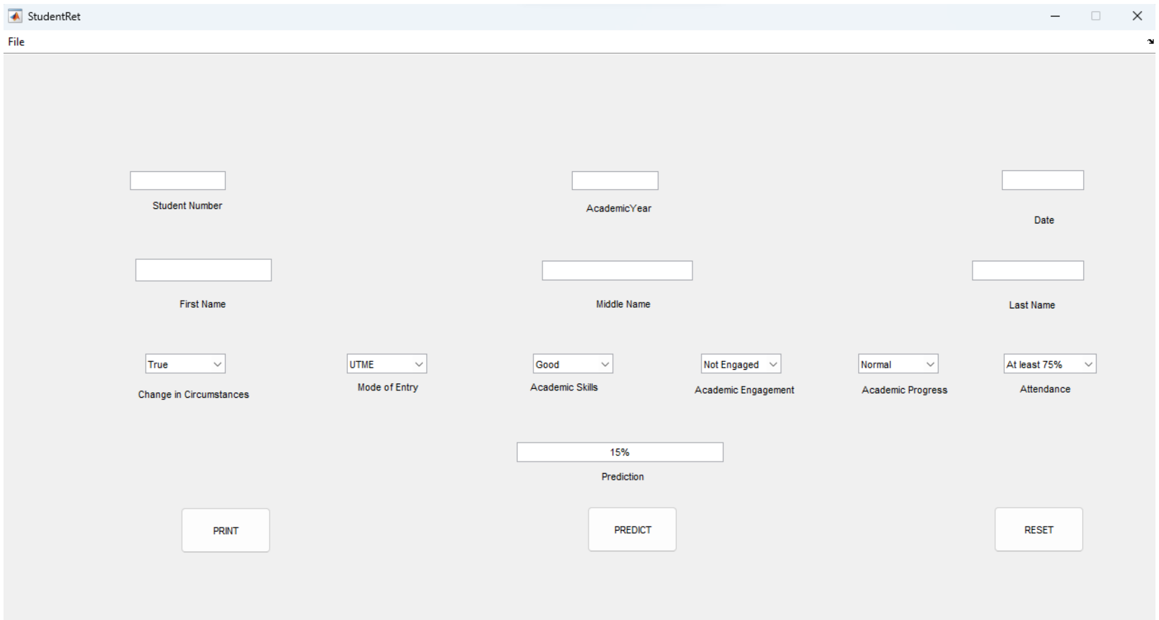


Figure 4-73 (Output for Figure 4-7): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

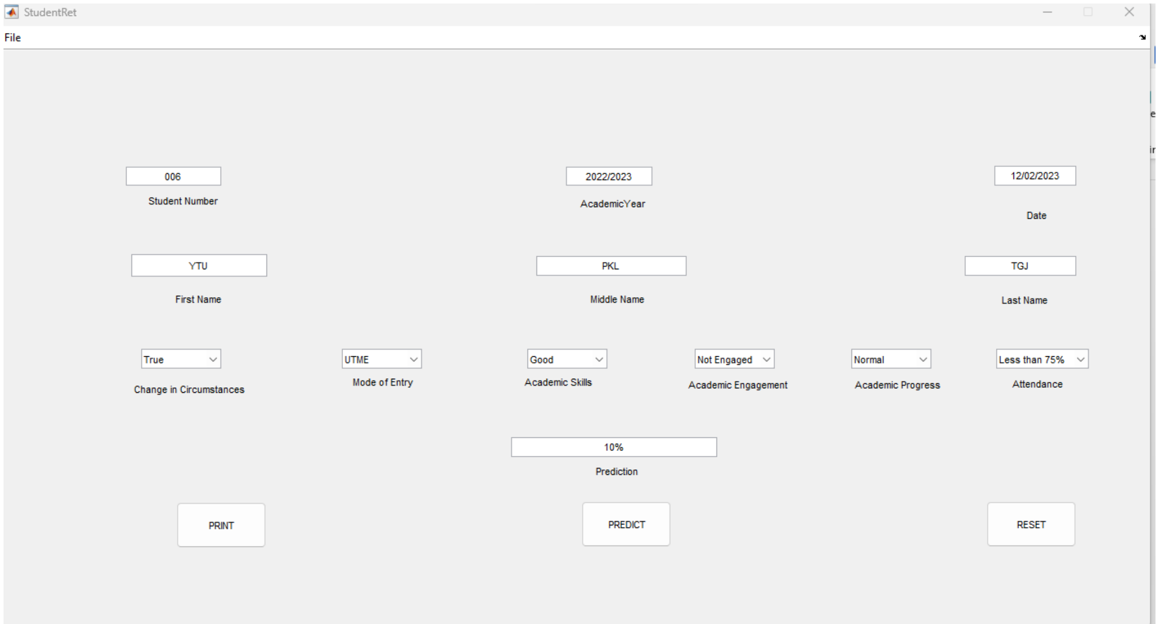


Figure 4-74 (Output for Figure 4-8): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 10%.

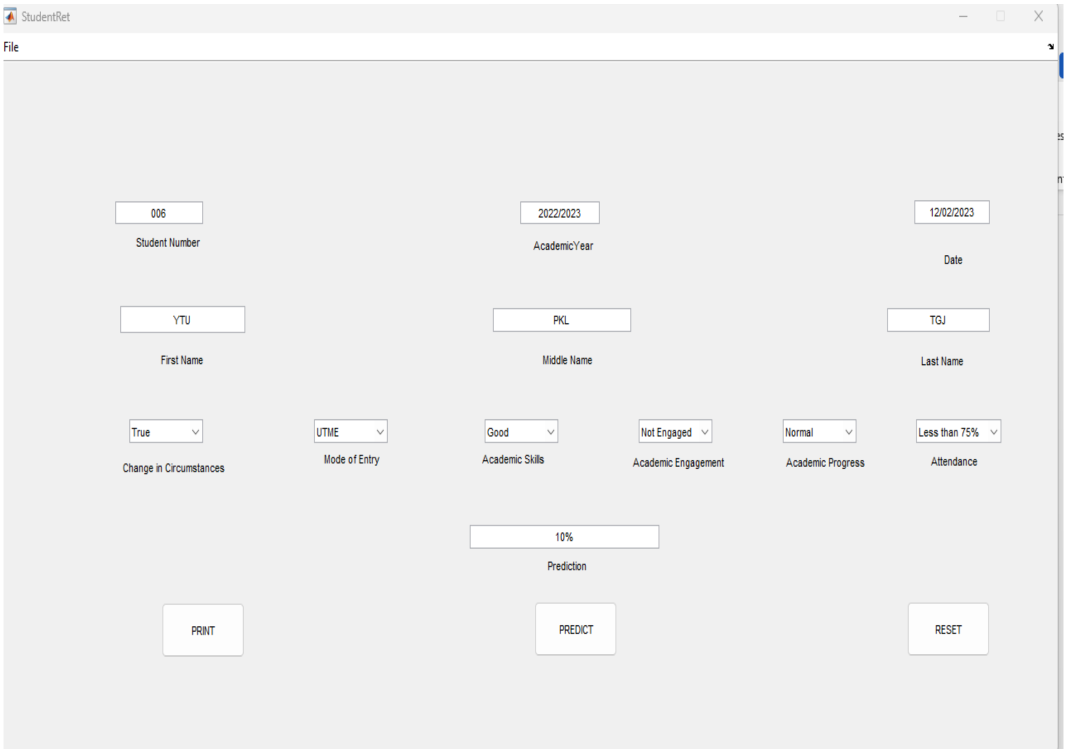


Figure 4-75 (Output for Figure 4-9): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

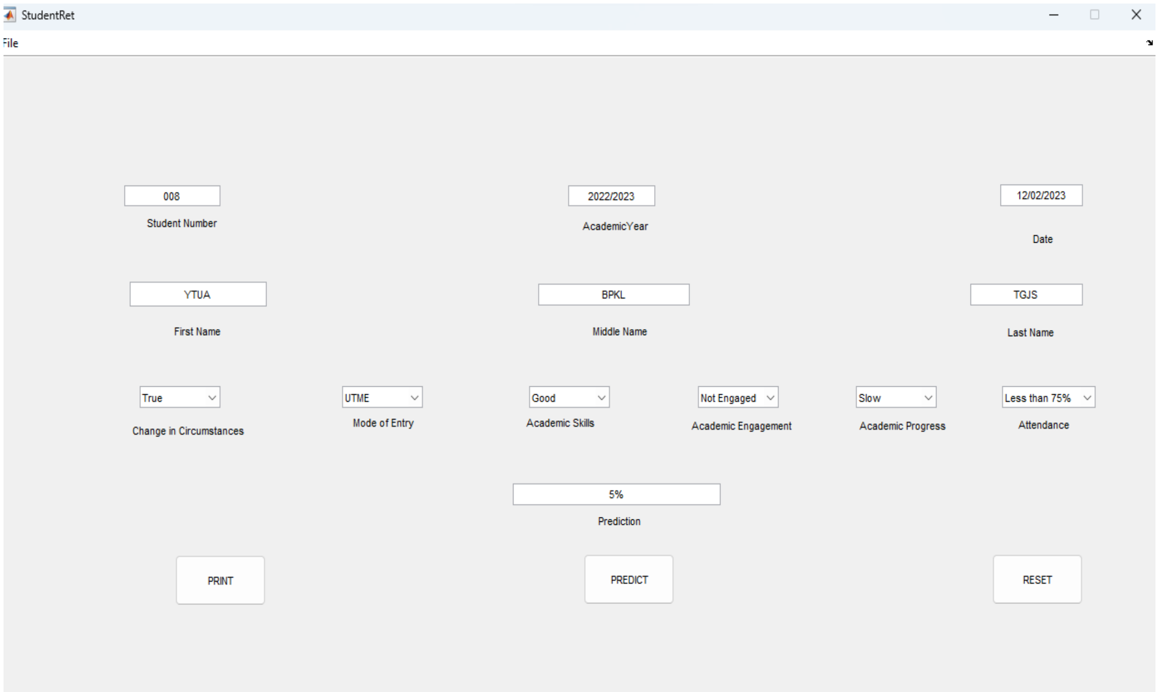


Figure 4-76 (Output for Figure 4-10): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

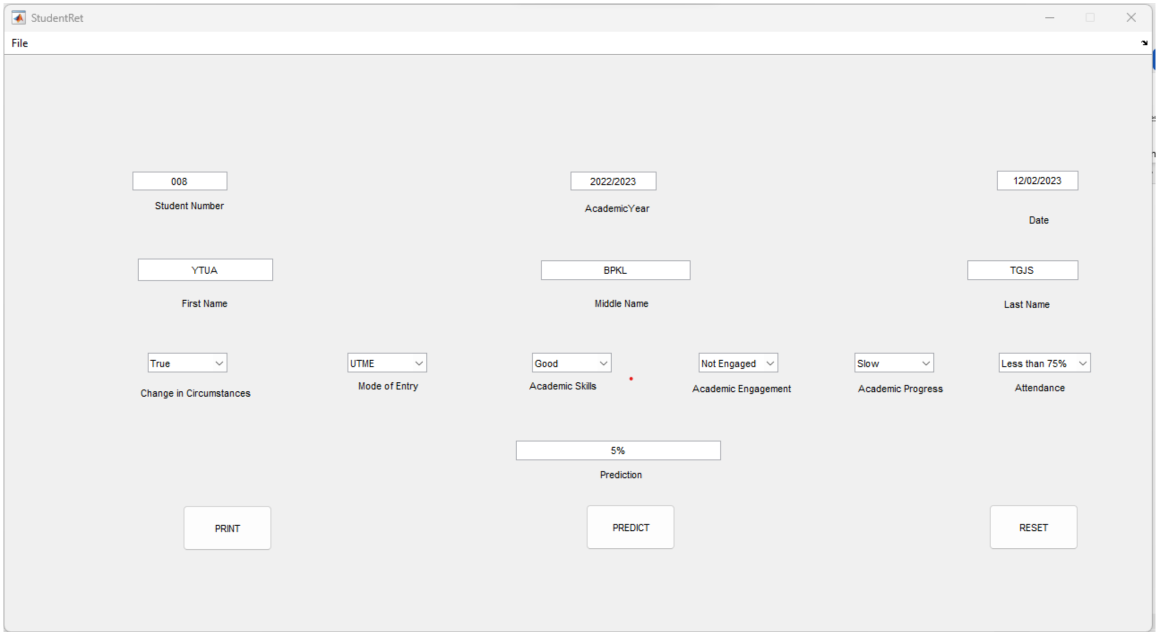


Figure 4-77 (Output for Figure 4-11): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'good' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 70%.

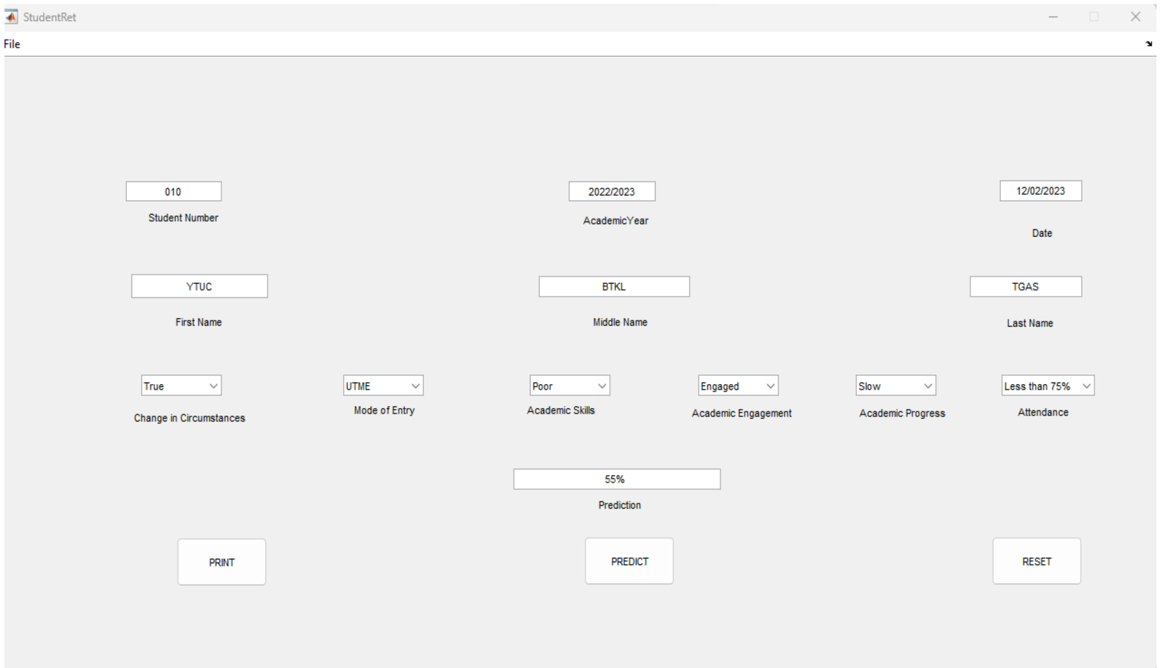


Figure 4-78 (Output for Figure 4-12): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a fairly high risk of attrition as his/her probability of being retained is 55%.

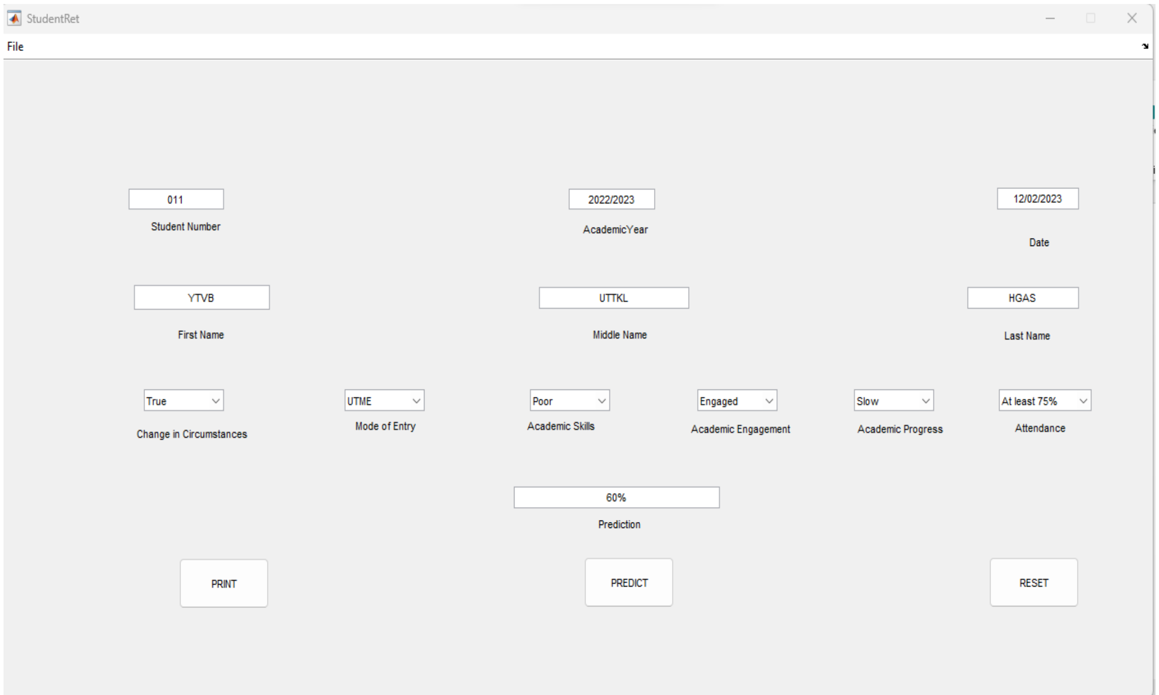


Figure 4-79 (Output for Figure 4-13): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a fairly high risk of attrition as his/her probability of being retained is 60%.

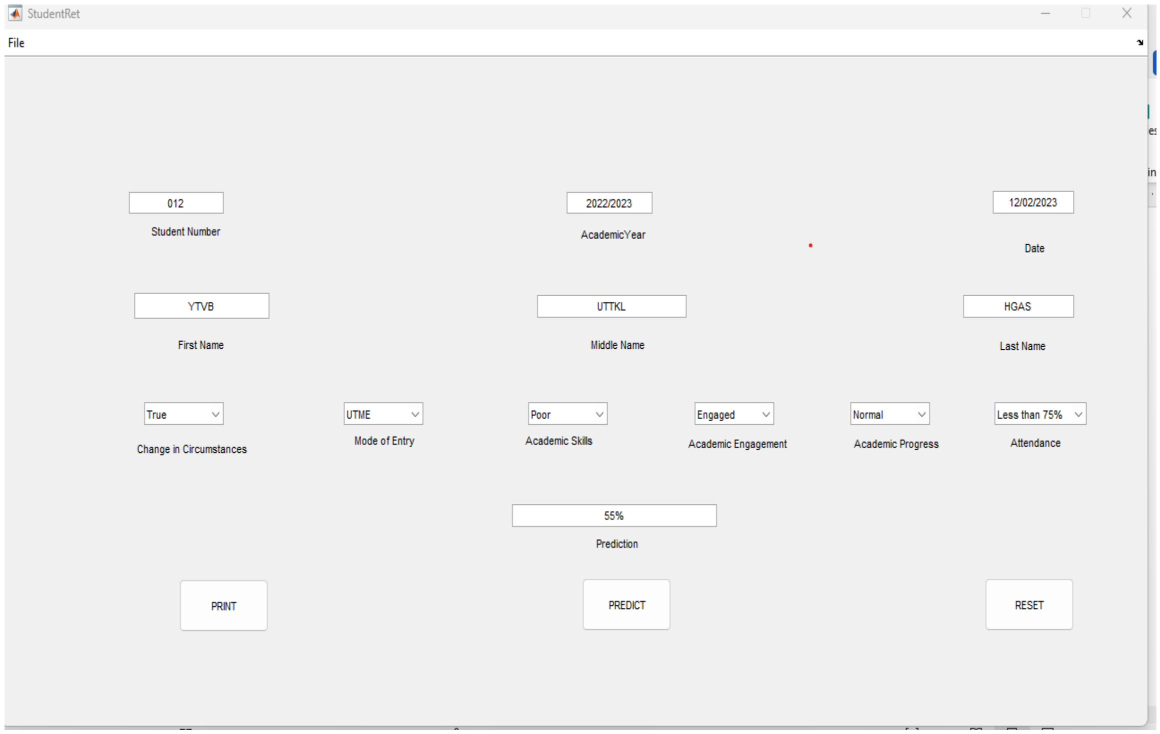


Figure 4-80 (Output for Figure 4-14): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a fairly high risk of attrition. as his/her probability of being retained is 55%.

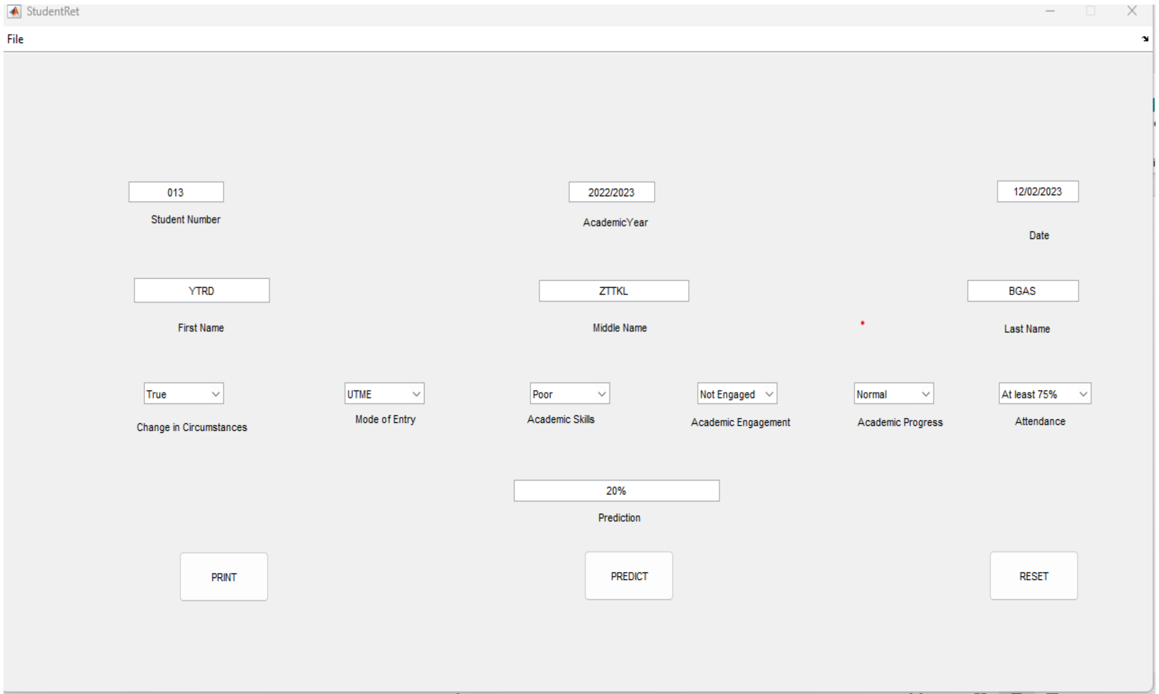


Figure 4-81 (Output for Figure 4-15): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 20%.

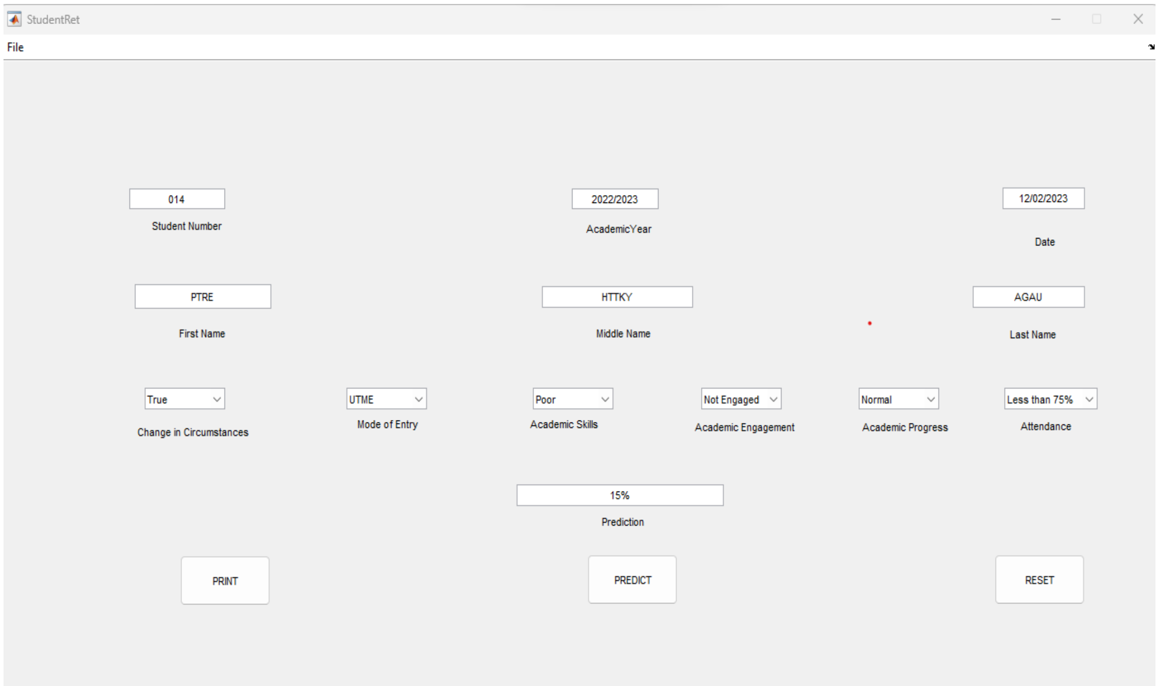


Figure 4-82 (Output for Figure 4-16): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

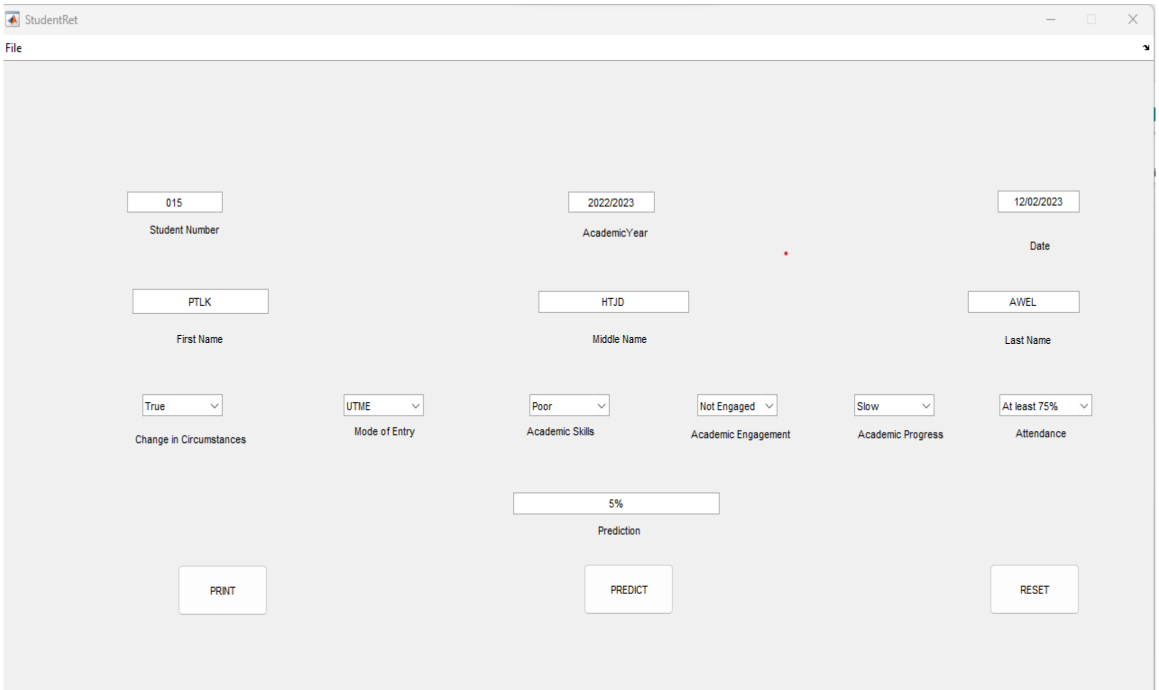


Figure 4-83 (Output for Figure 4-17): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

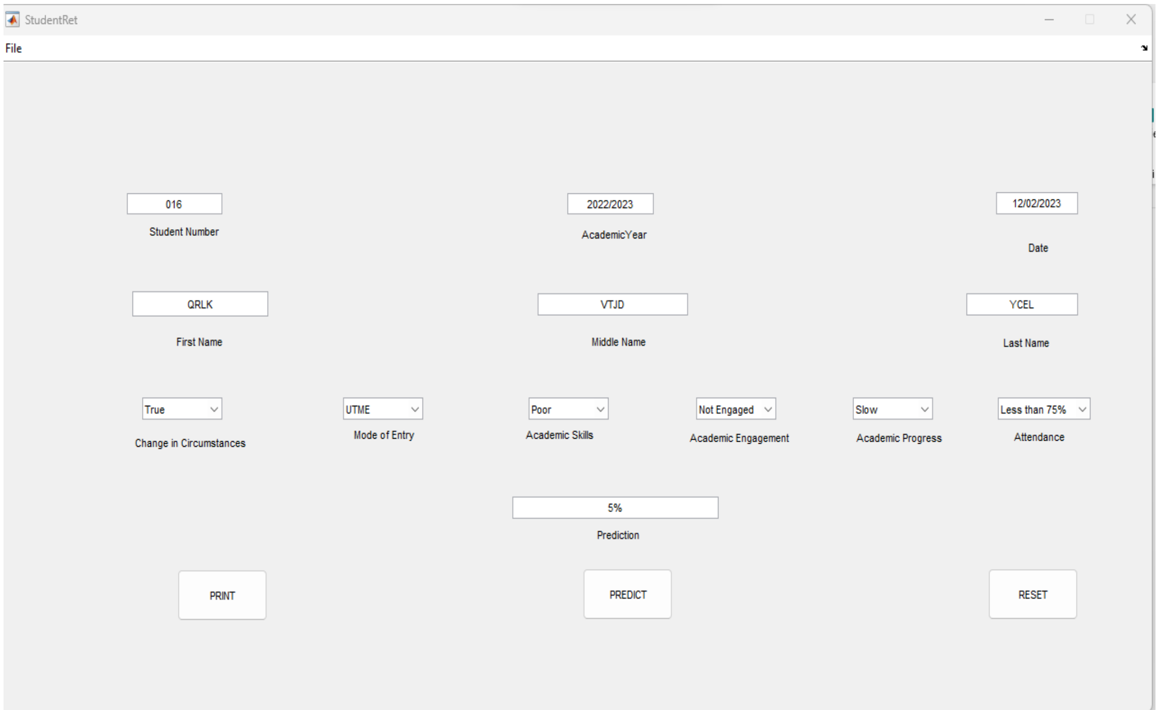


Figure 4-84 (Output for Figure 4-18): Bayesian network prediction for a student whose Change in circumstances is ‘true’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

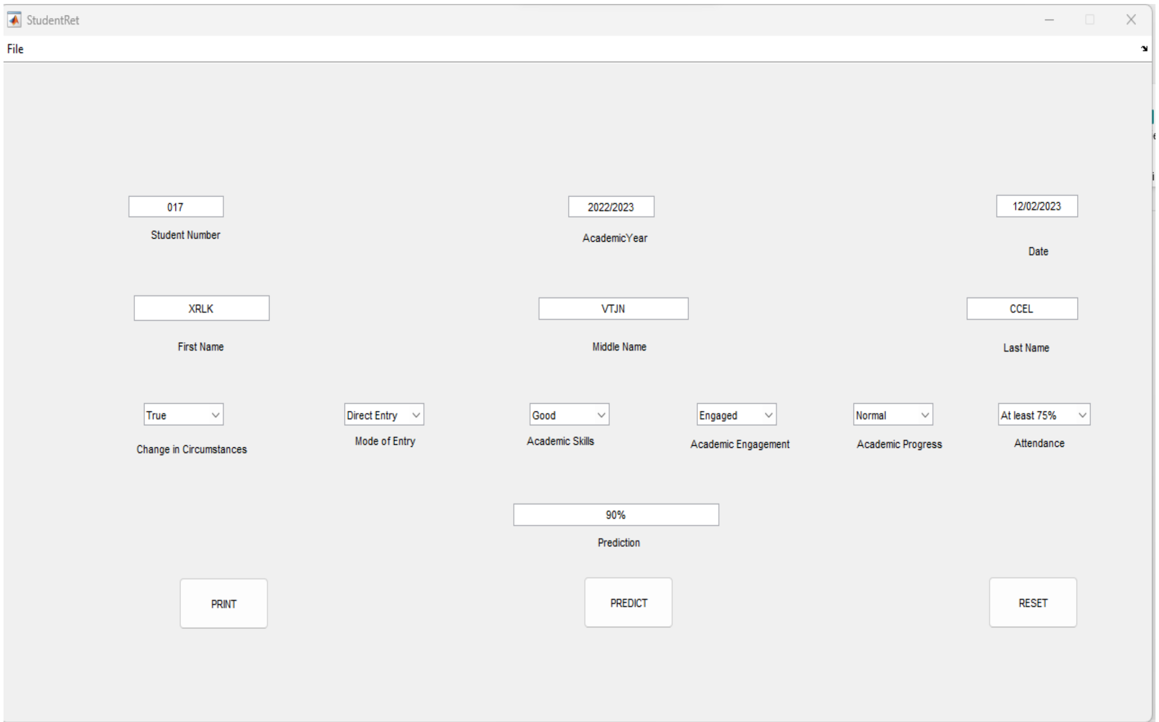


Figure 4-85 (Output for Figure 4-19): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 90%.

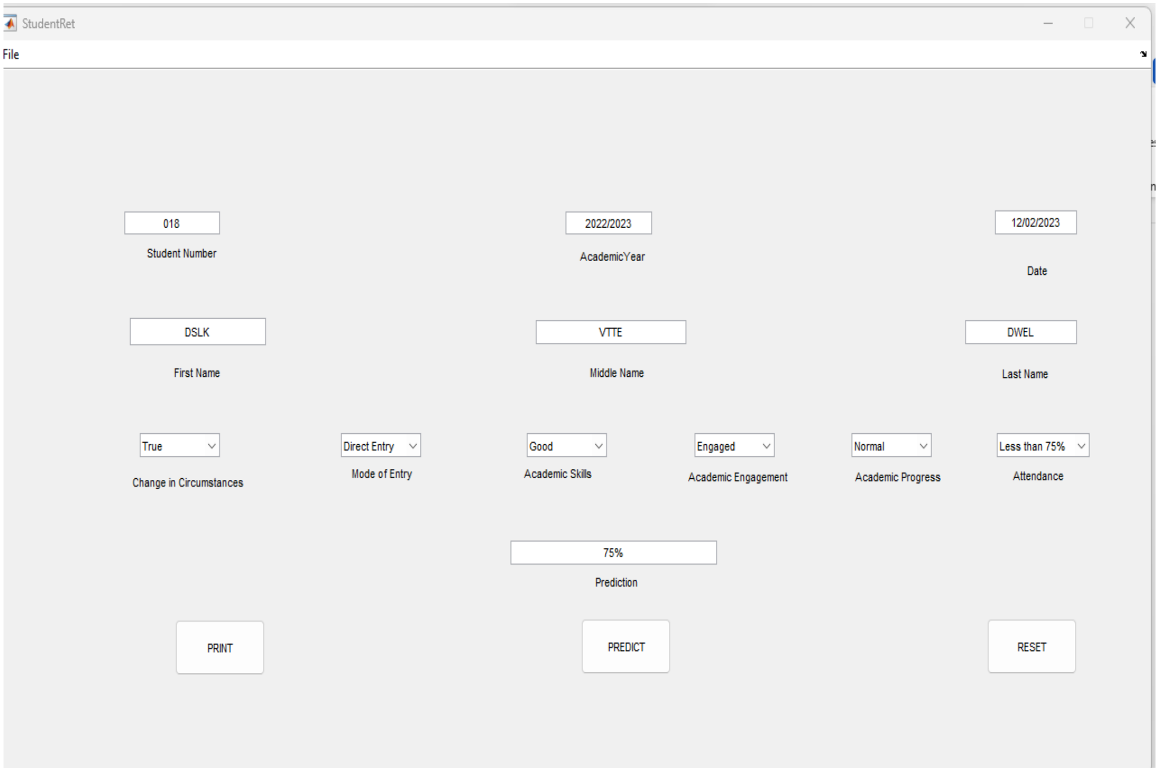


Figure 4-86 (Output for Figure 4-20): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a low risk of attrition as his/her probability of being retained is 75%.

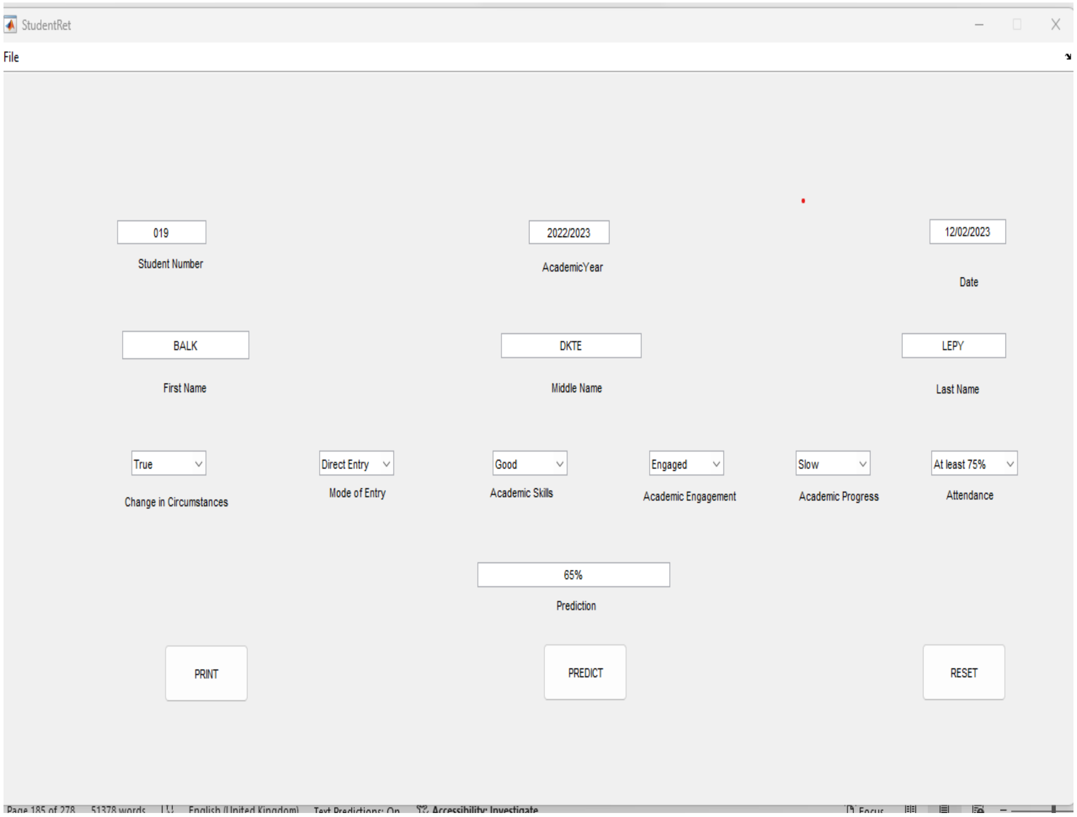


Figure 4-87 (Output for Figure 4-21): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a low risk of attrition as his/her probability of being retained is 65%.

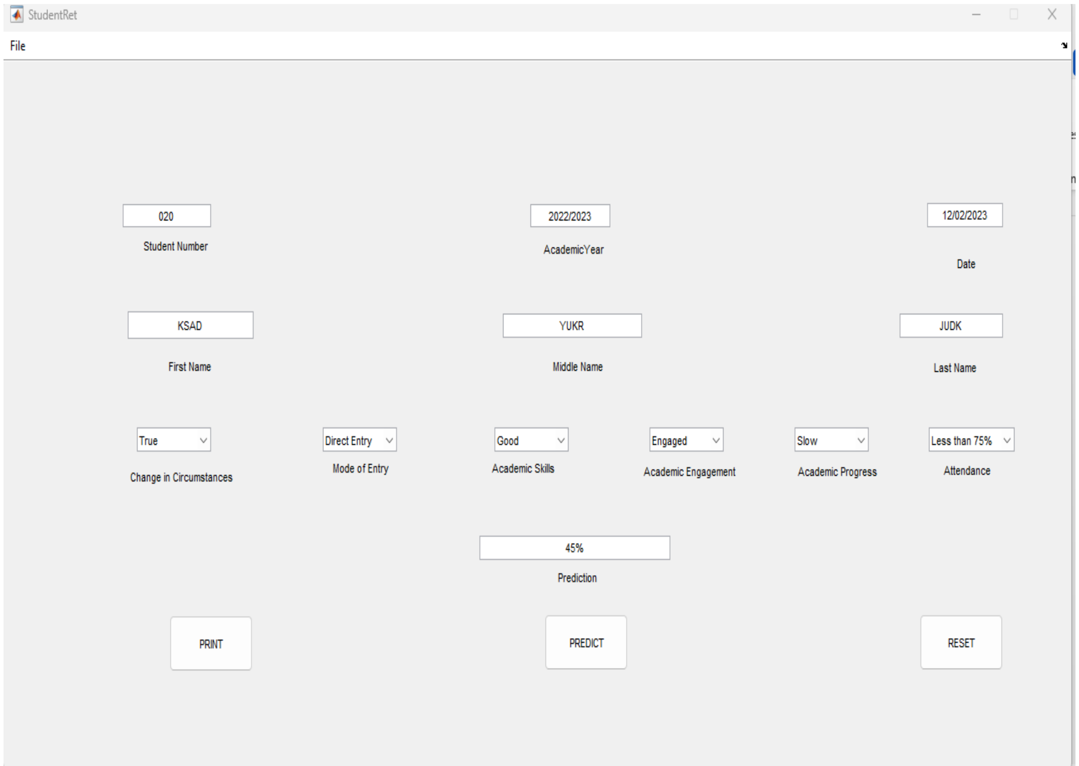


Figure 4-88 (Output for Figure 4-22): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a fairly high risk of attrition as his/her probability of being retained is 45%.

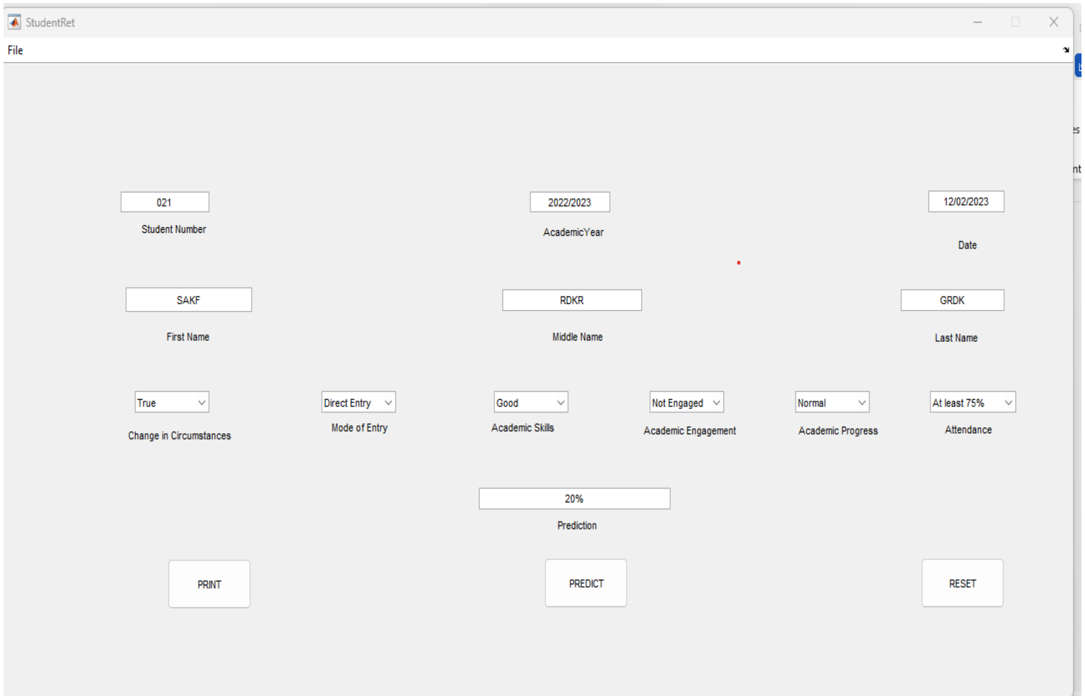


Figure 4-89 (Output for Figure 4-23): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 20%.

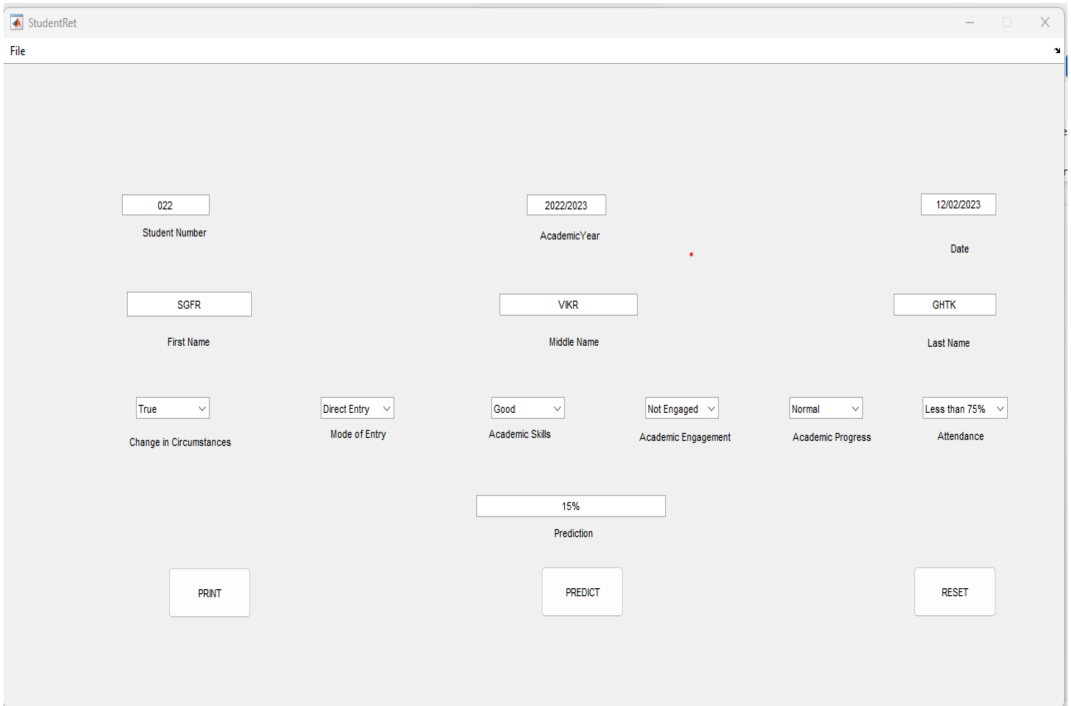


Figure 4-90 (Output for Figure 4-24): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

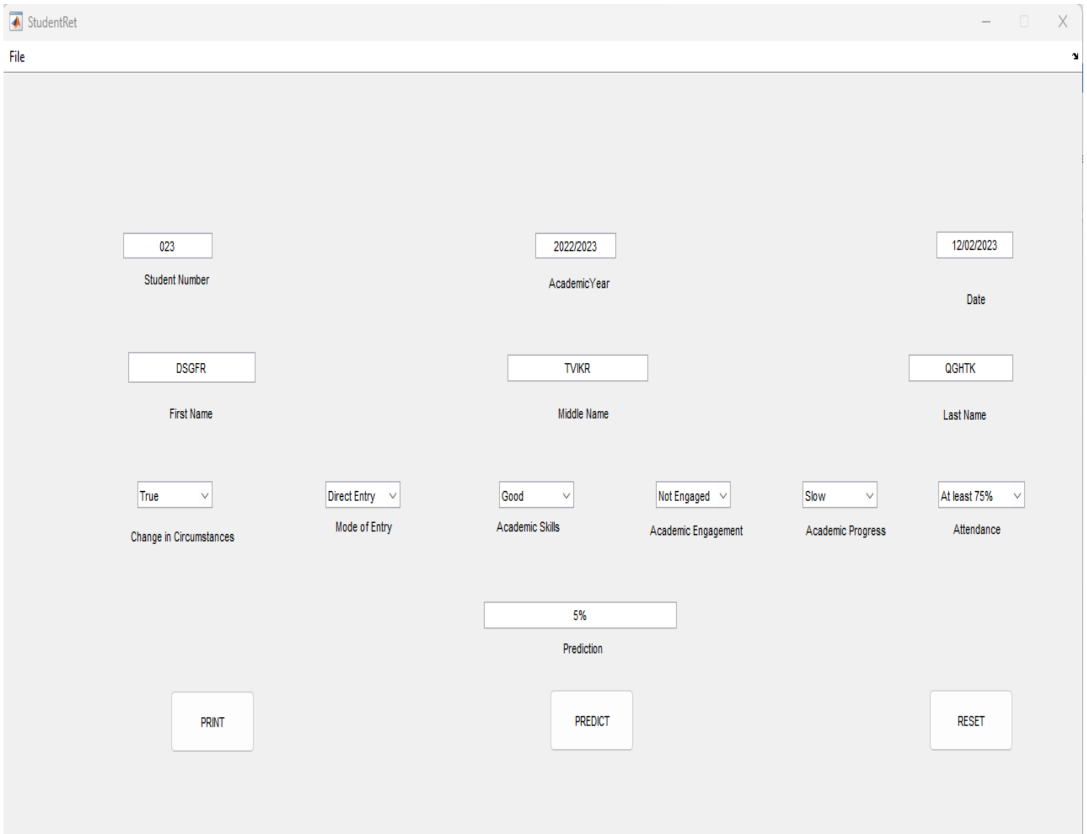


Figure 4-91 (Output for Figure 4-25): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

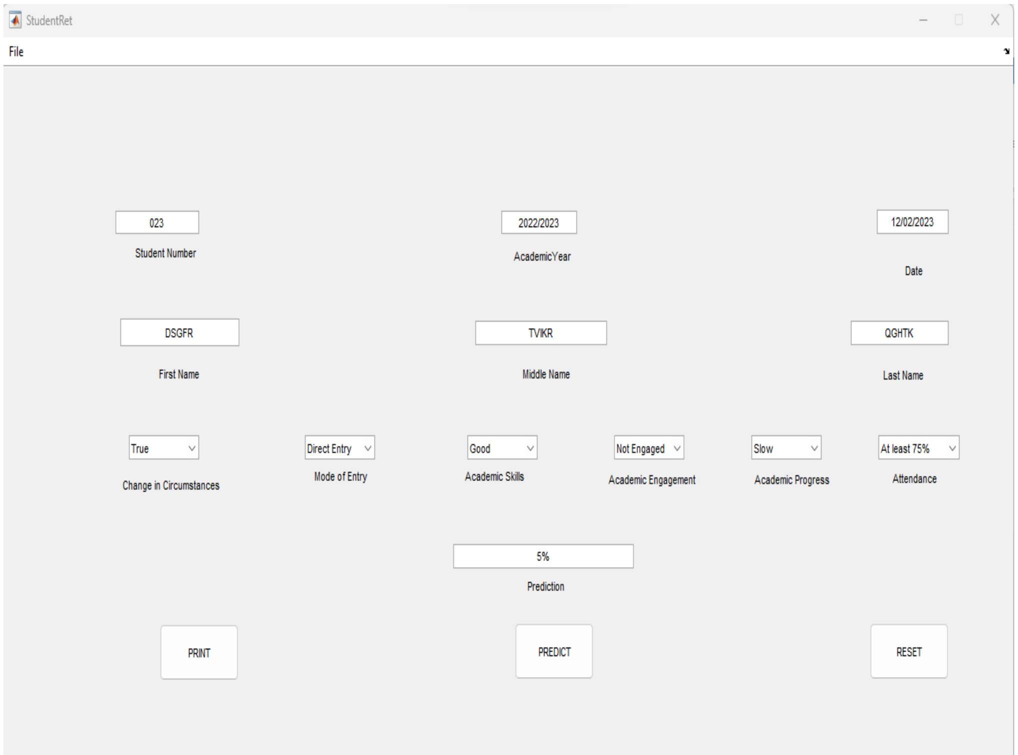


Figure 4-92 (Output for Figure 4-26): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

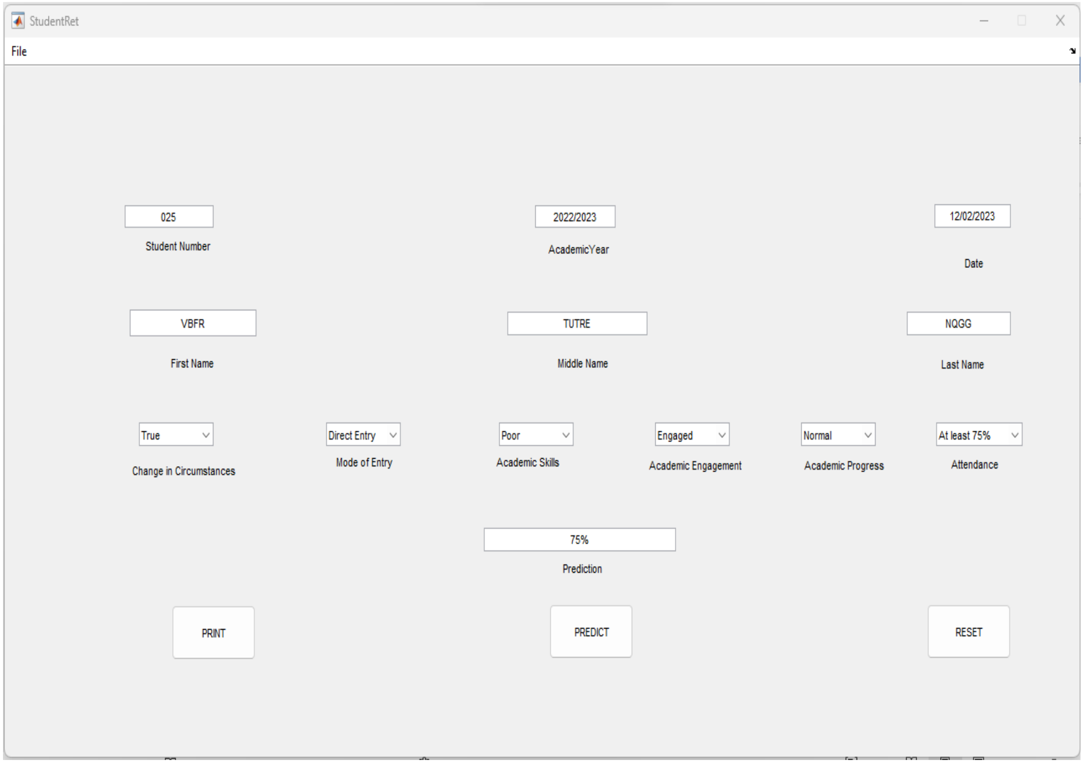


Figure 4-93 (Output for Figure 4-27): tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 75%.

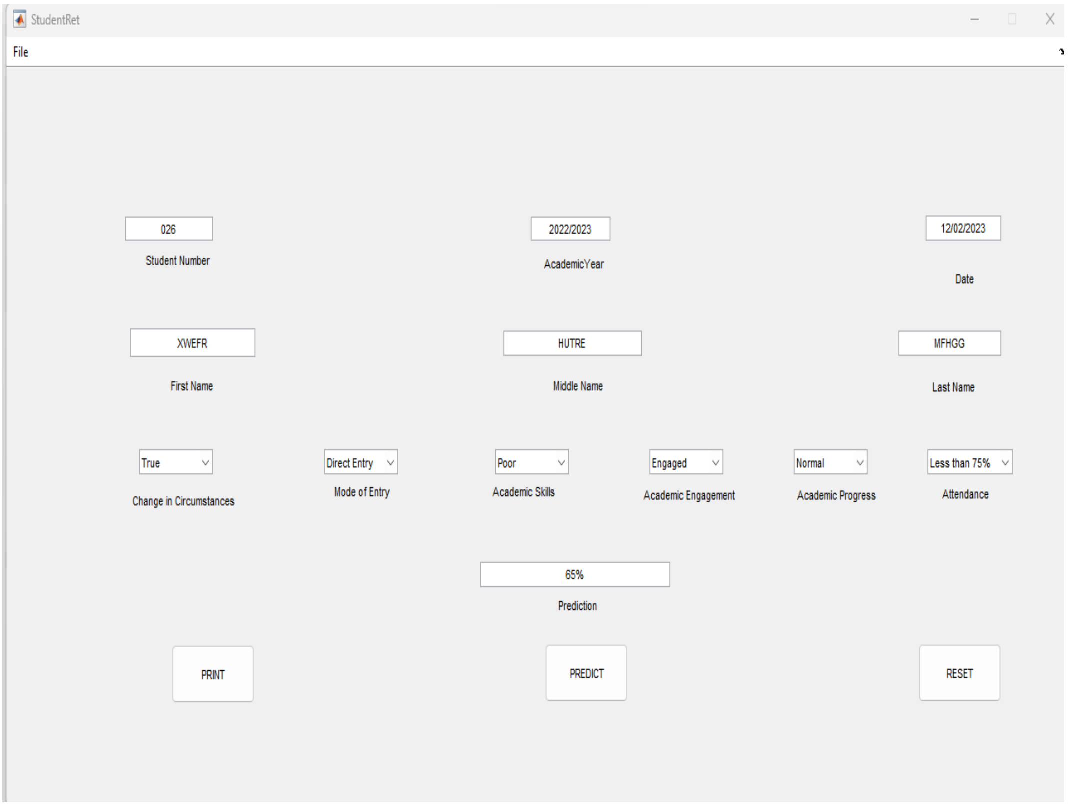


Figure 4-94 (Output for Figure 4-28): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a low risk of attrition as his/her probability of being retained is 65%.

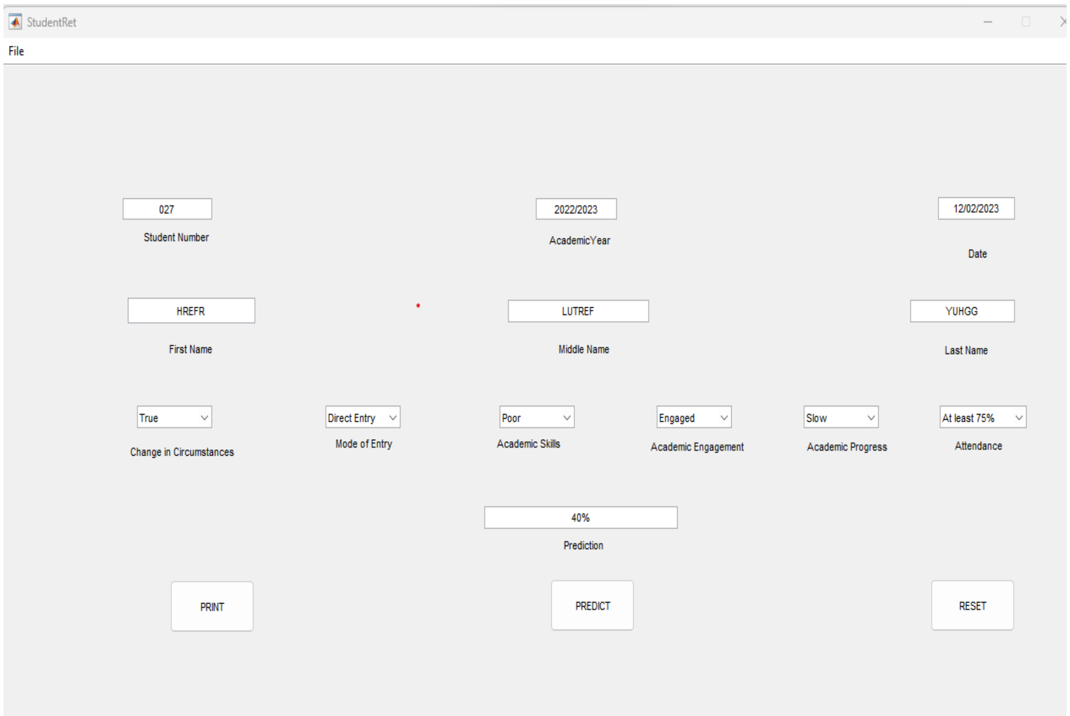


Figure 4-95 (Output for Figure 4-29): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 40%.

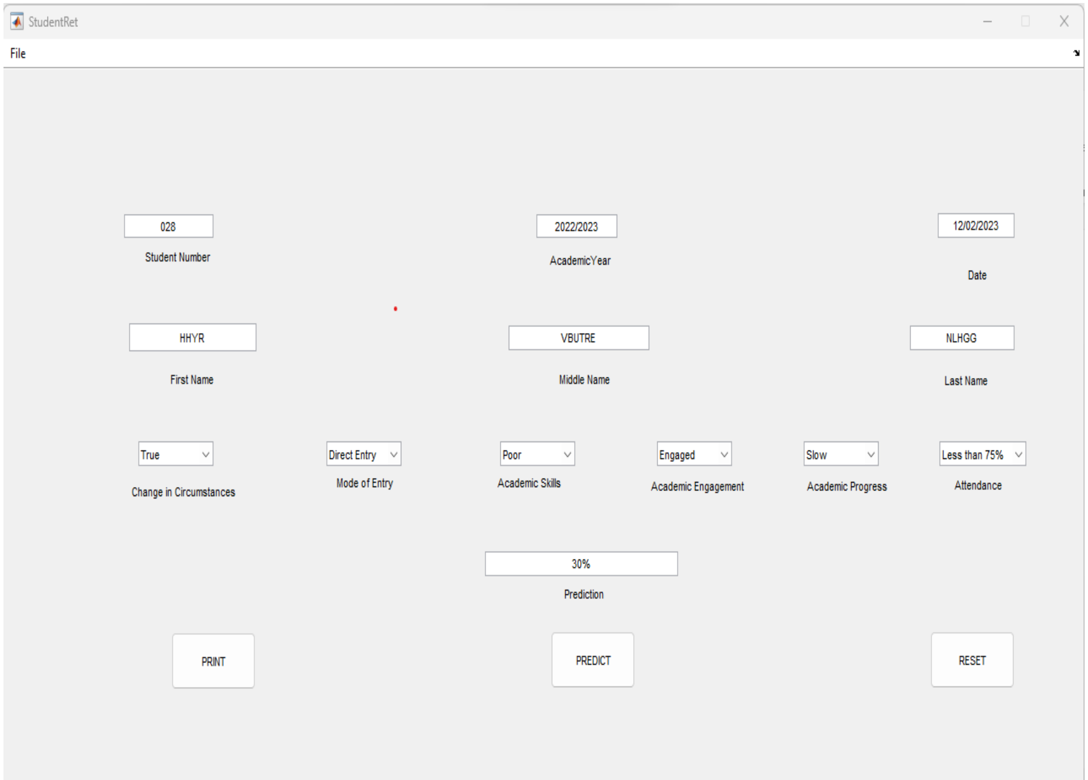


Figure 4-96 (Output for Figure 4-30): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 30%.

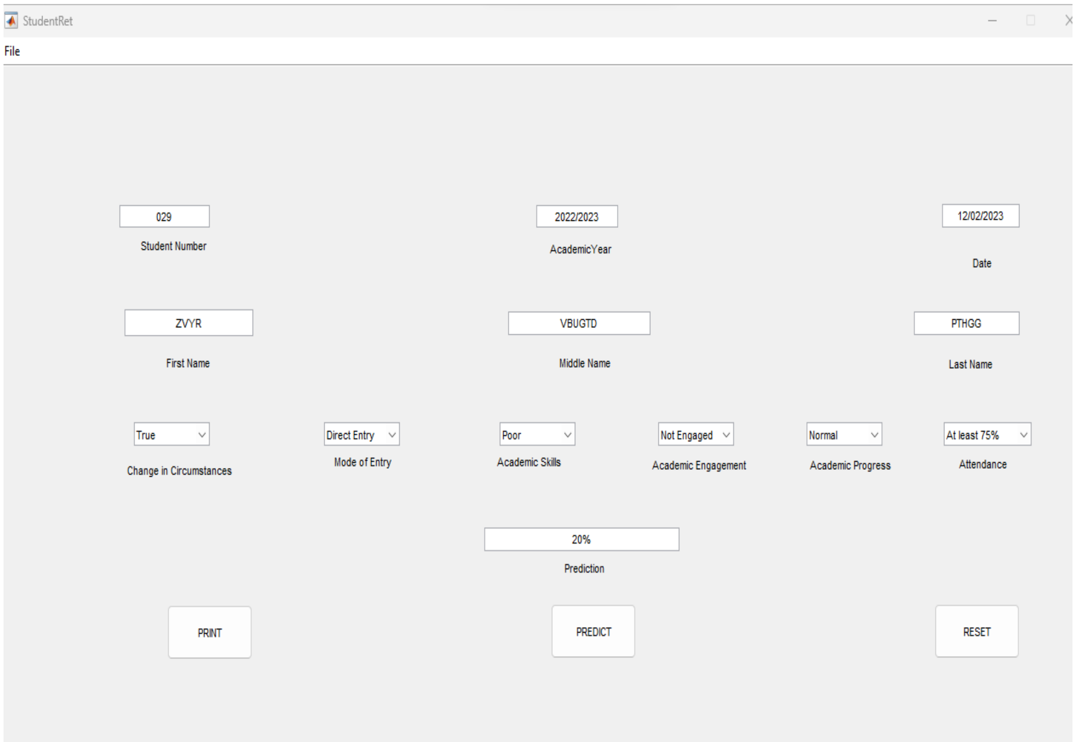


Figure 4-97 (Output for Figure 4-31): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 20%.

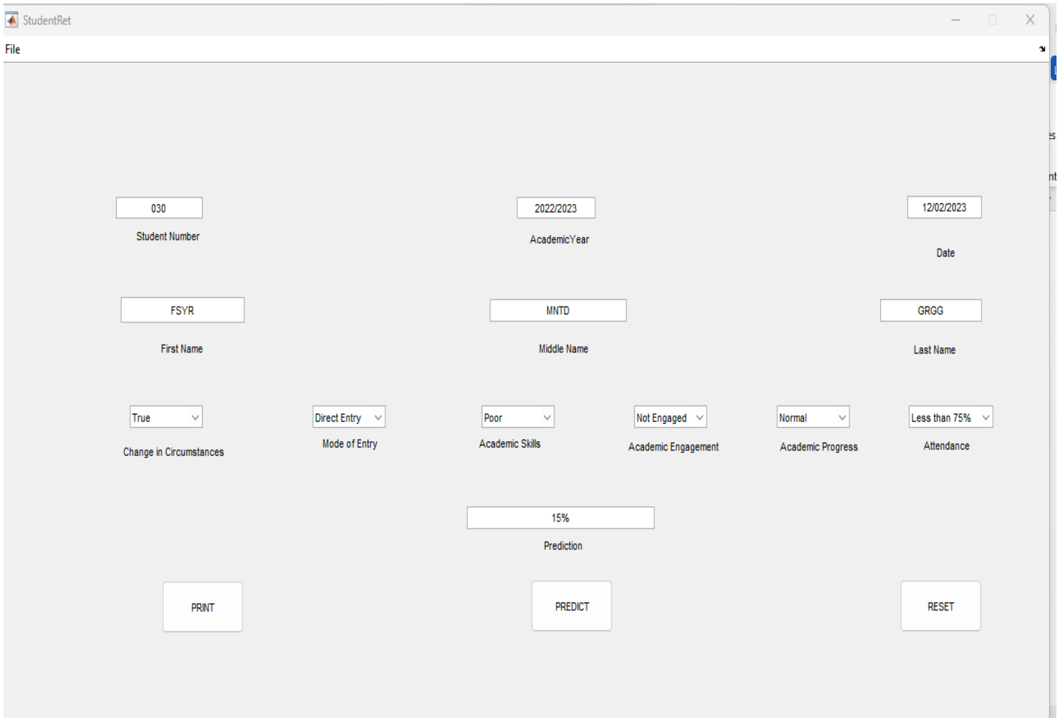


Figure 4-98 (Output for Figure 4-32): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

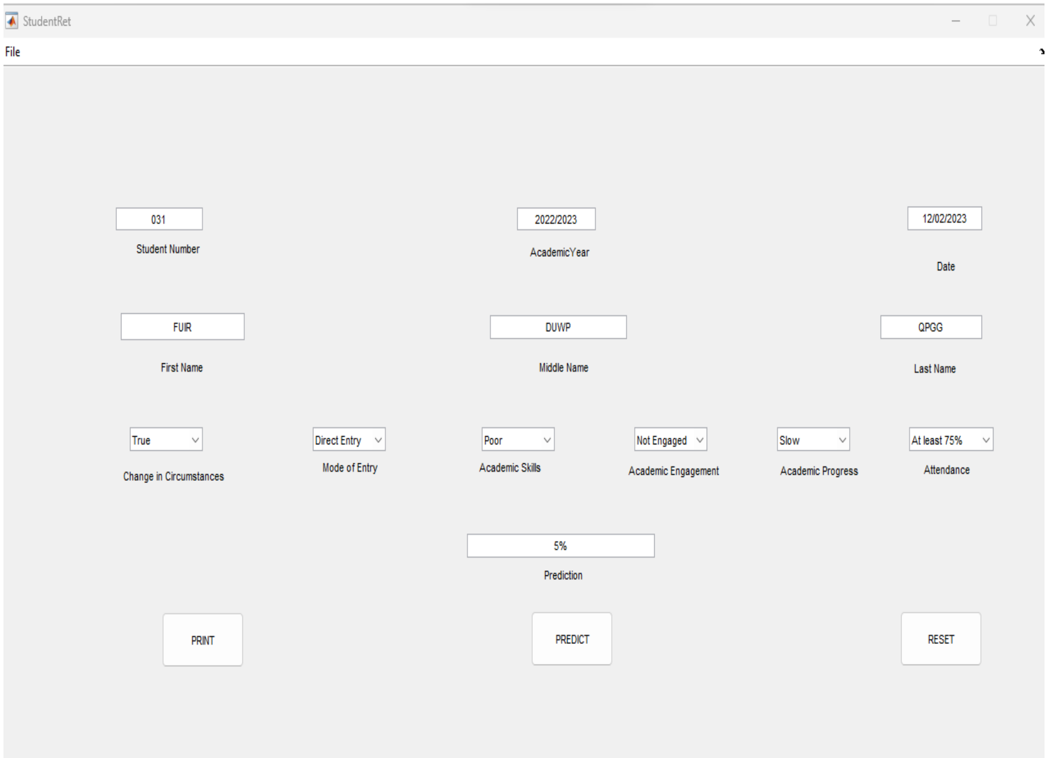


Figure 4-99 (Output for Figure 4-33): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

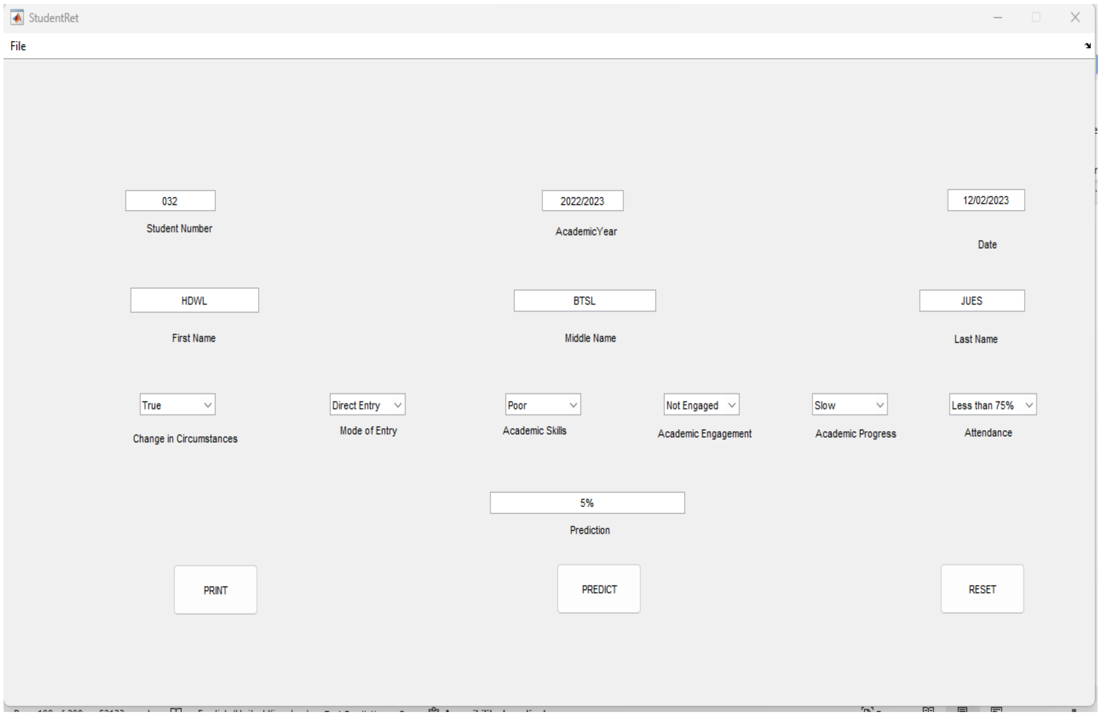


Figure 4-100 (Output for Figure 4-34): Tool's prediction for a student whose Change in circumstances is 'true', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

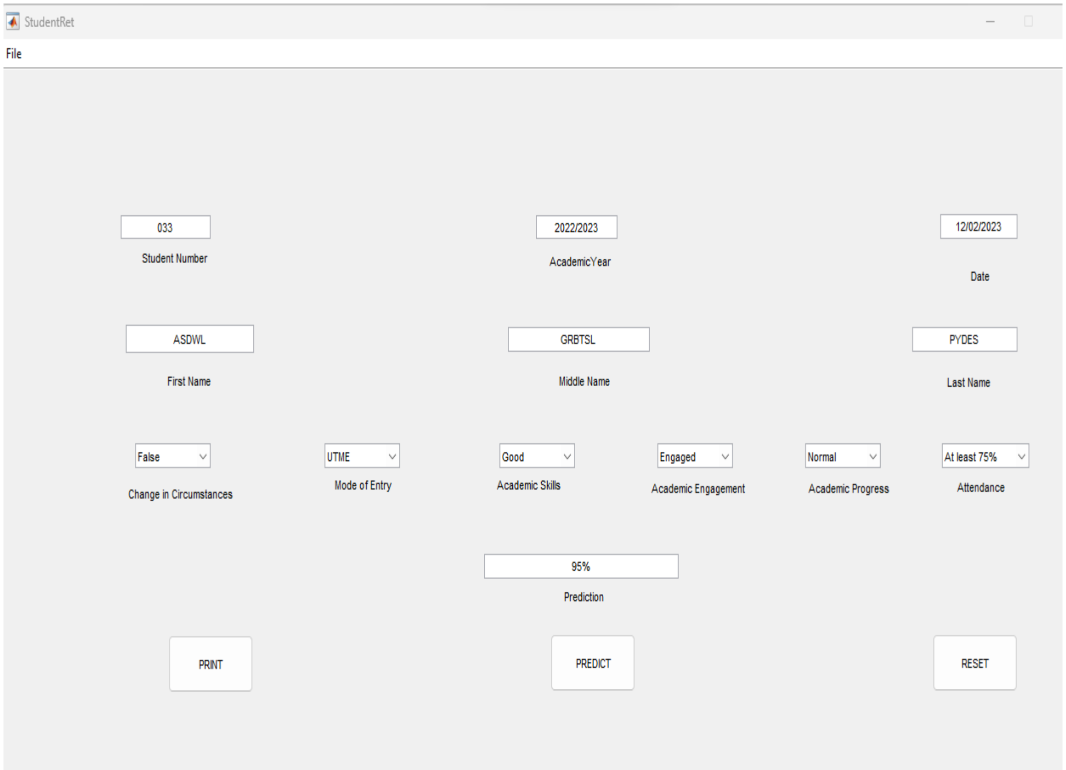


Figure 4-101 (Output for Figure 4-35): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 95%.

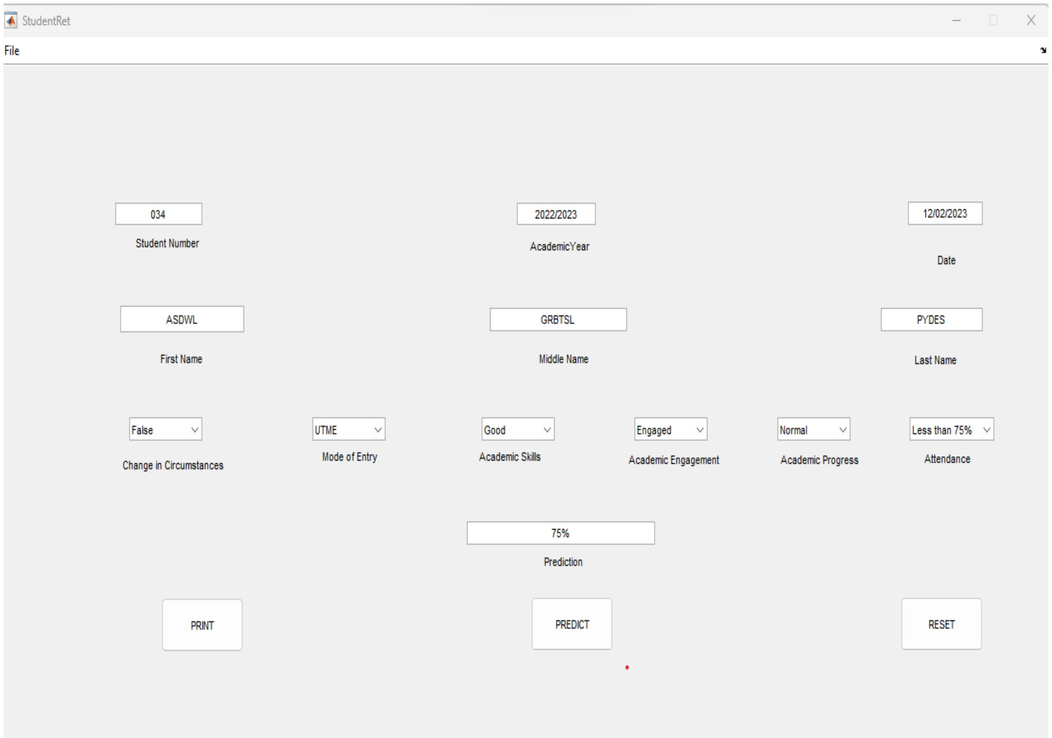


Figure 4-102 (Output for Figure 4-36): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very low risk of attrition as his/her probability of being retained is 75%.

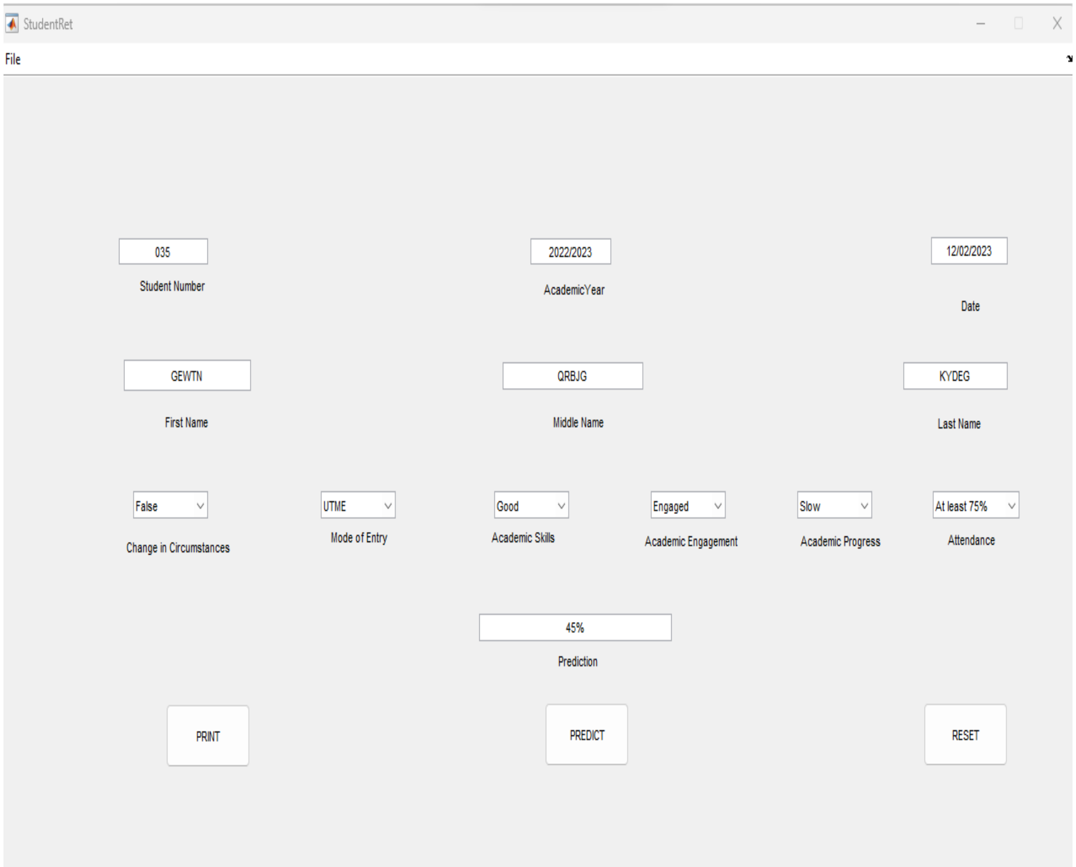


Figure 4-103 (Output for Figure 4-37): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a fairly high risk of attrition as his/her probability of being retained is 45%.

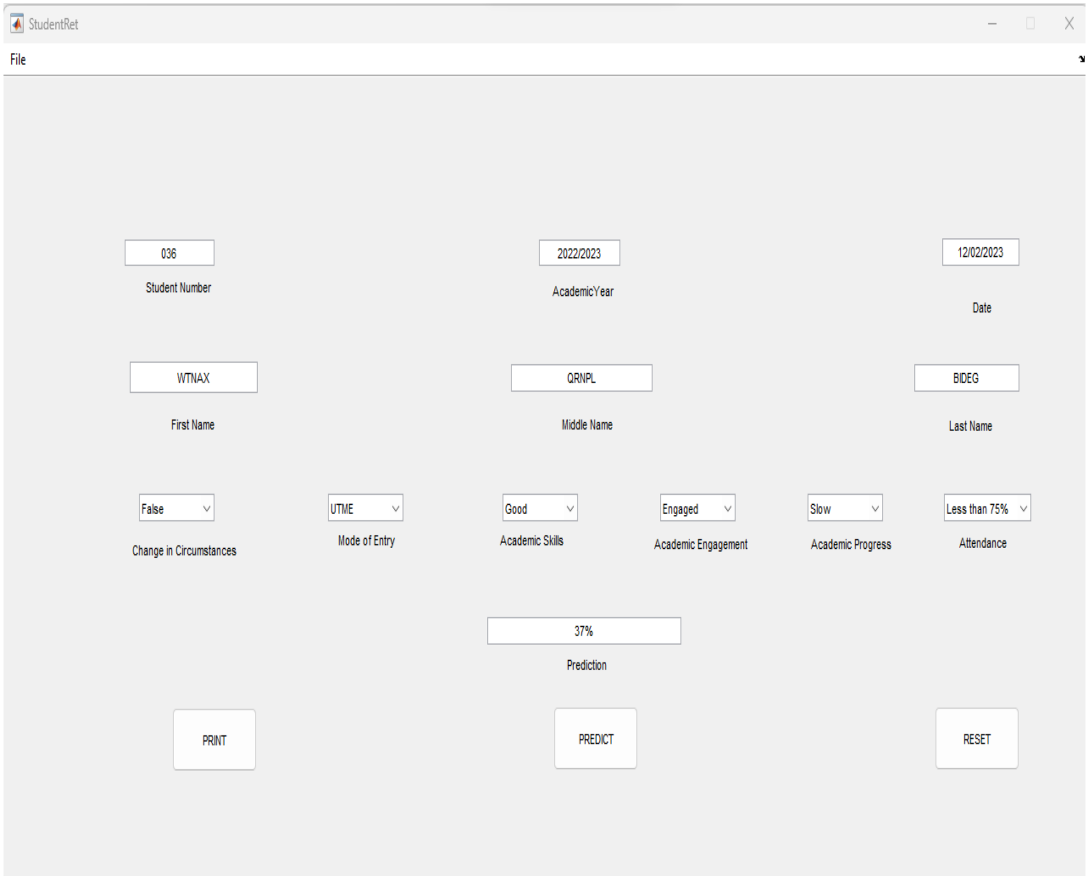


Figure 4-104 (Output for Figure 4-38): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 37%.

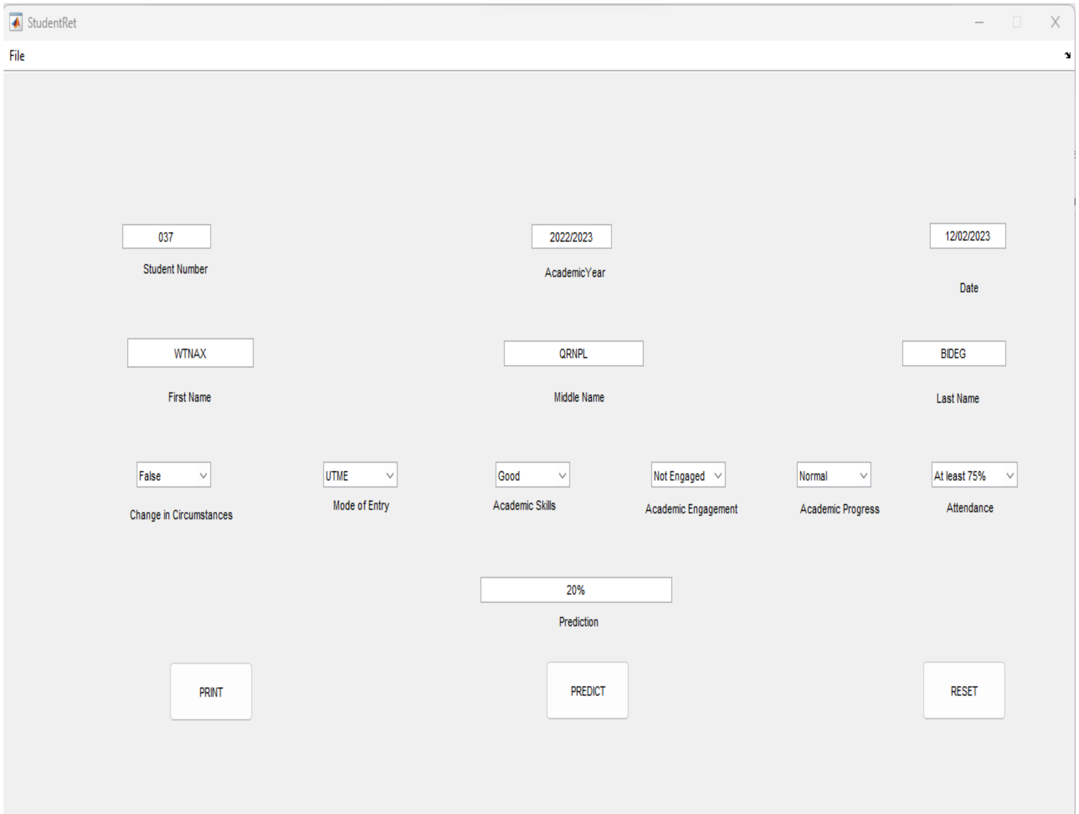


Figure 4-105 (Output for Figure 4-39): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 20%.

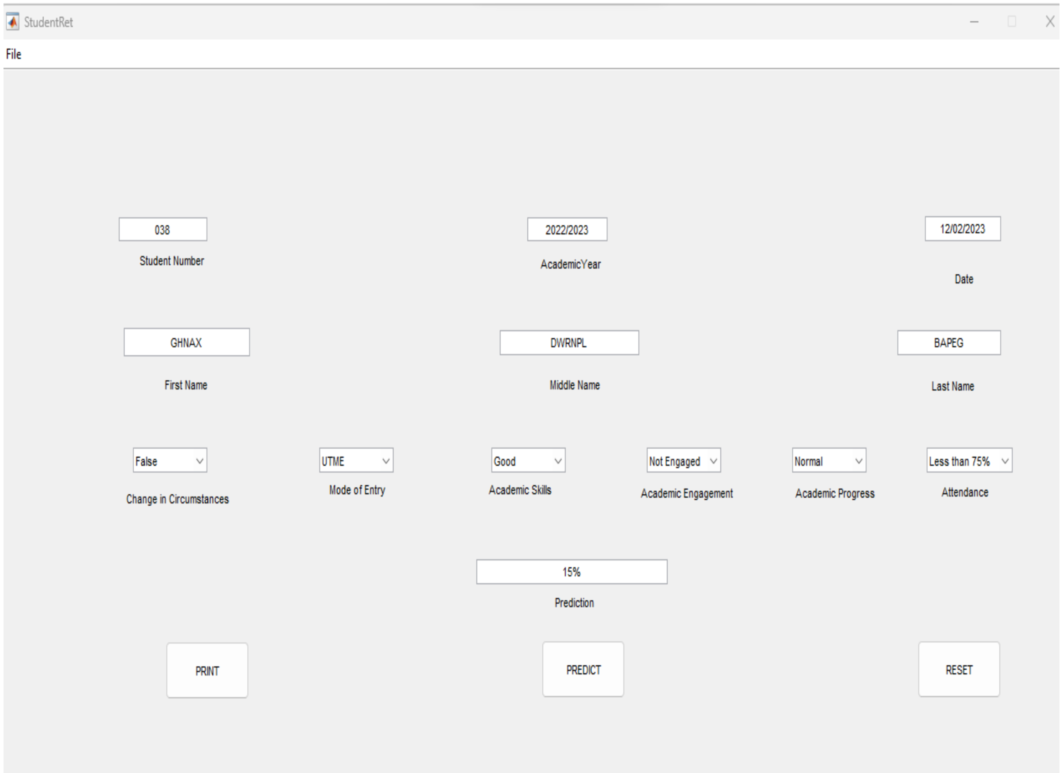


Figure 4-106 (Output for Figure 4-40): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

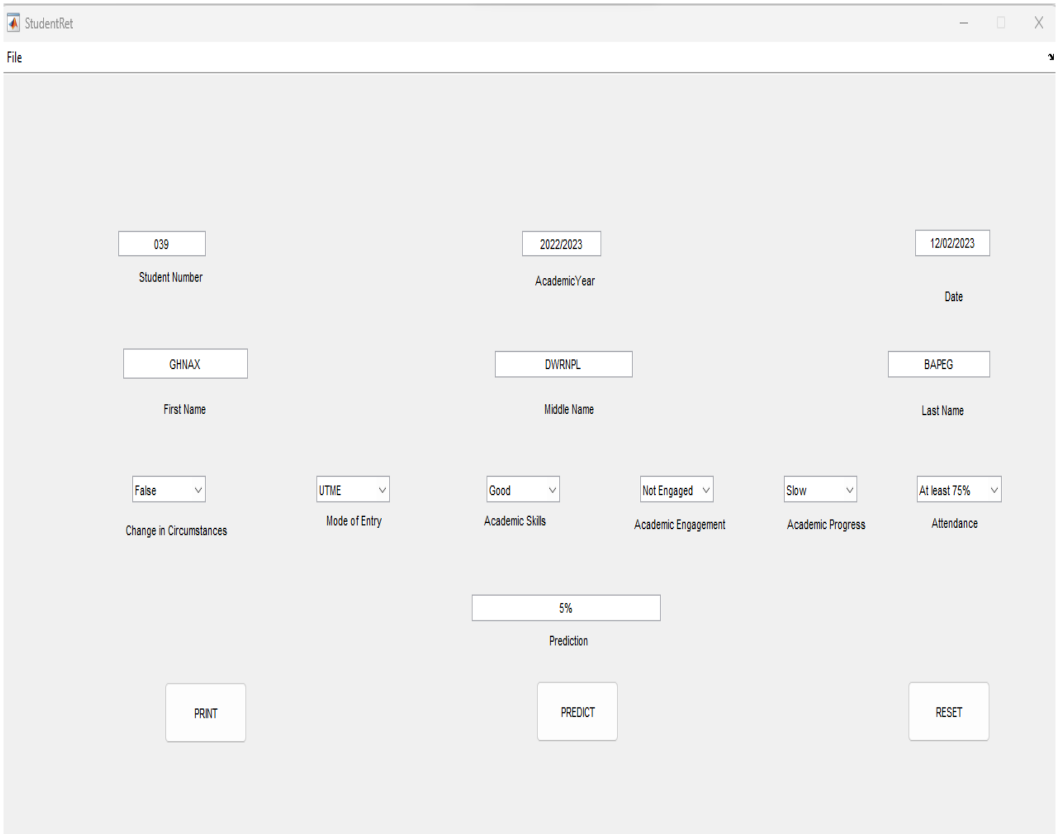


Figure 4-107 (Output for Figure 4-41): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

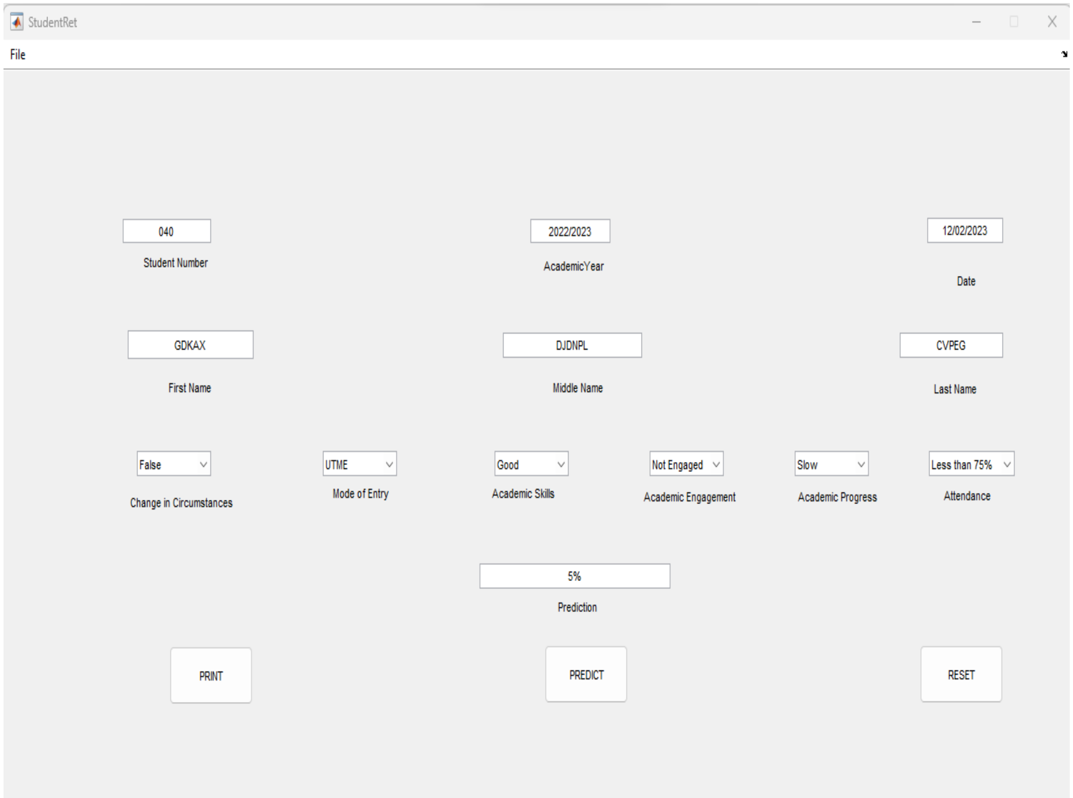


Figure 4-108 (Output for Figure 4-42): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

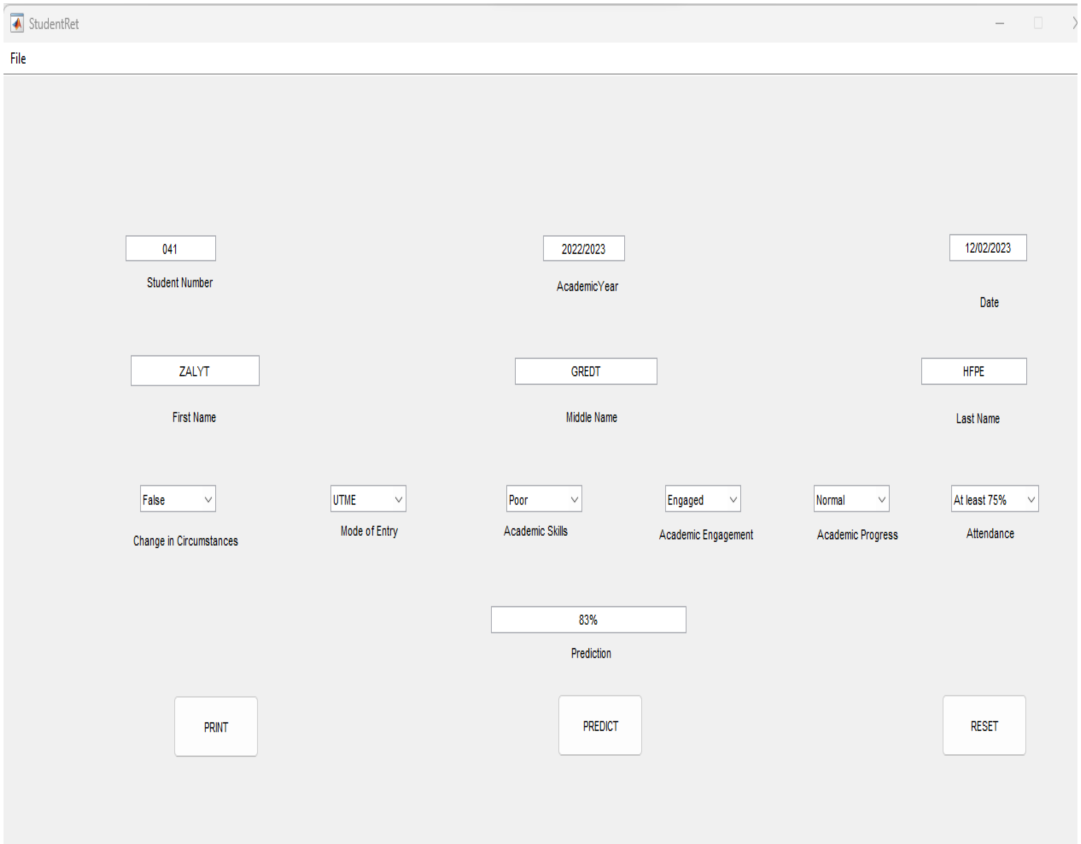


Figure 4-109 (Output for Figure 4-43): Tool's network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 83%.

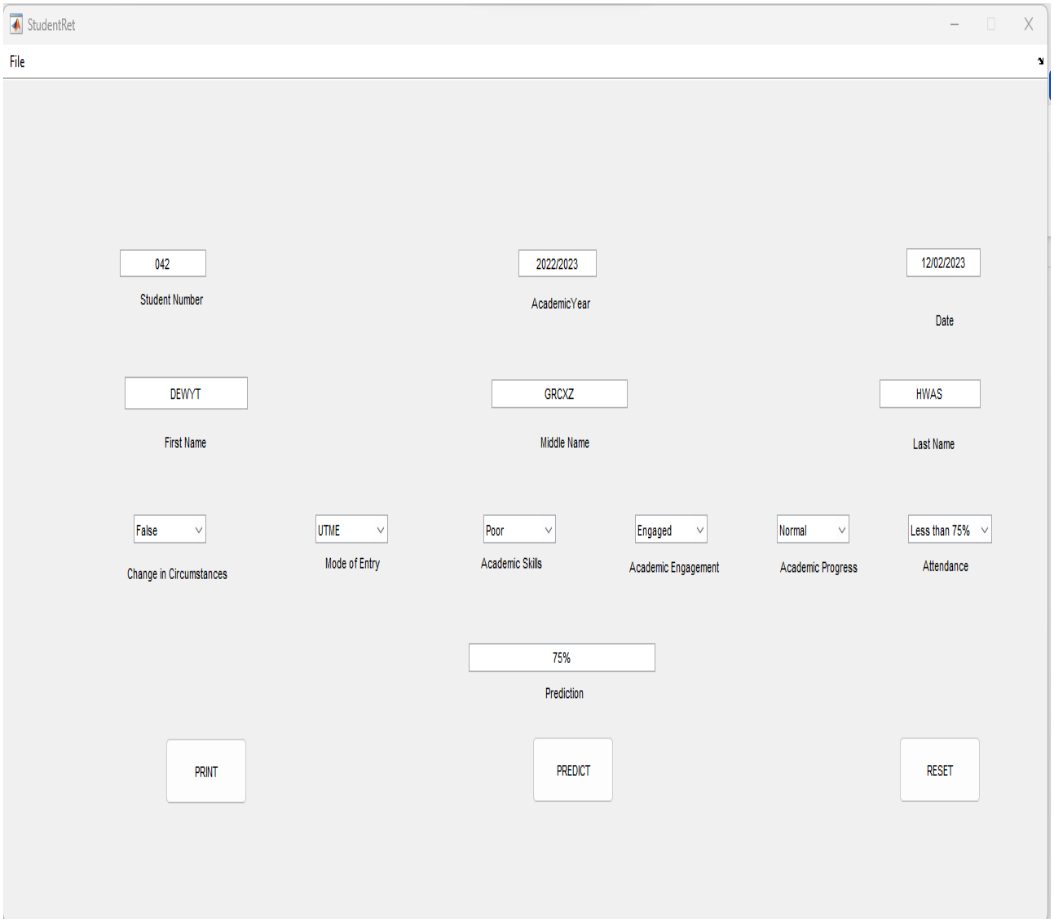


Figure 4-110 (Output for Figure 4-44): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a low risk of attrition as his/her probability of being retained is 75%.

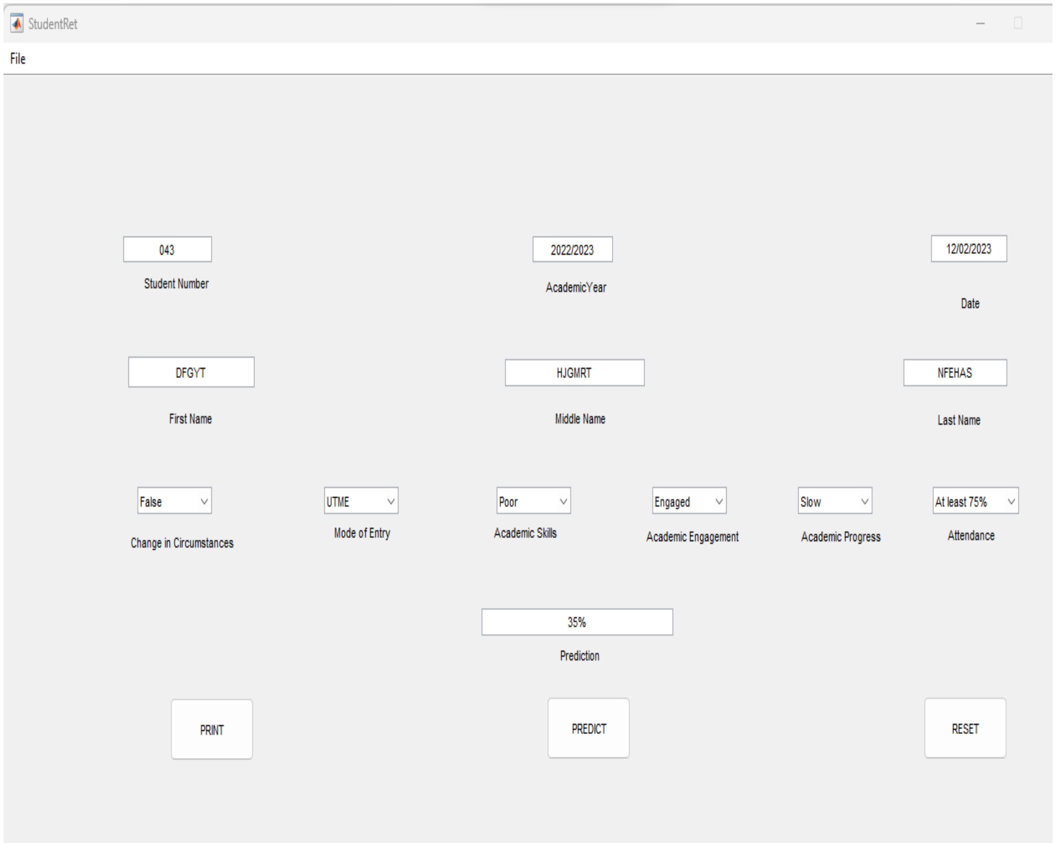


Figure 4-111 (Output for Figure 4-45): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 35%.

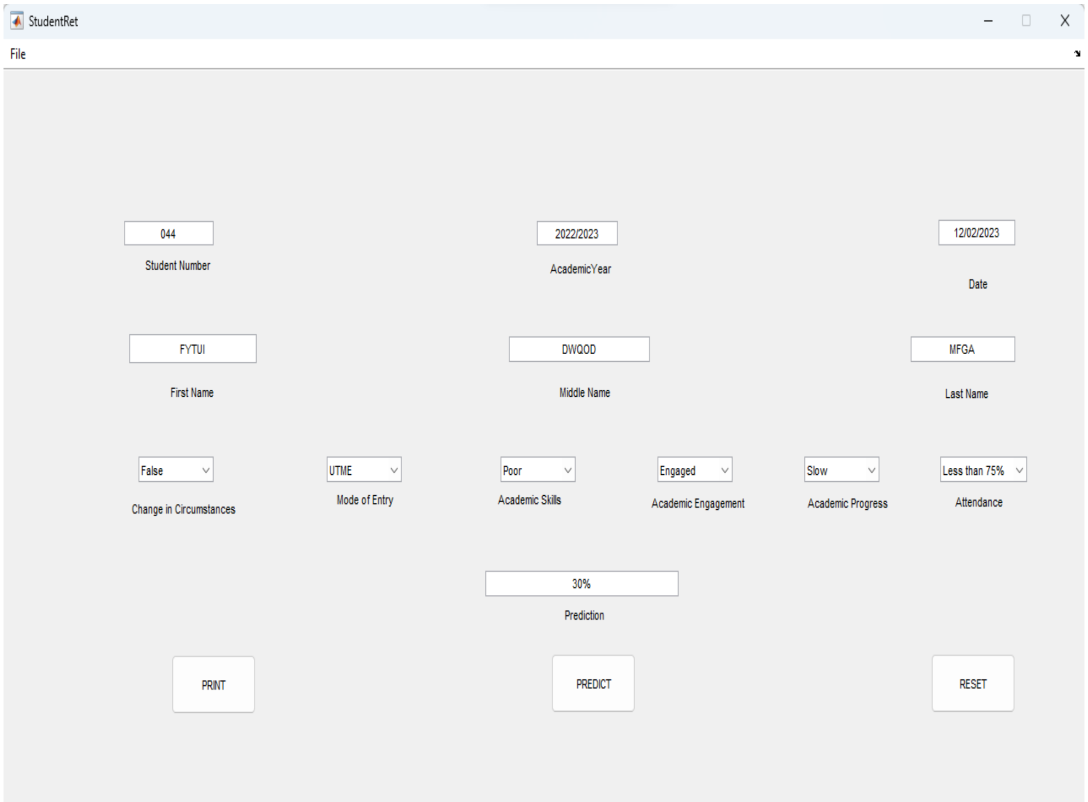


Figure 4-112 (Output for Figure 4-46): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 30%.

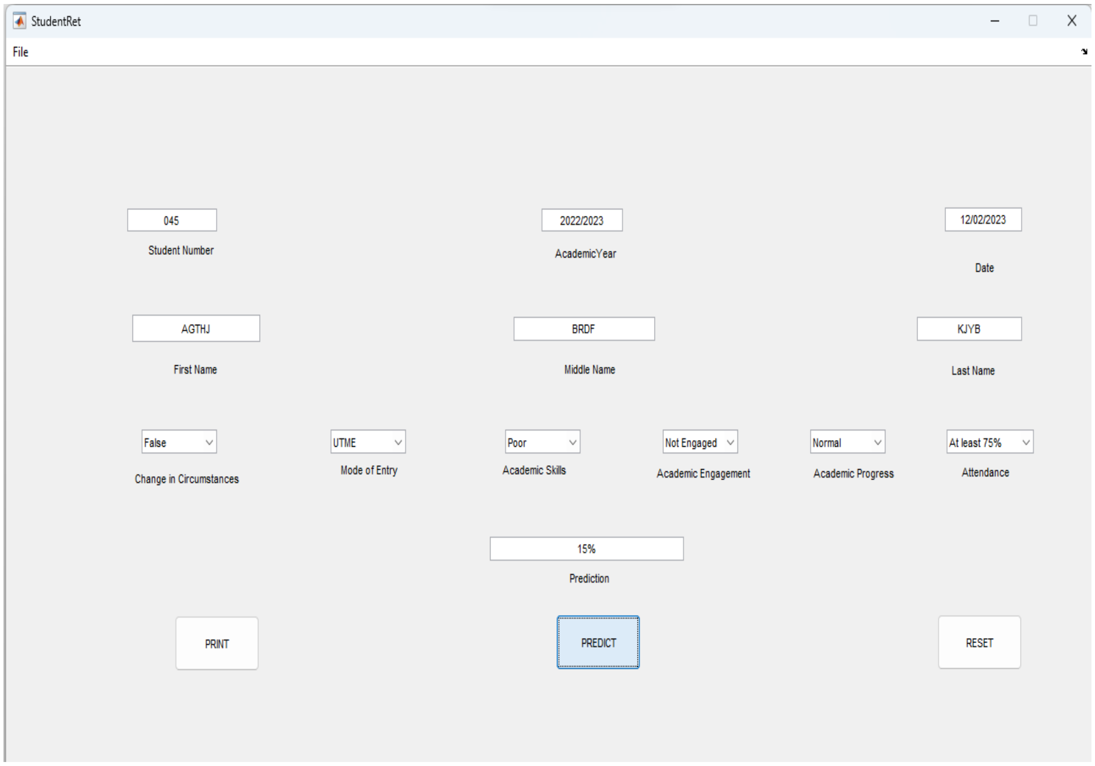


Figure 4-113 (Output for Figure 4-47): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

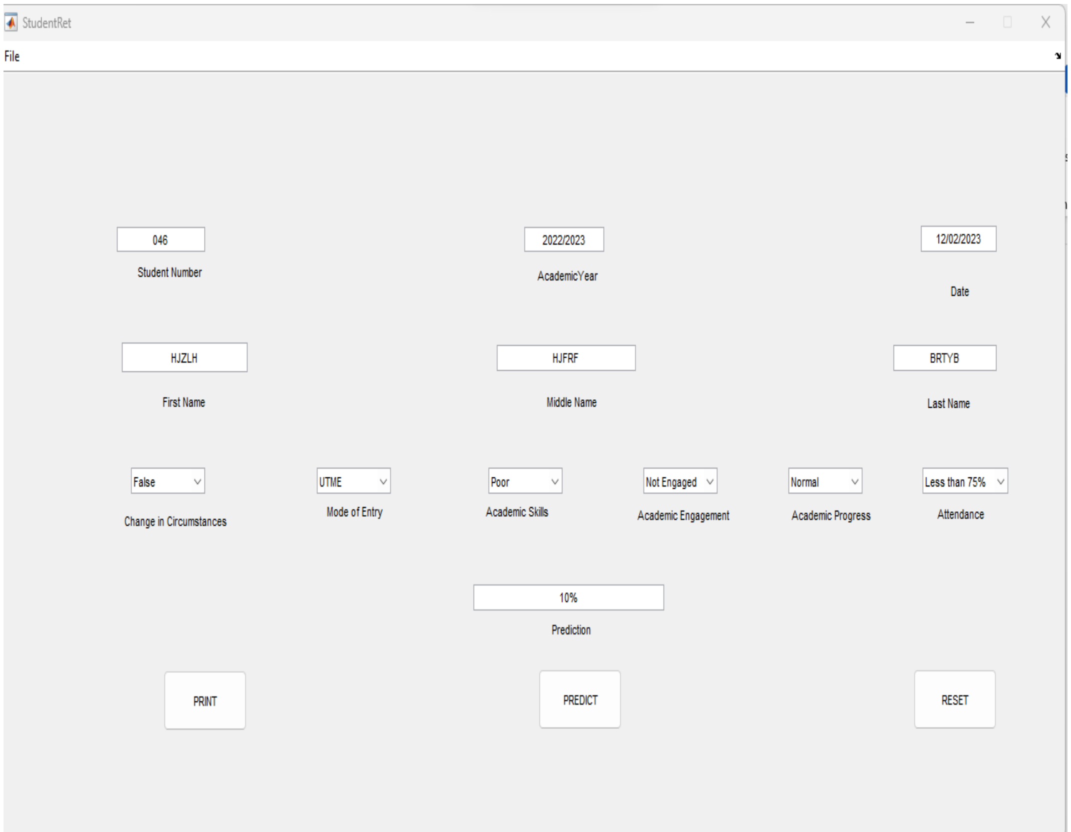


Figure 4-114 (Output for Figure 4-48): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 10%.

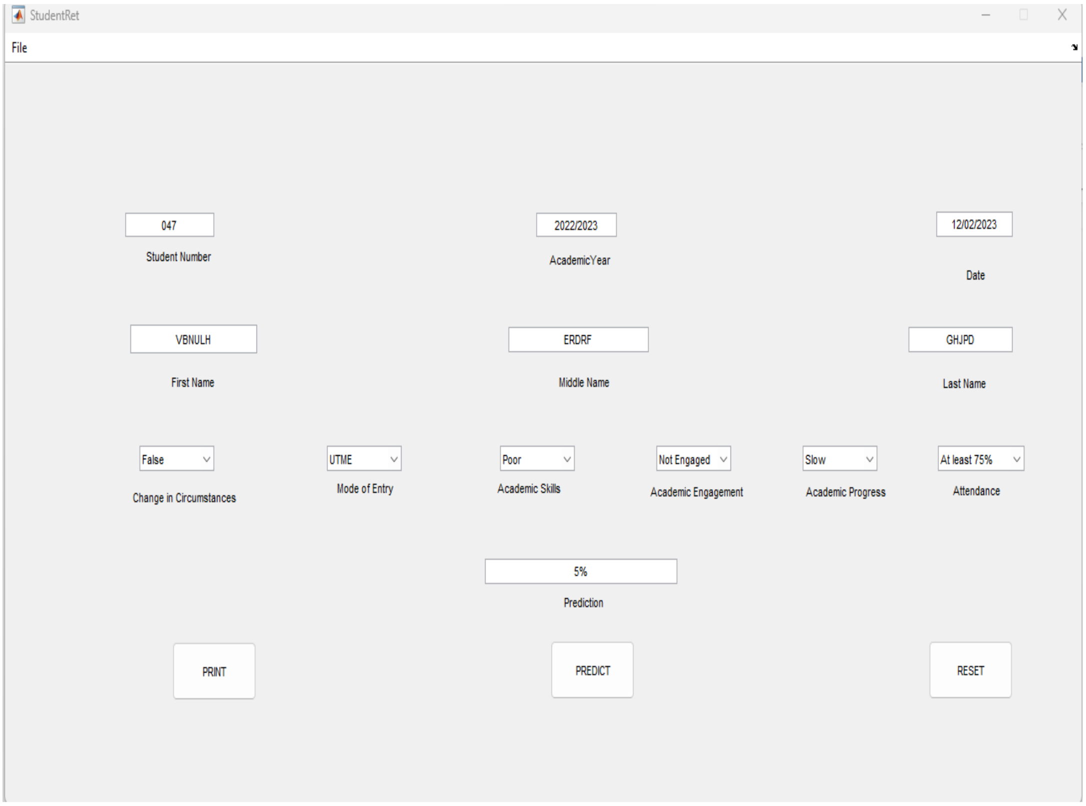


Figure 4-115 (Output for Figure 4-49): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

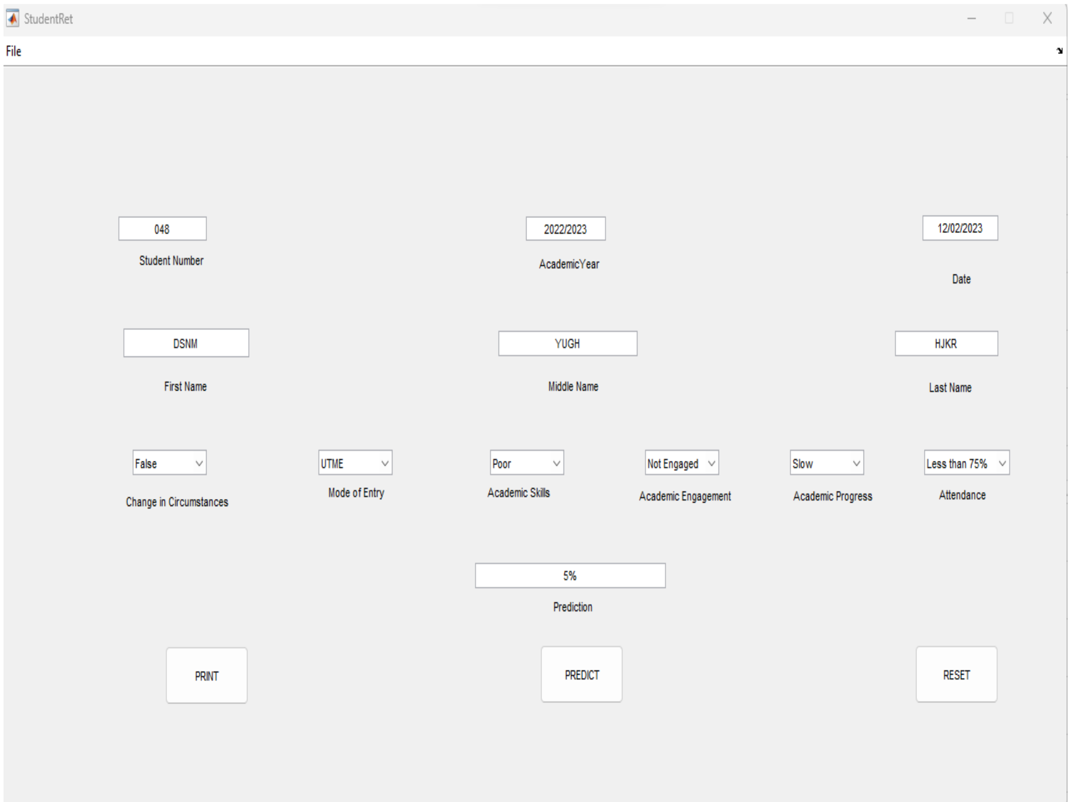


Figure 4-116 (Output for Figure 4-50): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

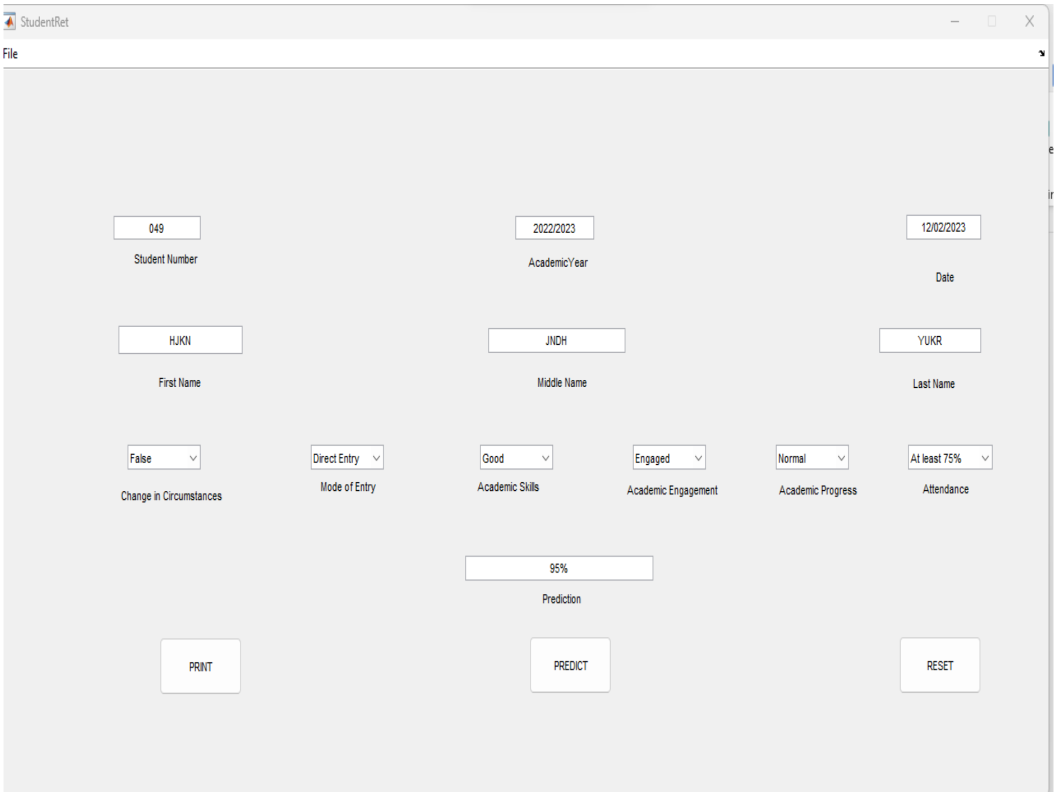


Figure 4-117 (Output foe Figure 4-51): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 95%.

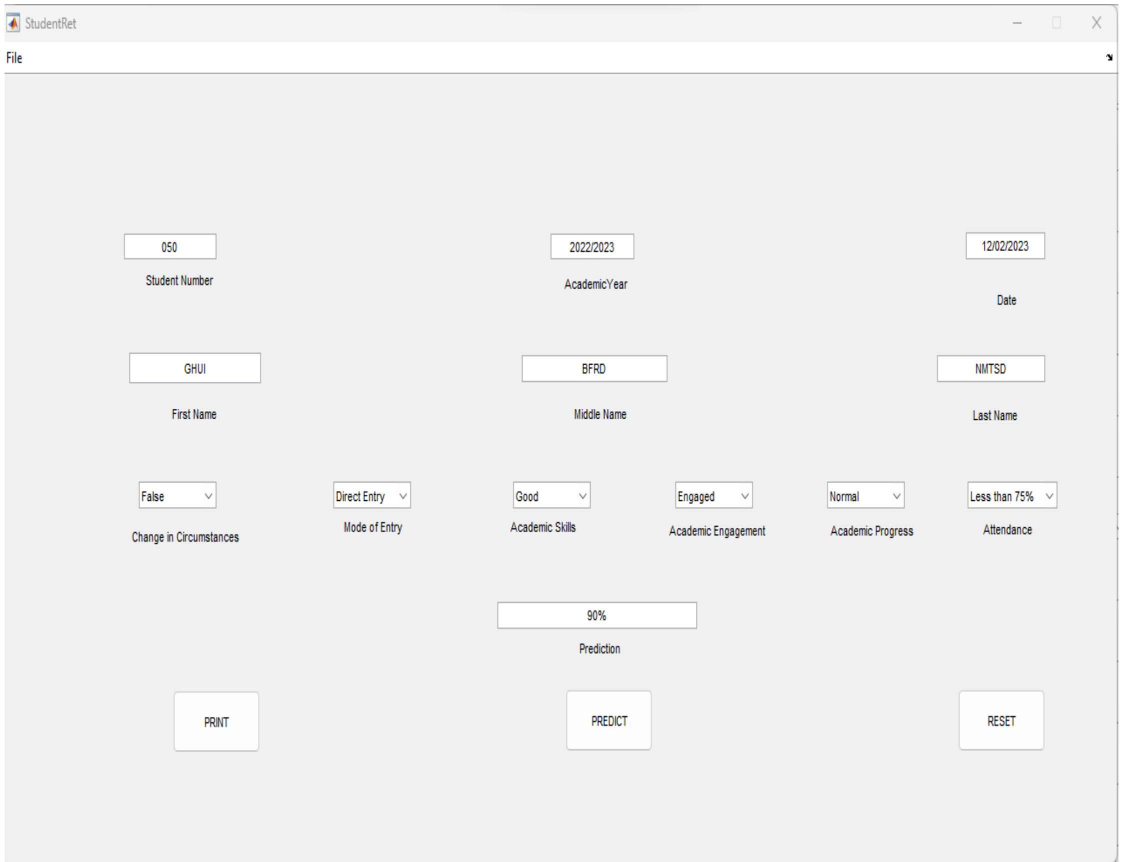


Figure 4-118 (Output for Figure 4-52): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very low risk of attrition as his/her probability of being retained is 90%.

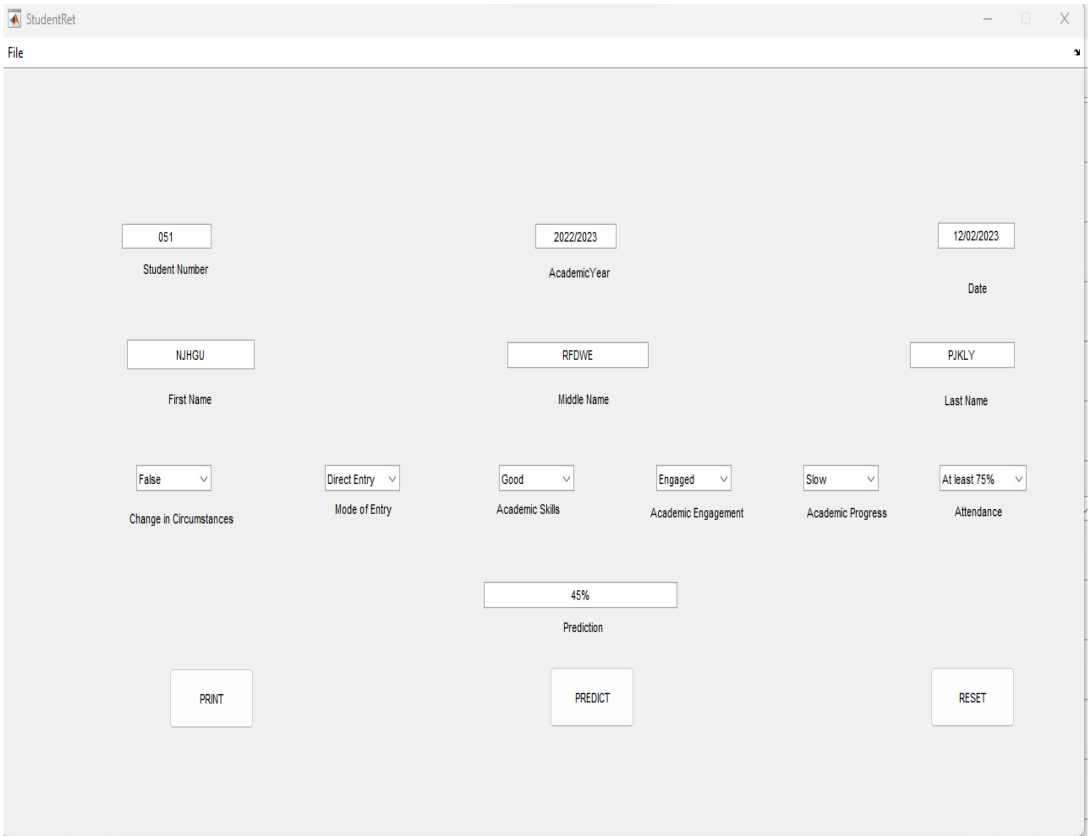


Figure 4-119 (Output for Figure 4-53): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a fairly high risk of attrition as his/her probability of being retained is 45%.

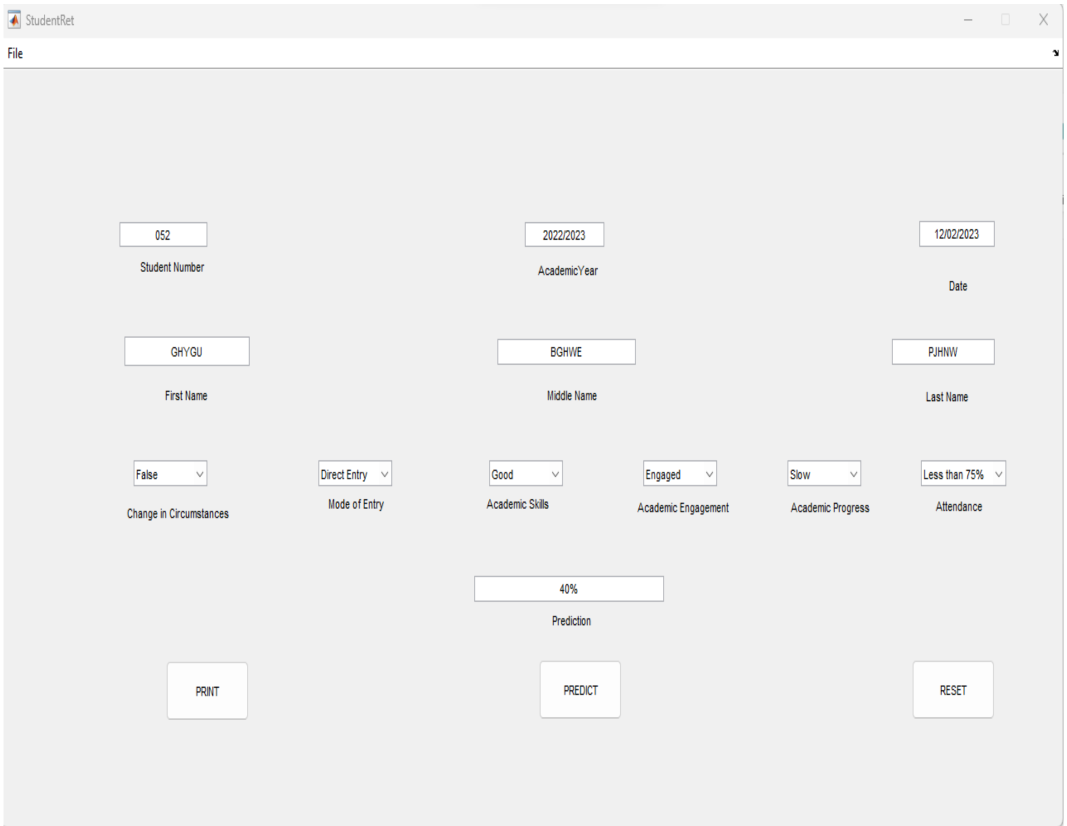


Figure 4-120 (Output for Figure 4-54): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 40%.

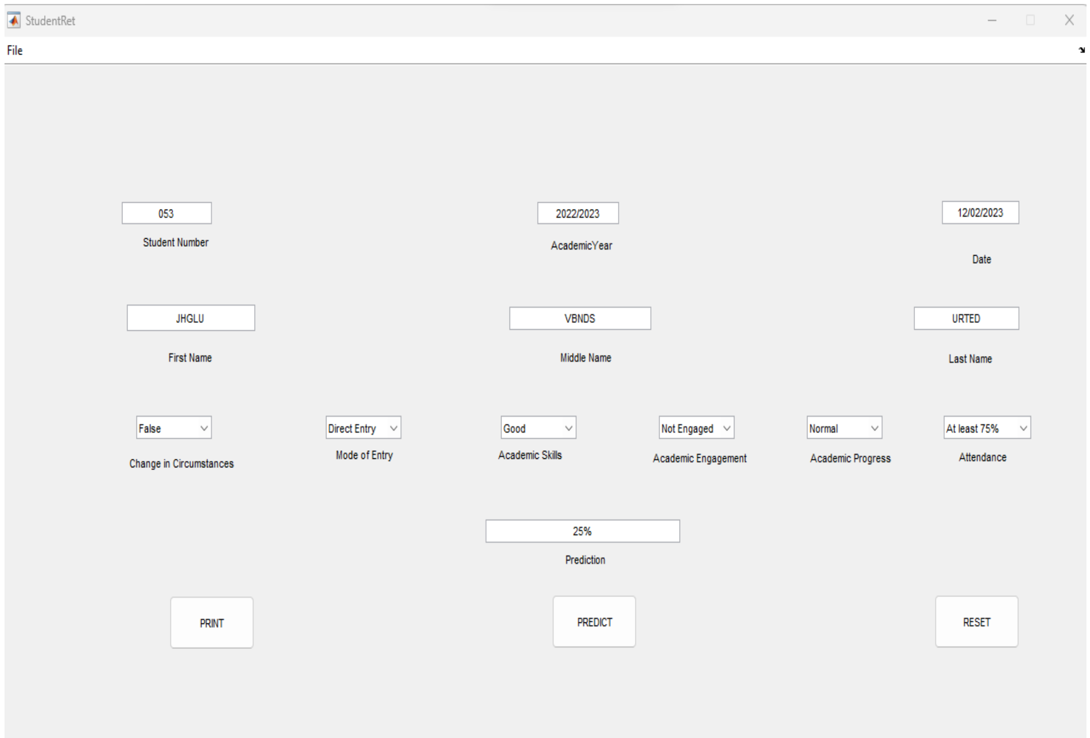


Figure 4-121 (Output for Figure 4-55): Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘good’, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 25%.

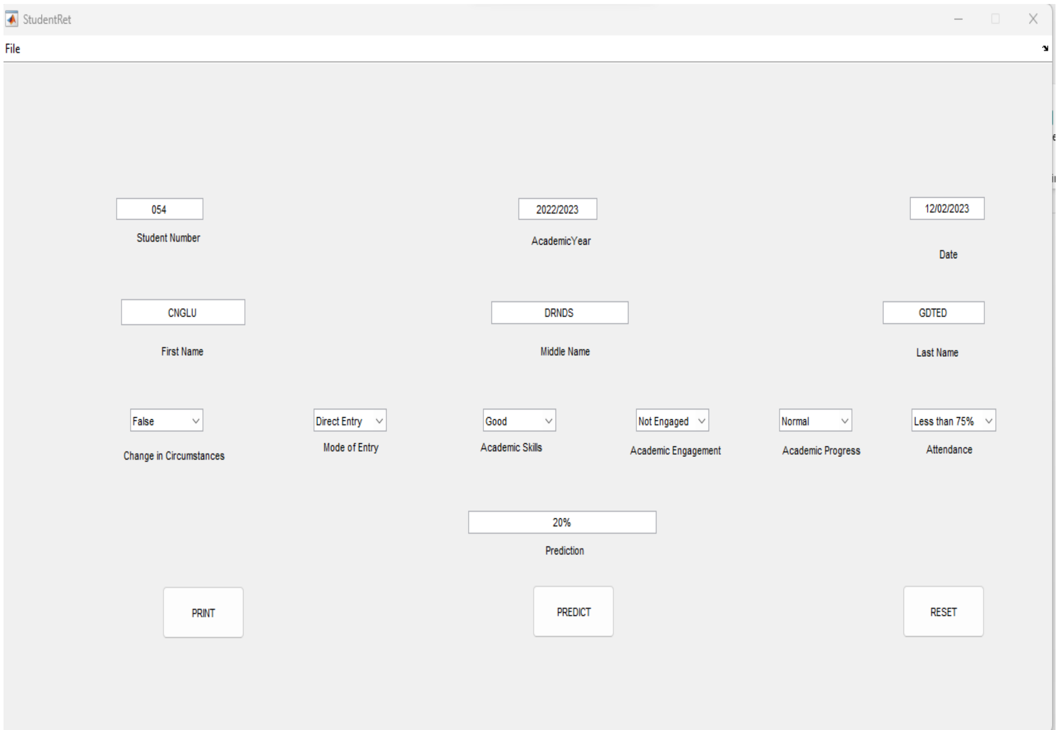


Figure 4-122 (Output for Figure 4-56): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition. as his/her probability of being retained is 20%.

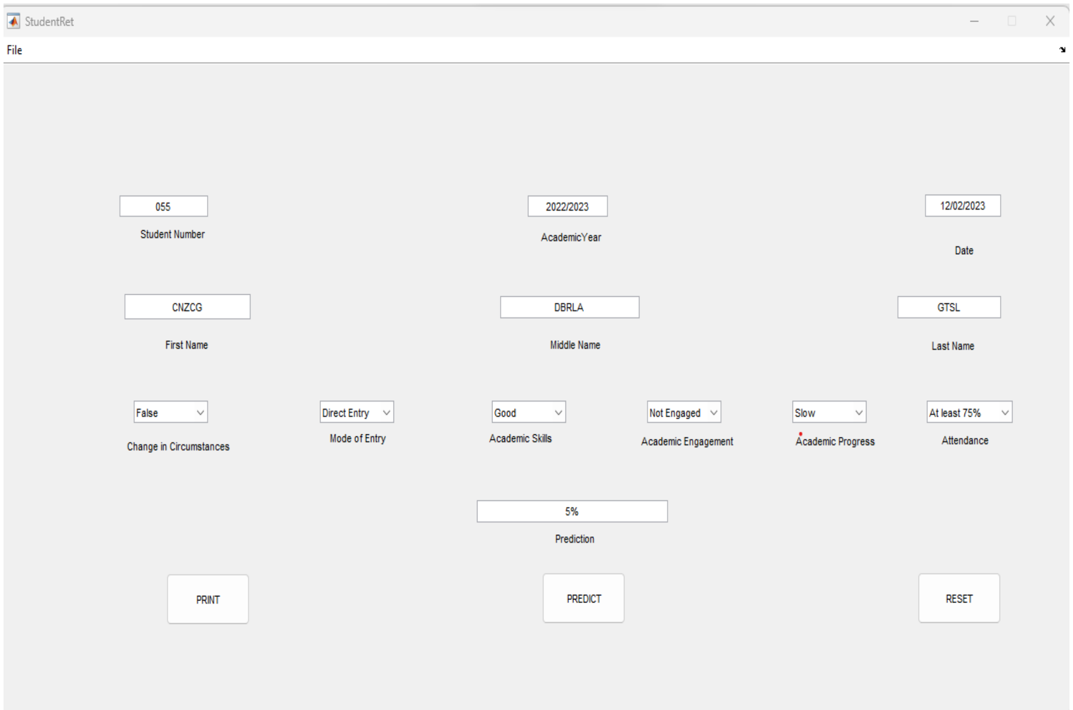


Figure 4-123 (Output for Figure 4-57): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

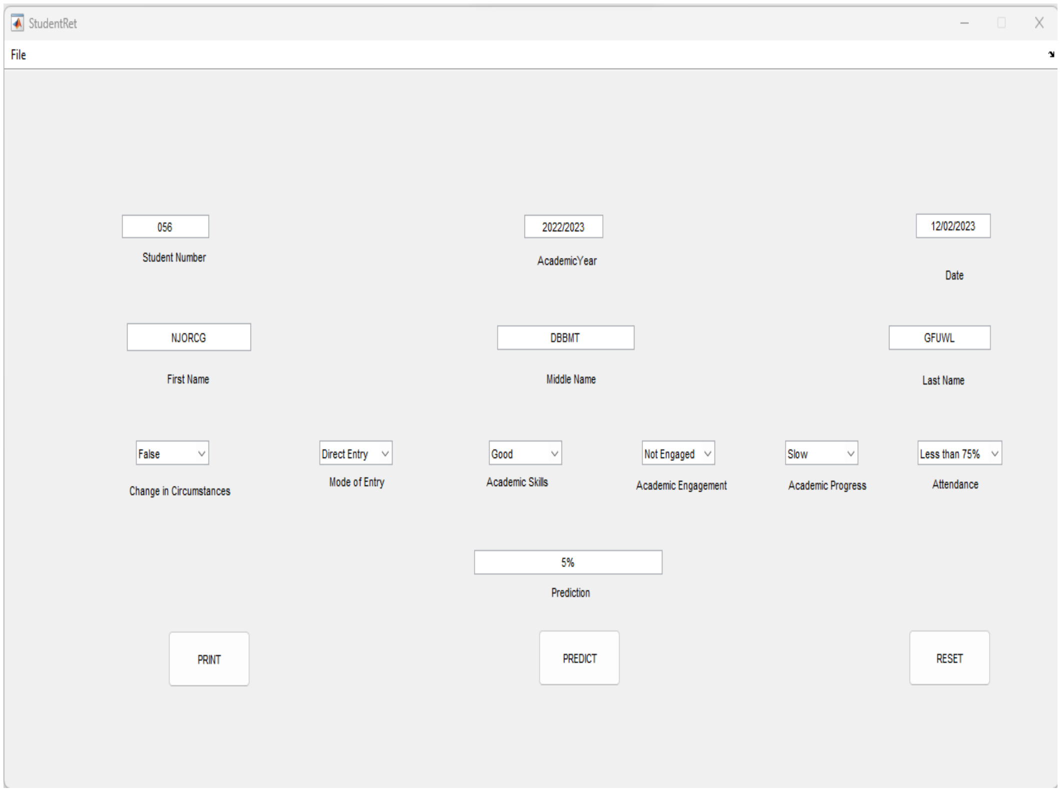


Figure 4-124 (Output for Figure 4-58): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'good', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

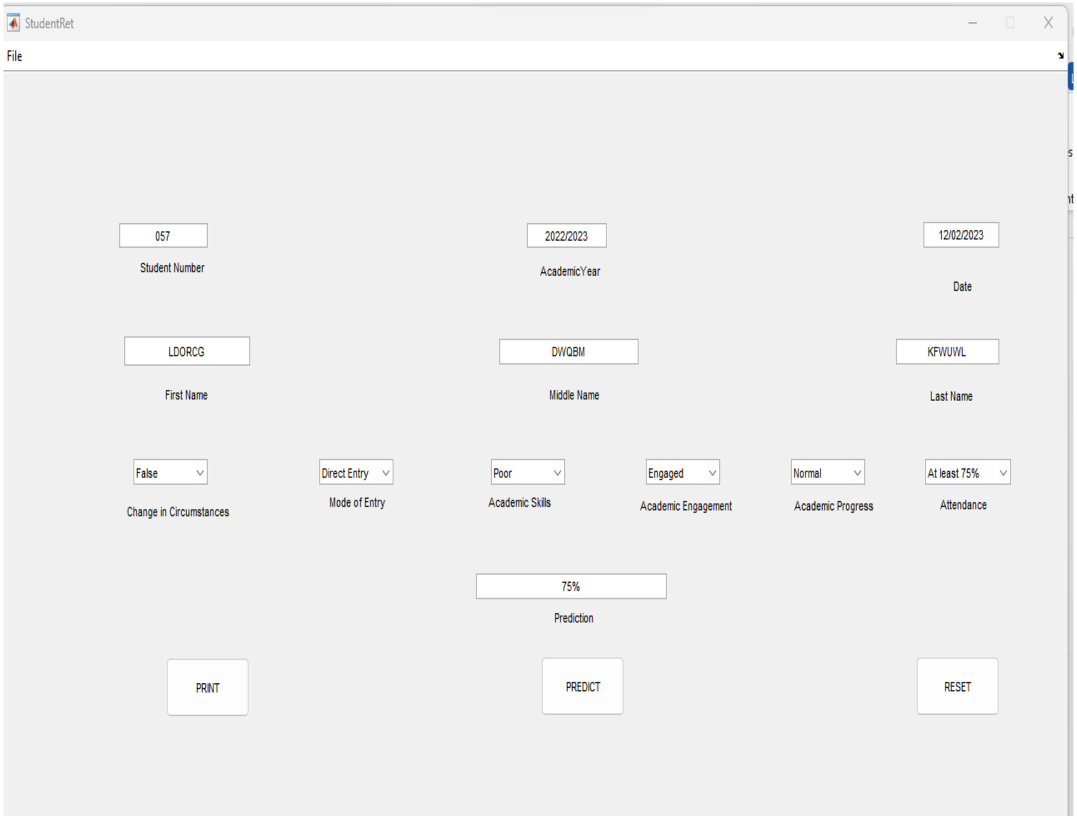


Figure 4-125 (Output for Figure 4-59): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very low risk of attrition as his/her probability of being retained is 75%.

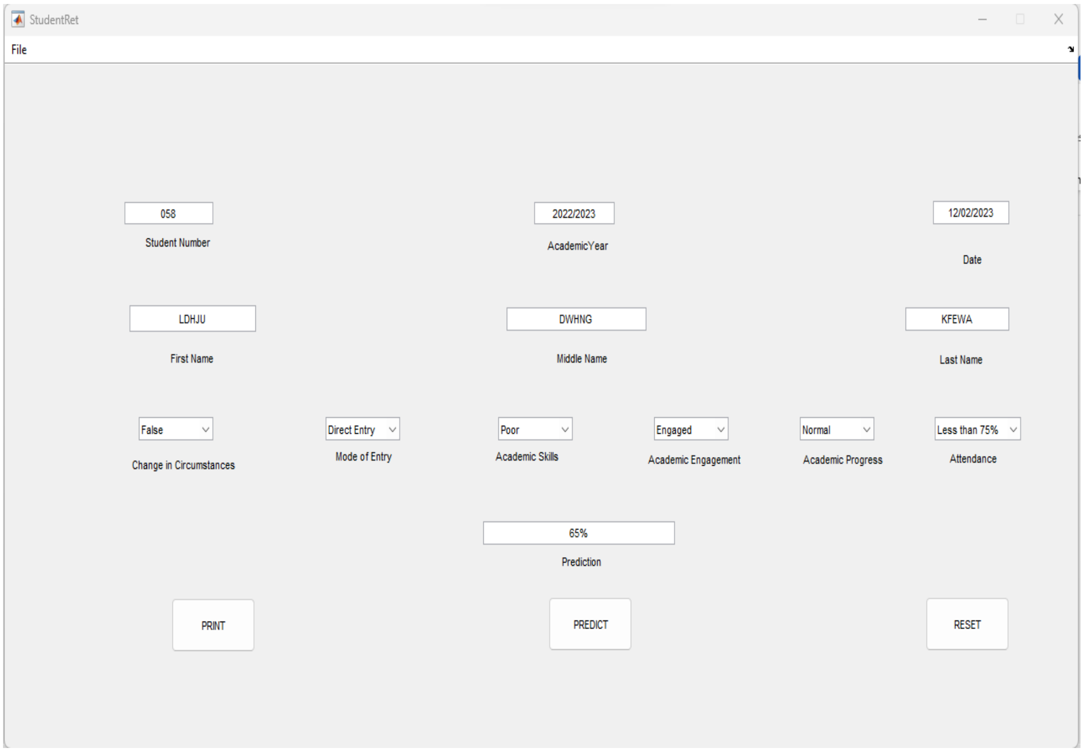


Figure 4-126 (Output for Figure 4-60): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a low risk of attrition as his/her probability of being retained is 65%.

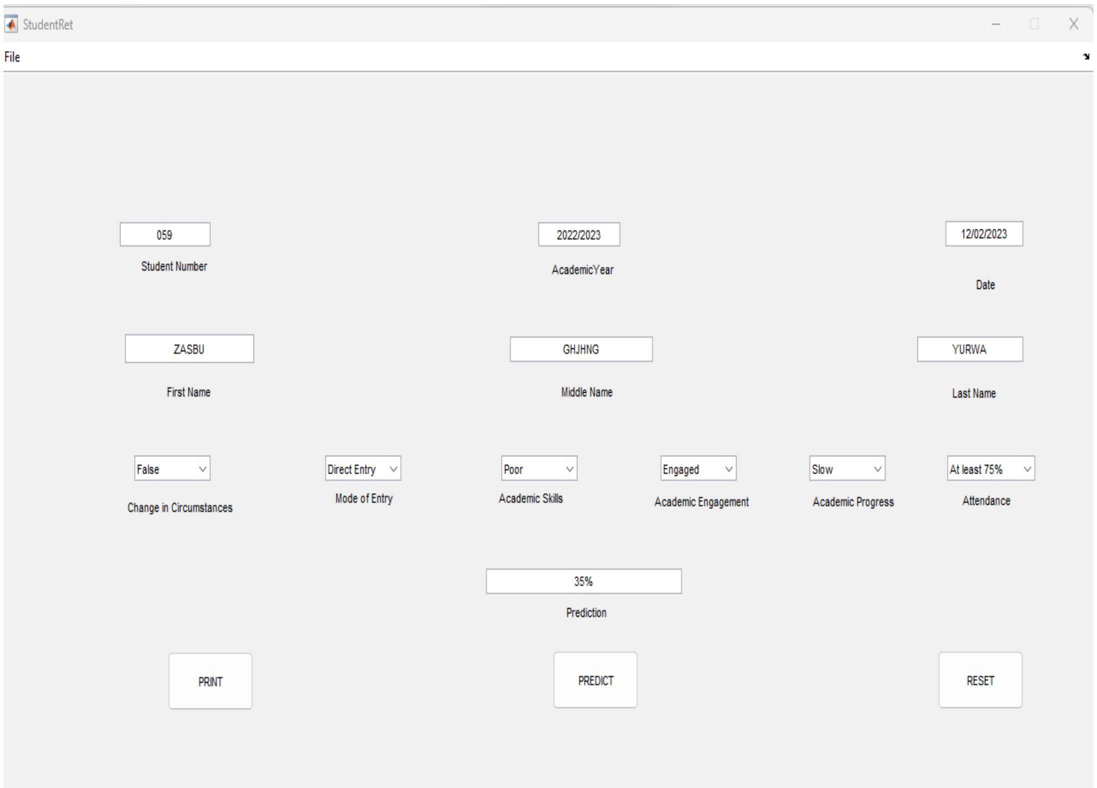


Figure 4-127 (Output for Figure 4-61): Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. This student represents a high risk of attrition as his/her probability of being retained is 35%.

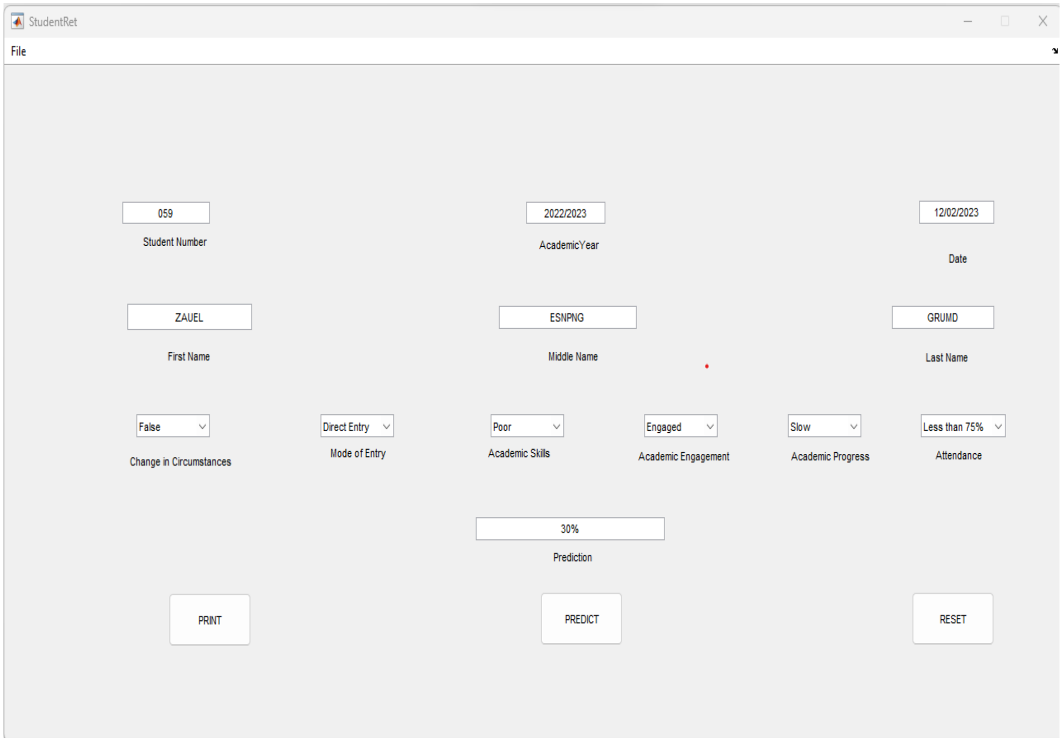


Figure 4-128 (Output for Figure 4-62): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a high risk of attrition as his/her probability of being retained is 30%.

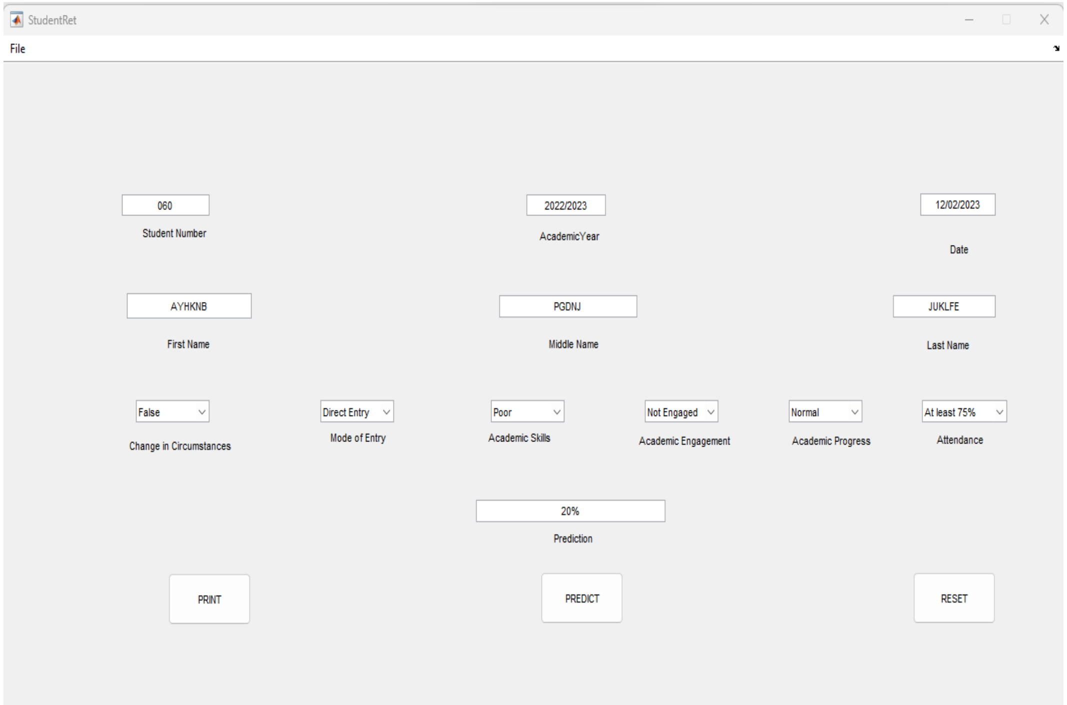


Figure 4-129 (Output for Figure 4-63): Bayesian network prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 20%.

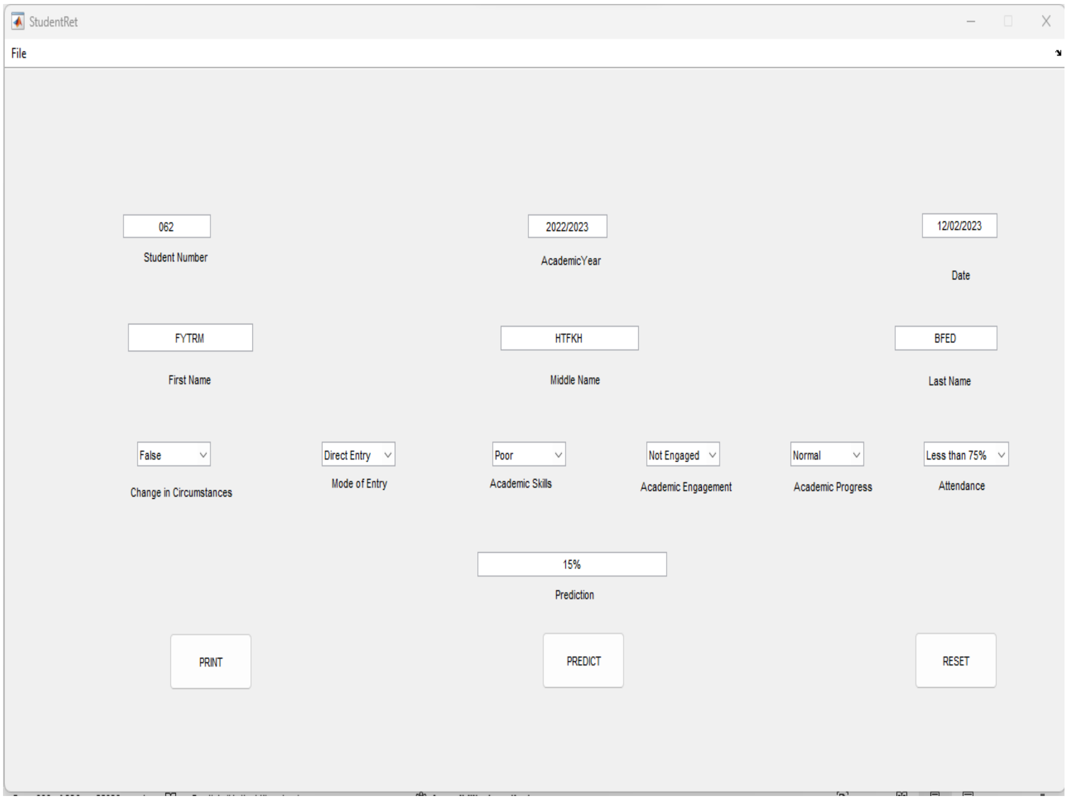


Figure 4-130 (Output for Figure 4-64): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 15%.

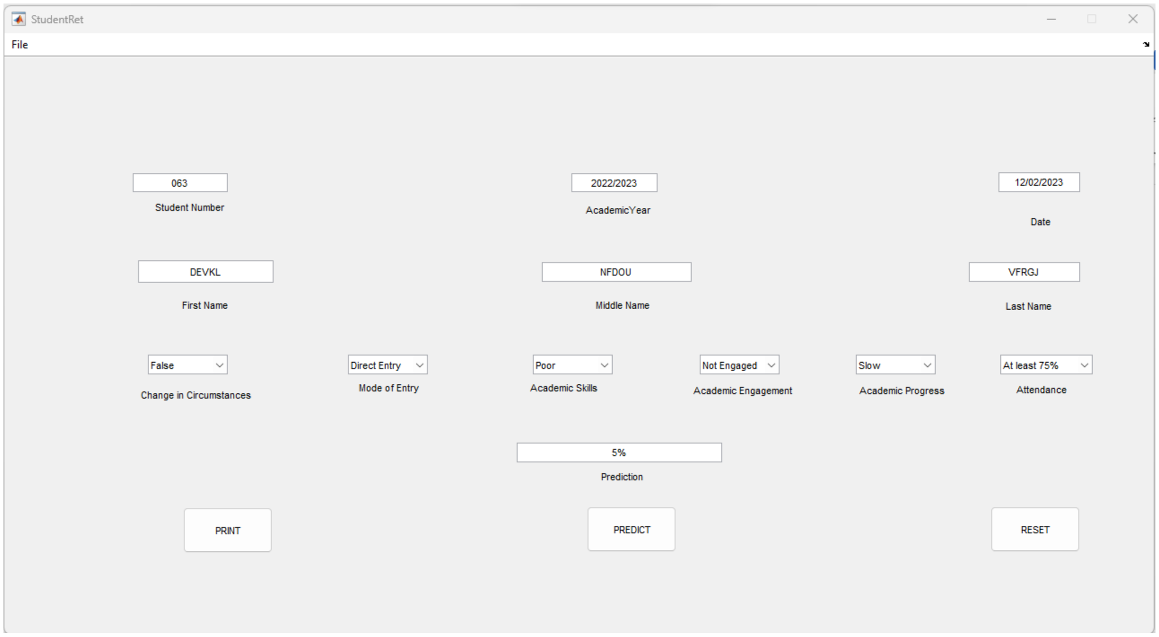


Figure 4-131 (Output for Figure 4-65): Bayesian network prediction for a student whose Change in circumstances is ‘false’, Mode of entry is through ‘Direct entry’, Academic skills is ‘poor’, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is 75% or above. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

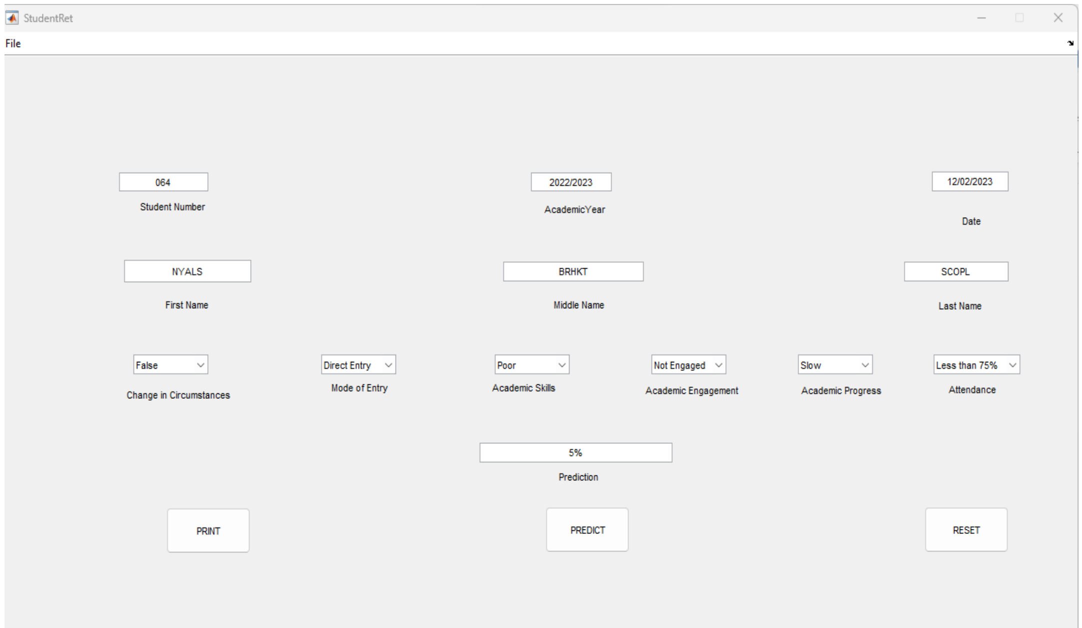


Figure 4-132 (Output for Figure 4-66): Tool's prediction for a student whose Change in circumstances is 'false', Mode of entry is through 'Direct entry', Academic skills is 'poor', Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. This student represents a very high risk of attrition as his/her probability of being retained is 5%.

4.20 Summary and Conclusion

This chapter has discussed the challenges and some solutions associated with the construction of Bayesian networks and elicitation of estimates of probabilities from domain expert(s). If elicitation of estimates of probabilities is not performed carefully, it can lead to obtaining inaccurate estimates of probabilities from domain expert(s). A protocol for an elicitation exercise should be formulated before the exercise itself and be used to conduct the session. The protocol should cover various stages such as background and preparation, identifying and recruiting the expert(s), motivating and training the experts, elicitation of causal structure, elicitation of estimates of probabilities, and recording of the elicitation exercise.

Human judgement in eliciting estimates of probabilities is prone to heuristic and biases. Consequently, a facilitator must check the expert(s) against various types of biases by using debiasing strategies in order to obtain better estimates.
It is important to frame the questions to be answered by the expert(s) properly in order for them to understand the questions well and consequently provide better estimates of probabilities. If questions are framed well, it will motivate the domain experts to provide better estimates of probabilities in an elicitation session.

An elicitation tool typically assists the expert(s) in providing estimates of probabilities during an elicitation session. Furthermore, it improves the quality of the elicitation and saves time as well. The tool should be interactive as well as user-friendly so as to make it easy for the expert(s) to provide their estimates of probabilities and for the end-users to utilise.

The Bayesian networks (BNs) predictions and the software tool predictions can be used by a university to determine the chance of a student of being retained. They can also be used to determine the support(s) needed by students at risk from becoming a dropout.

The results show that in a situation where domain data is sparse or not available at all, the knowledge-driven approach is suitable to be used to elicit estimates of probabilities from domain experts.

Chapter 5

Data Analysis and Results

5.1 Introduction

This chapter aims to present and analyse the estimates of probabilities elicited from the domain experts. These estimates of probabilities served as input data that was analysed. The descriptive analyses were performed on the data to analyse them from the questionnaire designed for this study (see Appendix 6). The descriptive analysis describes or summarizes the characteristics of the data in the experiment while the data were represented in form of Bar charts.

The domain experts were asked to provide the chances of students being retained and consequently complete their study in a university given certain conditions. For instance, how many students are likely to be retained out of every 100 students whose conditions are the same. The domain experts expressed their estimates in form of percentages since they feel comfortable with this format. The variables in the model are qualitative in nature hence descriptive data analysis was employed.

The causal structure for the Bayesian network (BN) was constructed based on knowledge elicited from the domain experts. The estimates of probabilities obtained from the expert was used as parameters for the Bayesian network (BN).

5.2 Data Analysis

The estimates of probabilitiesthat were elicited from the domain experts for this study which served as the data for the study, were entered into MS-Excel software and analysed in Statistical Package for Social Sciences (SPSS) software. In the data analysis, the summary of the estimates of probabilities was conducted for the study. See Table 5-1 to 5-64 in Appendix 10, for the frequency tables.

The Bar charts visually represent and compare the ratio of the students that would be retained to those that would not be retained. Since qualitative responses allow for analysis such as frequency counts and percentages hence they are plotted on the bar charts. Bar charts are useful for presenting and comparing data in a visually clear and easy-to-understand manner when interpreting them. By visually comparing the lengths of the bars, one can quickly identify which category, retained or unretained, has higher or lower value. The comparison helps one to understand relative differences and make informed decisions. For the bar charts representing the data in this study, see Figure 5-1 to 5-64 which can be found in Appendix 10. The tables and bar charts indicate that the more engaged a student is with his or her studies, the higher the probability of being retained.

In order to interpret the Bar charts, the number of categories was determined to be two in number, namely: retained and unretained, Since number of groups is not applicable to the study, this is not determined.at all. The retained category has the highest frequency (95%) when a student's: circumstances have not changed, enters the university with a qualification which is higher than GCSE, possesses academic skills, engages with his or her study, makes a normal progress, and achieves at least 75% attendance, while the unretained category has the highest frequency (95%) when a student's: circumstances have changed, enters the university with a GCSE, lacks academic skills, unengages with his or her study, makes a slow progress, and unable to achieve 75% attendance.

Conversely, the retained category has the lowest frequency (5%) when a student's: circumstances have changed, enters the university with a GCSE, lacks academic skills, unengages with his or her study, makes a slow progress, and achieves lower than 75% attendance, while the unretained category has the lowest frequency (5%) when a student's: circumstances have not changed, enters the university with a qualification which is higher than GCSE, possesses academic skills, engages with his or her study, makes a normal progress, and able to achieve at least 75% attendance.

5.3 Results

The domain experts were elicited to provide estimate of probability for each row which consists the states (values) of the nodes (variables) in the Bayesian network (BN). The estimate of probability for each row serves as the probability for the combination of various states (values) in that particular row. Each of the rows in the questionnaire (see Questionnaire 6) is called a scenario.

The probability that a student would be retained or otherwise is used to discover students that are at risk at a university so that the university will determine the corrective measures to take to ensure the students are retained, e. g, staff/student relationship, pastoral care, mentorship, peer relationship, peer mentorship.

If a student is able to engage in his or her studies, and also combined this with acadmic skills, the peobability that such a student will be retained and graduate is very high. Therefore, academic engagement is directly proportional to retention.

Attendance has a direct bearing on retention as any student who attends lectures regularly and punctually, attend seminars regularly, and also turns in his or her assignments and term papers by the deadline is more likely to be retained and graduate. In the case of any students whose circumstances have changed due to career, family ties, missing home/friends, religion, dietary, etc, lack academic skills, and unengaged with his or her studies and also exhibiting poor attendance, such a student is more unlikely to be retained.

If a student acquires academic skills he or she stands a better chance of being retained and graduate as academic skills is directly proportional to retention.

The mode of entry is another important factor since a student that entered a university with a qualification higher than GCSE (e.g A/L, BTEC, etc.) is more likely to be retained and graduate than another student who entered with a GCSE. This is due to the previous learning at his or her former college.

Change in circumstances of a student have an adverse effect on retention since this factor usually cause students to abandon their studies. Change in circumstances could be as a result of career, family ties, missing home, religion, dietary, etc. In the case of a student whose circumstances have changed, unengaged in his or her studies and does not acquire enough academic skills, such a student has a limited chance of being retained.

In a situation where a student’s circumstances have not changed, possesses academic skills, engaged with his or her studies coupled with required attendance, such a student stands a better chance of being retained and graduate.

During the elicitation of estimates of probabilities from the domain experts in the session, they were debiased against heuristics and biases such as: availability, anchoring and adjustment, and representativeness. The types of heuristics and biases enumerated above usually affect the accuracy of the estimates provided by domain experts, which in turn might have affected the results shown in the tables, bar charts, and the outputs from the software tool.

Furthermoore, the domain experts were asked to justify their estimates by asking them how they arrived at the given estimates, and they claimed that they based their estimates on experiences gained in previous elicitation sessions.

The challenge of translating domain experts' qualitative assessments into quantitave assessments was overcome by assigning numerical values or codes to qualitative responses as this approach allows for quantitative analysis such as frequency counts and percentage. This translation had no adverse effect on the the accuracy of the Bayesian network (BN) model and results from the tool.

5.4 Discussion of Findings

The study found out that when staff-student relationship is good, a student is encouraged to complete his study and graduate. In the same vein, when peer relationship is good a student relates with his colleagues and succeed.

When a student enters a university with a minimum qualification of advanced level (A/L) such a student is more likely to perform academically better than a student that enters with GCE ordinary level (GSCE). Hence, a student with a previous academic qualification that is higher than the latter is more likely to complete his programme of study.

Similarly, a student who studies on full-time basis is more likely to perform academically better than a student who studies on part-time basis. Since a full-time student would devote more time to his study and concentrate on it as well, he is more likely to complete his programme of study than a student who studies on part-time basis.

When a student develops academic skills such as critical thinking, reflection, verbal, and written communications, the student performs well academically and this prevents the student from leaving the school before graduating.

If a student is confronted with financial problems, he will find it difficult to pay tuition fees as well as to buy educational materials for his needs and this will affect his academic performance as he will not be able to concentrate on his studies. This situation can be stopped by receiving financial support from partner, family and friends, as well as receiving loan, aid, bursary or scholarship from government and/or university.

Journey time is another factor that hinders a student from performing well and consequently complete his study. It is advisable a student travels within one hour from his university in order to minimise cost.

When a student is healthy, he will be able to attend classes, engage with his studies, participate, and consequently graduate. Health reasons could be either physical, mental or both.

When a student attends lectures regularly coupled with other factors, he is more likely to make substantial academic progress. Hence, it is a requirement that a student needs to achieve at least 75% attendance before sitting for an examination at the end of a semester.

Due to the motivation received from the researcher in terms of the protocol used for the elicitation, the way questions were framed and the usage of a software tool during the elicitation of the Bayesian network (BN) causal structure and estimates of probabilities, this results in provision of better estimates of probabilities by the domain experts as evidenced by the predictions from the Bayesian network (BN) and the software tool.

Motivation is a function of the anticipated likelihood that domain experts' effort will lead to provision of better estimates of probabilities in an elicitation session, and that performance will lead to certain outcomes that are valuable. The likelihood that domain experts' effort will result in performance is based in part on previous expreciences in similar elicitation sessions. The expectation that a particular level of performance would result in a given outcome is known as performance to outcome expectancy.

Domain experts who set objectives are more likely to perform at greater levels in an elicitation session than those who do not set goals. Setting goals entails more than simply urging domain experts to try their best. Setting tough goals and engaging in the goal-setting process have also been demonstrated to improve domain experts performance in an elicitation session.

Conclusively, motivation normally causes domain experts to think hard in order to provide better estimates of probabilities in an elicitation sesion. Furthermore, motivation is capable of overcoming heuristics and biases thereby resulting in better results as well.

Chapter 6

Summary, Conclusion and Recommendations

6.1 Introduction

This chapter presents the summary of the study, conclusions, discusses the limitations of this study, problems encountered and also enumerates suggestions for further research.

6.2 Summary of the Study

This thesis has described the previous work done in the areas related to this study, namely elicitation of causal structure (graph) and estimates of probabilities from domain experts, student retention as a challenge in higher education institutions, and Bayesian network (BN).
The chapter has enumerated various areas where elicitation of estimates of probabilities had been used as a source of data in situations where there was little or no data.
Furthermore, it discusses student retention as a challenge in higher education institutions. This study used student retention as an application area for the combination of expert knowledge and Bayesian network (BN).
Conclusively, the combination of elicitation of causal structure and estimates of probabilities from domain experts, and Bayesian network (BN) have been and can still be deployed successfully in many applications.
If elicitation of estimates of probabilities is not performed carefully, it can lead to inaccurate estimates of probabilities from domain expert(s). A protocol for an elicitation exercise should be formulated before the exercise itself and be used to conduct the session. The protocol should cover various stages such as background and preparation, identifying and recruiting the expert(s), motivating and training the experts, elicitation of causal structure, elicitation of estimates of probabilities, and recording of the elicitation exercise.

Human judgement in eliciting estimates of probabilities is prone to heuristic and biases. Consequently, a facilitator must check the expert(s) against various types of biases by using debiasing strategies in order to obtain better estimates.
It is important to frame the questions to be answered by the expert(s) properly in order to obtain accurate estimates of probabilities from them.

An elicitation tool typically assists the expert(s) in providing estimates of probabilities during an elicitation session. Furthermore, it improves the quality of the elicitation and saves time as well. The tool should be interactive as well as user-friendly so as to make it easy for the expert(s) and end-users to utilise.

The study shows that in a situation where domain data is sparse or not available at all, the knowledge-driven approach is suitable for obtaining estimates of probabilities for creating models.

Due to the motivation received from the researcher in terms of the protocol used for the elicitation, the way questions were framed and the usage of a software tool during the elicitation of the Bayesian network (BN) causal structure and estimates of probabilities, this results in provision of better estimates of probabilities by the domain experts as evidenced by the predictions from the Bayesian network (BN) and the software tool.

Motivation normally causes domain experts to think hard in order to provide better results. Motivation is capable of overcoming heuristics and biases thereby resulting in better results as well. The impact of motivation on elicitation of Bayesian network (BN) causal structure and estimates of probabilities cannot be overemphasised.

Predicting student retention in higher education institutions (HEIs) is paramount because it allows the educational institutions to give them necessary supplementary support, such as personalized personal assistance and tutoring resources. Furthermore, the findings of the predictions can be used by the lecturers to detect the most relevant teaching materials and actions for each set of students to meet their needs.

Conclusively, predicting student retention in order to detect students at risk in the early stages of education is essential so that higher education institutions (HEIs) can minimise students not graduating on time or drop out outrightly.

6.3 Limitations of the Study

First, this study is restricted by scarcity of literature involving the trio : estimates of causal structure and estimates of probabilities; Bayesian network (BN); and student retention.

on impact of motivation on elicitation. Second, student retention experts who participated in the study were few in number.

The impromptu nature by which the student retention experts were given the research instrument (questionnaire) see Appendix 6 to respond, and return could result to possibility of affecting their estimates of probabilities.

The result of this study cannot be generalized as student retention varies from one country to another as well as from one institution to another.

6.4 Problems encountered

First, the experimenter encountered the problem of getting people to engage in this study, that is domain experts. Second, only ten participants out of forty that were recruited for this study eventually took part. Furthermore, the study is iterative in nature as it involves contacting the domain experts on several occasions.

From the literature, the reasons behind domain experts' unwillingness to participate in probability estimates elicitation session (probability encoding session) are listed below (in paragraph four of this section): Also, we have designed a questionnaire and sent it to those who have participated in probabilistic estimates elicitation session either as a domain expert or as a facilitator in order for them to share their experience in elicitation with us. Some of the reasons obtained from the literature and domain experts as to why domain experts do not like to participate in encoding process are enumerated below:

An expert may know that he or she performs the task in simple, intuitive ways that, if revealed, would reduce the esteem others hold for him or her.

The expert may not know how he/she performs that task and may be reluctant to express the uncertainty, thinking that experts are expected to be rational and articulate.

The expert may believe that if his/her expertise can be captured in a computer, this might cause him/her to lose his/her job.

The expert may think that a company is willing to invest the time and money to clone his or her expertise, to allow many more problems to be solved with the knowledge he or she has gained (Olson, 1987).

Lack of previous experience in encoding process: Experts make predictions based on analogies to previously experienced process(es) (Cooke, 1985). In this case, an expert who had not participated in an encoding process in the past, will not like to take part in an elicitation session.

Failure to take responsibility for their estimates (De Finetti, 1974). An expert may not have confidence in providing estimates hence he or she may decline to provide probability distributions.

An expert would not like to participate in encoding sessions as a result of epistemic uncertainty due to lack of knowledge about the proportion of the population (Oakley, 2010).

A domain expert would not like to participate in encoding sessions as a result of aleatory uncertainty due to randomly sampling of the proportion of the population (Oakley, 2010).

Time: Experts are busy people, and their time is precious hence they may not like to participate in probabilistic encoding sessions as this may conflict with their personal/work schedule.

Data Protection Act: Due to ethical issues about organisational data, an expert may decide to not provide probability distributions in order not to violate data protection act (The UK Department for Digital, 2018).

Incentivisation: An expert might not like to participate in an encoding session if he or she feels that there is nothing for him or her to gain from the exercise (Deci, 2000).

6.5 Suggestions for further research

The researcher is with the opinion that all prospective researchers in this area should undertake the following for further investigation into the multidisciplinary research area which involves elicitation of causal structure and estimates of probabilities, and Bayesian network (BN) using student retention as an application area.

1. The study should be replicated using more variables (nodes) and states.
2. The study should be expanded to cover students at postgraduate level.
3. In order to maintain the validity of the result of this research, this could be repeated at interval of 3 years.

References

Aczel, B., Szollosi, A., Bago, B. and Foldes, A. (2015). Is it time for studying real-life debiasing? Evaluation of the effectiveness of an analogical intervention technique. Journal of Frontiers in Psychology. pp. 1-13.

Aguilera, P.A., Fernandez, A., Fernandez, R., Rumi, R., Salmeron, A. (2011). Bayesian networks in environmental modelling. Environmental Modelling & Software 26 (12), 1376-1388.

Ahmad, F., Ismail, N. H., and Aziz, A. A. (2015). The Prediction of Students' Academic Performance Using Classification Data Mining Techniques. J. of Applied Mathematical Sciences, Vol. 9, pp. 6415-6426.

Al-Akwaa, A. M., Siddiqui, N., and Al-Mofleh, I. A. (2004). Primary gastric lymphpoma. World Journal of Gastroenterology, 10(1): 5-11.

Alhussain, Z. A. and Oakley, J. E. (2020a). Assurance for clinical trial design with normally distributed outcomes: "Eliciting uncertainty about variances." Pharmaceutical Statistics, 19(6): 827-839. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/pst.2040.

Ames, D.P., Neilson, B.T., Stevens, D.K., Lall, U. (2005). Using Bayesian networks to model watershed management decisions: An East Canyon Creek case study. Journal of Hydroinformatics 7 (4), 267-282.

Ansari, M., Nasrolahi, H., Kani, A. M., Hamedi, S. H., Razzaghi, S., Ahmadloo, N., Mohammadianpanah, M., Omidvari, S., and Mosalaei, A. (2013). Primary Non-Hodgkin's usLymphoma of Stomach: To report 54 patients and Analysis of Major Reported Series. [online] Available from: http://dx.doi.org/10.1155/2013/583826. [Accessed on: 30 December 2020].

Banegas, D. L., and Villacañas de Castro, L. S. (2015). A look at ethical issues in action research in education. *Argentinian Journal of Applied Linguistics*, 3(1), pp. 58-67.

Best, N., Dallow, N., and Montague, T. (2020). Prior elicitation." Bayesian methods in

Bilmes, J. A. (2004) Graphical models and automatic speech recognition. In: *Mathematical foundations of speech and language processing*, Springer New York, pp. 91-245.

Boets, P., Landuyt, D., Everaert, G., Broekx, S. and Goethals, P. L. M. (2015). Evaluation and comparison of data-driven and knowledge-supported Bayesian Belief Networks to assess the habitat suitability for alien macroinvertebrates. Environmental Modelling and Software 74, pp. 92-103.

Boiney, L. G., Kennedy, J., and Nye, P. (1997). Instrumental bias in motivated reasoning: more when more is needed. Journal of Organizational Behaviour and Human Decision Processes. 72(1) pp. 1-24.

Bonano, E. J., Hora, S. C., Keeny, R. L., and von Winterfeldt, D. (1989). Elicitation and use of expert judgement in performance assessment for high-level radioactive waste repositories NUREG/CR-5411. Washington: U> S. Nuclear Regulatory Commission.

Brase, G. L. (2009). How different types of participant payments alter task performance. Judgement and Decision Making, Vol. 4(5), pp 419-428.

Brase, G. L. (2009). How different types of participant payments alter task performance. Judgement and Decision Making, Vol. 4(5), pp 419-428.

Buehler, R., Griffin, D. and MacDonald, H. (1997). The role of motivated reasoning in optimistic time predictions. Journal of Society for Personality and Social Psychology 23(5) pp 238-247.

Buehler, R., Griffin, D. and MacDonald, H. (1997). The role of motivated reasoning in optimistic time predictions. Journal of Society for Personality and Social Psychology 23(5), pp 238-247.

Burton, D. and Bartlett, S. (2016). Introduction to Education Studies: Fourth Edition. London: SAGE Publications.

Bytheway, A. & Venter, I. M. (2014). Strategies for information management in education: ome international experience. South African Journal of Information Management, 16(1), 1-11.

Caruth, G. D. (2018). Student engagement, retention, and motivation: Assessing academic success in today's college students," *Participatory Educ. Res.*, vol. 5, no. 1, pp. 17_30, Dec. 2018, doi: 10.17275/per.18.4.5.1.

Chen, X. (2009). Cognitive and motivational parameters in motivated biases in human judgement. Unpublished manuscript, University of Maryland, USA.

Chickering, D. M. and Heckerman, D. (2000). A comparison of scientific and engineering criteria for Bayesian model selection. *J. of Statistics and Computing* **10,** 55–62.

Cook, D. C., Sheppard, A., Liu, S. and Lonsdale, W. M. (2015). Predicting the economic impacts of inverse species: The eradication of the giant sensitive plant from Western Australia. In Pest Risk Modelling and Mapping for Invasive Alien Species. (Ed. Venette, R. C.) Oxford: Wallinford, pp 145-161.

Cooke, N. M. (1985). Modelling Human Expertise in Expert Systems. Technical report, MCCS-85-12, Computing Research Laboratory, New Mexico State University, USA.

Coupé, V.M.H. and Van De Gaag, L. C. (2002). Properties of Sensitivity Analysis of Bayesian Belief Networks. *Annals of Mathematics and Artificial Intelligence, 36, 323-356.*

Creswell, J. W. (2014). Research design: qualitative, quantitative, and mixed methods approaches (4th ed.). Thousand Oaks: SAGE Publications. Festinger & DeMatteo.

Daly, R., Shen, Q., and Aitken, S. (2011). Learning Bayesian networks: approaches and issues. *The knowledge engineering review*, *26*(02), 99-157.

Darlington, K. (2000). The Essence of Expert Systems. London: Pearson Education limited. pp 43-52. ISBN 0-13-022774-9.

David, H. A. (1988). The method of paired comparisons. New York: Oxford University Press.

De Finetti, B. (1974). Theory of Probability: A critical Introductory Treatment, Vol. 1, London: John Wiley & Sons.

Deci, E. L. and Ryan, A. M. (2000). The “what and why” of goal pursuits: Human needs and self-determination of behaviour. Psychological Inquiry, 11: 227-268.

Dunn, J. C. (2016). Bayesian Networks with Expert Elicitation as Applicable to Student Retention in Institutional Research. PhD. Thesis, Georgia State University, USA.

elicit many probabilities.
https://www.researchgate.net/publication/221404281_
How_to_Elicit_Many_Probabilities. Accessed on: 10th October 2017.

Feamster, N and Balakrishnan, H. (2004). The case for separating routing from routers.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978), Fault trees: Sensibility of estimated failure probabilities to problem representation, *Journal of Experimental Psychology: Human Perception and Performance*, 4, 330–344.

Forio, M. A. E., Landuyt, D., Bennetsen, E., Lock, K., Boets, P., Everaert, G., Dominguez-Granda, L. and Goethals, P. L. M. (2015). Bayesian belief network models to analyse

FTU (2020). The University Calendar (1st ed.). A Publication of the Office of the Vice-Chancellor, First Technical University, Ibadan, Nigeria, p.

Gardener, D. (2008) Risk. The science and politics of fear. London: Virgin Books Ltd.

Garthwaite, P. H., Kadane, J. B. and O'Hagan, A. (2005). Statistical Methods for Eliciting Probability Distributions. Pittsburgh: Carnegie Mellon University.

Gosling, J. P. (2014). Methods for eliciting expert opinion to inform health technology assessment. Journal of Health.

Grist, E.P.M, O'Hagan, A., Crane, M., Sorokin, N., Sims, I. and Whitehouse, P. (2006). Bayesian and Time-Independent Species Sensitivity Distributions for Risk Assessment of Chemicals. Journal of Environmental Science and Technology 2006(2).

Harper, F.T., Hora, S.C., Young, M.L. Miller, L.A. Lui, C.H. McKay, M.D. Helton J.C., Goossens, L.H.J., Cooke, R.M., Pasler-Sauer, J., Kraan, B. & Jones, J.A. (1994). *Probability Accident Consequence Uncertainty Analysis, Vols. 1-3, (NUREG/ CR-6244, EUR 15855 EN)* Brussels: USNRC and CEC DG XII.

Hasannejad, M. R., Zoghi, M., & Asl, H. D. (2017). Motivation, tasks, attitudes: The influence of motivational pre-task strategies on tasks performance and tasks engagement. *The Journal of Language Teaching and Learning*, 7(2), 113-128.

Hasannejad, M. R., Zoghi, M., & Asl, H. D. (2017). Motivation, tasks, attitudes: The influence of motivational pre-task strategies on tasks performance and tasks engagement. *The Journal of Language Teaching and Learning*, 7(2), 113-128.

Haug, P., Koehler, S., Christensen, L., Gundersen, M. & Van Bree, R. (2001). *Probabilistic method for natural language processing and for encoding free-text data into a medical database by utilizing a Bayesian network to perform spell checking of words,* U.S. Patent No. 6,292,771. Washington, DC: U.S. Patent and Trademark Office.

Haverila, M., Haverila, K. and McLaughlin, C. Variables affecting the retention intentions of students in higher education institutions, *J. Int. Students*, vol. 10, no. 2, pp. 358_382, May 2020, doi: 10.32674/jis.v10i2.1849.

Heckerman D. (1998) An empirical comparison of three inference methods. The Proceedings of the Fourth Workshop on Uncertainty in Artificial Intelligence.

Heckerman, D. (1991). Probabilistic Similarity Networks. MIT Press, Cambridge, Massachusetts.

Mwiya, B., Bwalya, J., Siachinji, B., Sikombe, S., Chanda, H. and Chawala, M. (2017). Higher Education Quality and Student Satisfaction Nexus: Evidence from Zambia, Journal of Creative Education, Vol. 8 No.7, pp. 1044 – 1068.

Hora, S. C. and Winterfeldt, D. V. (1977). Nuclear waste and future societies: A look into the deep future. Journal of Technological Forecasting and Social Change, Volume 56, Issue 2, Pp. 155-170.

Hora, S. C., and M. Jensen (2002), Expert judgement elicitation, *Tech. Rep. SSI 2002:19*, Swedish Radiation Protection Authority, Stockholm, Sweden.

Hora, S.C. & Iman, R.L. (1989). Expert Opinion in Risk Analysis: The NUREG-1150 Experience, *Nuclear Science and Engineering*, 102, 323-331.

Hossain, N. U. I., Nagahi, M., Jaradat, R., Shah, C., Buchanan, R., and Hamilton, M. (2020). Modeling and assessing cyber resilience of smart grid using Bayesian network-based approach: a system of systems problem. Journal of Computational Design and Engineering, 7(3), 352–366.

Hossain, N. U. I., Nur, F., Hosseini, S., Jaradat, R., Marufuzzaman, M., Puryear, S. M. (2019). A Bayesian network-based approach for modeling and assessing resilience: A case study of a full-service deepwater port. *Reliability Engineering & System Safety*, 189, 378–396.

Howitt, J. A., Baldwin, D. S., Rees, G. N., and Williams, J. L. (2007). Modelling blackwater: Predicting water quality during flooding of lowland river forests. Ecological Modelling 203, 229-242.

Huang, C. and Darwiche, A. (1994). Inference in Belief Networks: A Procedural Guide. International Journal of Approximate Reasoning. New York: Elsevier Science Inc. 1994 11:1-158.

Hurd, M.D. and McGarry, K. (2002). The predictive validity of subjective probabilities of In Shanteau, J. (Ed.) Decision Science and Technology: Reflections on the contributions of Ward Edwards. Norwell, MA: Kluwer Academic Publishers. pp. 313-300. International Journal of Approximate Reasoning. New York: Elsevier Science Inc. 1994 11:1-158.

Jenkinson, D. (2005). The Elicitation of Probabilities – A Review of the Statistical Literature. [Online] Available: http://citeseerx.ist.psu.edu./viewdoc/download?doi=10.1.1.106.6173&rep=rep1&type=pdf. [Accessed on 26th September 2019].

Jensen, F. V. (1996). An introduction to Bayesian networks. UCL Press. London.

Jensen, F.V., (2001). Bayesian Networks and Decision Graphs. Springer-Verlag, New York, ISBN 0-387-95259-4.

Jeremy, E. O., Daneshkhah, A. and O'Hagan, A. (2014). Nonparametric Prior Elicitation using the Roulette Method. [online] Available from: www.tonyohagan.co.uk/academic/pdf/elic-roulette.pdf. [Accessed on: 19 October 2018].

Johansson, F. and Falkman, G. (2008). A Bayesian network approach to threat evaluation with application to an air defence scenario. In: *11th International Conference on Information fusion, Cologne, Germany*.

Johnson, S, R., Tomlinson, G. A., Hawker, G. A., Granton, J. T., Grosbein, H. A. and Feldman, B. M. (2010). A valid and reliable belief elicitation method for Bayesian priors. *Journal of Clinical Epidemiology*, **63**, 370 – 383.& *Journal of Statistical Planning and Inference*, **40**, 221–232.

Kadane, J.B. (1994). An application of robust Bayesian analysis to a medical experiment.

Kahneman, D. (2011). Thinking, fast and slow. London: Penquin Books Ltd.

Kapur, R. (2018). Educational Wastage: A Major Hindrance within Progression of Individuals, Communities and the Nation, [Online] Available: https://www.researchgate.net/publication/324029740_Educational_Wastage_A_Major_Hindrance_within_Progression_of_Individuals_Communities_and_the_Nation. [Accessed on 6th February 2023].

Kendall, M. G. (1962). Rank Correlations method. London: Griffin.

Knol, A. B., Slottje, P., Sluijs, J.,Labret, E. (2010). The use of expert elicitation in environmental health impact assessment: a seven step procedure.

Korb, K. B., Nicholson, A. E (2004). Bayesian Artificial Intelligence. London: Chapman and Hall/CRC Press. pp. 48-50. ISBN: 1-58488-387-1.

Kruglanski, A. W. (1989). Lay epistemology process and contents. Psychological Review, 87, 70 – 87.

Kruglanski, A. W. and Freund, T. (1983). The freezing and unfreezing of lay inferences: Effects on impressional primacy, ethnic stereotype, and numerical anchoring. Journal of Experimental Social Psychology, 19, 448-468.

Kruglanski, A. W., & Klar, Y. (1987). A view from a bridge: Synthesizing the consistency and attribution paradigms from a lay epistemic perspective. European Journal of Social Psychology, 17, 211-241.

Kuhnert, P. M., Martin, T. G. and Griffiths, S. P. (2010). A guide to eliciting and using expert knowledge in Bayesian ecological models. Blackwell Publishing Ltd/CNRS.

Kuikka S., Hilden M., Gislason H., Hansson S., Sparholt H. and Varis O. (1999) Modeling environmentally driven uncertainties in Baltic cod (Gadus morhua) management by Bayesian influence diagrams. Canadian Journal of Fisheries and Aquatic Sciences, 56, 629–641.

Kunda, Z. (1990). The case for motivated reasoning. Psychological Bulletin, 108 (3), pp. 480-498.

Kuvaas, B. Buch, R. and Dysvik, A. (2018). Individual Variable Pay for Performance, Incentive Effects, and Employee Motivation. Academy of Management Annual Meeting Proceedings 2018(1):12393

Lauritzen S, and Spiegelhalter D (1988). "Local Computation with Probabilities on Graphical Structures and their Application to Expert Systems (with discussion)". Journal of the Royal Statistical Society: Series B (Statistical Methodology), 50(2), 157-224.

Liu, A. X. and Khakpour, A. R. (2013). Quantifying and Verifying Reachability for Access Controlled Networks. IEEE/ACM Transactions on Networking. Vol. 21. No 2.

Lucas, P., Boot, H. and Taal, B. (1998). Computer-based decision support in the management of primary gastric non-Hodgkin lymphoma. *Methods of Information in Medicine*, 37 206-219.

Maheady, L., Michielli-Pendl, J., Harper, G. F. and Mallette, B. (2006). The effects of numbered heads together with and without an incentive package on the science test performance of a diverse group of sixth graders. Journal of Behavioural Education, Vol. 15, No. 1, pp. 25-39.

Mahmoud, L. & Zohair, A. (2019). Prediction of Student's performance by modelling small dataset size. [online] Available from: https://www.researchgate.net/publication/334957883_Prediction_of_Student's_performance_by_modelling_small_dataset_size [Accessed on: 20 January 2022].

Malbasic, S. B. and Duric, S. V. (2019). Risk assessment methodology - application of Bayesian probability networks in an ammunition delaboration project. Military Technical Bulletin, 67(3), 614-641.

Martin, T. G., Burgman, M. A., Fidler, F., Kuhnert, P. M., Low-Coy, S., Mcbride, M. and Mengersen, K. (2011). Eliciting Expert Knowledge in Conservation Science. Journal of Conservation Biology, 26(1), pp 29-38.

Martin, T. G., Burgman, M. A., Fidler, F., Kuhnert, P. M., Low-Coy, S., Mcbride, M. and Mengersen, K. (2011). Eliciting Expert Knowledge in Conservation Science. Journal of Conservation Biology, 26(1), pp 29-38.

Matthews, R. A., Buikema, A. L., Jr., Caims, J., Jr., and Rogers, J. H., Jr. (1982). Biological monitoring. Part IIA. Receiving System functional methods, relationships and indices. *Water Research* 16(2), pp. 129-139.

Meyer, M. A. and Booker, J. M. (2001). Eliciting and Analysing Expert Judgement: A Practical Guide. SIAM edition, Philadelphia: ISBN 0-89871-474-5.

Montibeller and Winterfeldt. (2015). Cognitive and motivational biases in decision and risk Analysis. Journal of Risk Analysis, 35(7) pp.1230-1252.

Mwiya, B., Bwalya, J., Siachinji, B., Sikombe, S., Chanda, H. and Chawala, M. (2017). Higher Education Quality and Student Satisfaction Nexus: Evidence from Zambia, Journal of Creative Education, Vol. 8 No.7, pp. 1044 – 1068.

Nicholson, A., Woodberry, O., and Twardy, C. (2010). The native "Fish" Bayesian networks. Bayesian Intelligence Technical Report 2010/3. pp. 1-30.

Nikovski, D. (2000) Constructing Bayesian networks for medical diagnosis from incomplete and partially correct statistics. *Knowledge and Data Engineering, IEEE Transactions on*, 12(4), pp.509-516.

NSUK (2019). The University Calendar (5th ed.). A Publication of the Office of the Vice-Chancellor Nasarawa State University, Keffi, Nigeria, p. 253.

O'Leary R. A. (2015). Characterising Uncertainty in Expert Assessments. PLOS ONE.

Oakley, J. E. (2010). Eliciting Univariate Probability Distributions. [online] Available from http://www.jeremy-oakley.staff.shef.ac.uk/Oakley_elicitation.pdf [Accessed on: 22nd December, 2018].

Ochuba, O. V. (2005) An analysis of wastage among fresh students in Nigerian Universities: A case study of the University of Benin. Unpublished PhD Thesis, University of Benin, Nigeria.

Odekunle, K. S. (2007). Educational wastage: Causes, costs, and strategies for education. Lagos Journal of Educational Administration and Planning, 1(1), pp. 60-65.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser , J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E. and Rakow, T. (2006). Uncertain Judgements: Eliciting Experts' Probabilities. Wiley: ISBN: 978-0-470-02999-2.

Olson, J. R. and Rueter, H. H. (1987). Extracting expertise from experts: Methods for knowledge acquisition. Journal of Expert Systems, 4(3), pp 152-168.

Olszewska-Szopa, M. and Wróbel, T. (2019). Gastrointestinal non-Hodgkin lymphomas. J. of Adv Clin Exp Med. 28(8):1119-1124.

Oragwu, A. A. (2020). Minimizing Students' Dropout Rates in Universities for Sustainable Development. International Journal of Educational Research and Policy Making (IJERPM), 3(1), pp. 511-521.

Pearl, J. (1998). Probabilistic reasoning in intelligent systems: Networks of plausible inference. Morgan Kaufmann Publishers. San Mateo California.

Pharmaceutical Research, pp. 87-109.

Pollino C.A., Woodberry O., Nicholson A., Korb K. & Hart, B.T. (2007) Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. Environmental Modelling and Software 22, 1140–1152.

Pollino, C. A. (2008). Application of Bayesian Networks in Natural Resource Management (SRES3035). 11-22 February 2008. Canberra, Australian National University.

Ritchie, J, Lewis, J and Elam, G (2003). Designing and selecting samples. In Jane Ritchie and Jane Lewis *(Eds.), Qualitative research practice*. A guide for social science students and researchers (pp.77-108) Thousand Oaks, CA: Sage.

Ritchie, J, Lewis, J and Elam, G (2003). Designing and selecting samples. In Jane Ritchie and Jane Lewis *(Eds.), Qualitative research practice*. A guide for social science students and researchers (pp.77-108) Thousand Oaks, CA: Sage.

Riyadi, T. (2020). Classification of Student Academic Performance using Fuzzy Soft Set. 2020 International Conference on Smart Technology and Applications, pp 1-6. DOI: 10.1109/ICoSTA48221.2020.1570606632.

Rojas, J. A, Espitias, H. E. & Bejanaro, L. A. (2021). Design and Optimization of a Fuzzy Logic System for Academic Performance Prediction. *Symmetry*. 2021; 13(1):133. https://doi.org/10.3390/sym1301013.

Russell, S. J. and Norvig, P. (2010). Artificial Intelligence: A Modern Approach, (3rd ed.). Upper Saddle River, NJ: Prentice Hall. ISBN -13:978-0-13-207148-2.

Saa, A. A. (2016). Educational Data Mining and Students' Performance Prediction. International Journal of Advanced Computer Science and Applications, 7(5), pp. 212-220.

Sadeghi, R., Zarkami, R., Sabetraftar, K., and Damme, P.V. (2012). Application of classification trees to model the distribution pattern of a new exotic species Azollafiliculoides (Lam.) at Selkeh Wildfire Refuge, Anzali wetland, Iran. Ecol. Modell.243, 8–17.

Sani, N. S., Fikri, A., Ali, Z., Zakree, M. and K. Nadiyah, K. ``Dropout prediction in higher education among B40 students," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 11, pp. 550_559, 2020, doi: 10.14569/IJACSA.2020.0111169.

Saxena, U. R., & Singh, S. P. (2012). Integrating Neuro-fuzzy Systems to Develop Intelligent Planning systems for predicting Students' Performance. Int J Eval Res Educ., 1(2), 61-66.

Shephard, G. G. and Kirkwood, C. W. (1994). Managing the Judgemental Probability Elicitation Process: A case study of Analyst/Manager Interaction, J. of IEEE Transactions on Engineering Management, 41(414-425).

Shirley, R. B. and Smidts, C. Bridging the simulator gap: Measuring motivational bias in digital nuclear power plant environments. Journal of Reliability Engineering and System Safety 177(2018) pp.191-209.

Shirley, R. B. and Smidts, C. Bridging the simulator gap: Measuring motivational bias in digital nuclear power plant environments. Journal of Reliability Engineering and System Safety 177(2018) pp.191-209.

Silas, G. A. (2003). An investigation of black students’ attrition at a large predominating white, midwestern university. Western Journal of Black Studies. 17 (14), 179-182.

Simpson, O. (2005). The costs and benefits of student retention for students, institutions and governments. Studies in Learning, Evaluation Innovation and Development, 2(3), pp. 34–43.
(PDF) The costs and benefits of student retention for students, institutions and governments. Available from: https://www.researchgate.net/publication/42792577_The_costs_and_benefits_of_student_retention_for_students_institutions_and_governments [accessed April 02 2022].

Singh, W. and Kaur, P. (2016). Comparative Analysis of Classification Techniques for predicting Computer engineering Students' Academic Performance. International Journal of Advanced Research in Computer Science. Vol. 7 Issue 6, p31-36.

Soll, J.B., Milkman K. L. .and Payne, J. W. (2013). A User's guide to Debiasing. Wiley-Blackwell Handbook of Judgement and decision Making. ISBN 978-1-118-46839-5, USA.

Soll, J.B., Milkman K. L. .and Payne, J. W. (2013). A User's guide to Debiasing. Wiley-Blackwell Handbook of Judgement and decision Making. ISBN 978-1-118-46839-5, USA.

Stat Trek. (2019). Probability Distributions. [online] Available from: http://stattrek.com/probability-distributions/normal.aspx. [Accessed on: 22 September 2019].

Stewart-Koster B., Bunn S.E., Mackay S.J., Poff, N. L., Naiman, R.J. & Lake P.S. (2005) The use of Bayesian networks to guide investments in flow and catchment restoration for impaired river ecosystems. Freshwater Biology, 55, 243–260.

Teixeira, R. and Rexford, J. (2004). A measurement framework for pin-pointing routing changes. In ACM SIGC0MM Workshop on Network Troubleshooting.

The UK Department for Digital, Culture, Media and Sport. (2018). The UK Data Protection Act 2018. [Online] Available: https://www.gov.uk/government/collections/data-protection-act-2018. [Accessed on: 23rd May 2021].

Ticehurst, J., Newham, L., Rissik, D., Letcher, R., & Jakeman, A. (2007) Bayesian network approach for assessing the sustainability of Coastal Lakes in New South Wales, Australia. Environmental Modelling and Software, pp.1129-1139.

Timothy, R. and Manley, K. (2011). Motivation toward financial incentive goals on construction projects. Journal of Business Research 64 (2011) pp. 765–773.

Timothy, R. and Manley, K. (2011). Motivation toward financial incentive goals on construction projects. Journal of Business Research 64 (2011) pp. 765–773.

Tinto, V. (2006). Research and practice of student retention: What next? *Journal of College Student Retention: Research, Theory and Practice*, 8(1), pp.1-19.

Trigg, D. J. (2004). The Development of an Expert System using Plausible Reasoning for the Diagnosis and Prognosis of River Quality. PhD. Thesis, Staffordshire University.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psych*. Vol 4, pp. 174-177.

Tversky, A. and Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, Vol. 185. pp. 1124-113. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/pst.2040.

URL https://onlinelibrary.wiley.com/doi/abs/10.1002/pst.2040.

Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. Ecological Modelling, 203, 312-318.

van der Gaag, L. C., Renooij, S., Witteman, C.L.M., Aleman, B.M.P and Taal, B.G. (1999). How to elicit many probabilities. [online] Available from: https://www.researchgate.net/publication/221404281_ How_to_ Elicit_Many_ Probabilities. [Accessed on: 5th October 2019].

van der Gaag, L.C., Renooij, S., Witteman, C.L.M., Aleman, B.M.P. and Taal, B.G. (2002). Probabilities for a probabilistic network: A case study in oesophagealcancer. *Artificial Intelligence in Medicine*, 25, 123–148.

Velikova, M., van Scheltinga, J., Lucas, P., & Spaanderman, M. (2014) Exploiting causal functional relationships in Bayesian network modelling for personalised healthcare. *International Journal of Approximate Reasoning*, 55(1), pp.59-73.

Willmot, L. and Lloyd, P. (2015). Improving the support and retention of students. [Online] Available: https://www.researchgate.net/publication/242563337 [Accessed on: 8th February, 2023].

Wooldridge, S. (2003). Bayesian Belief Networks. Centre for Complex System Science, CSIRO, Canberra.

Xie, G. G., Zhan, J., Maltz, D. A., Zhang, H., Greenberg, A., Hjalmtysson, G. and Rexford, J. (2005). Static Reachability Analysis of Internet Protocol Networks: in IEEE INFOCOMM 2005 Proceedings.

Yang, Y. (1994). Expert network: Effective and efficient learning from human decisions in text categorization and retrieval. In: *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, Springer-Verlag New York, Inc., pp. 13-22.

Yorke, M (1999), *Leaving Early: Undergraduate Non-completion in Higher Education*, London: Falmer.

Zapata-Vazquez, R. E., O'Hagan, A., and Bastos, L. S. (2012). Eliciting expert judgements about as set of proportions. *Journal of Applied Statistics* 13 (7) 1-15.

Zou, M. & Conzen, S. (2005) A New Dynamic Bayesian Network (DBN) Approach for Identifying Gene Regulatory Networks from Time Course Microarray Data. Bioinformatics, 21(1), pp.71-79.

APPENDICES

Appendix 1: Questionnaire 1 (Impact of motivation on probabilistic estimates in an elicitation session)

Dear Respondent,

I am a researcher who is investigating the impact of motivation on probabilistic estimates in an elicitation session. The information gathered from this study will be used strictly for academic purposes and be kept confidential.

Questions

Instructions:

(i) Could you provide your estimate (in percentage) for each of the following variables?

(ii) You can skip questions and return back to them later.

S/N	Question	(%)
1	If 88% of all adults go online while 87% of male adults also do so. Can you estimate the percentage of female adults that are likely to go online?	
2	98% of adults in the 16 to 24 years age group go online while 82% of adults in the 55 to 64 years age group also do so. Can you estimate the percentage of adults aged 75 and over that are likely to go online?	
3	73% of adults in the 25 to 34 years age belong to/use Facebook while 70% of those in the age group 55 to 64 years also belong to/use Facebook. Can you estimate the percentage of adults in the 45 to 54 years age group that are likely to belong to/use Facebook?	
4	67% of all adults go online using computers while 65% of female adults also do so. Can you estimate the percentage of male adults that are likely to go online using computers?	
5	88% of adults in the Skilled Working class (grade C2) socio-economic group use mobile phones to go online while 93% of adults in the Upper/Middle class (grade A or B) socio-economic group also use the same device to go online. Can you estimate the percentage of adults in the Lower Middle class (grade C1) socio-economic group that go online using mobile phones?	

6	34% of adults aged 75 years and over go online to look for public services information on government sites such as gov.uk, ni.direct or HMRC while 42% of adults in the 16 to 24 age group also do so. Can you estimate the percentage of adults in the 35 to 44 years age group that go online for the above purposes?	
7	33% of adults in the 16 to 24 years age group only use devices other than a computer to go online while 24% of adults in the 45 to 54 years age group also do so. Can you estimate the percentage of adults in the 25 to 34 years age group that only use devices other than a computer to go online?	
8	75% of adults in the 55 to 64 years age group go online to send or receive emails while 54% of adults aged 75 and over go online for the same purpose. Can you estimate the percentage of adults in the 65 to 74 years age group that go online to send or receive emails?	
9	20% of all adults in the Unskilled/Non-working class (grade D or E) socio-economic group go online using non-smartphones while 17% in the Lower Middle class (grade C1) socio-economic group also use the same type of phone. Can you estimate the percentage of adults in the Upper/Middle class (grade A or B) socio-economic group that go online using non-smartphones?	
10	2% of adults in the Unskilled/Non-working class (grade D or E) socio-economic group go online using a Smart TV set for gaming while 3% of the Lower Middle Class (grade C1) socio-economic group also do so. Can you estimate the percentage of adults in the Upper/Middle class (grade A or B) socio-economic group that are likely to go online using a Smart TV set for gaming?	

Many thanks for giving your invaluable time to complete this questionnaire.
Yours sincerely,

Gbolagade Kola Adegoke
Doctoral Student.

Appendix 2: Questionnaire 2 (Impact of motivation on probabilistic estimates in an elicitation session)

Dear Respondent,

I am a researcher who is investigating the impact of motivation on probabilistic estimates in an elicitation session. The information gathered from this study will be used strictly for academic purposes and be kept confidential.

Questions

Instructions:

(i) Could you provide your estimate (in percentage) for each of the following variables?

(ii) You can skip questions and return back to them later.

S/N	Question	(%)
1	89% of all adults use mobile phones to go online while 90% of female adults also do so. Can you estimate the percentage of male adults that are likely to go online using mobile phones?	
2	15% of adults in the 45 to 54 years age group belong to/use WhatsApp while 23% of adults in the 35 to 44 years age group also belong to/use WhatsApp. Can you estimate the percentage of adults aged 65 and over that are likely to belong to/use WhatsApp?	
3	58% of all adults go online using tablets while 60% of male adults also do so. Can you estimate the percentage of female adults that are likely to go online using tablets?	
4	54% of adults in the 35 to 44 years age group use a Smart TV set to go online. Can you estimate the percentage of adults in the 25 to 34 years age group that are likely to go online using the same device?	
5	91% of adults in the Lower Middle class (grade C1) socio-economic group go online using a mobile phone while 84% of the Unskilled/Non-working class (grade D or E) socio-economic group also do so. Can you estimate the percentage of adults in the Skilled Working class (grade C2) socio-economic group that are likely to go online using the same device?	
6	83% of adults in the Upper/Middle class (grade A or B) socio-economic group go online using computers while 62% of Skilled Working Class (grade C2)	

	socio-economic group also do so. Can you estimate the percentage of adults in the Lower Middle class (grade C1) socio-economic group that are likely to go online using computers?	
7	62% of all adults in the 16 to 64 age group use tablets to go online while 68% of the adults in the same age group that go online using tablets belong to the Upper/Middle, Lower Middle or Skilled Working class (grade A, B, C1 or C2) socio-economic group. Can you estimate the percentage of adults in the Unskilled/Non-working class (grade D or E) socio-economic group that go online using tablets?	
8	76% of adults in the 16 to 24 years age group use smartphones to complete a form or application while 75% of adults in the 35 to 44 years age group also use smartphones to perform the same activity. Can you estimate the percentage of adults aged 65 and over that use smartphones to complete a form or application?	
9	18% of adults in the 65 to 74 years age group communicate via instant messaging e.g. Facebook chat, Skype chat and Snapchat while 41% of adults in the 45 to 54 years age group also communicate via the above instant messaging. Can you estimate the percentage of adults in the 55 to 64 years age group that communicate via the above instant messaging?	
10	15% of Skilled Working class (grade C2) socio-economic group make voice calls e.g. via FaceTime and Skype while 17% of adults in the Unskilled/Non-working class (grade D or E) socio-economic group also do so. Can you estimate the percentage of adults in the Upper/Middle class (grade A or B) socio-economic group that make voice calls via the above means?	

Many thanks for giving your invaluable time to complete this questionnaire.
Yours sincerely,

Gbolagade Kola Adegoke
Doctoral Student.

Appendix 3: Information Form (Impact of motivation on probabilistic estimates in an elicitation session)

Project Title
Impact of motivation on probabilistic estimates in an elicitation session.

Invitation
You are being invited to take part in a research study. Before you decide to participate in this study, it is important that you understand why the research is being done and what it will involve. Please take the time to read the following information carefully. Please ask the researcher if there is anything that is not clear or if you need more information.

Project Summary
The Preliminary research study aims to investigate the impact of motivation on probabilistic estimates in an elicitation session. The goal of an elicitation session is to obtain accurate probabilistic estimates from the participants. In this study, we are going to use series of probability questions.

The study will involve eliciting probabilistic estimates on the Internet usage in the UK, using the questionnaires designed for this purpose.

The participants involved in the elicitation session will be students of the Staffordshire University, UK who will provide probabilistic estimates for the variables in Internet usage in the UK. It is envisaged that session will involve completion of questionnaires designed for this study. Sessions will involve a group of students.

Evaluation of the elicitation session will be based on comparing the estimates provided by the participants against the actual values in the survey data, in order to know how near the estimated values are, to the actual values.

What rights do I have?

- You may decide to stop being a part of the research study at any time or stage, without giving a reason.
- You have the right to ask that any data you have supplied to that point be withdrawn/destroyed.
- You have the right to omit or refuse to answer or respond to any question that you are asked.
- You have the right to have your questions about the procedures answered.

Thank you so much for your participation.

Yours sincerely,

Gbolagade Kola Adegoke.

Doctoral Student.

Appendix 4: Consent Form (Impact of motivation on probabilistic estimates in an elicitation session)

Impact of motivation on probabilistic estimates in an elicitation session.

Please check box

1. I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time or stage, without giving a reason. ☐

3. I agree to take part in the above study. ☐

4. I would only like to participate by completing Questionnaires for the study. ☐

5. I agree to the use of anonymised quotes in publications. ☐

Signature of Participant: ……………………………………………..Date ……………….

Signature of Researcher: ……………………………………………….Date ……………….

Thank you so much for your participation.

Yours sincerely,

Gbolagade Kola Adegoke

Doctoral Student.

Appendix 5: Ethical Approval (Impact of motivation on probabilistic estimates in an elicitation session)

Computing and Digital Technologies

PROPORTIONATE REVIEW APPROVAL FEEDBACK

Researcher Name:	Gbolagade Kola Adegoke
Title of Study:	Impact of motivation on probabilistic estimates
Status of approval:	Approved

Overall a sound submission that reads as well thought through. Three corrections are required with the questionnaire:

Question 4 – reword 'use Smart TV set' to 'use a Smart TV'. Same with Question 9 as well.

Question 16 'government site' change to 'government sites'

Make sure you provide the information sheet and consent forms to participants, after they are reviewed and approved by your supervisors

Action now needed:

Your project proposal has been approved by the Ethics Panel and you may commence the implementation phase of your study. You should do so in conjunction with your supervisor.

You should note that any divergence from the approved procedures and research method will invalidate any insurance and liability cover from the University. You should, therefore, notify the Panel of any significant divergence from this approved proposal.

When your study is complete, please send the ethics committee an end of study report. A template can be found on the ethics BlackBoard site.

Signed: **Date: 22.02.19**

Prof. Elhadj Benkhelifa

Chair of the Computing and Digital Technologies Ethics Panel

Appendix 6: Questionnaire (Elicitation of Probability Estimates for Student Retention)

Instructions:

(i) Could you provide your estimate (in percentage) for each of the following rows?

(ii) You can skip questions and return back to them later.

SN	Change in Circumstances	Mode of Entry	Academic Skills	Academic Engagement	Academic Progress	Attendance	Probability of being retained (%)
1	True	UTME	Good	Engaged	Normal	At least 75%	
2	True	UTME	Good	Engaged	Normal	Less than 75%	
3	True	UTME	Good	Engaged	Slow	At least 75%	
4	True	UTME	Good	Engaged	Slow	Less than 75%	
5	True	UTME	Good	Not engaged	Normal	At least 75%	
6	True	UTME	Good	Not engaged	Normal	Less than 75%	
7	True	UTME	Good	Not engaged	Slow	At least 75%	
8	True	UTME	Good	Not engaged	Slow	Less than 75%	
9	True	UTME	Poor	Engaged	Normal	At least 75%	
10	True	UTME	Poor	Engaged	Normal	Less than 75%	
11	True	UTME	Poor	Engaged	Slow	At least 75%	
12	True	UTME	Poor	Engaged	Slow	Less than 75%	
13	True	UTME	Poor	Not engaged	Normal	At least 75%	
14	True	UTME	Poor	Not engaged	Normal	Less than 75%	
15	True	UTME	Poor	Not engaged	Slow	At least 75%	
16	True	UTME	Poor	Not engaged	Slow	Less than 75%	
17	True	Direct entry	Good	Engaged	Normal	At least 75%	
18	True	Direct entry	Good	Engaged	Normal	Less than 75%	
19	True	Direct entry	Good	Engaged	Slow	At least 75%	
20	True	Direct entry	Good	Engaged	Slow	Less than 75%	
21	True	Direct entry	Good	Not engaged	Normal	At least 75%	
22	True	Direct entry	Good	Not engaged	Normal	Less than 75%	
23	True	Direct entry	Good	Not engaged	Slow	At least 75%	
24	True	Direct entry	Good	Not engaged	Slow	Less than 75%	

25	True	Direct entry	Poor	Engaged	Normal	At least 75%	
26	True	Direct entry	Poor	Engaged	Normal	Less than 75%	
27	True	Direct entry	Poor	Engaged	Slow	At least 75%	
28	True	Direct entry	Poor	Engaged	Slow	Less than 75%	
29	True	Direct entry	Poor	Not engaged	Normal	At least 75%	
30	True	Direct entry	Poor	Not engaged	Normal	Less than 75%	
31	True	Direct entry	Poor	Not engaged	Slow	At least 75%	
32	True	Direct entry	Poor	Not engaged	Slow	Less than 75%	
33	False	UTME	Good	Engaged	Normal	At least 75%	
34	False	UTME	Good	Engaged	Normal	Less than 75%	
35	False	UTME	Good	Engaged	Slow	At least 75%	
36	False	UTME	Good	Engaged	Slow	Less than 75%	
37	False	UTME	Good	Not engaged	Normal	At least 75%	
38	False	UTME	Good	Not engaged	Normal	Less than 75%	
39	False	UTME	Good	Not engaged	Slow	At least 75%	
40	False	UTME	Good	Not engaged	Slow	Less than 75%	
41	False	UTME	Poor	Engaged	Normal	At least 75%	
42	False	UTME	Poor	Engaged	Normal	Less than 75%	
43	False	UTME	Poor	Engaged	Slow	At least 75%	
44	False	UTME	Poor	Engaged	Slow	Less than 75%	
45	False	UTME	Poor	Not engaged	Normal	At least 75%	
46	False	UTME	Poor	Not engaged	Normal	Less than 75%	
47	False	UTME	Poor	Not engaged	Slow	At least 75%	
48	False	UTME	Poor	Not engaged	Slow	Less than 75%	
49	False	Direct entry	Good	Engaged	Normal	At least 75%	
50	False	Direct entry	Good	Engaged	Normal	Less than 75%	
51	False	Direct entry	Good	Engaged	Slow	At least 75%	
52	False	Direct entry	Good	Engaged	Slow	Less than 75%	
53	False	Direct entry	Good	Not engaged	Normal	At least 75%	
54	False	Direct entry	Good	Not engaged	Normal	Less than 75%	
55	False	Direct entry	Good	Not engaged	Slow	At least 75%	

56	False	Direct entry	Good	Not engaged	Slow	Less than 75%	
57	False	Direct entry	Poor	Engaged	Normal	At least 75%	
58	False	Direct entry	Poor	Engaged	Normal	Less than 75%	
59	False	Direct entry	Poor	Engaged	Slow	At least 75%	
60	False	Direct entry	Poor	Engaged	Slow	Less than 75%	
61	False	Direct entry	Poor	Not engaged	Normal	At least 75%	
62	False	Direct entry	Poor	Not engaged	Normal	Less than 75%	
63	False	Direct entry	Poor	Not engaged	Slow	At least 75%	
64	False	Direct entry	Poor	Not engaged	Slow	Less than 75%	

Many thanks for giving your invaluable time to complete this questionnaire.

Yours sincerely,

Gbolagade Kola Adegoke

Doctoral Student.

Appendix 7: Information Sheet for Domain Experts (Elicitation of Probability Estimates for Student Retention)

Title of Project: Bayesian Network for Predicting Student Retention Based on Expert Elicitation

Researcher: Gbolagade Kola Adegoke

You are being invited to take part in this research project titled "Bayesian Network for Predicting Student Retention Based on Expert Elicitation" which forms part of my PhD. Before you decide to participate in this study, it is important that you understand why the research is being done and what your participation will involve. Please take the time to read the following information carefully. Please ask me if there is anything that is not clear or if you would need more information.

1. What is the purpose of the study?

I would like to elicit a causal structure (a diagram that depicts topology of the nodes (variables) in the domain and their respective states/categories) in Student Retention domain as well as the estimates of probabilities (probability distributions) associated with the variables and states (categories) from you being an expert in this domain. The nodes (variables) and estimates of probabilities are needed for construction of a Bayesian network (BN), which is a combination of a diagram that depicts the topology of the nodes (variables) in the domain and their respective states/categories. The BN would be capable of predicting the percentage chance of a student remaining in the university the following academic session without dropping out.

2. Why have I been invited to take part?

You have been invited to take part in this study because you are a domain expert in student retention domain, and we want to acquire knowledge from you about the domain.

3. What will happen if I take part?

This study will adopt the mixed methods approach. Elicitation of causal structure is qualitative in nature while elicitation of estimates of probabilities is quantitative in nature. You will be asked to topologically list the variables and states (categories) in student retention as well as to provide estimates of probabilities for the relationships between the variables. I would like to elicit the variables from you by conducting an interview in an elicitation session. The examples of questions you will be asked are; to topologically list all of the variables and states (categories) in the domain, the relationships between the variables, and the number of states (categories) that each variable consists of.

The interview is expected to take the face-to-face format. The interview will last no more than one hour. Thereafter, you will be given a probability training that will guide you how to provide estimates of probabilities as well as how to avoid heuristics and biases while providing estimates of probabilities, the probability training will last no more than one hour as well.

Thereafter, you will be given a questionnaire into which you need to fill in your estimates of probabilities in form of percentages. You are expected to complete the questionnaire in the comfort of your home or office and make sure you send the completed questionnaire back to myself three (3) weeks after you have received it.

You will be asked to provide the estimates of probabilities between those variables in form of percentages hence the nature of data you are expected to provide lie between 0% and 100%, e. g., 20%, 15%, 60%, etc. The percentages will serve as input data into the Bayesian network software that will predict the chance of a student remaining in the university the following academic session without dropping out. You are at the very least expected to complete 150 rows of the questionnaire.

4. Do I have to take part?

No. It is up to you to decide whether or not you want to participate in this study. Once you have read the information sheet, please contact us if you have any questions that will help you make a decision about taking part. If you decide to take part, we will ask you to sign a consent form and you will be given a copy of this information sheet to keep. You are free to change your mind and withdraw without giving any reasons.

5. What are the possible risks of taking part?

This study will take the questions and answers format as well as completion of questionnaire. It will not involve selection of participants' records / tissue or bodily samples. Also, it will not involve experimental group or control group. For the purposes of ethical approval, we have completed the Research Ethics Proportionate Review form since the research raised only minimal ethical risk and it directly engaged human participants (that is, the student retention domain experts).

6. What are the possible benefits of taking part?

The possible benefits of taking part in this study are:

This project will benefit the university administrators as it will support their decisions in predicting the chance of a student remaining in university the following academic session without having to drop out. It ensures support systems are in place to enable students to remain at university and succeed. Student retention improves graduate rates, decreases loss of tuition revenue from students that either drop out or transfer to another institution, and brings reputation to an institution as well.

Futhermore, this project will benefit students, their families as well as their communities as it will improve the lives of the students, the lives of their families as well as affording them the opportunity to make a positive contribution to their local community. It will also allow the students to achieve their full potential if they did not dropout of school.

7. Data handling and confidentiality

Your data will be processed in accordance with the data protection law and will comply with the General Data Protection Regulation 2016 (GDPR).

8. **Data Protection Statement**

The data controller for this project will be Staffordshire University. The University will process your personal data for the purpose of the research outlined above. The legal basis for processing your personal data for research purposes under the data protection law is a 'task in the public interest' You can provide your consent for the use of your personal data in this study by completing the consent form that has been provided to you.

9. **What if I change my mind about taking part?**

You are free to withdraw at any point of the study, without having to give a reason. Withdrawing from the study will not affect you in any way. You are able to withdraw your data from the study up to two weeks after your participation in the elicitations, after which withdrawal of your data will no longer be possible because the data would have been processed.

If you choose to withdraw from the study, we will not retain any information that you have provided us as a part of this study.

10. **How is the project being funded?**

This research is unfunded.

11. **What will happen to the results of the study?**

The results of the study will be available after it finishes and will usually be published in a scientific journal and / or be presented at a scientific conference. The data will be anonymous and none of the participants involved in the study will be identifiable in any report or publication.

12. **Time commitment for this study**

The time commitment for this study is about 4.5 hours of your time.

13. **Checking of the appropriate box for each of the items on the consent form**

Kindly check the $\boxed{\surd}$ appropriate box for each of the items listed on the consent form.

14. **Statement of pseudonym on the consent form**

Please fill in your desired pseudonym in the space provided on the consent form.

15. **Who should I contact for further information?**

If you have any questions or require more information about this study, please contact the researcher using the following contact details:

Name: Gbolagade Kola Adegoke
Doctoral Student

Tel. number: +2348076524815
Email: gbolagade.adegoke@research.staffs.ac.uk

16. What if I have further questions, or if something goes wrong?

If this study has harmed, you in any way or if you wish to make a complaint about the conduct of the study you can contact the study supervisor or the Chair of the Staffordshire University Ethics Committee for further advice and information:

Contact information:

Chair, University Ethics Committee
Dr Tim Horne
Cadman Building
Staffordshire University
College Road
Stoke-on-Trent ST4 2DE
Tel: +44 (0) 1782295722
Email: tim.horne@staffs.ac.uk

Thank you for taking the time to read this information sheet and for considering taking part in this research.

Appendix 8: Research Project Consent Form (Elicitation of Probability Estimates for Student Retention)

Title of Project: Bayesian Network for Predicting Student Retention Based on Expert Elicitation

Researcher: Gbolagade Kola Adegoke

Kindly check [√] the appropriate box for each of the items listed below and fill in your desired pseudonym in the space provided below as well.

I have read and understood the information sheet.	Yes ☐	No ☐
I have been given the opportunity to ask questions, and I have had any questions answered satisfactorily.	Yes ☐	No ☐
I understand that my participation in this study is entirely voluntary and that I can withdraw at any time without having to give an explanation without this in any way affecting my treatment now or in the future.	Yes ☐	No ☐
I would like to participate in this study by attending an interview (elicitation session) followed by completion of Questionnaire.	Yes ☐	No ☐
I consent that data collected could be used for publication in scientific journals or could be presented in scientific forums (conferences, seminars, workshops) and understand that all data will be presented anonymously.	Yes ☐	No ☐
I agree that data will only be used for this project titled "Bayesian Network for Predicting Student Retention Based on Expert Elicitation", although the data may also be audited for quality control purposes.	Yes ☐	No ☐
All data will be stored safely on a password protected computer (electronic data), or locked away securely (hard copies of data) for 10 years before being destroyed	Yes ☐	No ☐
I understand that I can withdraw my consent from this study at any time, up to two weeks after I have participated in both elicitations without having to give an explanation, and this will also mean that my data is subject to removal from the project.	Yes ☐	No ☐

My desired pseudonym is: ---

I hereby give consent to take part in this study Yes ☐ No ☐

________________________	________________	____________________
Name of Participant (print)	Date	Signature
Gbolagade Kola Adegoke Doctoral Student	……………………… Date	………………………. Signature

Appendix 9: Ethical Approval (Elicitation of Probability Estimates for Student Retention)

STAFFORDSHIRE UNIVERSITY

School of Digital, Technologies and Arts

ETHICAL APPROVAL FEEDBACK

Researcher name:	Gbolagade Kola Adegoke
Title of Study:	SU_21_148 Bayesian Network for Predicting Student Retention Based on Expert Elicitation
Award Pathway:	PhD
Status of approval:	Approved

Your project ***proposal has been approved*** by the Ethics Panel and you may commence the implementation phase of your study. You should note that any divergence from the approved procedures and research method will invalidate any insurance and liabilitycover from the University. You should, therefore, notify the Panel of any significant divergence from this approved proposal.

You should arrange to meet with your supervisor for support during the process of completing your study and writing your dissertation.

When your study is complete, please send the ethics committee an end of study report. A template can be found on the ethics BlackBoard site.

The Ethics Committee wish you well with your research.

Signed: **Date: 3rd May 2022**

Prof. Elhadj Benkhelifa

Chair of the Digital Technologies Ethics Panel

Appendix 10

Frequency Tables and their respective Bar Charts

Table 5-1: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or more.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retained	80	80.0	80.0	80.0
	Unretained	20	20.0	20.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above estimates as elicited from the domain experts in this study is displayed in Table 5-1 above. Results in Table 5-1 show that this student has an 80% (0.8) chance of being retained and consequently complete his/her study therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0) therefore a student in this category has a 20% (0.2) chance of being dropped out of the university. The Figure 5-1 below is a visual representation of the Table 5-1 above.

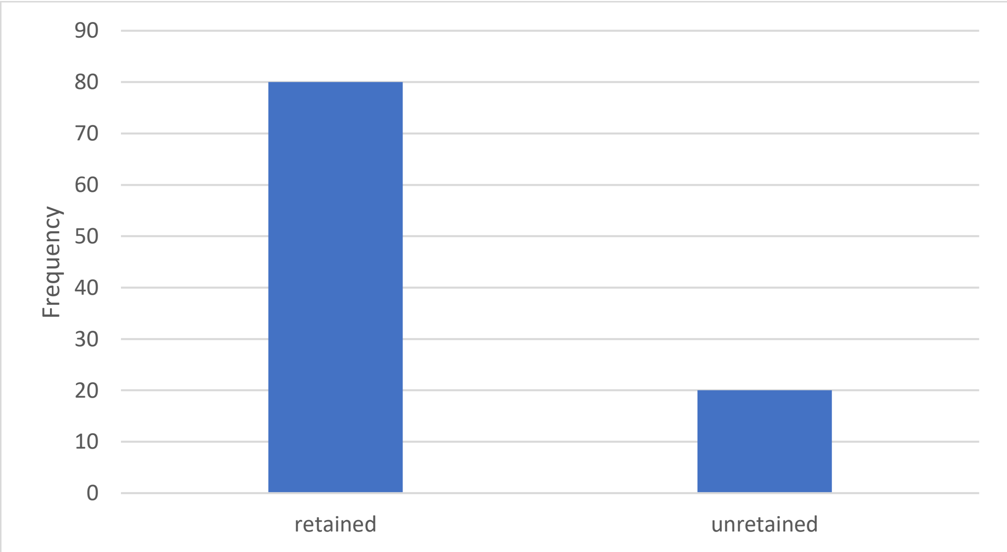


Figure 5-1: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or more. His or her chance of being retained is 80%. The Figure 5-1 is a visual representation of the Table 5-1 above.

Table 5-2: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retained	65	65.0	65.0	65.0
	Unretained	35	35.0	35.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-2 above. Results in Table 5-2 show that this student has a 65% (0.65) chance of being retained and consequently complete his/her study therefore a student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0) therefore a student in this category has a 35% (0.35) chance of being dropped out of the university. The Figure 5-2 below is a visual representation of the Table 5-2 above.

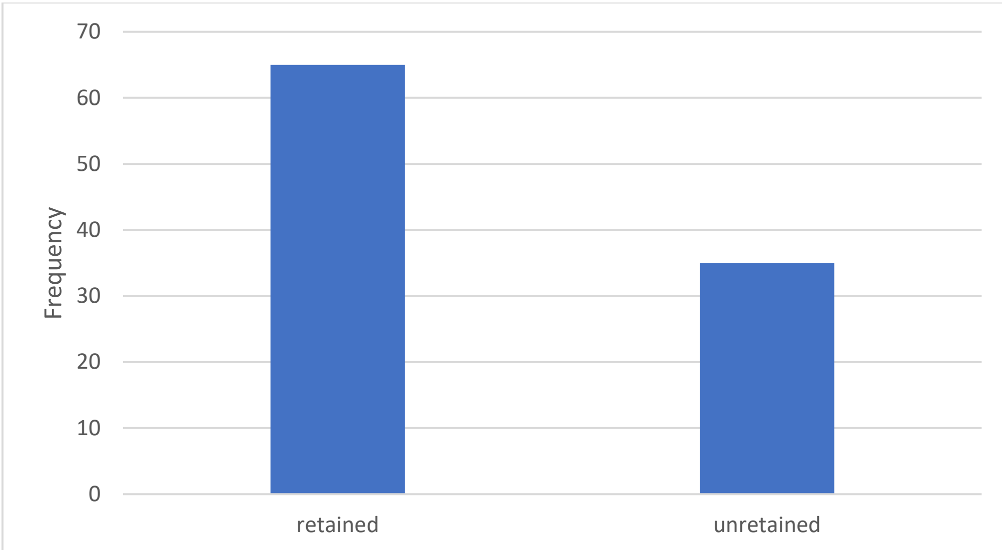


Figure 5-2: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or more. His or her chance of being retained is 65%. The Figure 5-2 is a visual representation of the Table 5-2 above.

Table 5-3: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retained	55	55.0	55.0	55.0
	Unretained	45	45.0	45.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-3 above. Results in Table 5-3 show that this student has a 55% (0.55) chance of being retained and consequently complete his/her study therefore a student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0) therefore a student in this category has a 45% (0.45) chance of being dropped out of the university. The Figure 5-3 below is a visual representation of the Table 5-3 above.

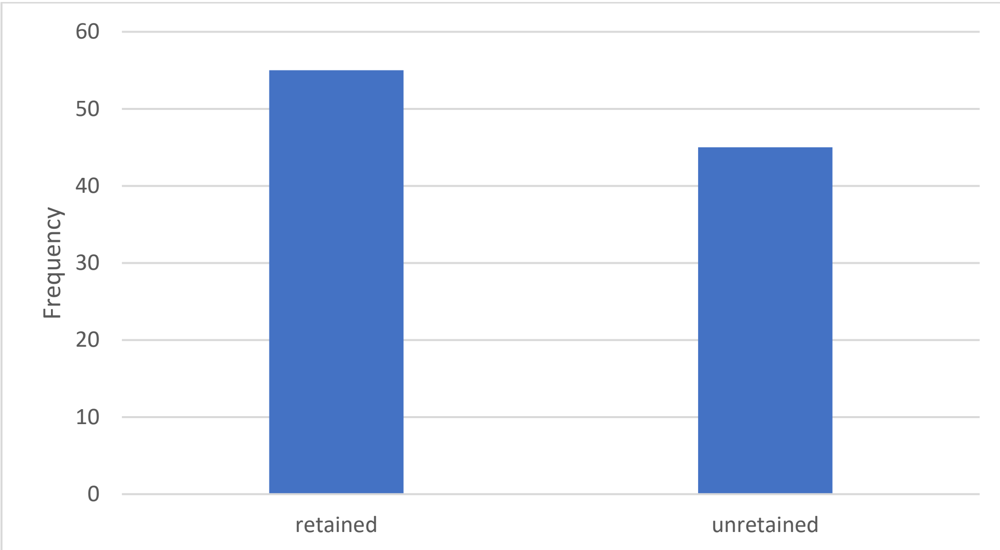


Figure 5-3: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 55%. The Figure 5-3 is a visual representation of the Table 5-3 above.

Table 5-4: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retained	32	32.0	32.0	32.0
	Unretained	68	68.0	68.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-4 above. Results in Table 5-4 show that this student has a 32% (0.32) chance of being retained therefore a student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0) therefore a student in this category has a 68% (0.68) chance of being dropped out of the university. The Figure 5-4 below is a visual representation of the Table 5-4 above.

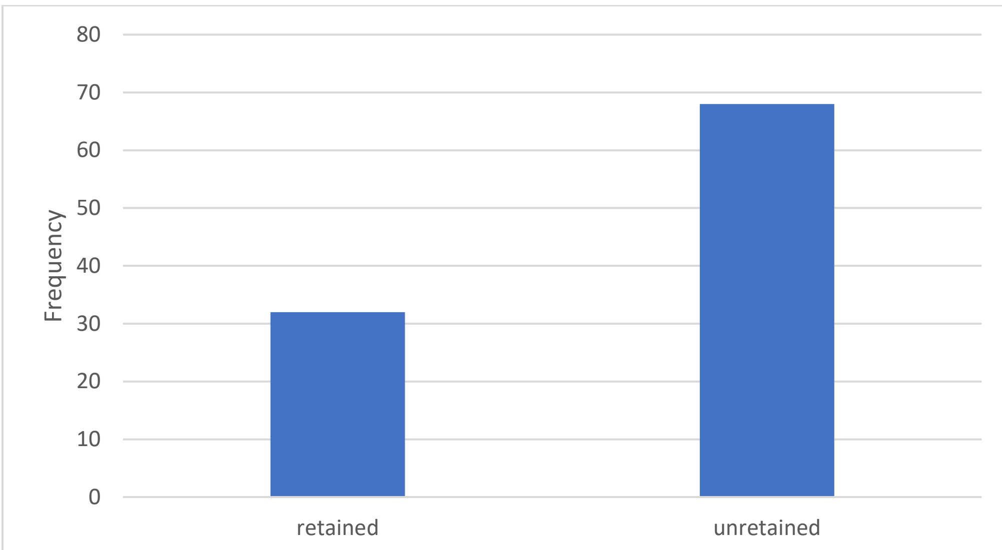


Figure 5-4: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is less than 75%. His or her chance of being retained is 32%. The Figure 5-4 is a visual representation of the Table 5-4 above.

Table 5-5: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retained	15	15.0	15.0	15.0
	Unretained	85	85.0	85.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-5 above. Results in Table 5-5 show that this student has a 15% (0.15) chance of being retained therefore a student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore a student in this category has an 85% (0.85) chance of being dropped out of the university. The Figure 5-5 below is a visual representation of the Table 5-5 above.

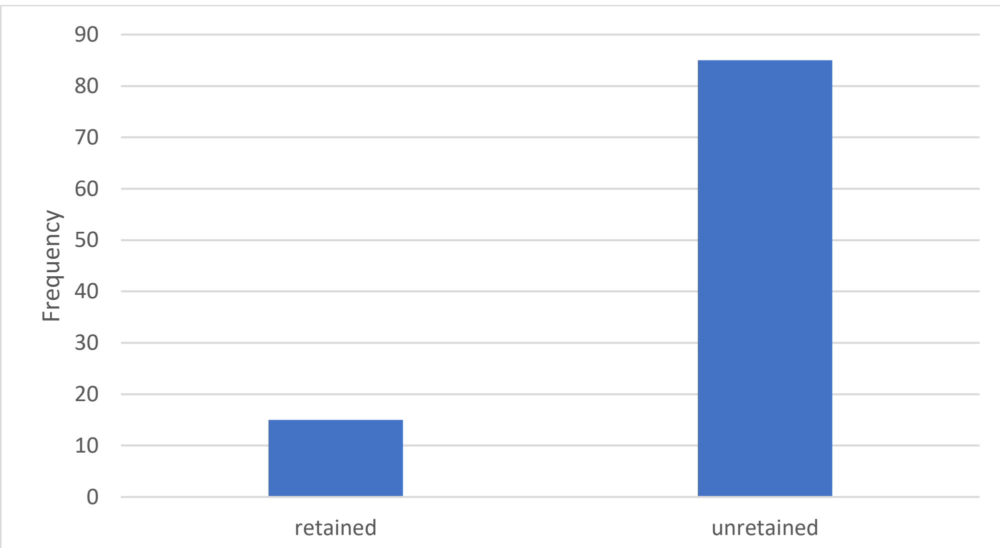


Figure 5-5: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. His or her chance of being retained is 15%. The Figure 5-5 is a visual representation of the Table 5-5 above.

Table 5-6: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	10	10.0	10.0	10.0
	unretained	90	90.0	90.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-6 above. Results in Table 5-6 show that this student has a 10% (0.1) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 90% (0.9) chance of being dropped out of the university. The Figure 5-6 below is a visual representation of the Table 5-6 above.

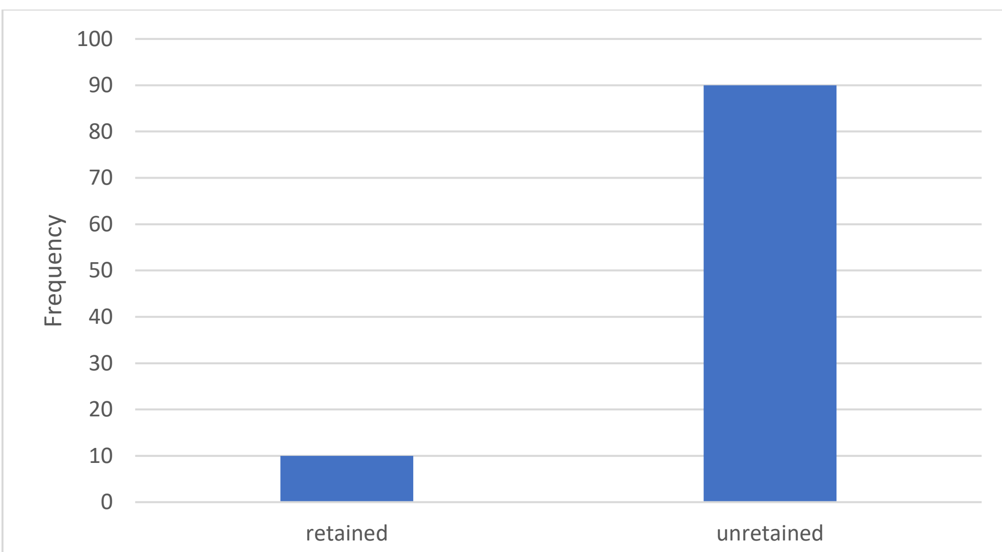


Figure 5-6: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is less than

75%**.** His or her chance of being retained is 10%. The Figure 5-6 is a visual representation of the Table 5-6 above.

Table 5-7: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-7 above. Results in Table 5-7 show that this student has a 5% (0.05) chance of being retained therefore a student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore a student in this category has a 95% (0.95) chance of being dropped out of the university. The Figure 5-7 below is a visual representation of the Table 5-7 above.

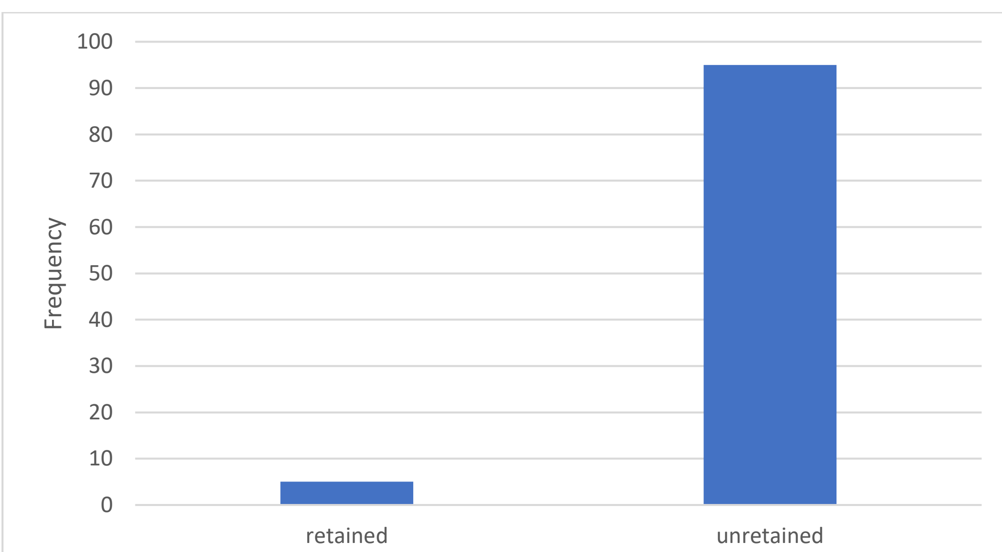


Figure 5-7: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 5.0%. The Figure 5-7 is a visual representation of the Table 5-7 above.

Table 5-8: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-8 above. Results in Table 5-8 show that this student has a 5% (0.05) chance of being retained therefore a student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore a student in this category has a 95% (0.95) chance of being dropped out of the university. The Figure 5-8 below is a visual representation of the Table 5-8 above.

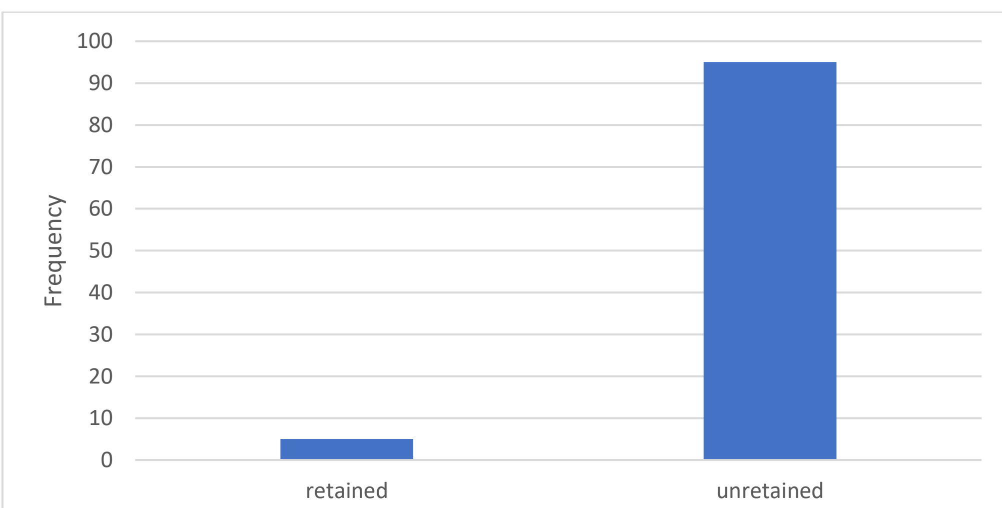


Figure 5-8: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is less than 75%. His or her chance of being retained is 5.0%. The Figure 5-8 is a visual representation of the Table 5-8 above.

Table 5-9: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or more.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	70	70.0	70.0	70.0
	unretained	30	30.0	30.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-9 above. Results in Table 5-9 show that this student has a 70% (0.7) chance of being retained and consequently complete his/her study therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 70% (0.7) chance of being dropped out of the university. The Figure 5-9 below is a visual representation of the Table 5-9 above.

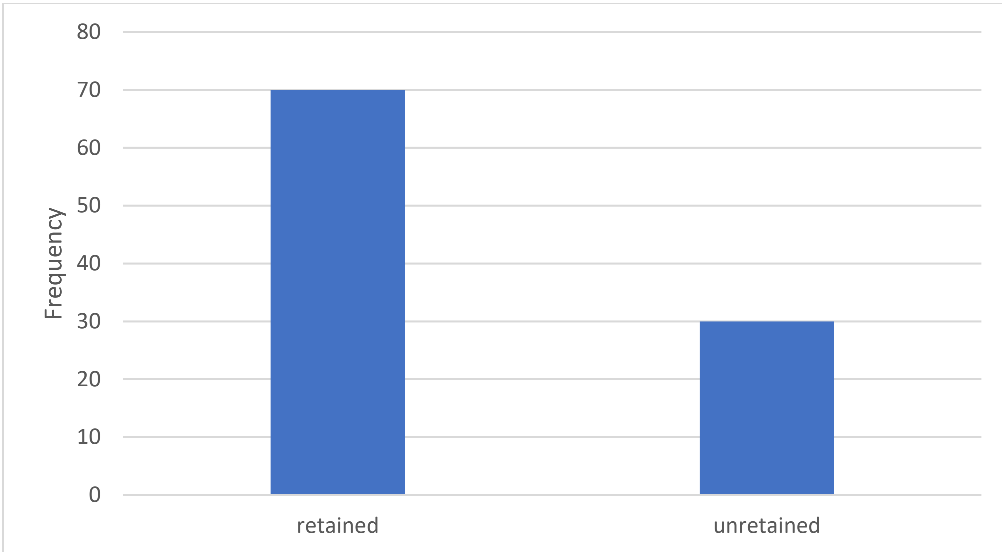


Figure 5-9: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or more. His or her chance of being retained is 70%. The Figure 5-9 is a visual representation of the Table 5-9 above.

Table 5-10: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	55	55.0	55.0	55.0
	unretained	45	45.0	45.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-10 above. Results in Table 5-10 show that this student has a 55% (0.55) chance of being retained and consequently complete his/her study therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0) therefore any student in this category has a 45% (0.45) chance of being dropped out of the university. The Figure 5-10 below is a visual representation of the Table 5-10 above.

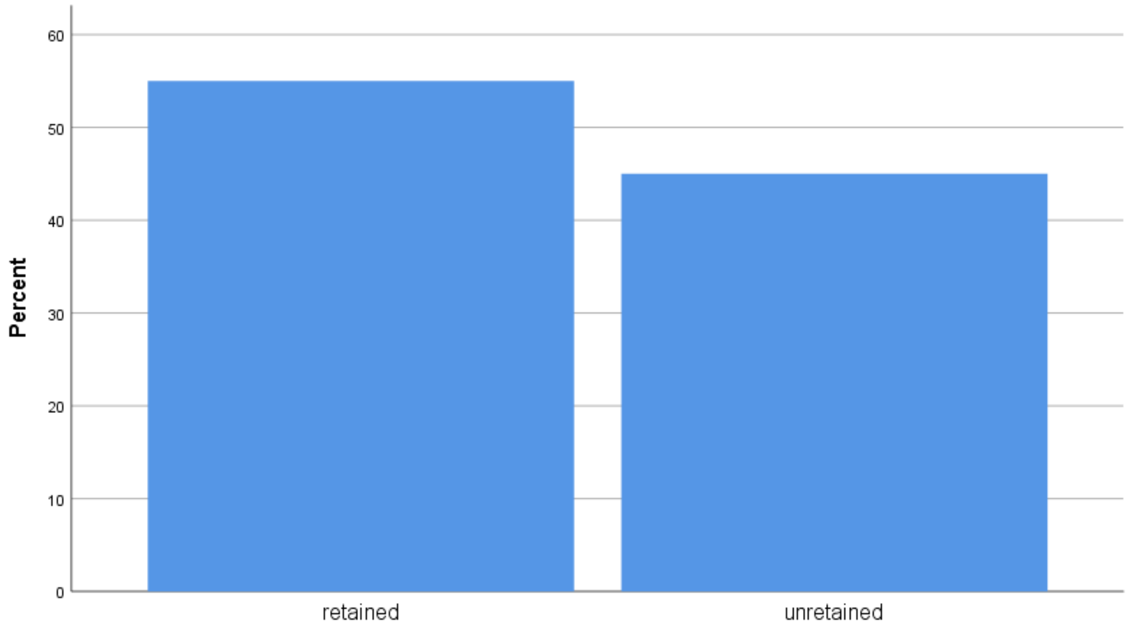


Figure 10: Retention given Change in circumstances is true, Mode of entry is UTME, Academic skills is poor,Academic engagement is engaged, Academic progress is normal and Attendance is less than 75%.

Figure 5-10: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is poor, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is less than 75%. His or her chance of being retained is 55%. The Figure 5-10 is a visual representation of the Table 5-10 above.

Table 5-11: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	60	60.0	60.0	60.0
	unretained	40	40.0	40.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-11 above. Results in Table 5-11 show that this student has a 60% (0.6) chance of being retained and consequently complete his/her study therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0) therefore any student in this category has a 40% (0.40) chance of being dropped out of the university. The Figure 5-11 below is a visual representation of the Table 5-11 above.

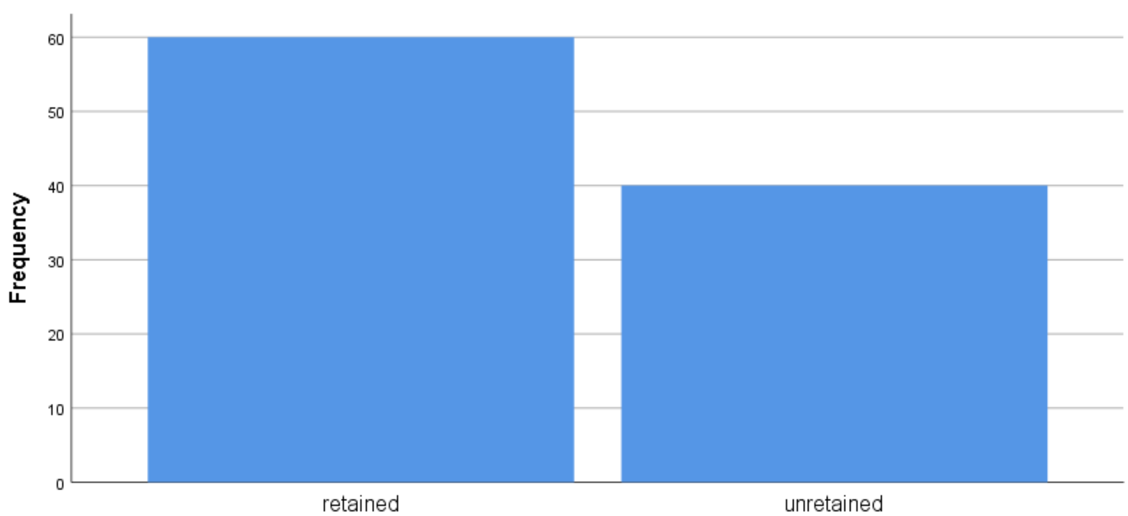


Figure 11: Retention given Change in circumstances is true, Mode of entry is UTME, Academic skills is poor,cademic engagement is engaged, Academic progress is slow and Attendance is 75% or above.

Figure 5-11: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is poor, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above. His or her chance of being retained is 60%. The Figure 5-11 is a visual representation of the Table 5-11 above.

Table 5-12: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is ‘UTME’, Academic skills is poor, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	55	55.0	55.0	55.0
	unretained	45	45.0	45.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-12 above. Results in Table 5-12 show that this student has a 55% (0.55) chance of being retained and consequently complete his/her study therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0) therefore any student in this category has a 45% (0.45) chance of being dropped out of the university. The Figure 5-12 below is a visual representation of the Table 5-12 above.

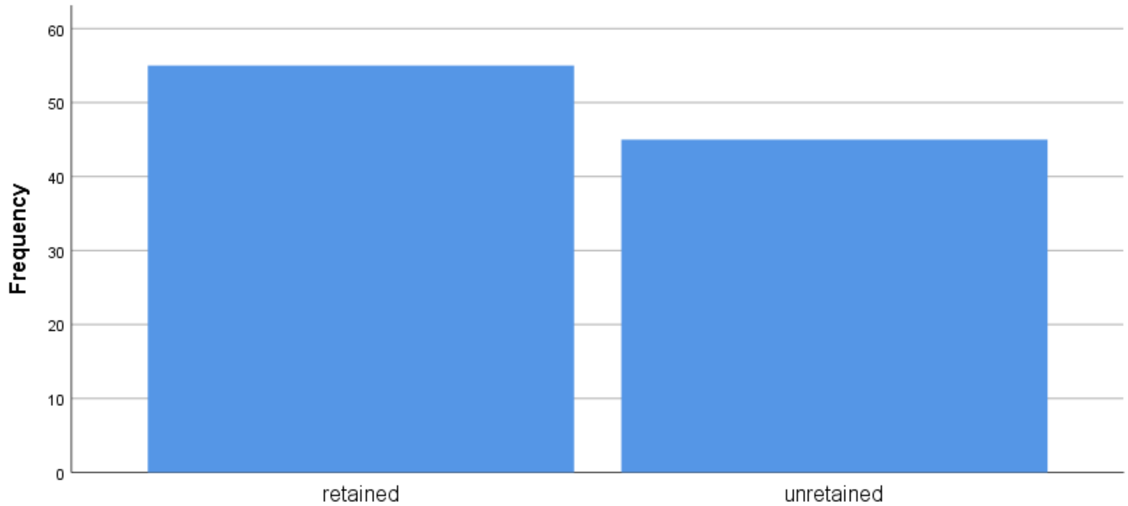


Figure 12: Retention given Change in circumstances is true, Mode of entry is UTME, Academic skills is poor,Academic engagement is engaged, Academic progress is slow and Attendance is less than 75%.

Figure 5-12: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%. His or her chance of being retained is 55%. The Figure 5-12 is a visual representation of the Table 5-12 above.

Table 5-13: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or more.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	20	20.0	20.0	20.0
	unretained	80	80.0	80.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-13 above. Results

in Table 5-13 show that this student has a 20% (0.2) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 80% (0.8) chance of being dropped out of the university. The Figure 5-13 below is a visual representation of the Table 5-13 above.

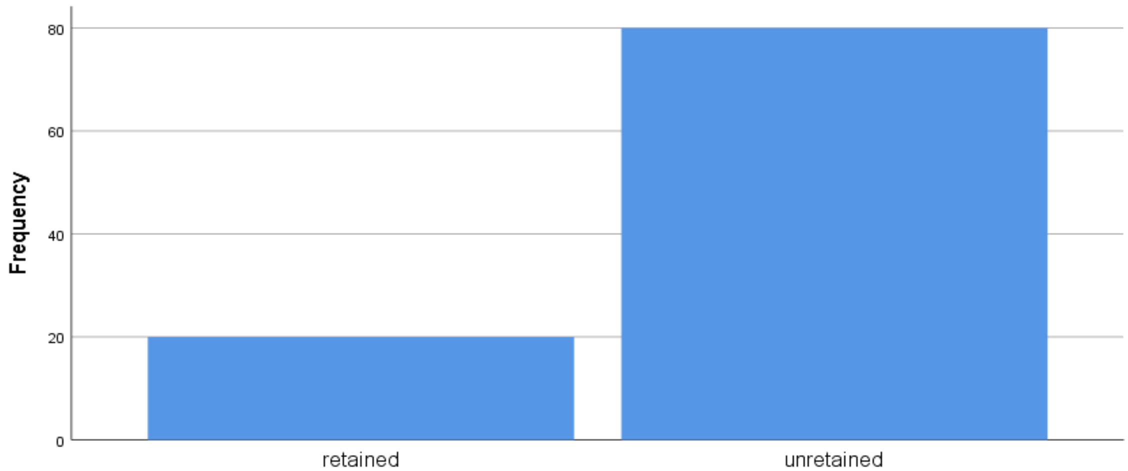


Figure 13: Retention given Change in circumstances is true, Mode of entry is UTME, Academic skills is poor,Academic engagement is unengaged, Academic progress is normal and Attendance is 75% or above.

Figure 5-13: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or more. His or her chance of being retained is 20%. The Figure 5-13 is a visual representation of the Table 5-13 above.

Table 5-14: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	15	15.0	15.0	15.0
	unretained	85	85.0	85.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-14 above. Results in Table 5-14 show that this student has a 15% (0.15) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 85% (0.85) chance of being dropped out of the university. The Figure 5-14 below is a visual representation of the Table 5-14 above.

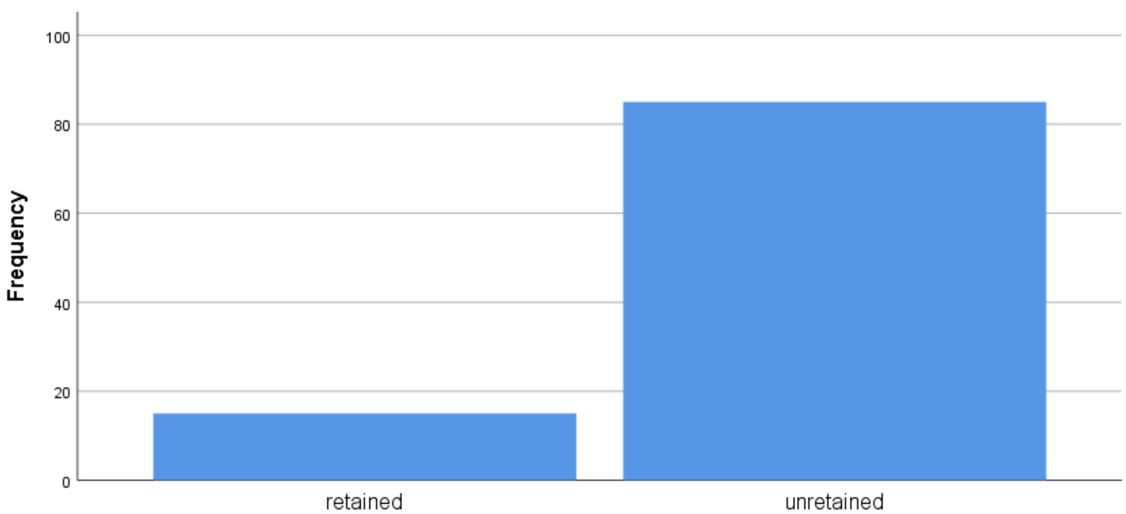


Figure 14: Retention given Change in circumstances is true, Mode of entry is UTME, Academic skills is poor,Academic engagement is unengaged, Academic progress is normal and Attendance is less than 75%.

Figure 5-14: Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is less than 75%. His or her chance of being retained is 15%. The Figure 5-14 is a visual representation of the Table 5-14 above.

Table 5-15: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-15 above. Results in Table 5-15 show that this student has a 5% (0.95) chance of being retained therefore a student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore a student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-15 below is a visual representation of the Table 5-15 above.

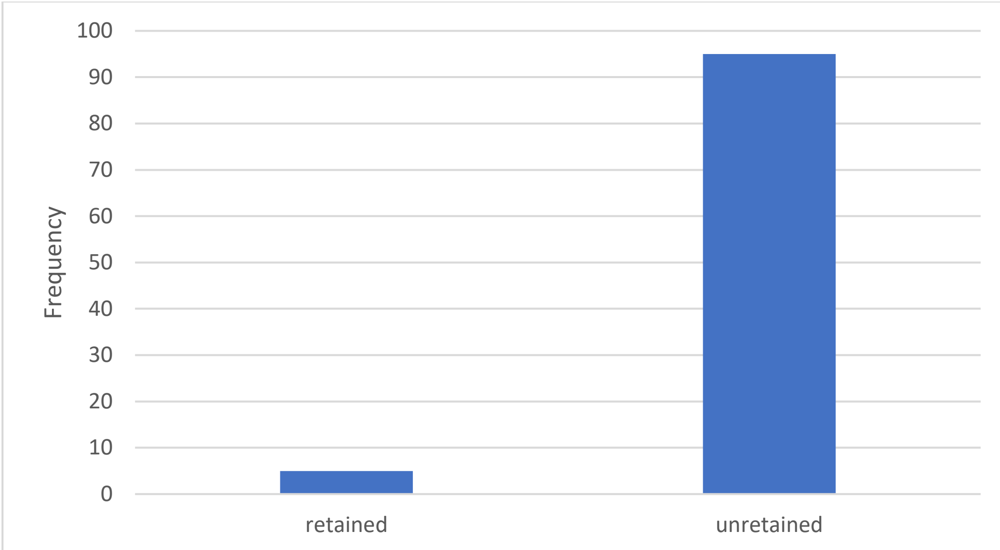


Figure 5-15: Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 5%. The Figure 5-15 is a visual representation of the Table 5-15 above.

Table 5-16: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-16 above. Results in Table 5-16 show that this student has a 5% (0.95) chance of being retained therefore a student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore a student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-16 below is a visual representation of the Table 5-16 above.

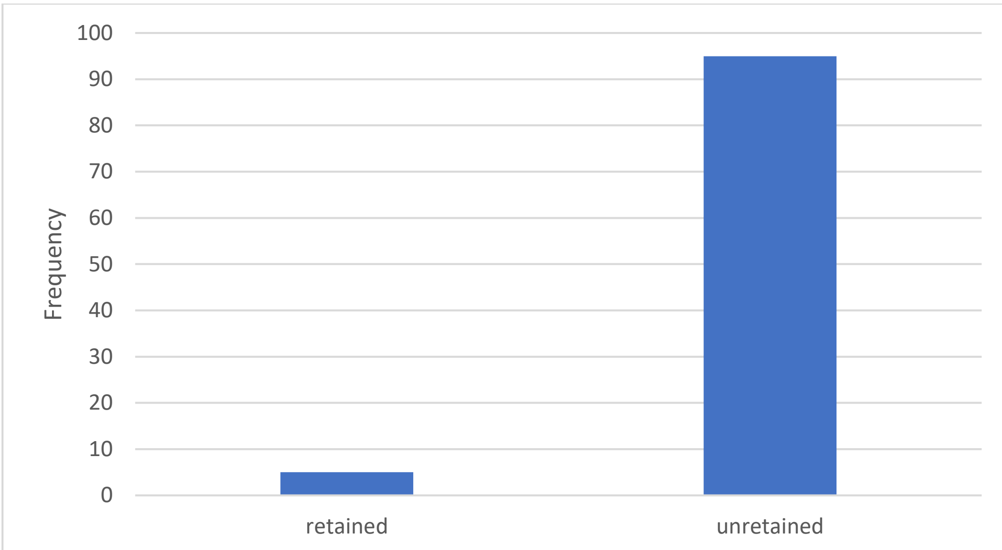


Figure 5-16: Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘UTME’, Academic skills is poor, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is less than 75%. His or her chance of being retained is 5%.. The Figure 5-16 is a visual representation of the Table 5-16 above.

Table 5-17: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘DE’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	90	90.0	90.0	90.0
	unretained	10	10.0	10.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-17 above. Results in Table 5-17 show that this student has a 90% (0.0) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore a student in this category has a 10% (0.1) chance of being dropped out of the university. The Figure 5- 17 below is a visual representation of the Table 5-17 above.

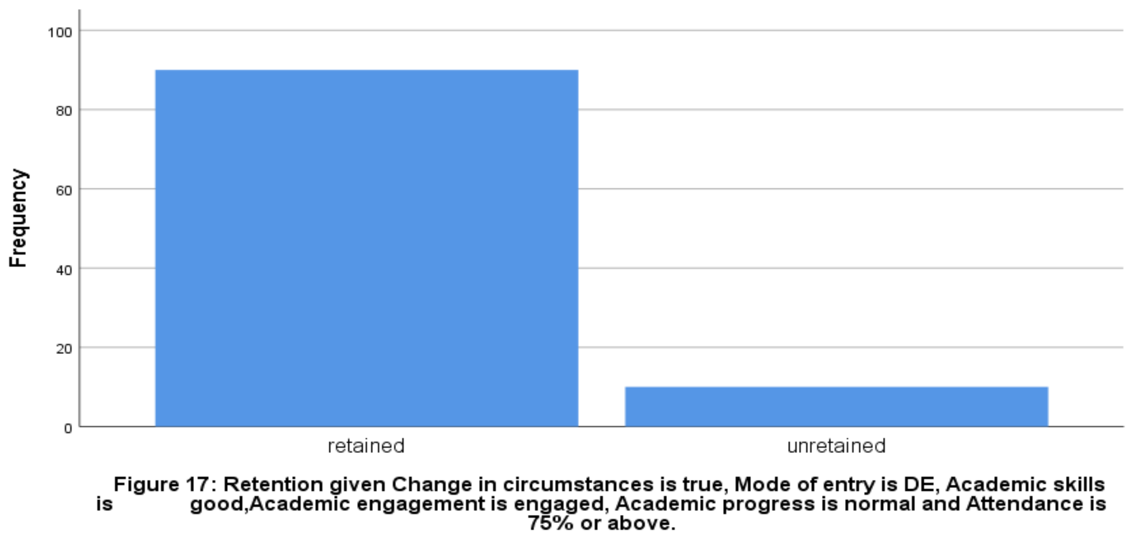


Figure 5-17: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. His or her chance of being retained is 90%. The Figure 5-17 is a visual representation of the Table 5-17 above.

Table 5-18: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	75	75.0	75.0	75.0
	unretained	25	25.0	25.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-18 above. Results in Table 5-18 show that this student has a 75% (0.75) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore a student in this category has a 25% (0.25) chance of being dropped out of the university. The Figure 5-18 below is a visual representation of the Table 5-18 above.

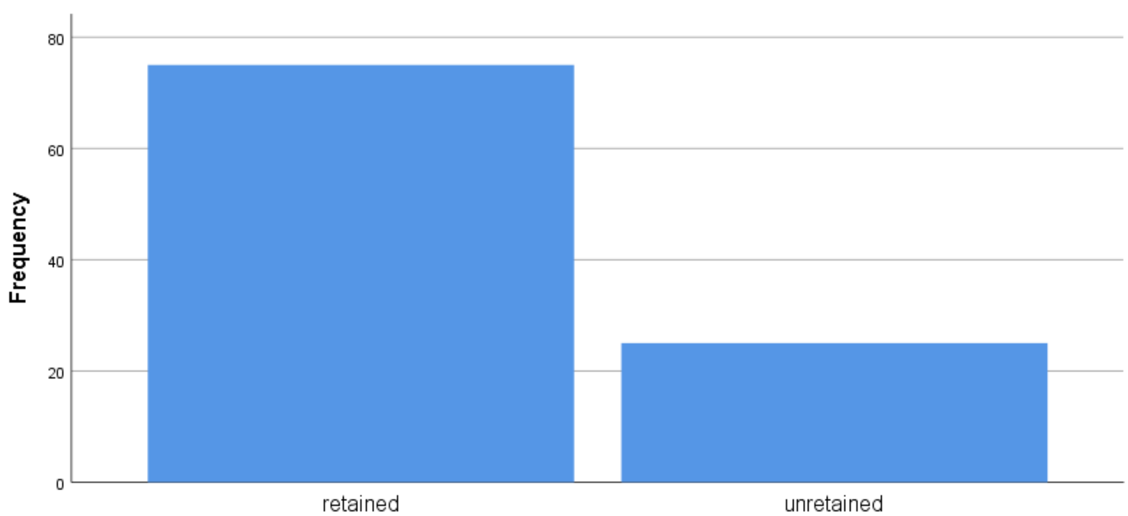


Figure 18: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is good,Academic engagement is engaged, Academic progress is normal and Attendance is less than 75%.

Figure 5-18: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘DE’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is less than 75%. His or her chance of being retained is 75%. The Figure 5-18 is a visual representation of the Table 5-18 above.

Table 5-19: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘DE’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	65	65.0	65.0	65.0
	unretained	35	35.0	35.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-19 above. Results in Table 5-19 show that this student has a 65% (0.65) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore a student in this category has a 35% (0.35) chance of being dropped out of the university. The Figure 5-19 below is a visual representation of the Table 5-19 above.

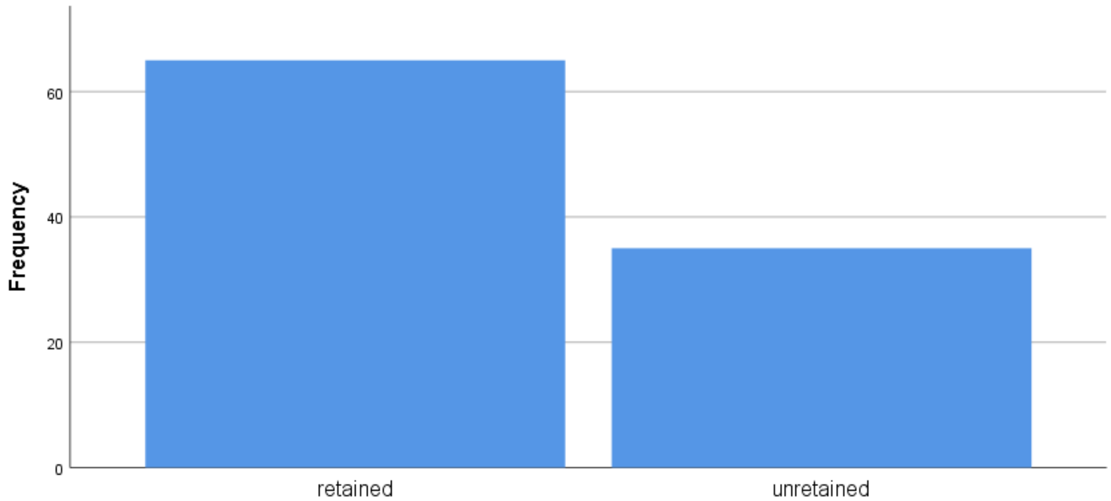


Figure 19: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is good,Academic engagement is engaged, Academic progress is slow and Attendance is 75% or above.

Figure 5-19: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 65%. The Figure 5-19 is a visual representation of the Table 5- 19 above.

Table 5-20: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	45	45.0	45.0	45.0
	unretained	55	55.0	55.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-20 above. Results in Table 5-20 show that this student has a 45% (0.45) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 55% (0.55) chance of being dropped out of the university. The Figure 5-20 below is a visual representation of the Table 5-20 above.

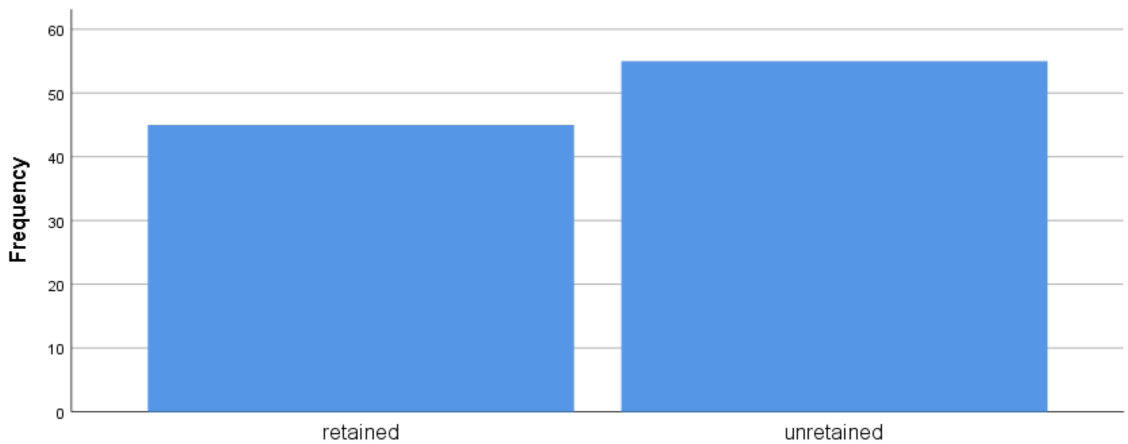


Figure 20: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is good, Academic engagement is engaged, Academic progress is slow and Attendance is less than 75%.

Figure 5-20: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%. His or her chance of being retained is 45%. The Figure 5-20 is a visual representation of the Table 5-20 above.

Table 5-21: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	20	20.0	20.0	20.0
	unretained	80	80.0	80.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-21 above. Results in Table 5-21 show that this student has a 20% (0.2) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 20% (0.2) chance of being dropped out of the university. The Figure 5-21 below is a visual representation of the Table 5-21 above.

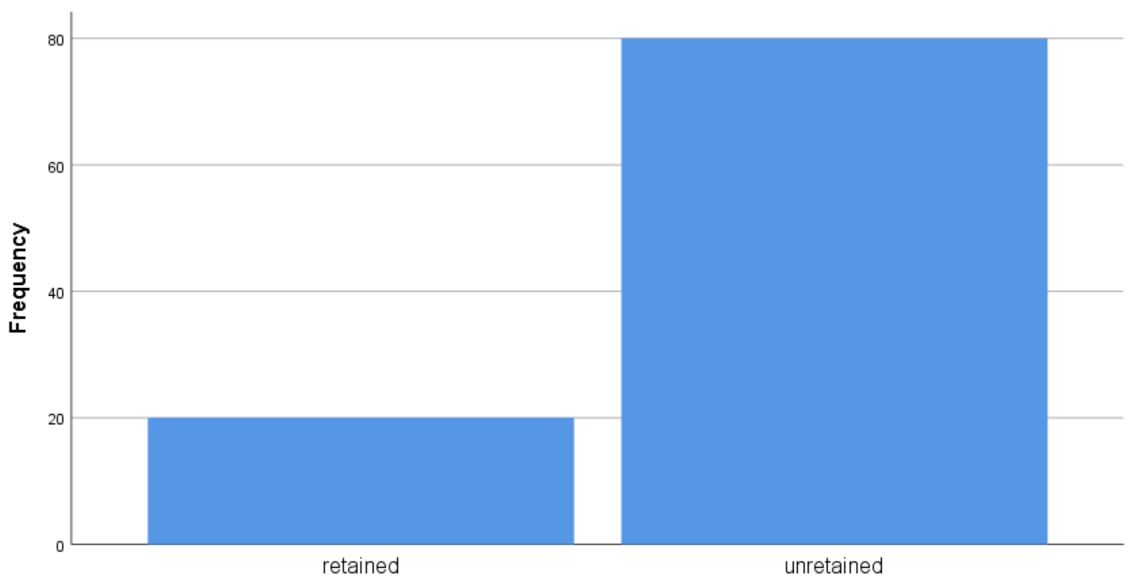


Figure 21: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is good, Academic engagement is not engaged, Academic progress is normal and Attendance is 75% or above.

Figure 5-21: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘DE’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. His or her chance of being retained is 20%. The Figure 5-21 is a visual representation of the Table 5-21 above.

Table 5-22: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘DE’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	15	15.0	15.0	15.0
	unretained	85	85.0	85.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-22 above. Results in Table 5-22 show that this student has a 15% (0.15) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 85% (0.85) chance of being dropped out of the university. The Figure 5-22 below is a visual representation of the Table 5-22 above.

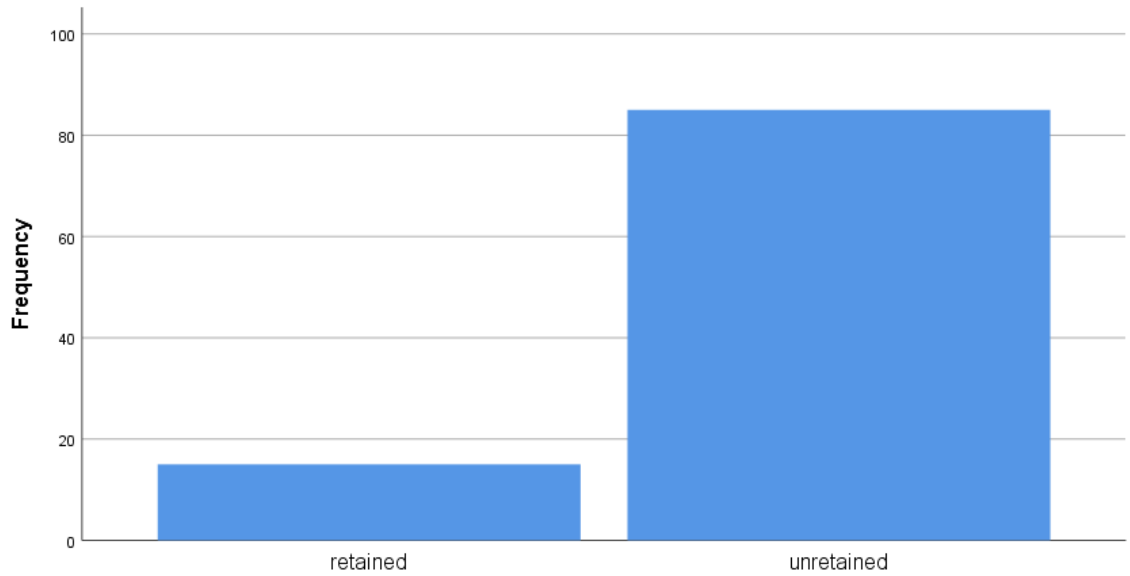


Figure 5-22: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is less than 75%. His or her chance of being retained is 15%. The Figure 5-22 is a visual representation of the Table 5-22 above.

Table 5-23: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-23 above. Results in Table 5-23 show that this student has a 5% (0.95) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-23 below is a visual representation of the Table 5-23 above.

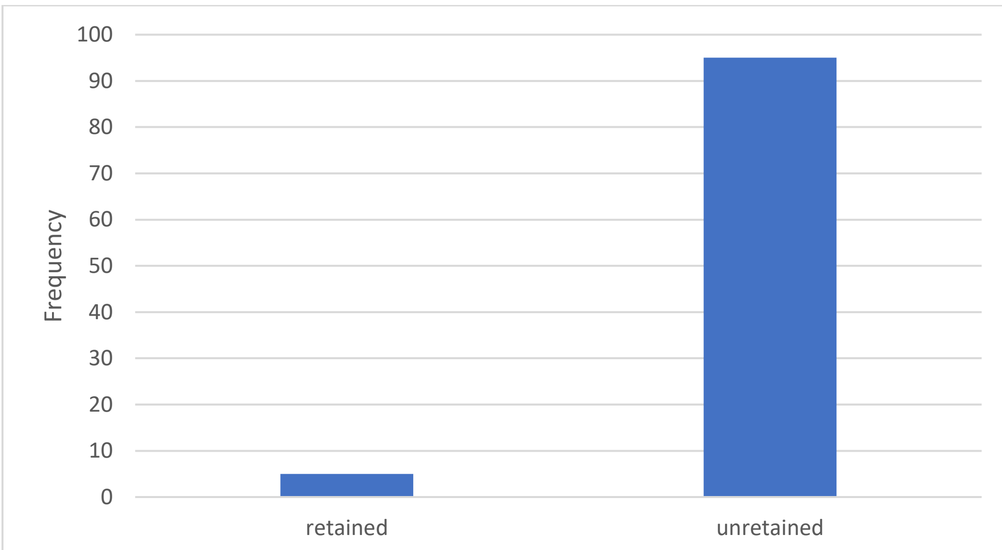


Figure 5-23: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 5%. The Figure 5-23 is a visual representation of the Table 5-23 above.

Table 5-24: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-24 above. Results in Table 5-24 show that this student has a 0% (0.0) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-24 below is a visual representation of the Table 5-24 above.

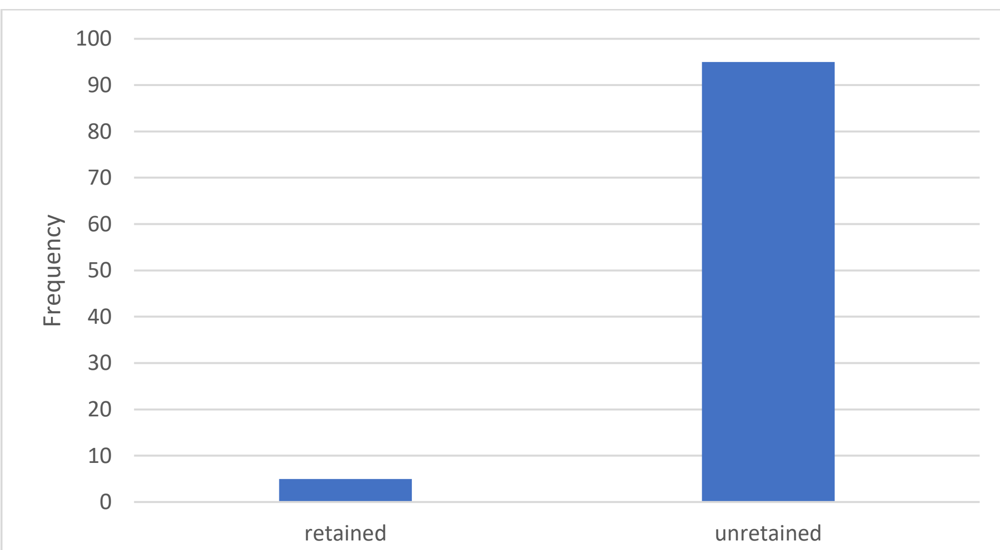


Figure 5-24: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is less than 75%. His or her chance of being retained is 5%. The Figure 5-24 is a visual representation of the Table 5-24 above.

Table 5-25: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	75	75.0	75.0	75.0
	unretained	25	25.0	25.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-s25 above. Results in Table 5-25 show that this student has a 75% (0.75) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 25% (0.25) chance of being dropped out of the university. The Figure 5-25 below is a visual representation of the Table 5-25 above.

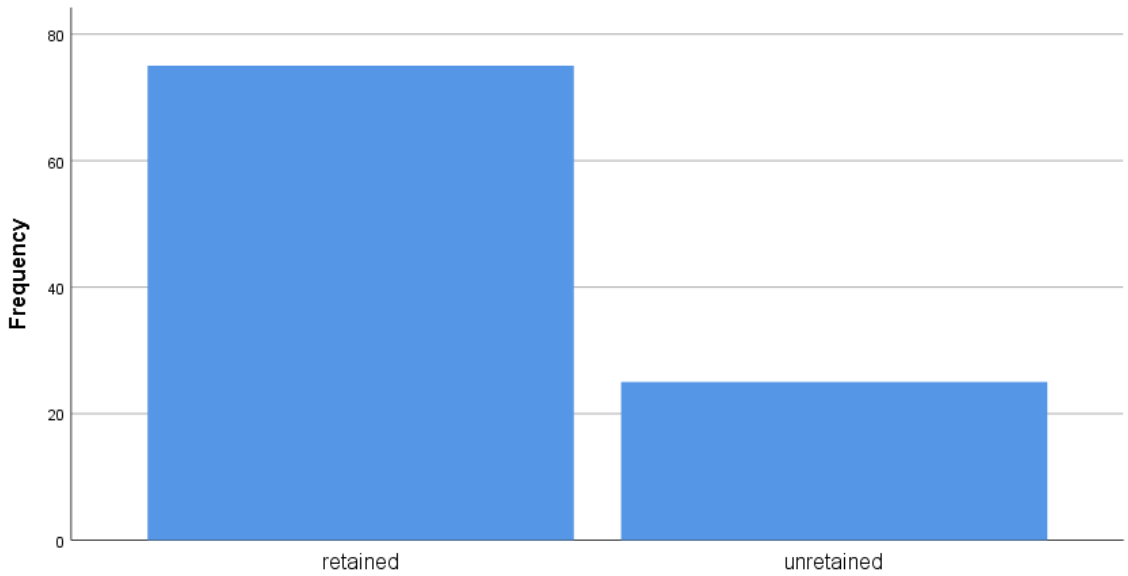


Figure 25: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is poor,Academic engagement is engaged, Academic progress is normal and Attendance is 75% or above.

Figure 5-25: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘DE’, Academic skills is poor, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. His or her chance of being retained is 75%. The Figure 5-25 is a visual representation of the Table 5-25 above.

Table 5-26: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘DE’, Academic skills is poor, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	65	65.0	65.0	65.0
	unretained	35	35.0	35.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-26 above. Results

in Table 5-26 show that this student has a 65% (0.65) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 35% (0.35) chance of being dropped out of the university. The Figure 5-26 below is a visual representation of the Table 5-26 above.

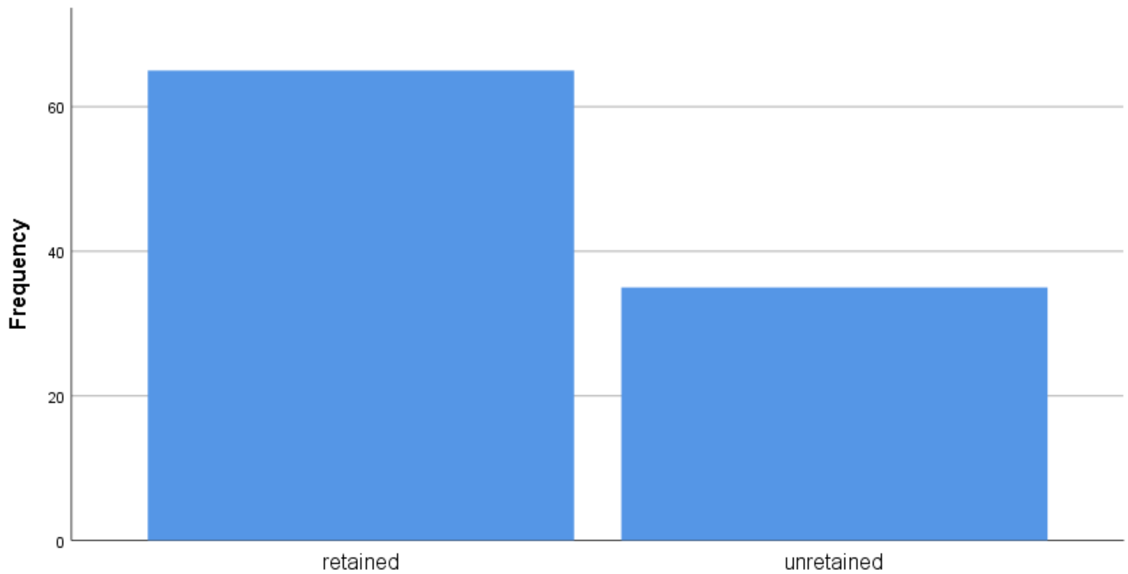


Figure 26: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is poor,Academic engagement is engaged, Academic progress is normal and Attendance is below 75%.

Figure 5-26: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘DE’, Academic skills is poor, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is less than 75%. His or her chance of being retained is 65%. The Figure 5-26 is a visual representation of the Table 5-26 above.

Table 5-27: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	40	40.0	40.0	40.0
	unretained	60	60.0	60.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-27 above. Results in Table 5-27 show that this student has a 40% (0.4) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 60% (0.6) chance of being dropped out of the university. The Figure 5-27 below is a visual representation of the Table 5-27 above.

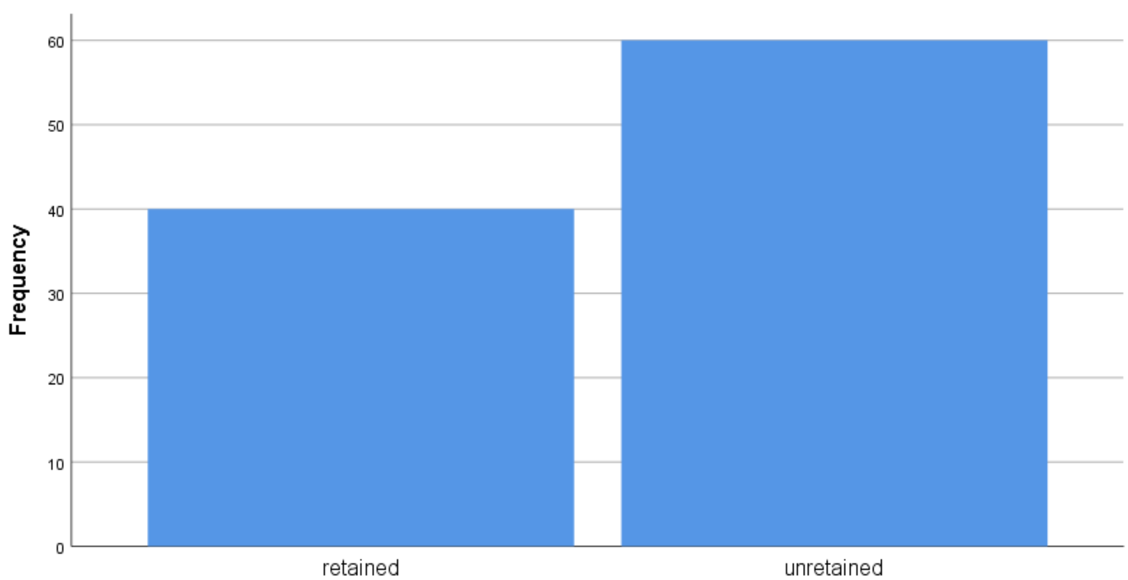


Figure 27: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is poor,Academic engagement is engaged, Academic progress is slow and Attendance is 75% or above.

Figure 5-27: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 40%. The Figure 5-27 is a visual representation of the Table 5-27 above.

Table 5-28: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	30	30.0	30.0	30.0
	unretained	70	70.0	70.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-28 above. Results in Table 5-28 show that this student has a 30% (0.3) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 70% (0.7) chance of being dropped out of the university. The Figure 5-28 below is a visual representation of the Table 5-28 above.

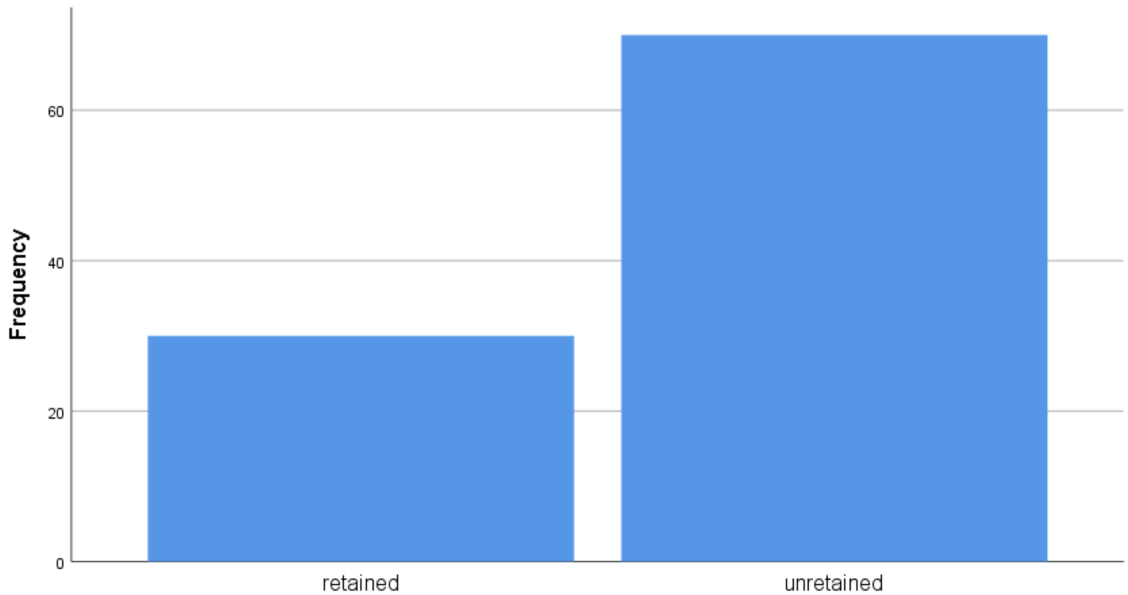


Figure 28: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is poor,Academic engagement is engaged, Academic progress is slow and Attendance is below 75%.

Figure 5-28: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%. His or her chance of being retained is 30%. The Figure 5-28 is a visual representation of the Table 5-28 above.

Table 5-29: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	20	20.0	20.0	20.0
	unretained	80	80.0	80.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-29 above. Results

in Table 5-29 show that this student has a 20% (0.2) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 80% (0.8) chance of being dropped out of the university. The Figure 5-29 below is a visual representation of the Table 5-29 above.

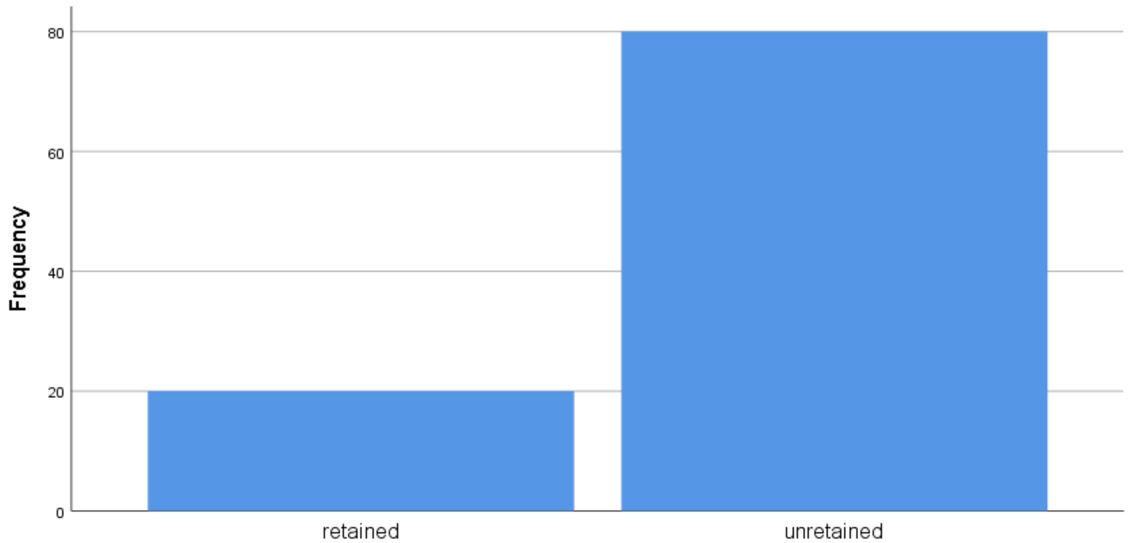


Figure 29: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is poor,Academic engagement is not engaged, Academic progress is normal and Attendance is 75% or above.

Figure 5-29: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above. His or her chance of being retained is 20%. The Figure 5-29 is a visual representation of the Table 5- 29 above.

Table 5-30: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	15	15.0	15.0	15.0
	unretained	85	85.0	85.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-30 above. Results in Table 5-30 show that this student has a 15% (0.15) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 85% (0.85) chance of being dropped out of the university. The Figure 5-30 below is a visual representation of the Table 5-30 above.

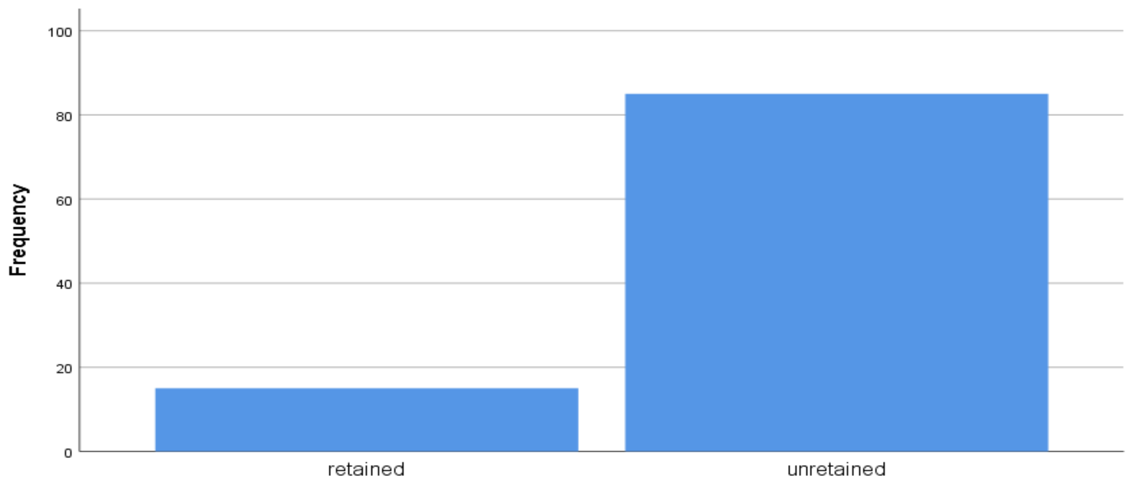


Figure 30: Retention given Change in circumstances is true, Mode of entry is DE, Academic skills is poor,Academic engagement is not engaged, Academic progress is normal and Attendance is below 75%.

Figure 5-30: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is less than 75%. His or her chance of being retained is 15%. The Figure 5-30 is a visual representation of the Table 5-30 above.

Table 5-31: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-31 above. Results in Table 5-31 show that this student has a 5% (0.95) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-31 below is a visual representation of the Table 5-31 above.

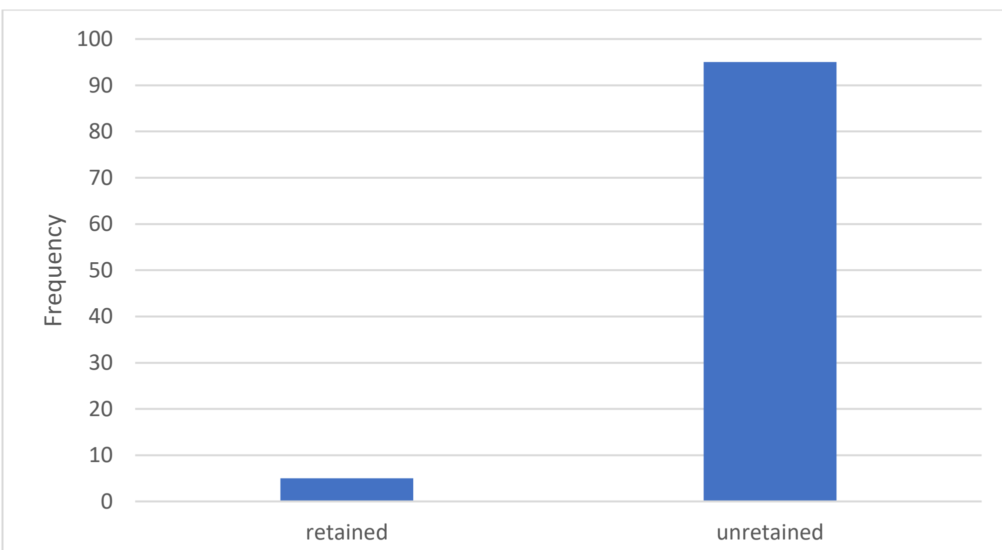


Figure 5-31: Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic

engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. His or her chance of being retained is 5%. The Figure 5-31 is a visual representation of the Table 5-31 above.

Table 5-32: Frequency table for the chance of a student being retained given: Change in circumstances is true, Mode of entry is through ‘DE’, Academic skills is poor, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-32 above. Results in Table 5-32 show that this student has a 5% (0.05) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 95% (0.95) chance of being dropped out of the university. The Figure 5-32 below is a visual representation of the Table 5-32 above.

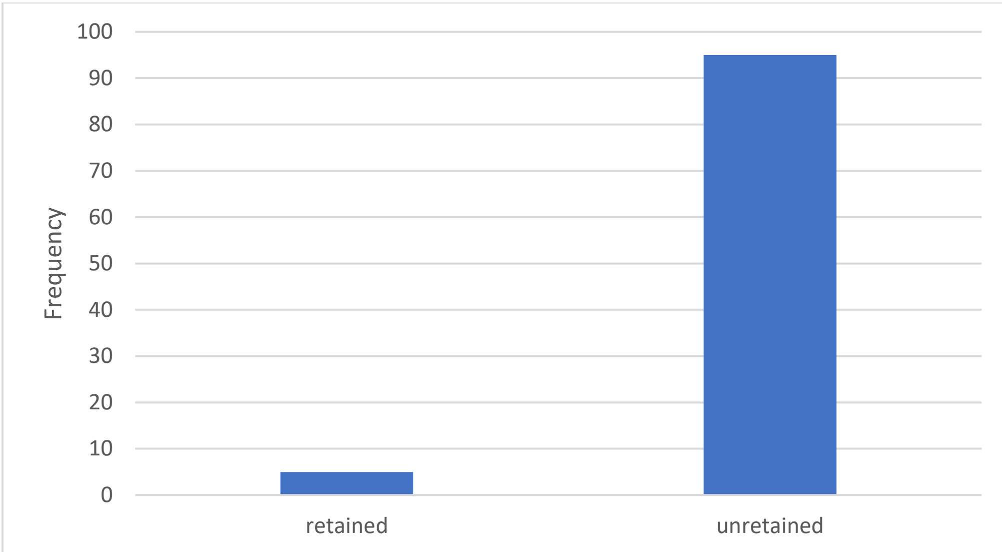


Figure 5-32: A Bar chart representing the chance of a student being retained given: Change in circumstances is true, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is less than 75%. His or her chance of being retained is 5%. The Figure 5-32 is a visual representation of the Table 5-32 above.

Table 5-33: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	90	90.0	90.0	90.0
	unretained	10	10.0	10.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-33 above. Results in Table 5-33 show that this student has a 95% (0.95) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 5% (0.05) chance of being dropped out of the university. The Figure 5-33 below is a visual representation of the Table 5-33 above.

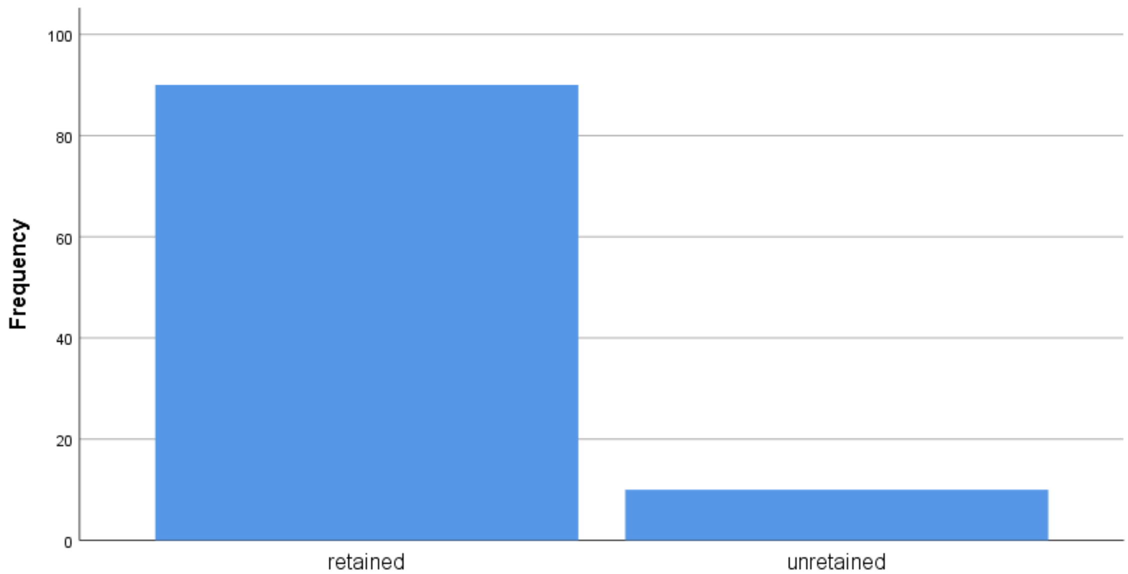


Figure 33: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is good,Academic engagement is engaged, Academic progress is normal and Attendance is 75% or above.

Figure 5-33: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is 75% or above. His or her chance of being retained is 90%. The Figure 5-33 is a visual representation of the Table 5-33 above.

Table 5-34: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	75	75.0	75.0	75.0
	unretained	25	25.0	25.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-34 above. Results in Table 5-34 show that this student has a 75% (0.75) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 25% (0.25) chance of being dropped out of the university. The Figure 5-34 below is a visual representation of the Table 5-34 above.

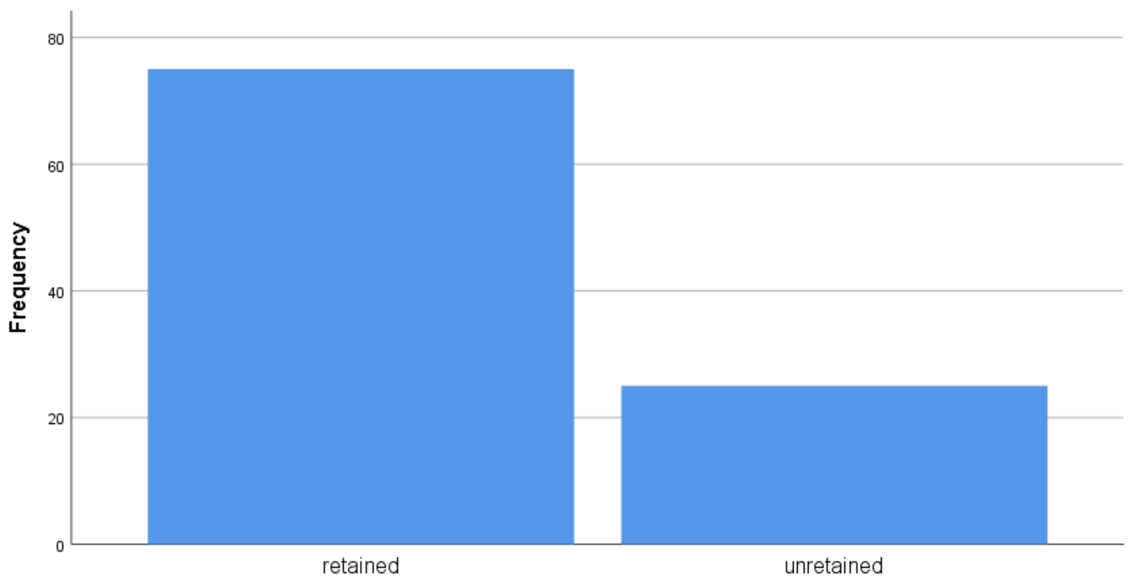


Figure 34: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is good,Academic engagement is engaged, Academic progress is normal and Attendance is less than 75%.

Figure 5-34: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is less than 75%. His or her chance of being retained is 75%. The Figure 5-34 is a visual representation of the Table 5-34 above.

Table 5-35: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	45	45.0	45.0	45.0
	unretained	55	55.0	55.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-35 above. Results in Table 5-35 show that this student has a 45% (0.45) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 55% (0.55) chance of being dropped out of the university. The Figure 5-35 below is a visual representation of the Table 5-35 above.

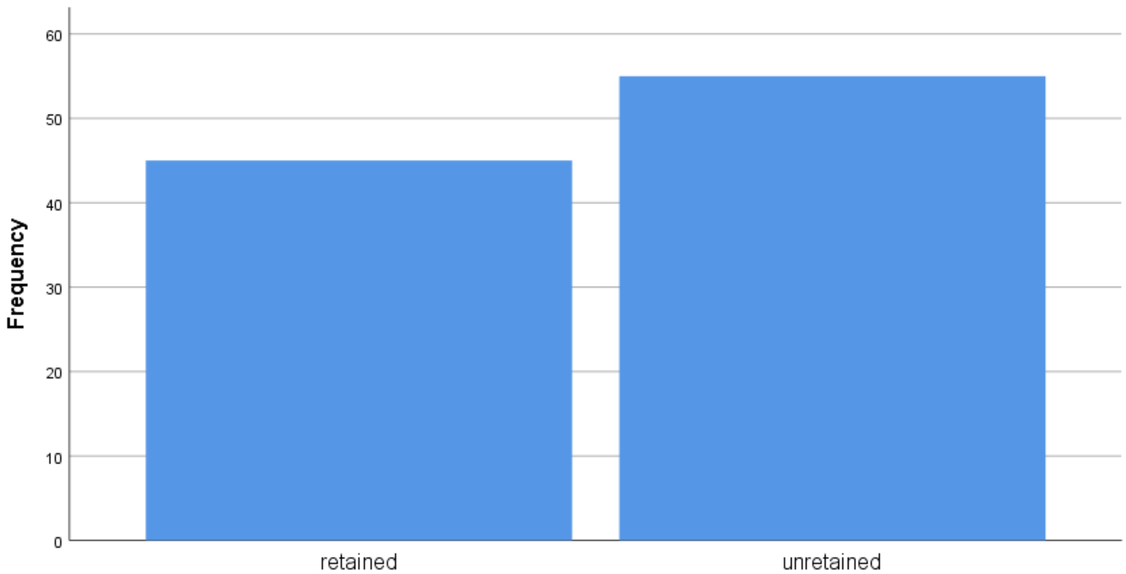


Figure 35: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is good,Academic engagement is engaged, Academic progress is slow and Attendance is 75% or above.

Figure 5-35: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 45%. The Figure 5-35 is a visual representation of the Table 5-35 above.

Table 5-36: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	37	37.0	37.0	37.0
	unretained	63	63.0	63.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-36 above. Results

in Table 5-36 show that this student has a 37% (0.37) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 63% (0.63) chance of being dropped out of the university. The Figure 5-36 below is a visual representation of the Table 5-36 above.

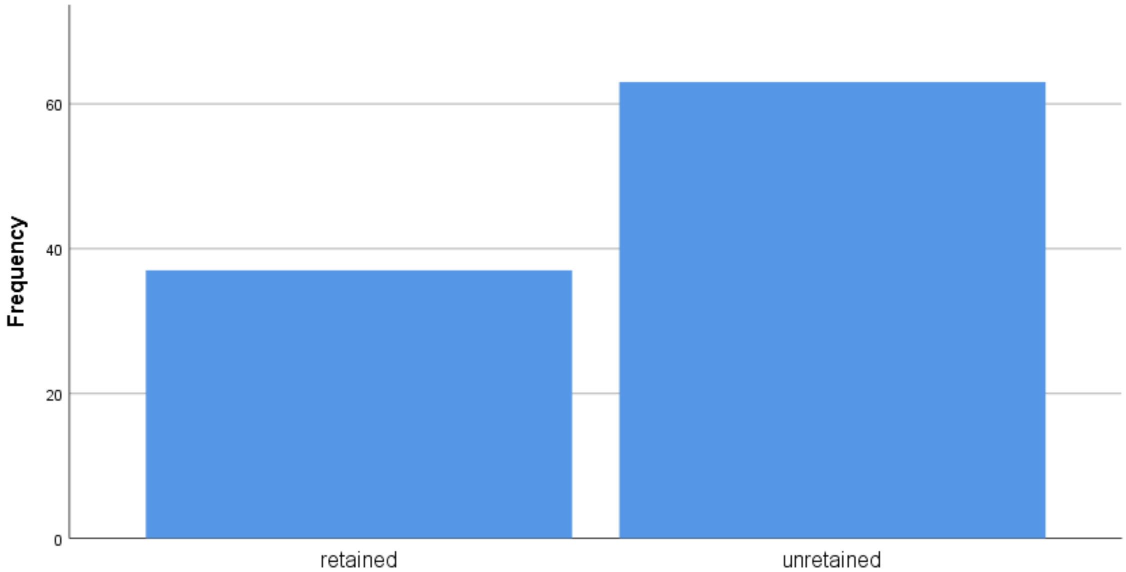


Figure 36: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is good,Academic engagement is engaged, Academic progress is slow and Attendance is less than 75%.

Figure 5-36: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%. His or her chance of being retained 37%. The Figure 5-36 is a visual representation of the Table 5-36 above.

Table 5-37: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	20	20.0	20.0	20.0
	unretained	80	80.0	80.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-37 above. Results in Table 5-37 show that this student has a 20% (0.2) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 80% (0.8) chance of being dropped out of the university. The Figure 5-37 below is a visual representation of the Table 5-37 above.

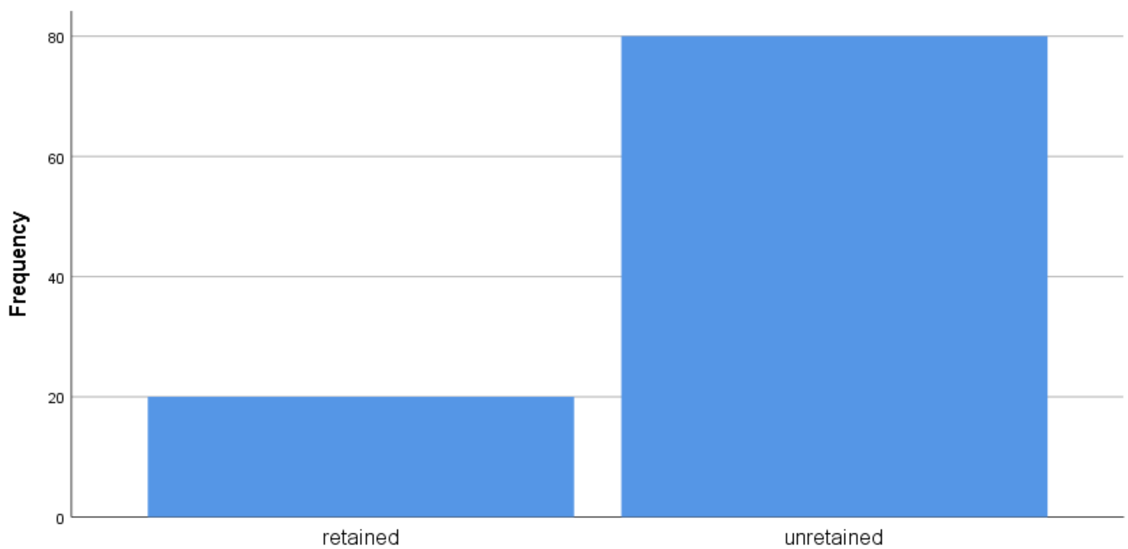


Figure 37: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is good,Academic engagement is not engaged, Academic progress is normal and Attendance is 75% or above.

Figure 5-37: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. His or her chance of being retained 20%. The Figure 5-37 is a visual representation of the Table 5-37 above.

Table 5-38: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	15	15.0	15.0	15.0
	unretained	85	85.0	85.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-38 above. Results in Table 5-38 show that this student has a 15% (0.15) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has 85% (0.85) chance of being dropped out of the university. The Figure 5-38 below is a visual representation of the Table 5-38 above.

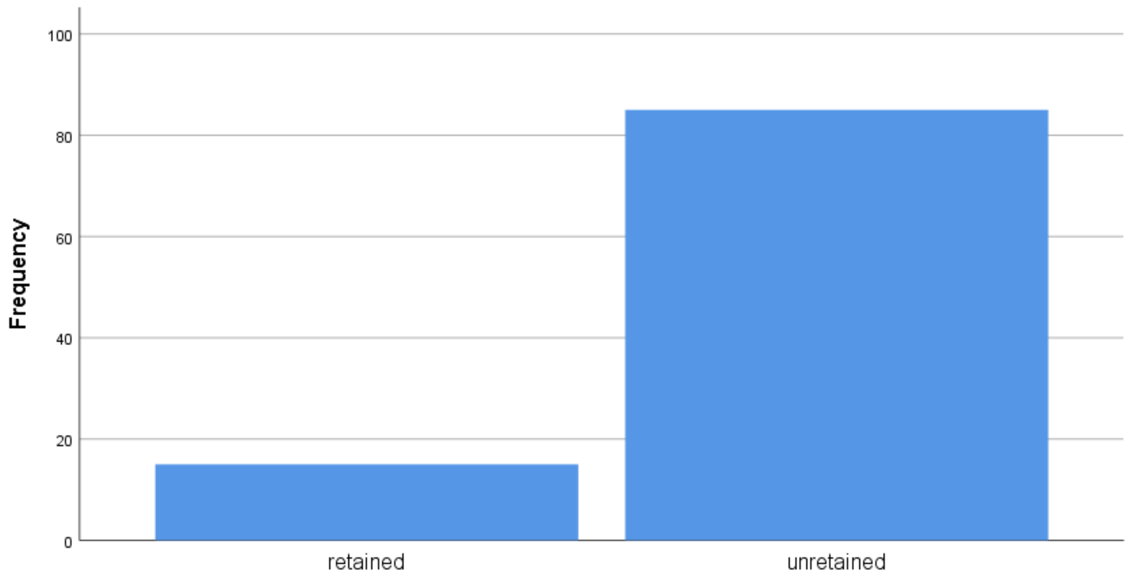


Figure 38: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is good,Academic engagement is not engaged, Academic progress is normal and Attendance is below 75%.

Figure 5-38: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. His or her chance of being retained 15%. The Figure 5-38 is a visual representation of the Table 5-38 above.

Table 5-39: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-39 above. Results

in Table 5-39 show that this student has a 5% (0.0) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 95% (0.95) chance of being dropped out of the university. The Figure 5-39 below is a visual representation of the Table 5-39 above.

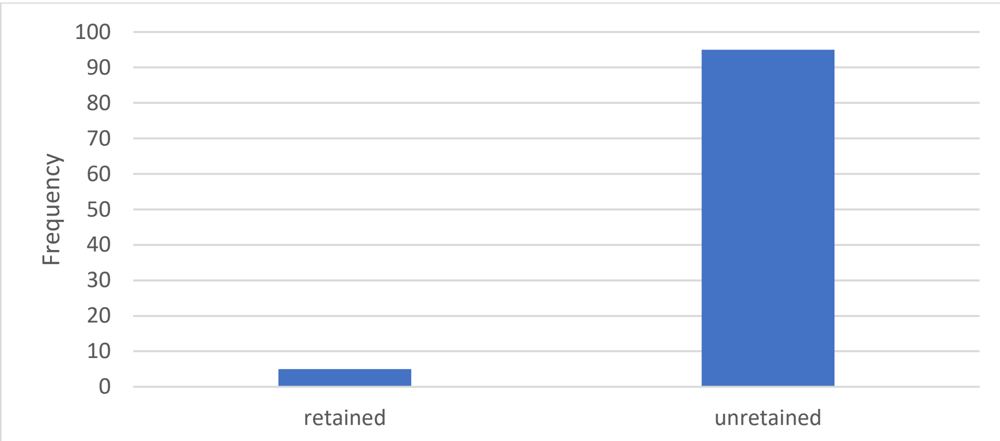


Figure 5-39: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is 75% or above. His or her chance of being retained is 5%. The Figure 5-39 is a visual representation of the Table 5-39 above.

Table 5-40: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-40 above. Results in Table 5-40 show that this student has a 5% (0.95) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 95% (0.95) chance of being dropped out of the university. The Figure 5-40 below is a visual representation of the Table 5-40 above.

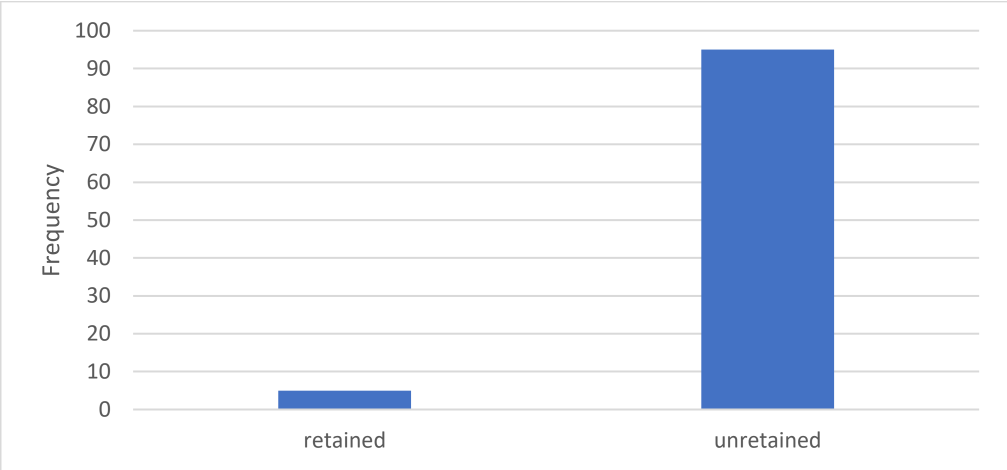


Figure 5-40: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. His or her chance of being retained is 5%. The Figure 5-40 is a visual representation of the Table 5-40 above.

Table 5-41: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	83	83.0	83.0	83.0
	unretained	17	17.0	17.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-41 above. Results in Table 5-41 show that this student has a 83% (0.83) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has 17% (0.17) chance of being dropped out of the university. The Figure 5-41 below is a visual representation of the Table 5-41 above.

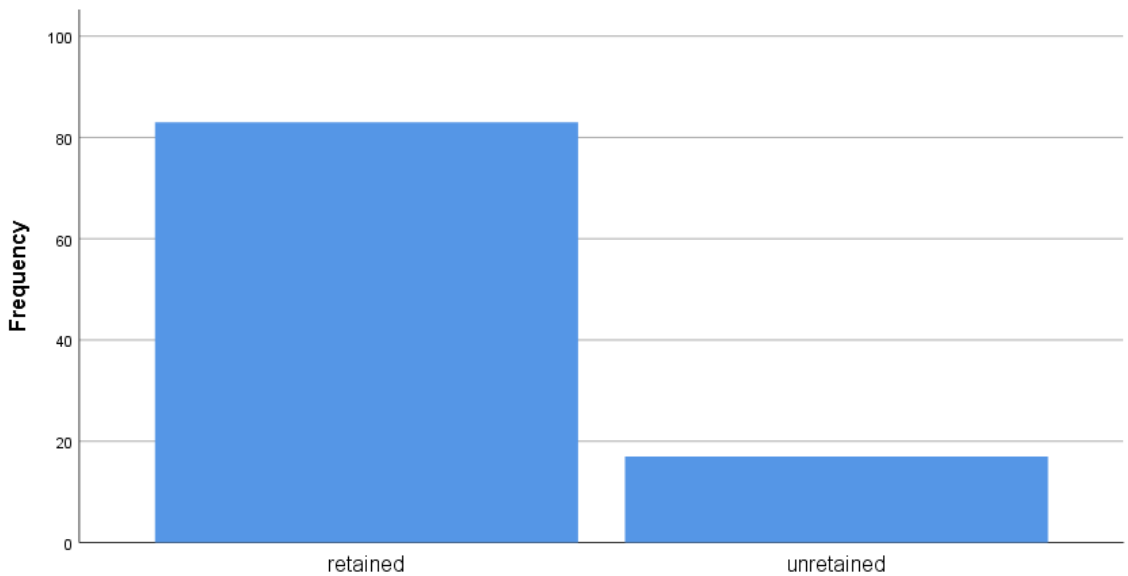


Figure 41: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is poor,Academic engagement is engaged, Academic progress is normal and Attendance is 75% or above.

Figure 5-41: Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. His or her chance of being retained is 83%. The Figure 5-41 is a visual representation of the Table 5-41 above.

Table 5-42: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	75	75.0	75.0	75.0
	unretained	25	25.0	25.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-42 above. Results in Table 5-42 show that this student has a 75% (0.75) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 25% (0.25) chance of being dropped out of the university. The Figure 5-42 below is a visual representation of the Table 5-42 above.

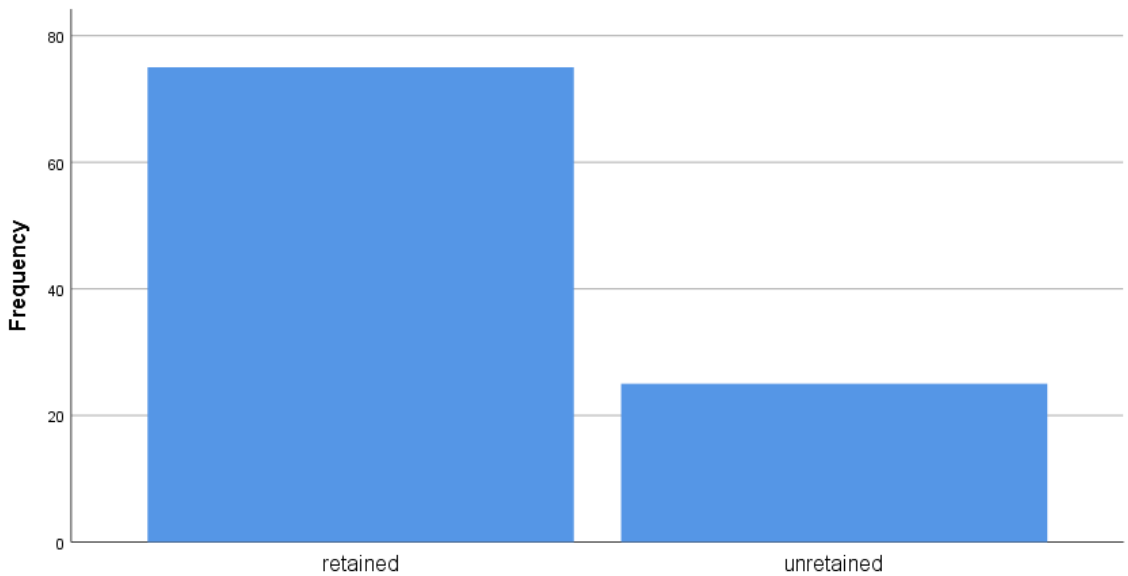


Figure 42: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is poor,Academic engagement is engaged, Academic progress is normal and Attendance is below 75%.

Figure 5-42: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is poor, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is below 75%. His or her chance of being retained is 75%. The Figure 5-42 is a visual representation of the Table 5-42 above.

Table 5-43: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is poor, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	35	35.0	35.0	35.0
	unretained	65	65.0	65.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-43 above. Results in Table 5-43 show that this student has a 35% (0.35) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 65% (0.65) chance of being dropped out of the university. The Figure 5-43 below is a visual representation of the Table 5-43 above.

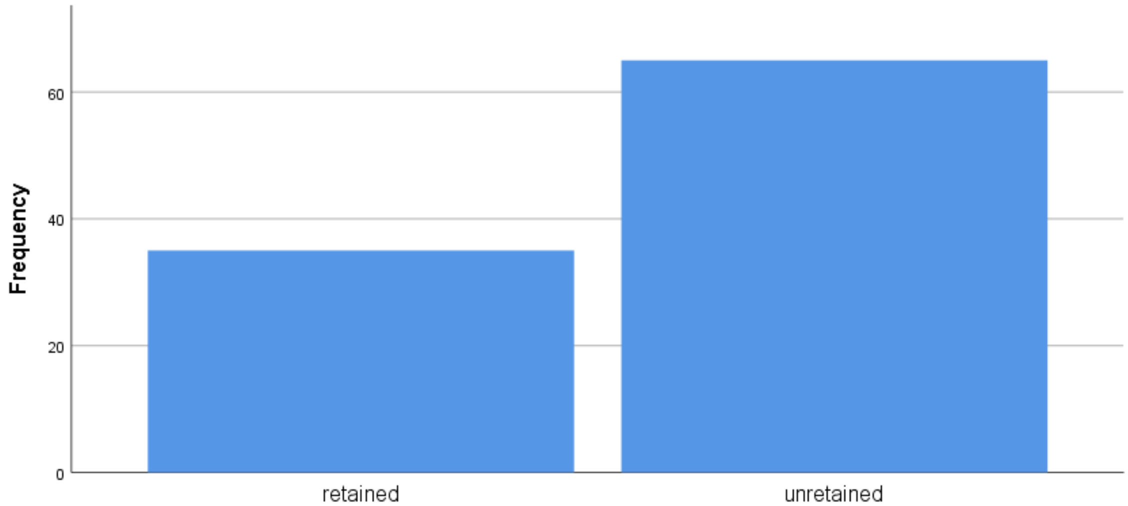


Figure 43: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is poor,Academic engagement is engaged, Academic progress is slow and Attendance is 75% or above.

Figure 5-43: A Bar chart representing the chance of a student being retained given: Change in circumstances is ‘false’, Mode of entry is through ‘UTME’, Academic skills is ‘poor’, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above. His or her chance of being retained is 35%. The Figure 5-43 is a visual representation of the Table 5-43 above.

Table 5-44: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is poor, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	30	30.0	30.0	30.0
	unretained	70	70.0	70.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability sas elicited from the domain experts in this study is displayed in Table 5-44 above. Results in Table 5-44 show that this student has a 30% (0.3) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 70% (0.7) chance of being dropped out of the university. The Figure 5-44 below is a visual representation of the Table 5-44 above.

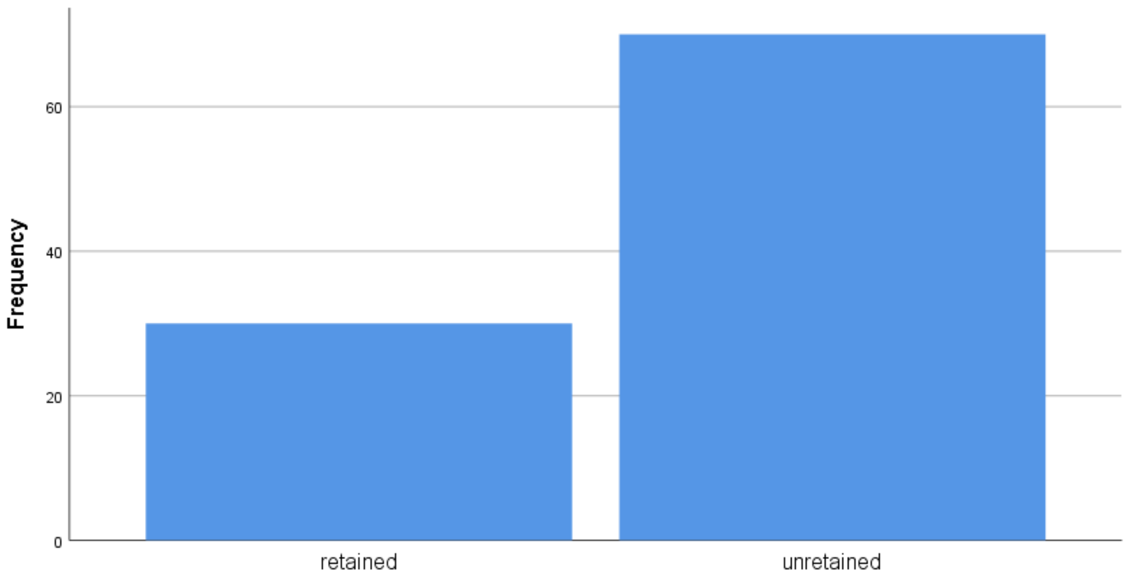


Figure 44: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is poor,Academic engagement is engaged, Academic progress is slow and Attendance is less than 75%.

Figure 5-44: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is 'poor', Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. His or her chance of being retained is 30%. The Figure 5-44 is a visual representation of the Table 5-44 above.

Table 5-45: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	15	15.0	15.0	15.0
	unretained	85	85.0	85.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-45 above. Results in Table 5-45 show that this student has a 15% (0.15) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 85% (0.85) chance of being dropped out of the university. The Figure 5-45 below is a visual representation of the Table 5-45 above.

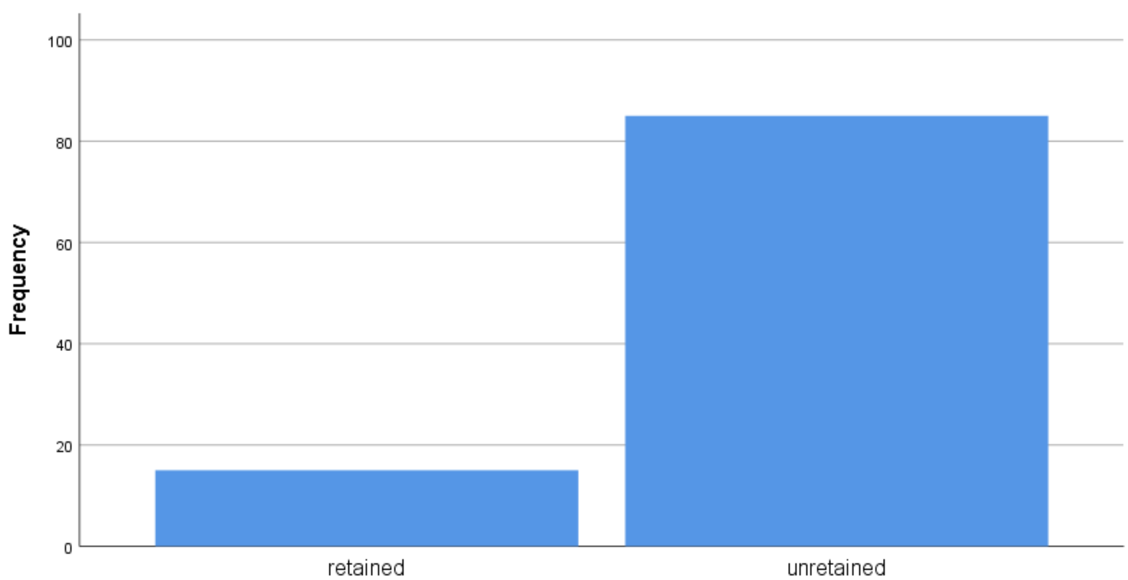


Figure 45: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is poor,Academic engagement is not engaged, Academic progress is normal and Attendance is 75% or above.

Figure 5-45: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is poor, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. His or her chance of being retained is 15%. The Figure 5-45 is a visual representation of the Table 5-45 above.

Table 5-46: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘UTME’, Academic skills is poor, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	10	10.0	10.0	10.0
	unretained	90	90.0	90.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-46 above. Results in Table 5-46 show that this student has a 10% (0.1) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 90% (0.9) chance of being dropped out of the university. The Figure 5-46 below is a visual representation of the Table 5-46 above.

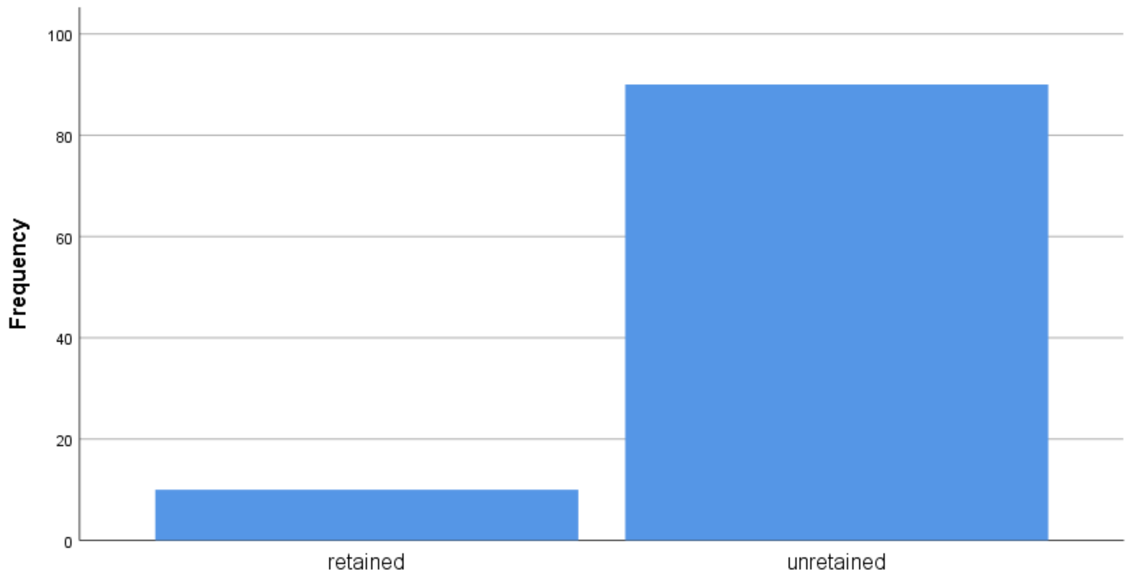


Figure 46: Retention given Change in circumstances is false, Mode of entry is UTME, Academic skills is poor,Academic engagement is not engaged, Academic progress is normal and Attendance is below 75%.

Figure 5-46: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. His or her chance of being retained is 10%. The Figure 5-46 is a visual representation of the Table 5-46 above.

Table 5-47: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-47 above. Results

in Table 5-47 show that this student has a 5% (0.05) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-47 below is a visual representation of the Table 5-47 above.

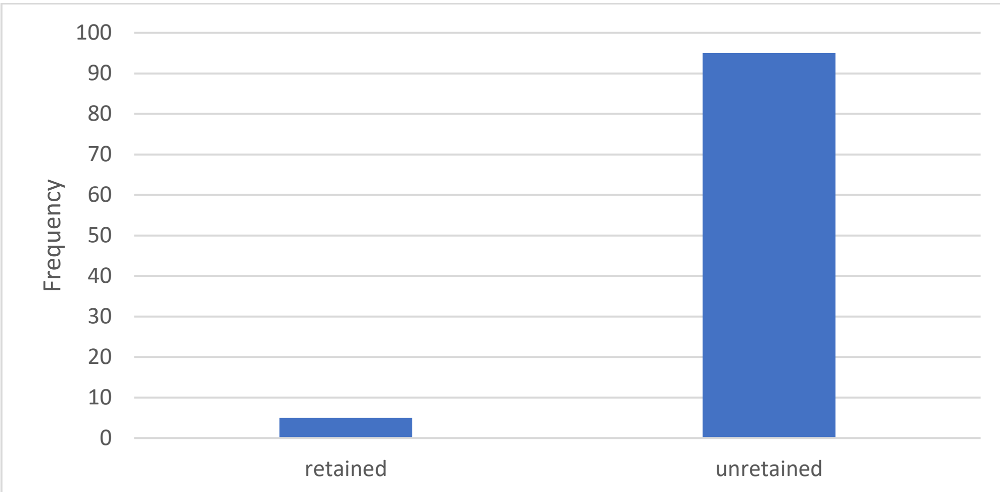


Figure 5-47: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 5%. The Figure 5-47 is a visual representation of the Table 5-47 above.

Table 5-48: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-48 above. Results in Table 5-48 show that this student has a 5% (0.05) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-48 below is a visual representation of the Table 5-48 above.

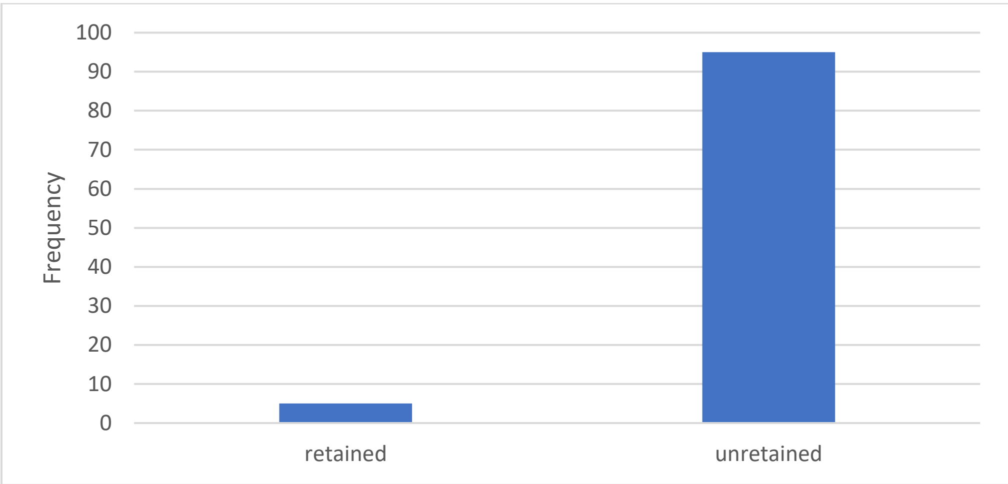


Figure 5-48: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'UTME', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. His or her chance of being retained is 5%. The Figure 5-48 is a visual representation of the Table 5-48 above.

Table 5-49: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	95	95.0	95.0	95.0
	unretained	5	5.0	5.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-49 above. Results in Table 5-49 show that this student has a 95% (0.95) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 5% (0.05) chance of being dropped out of the university. The Figure 5-49 below is a visual representation of the Table 5-49 above.

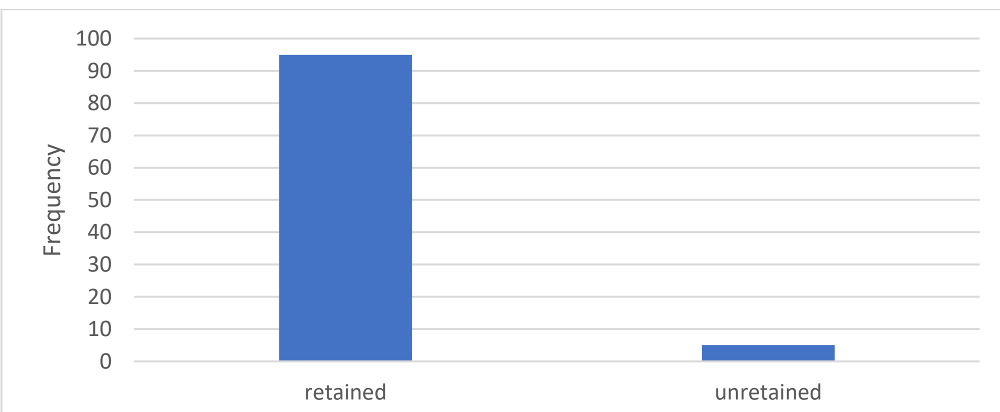


Figure 5-49: Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above. His or her chance of being retained is 95%. The Figure 5-49 is a visual representation of the Table 5-49 above.

Table 5-50: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	90	90.0	90.0	90.0
	unretained	10	10.0	10.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-50 above. Results in Table 5-50 show that this student has a 90% (0.9) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 10% (0.1) chance of being dropped out of the university. The Figure 5-50 below is a visual representation of the Table 5-50 above.

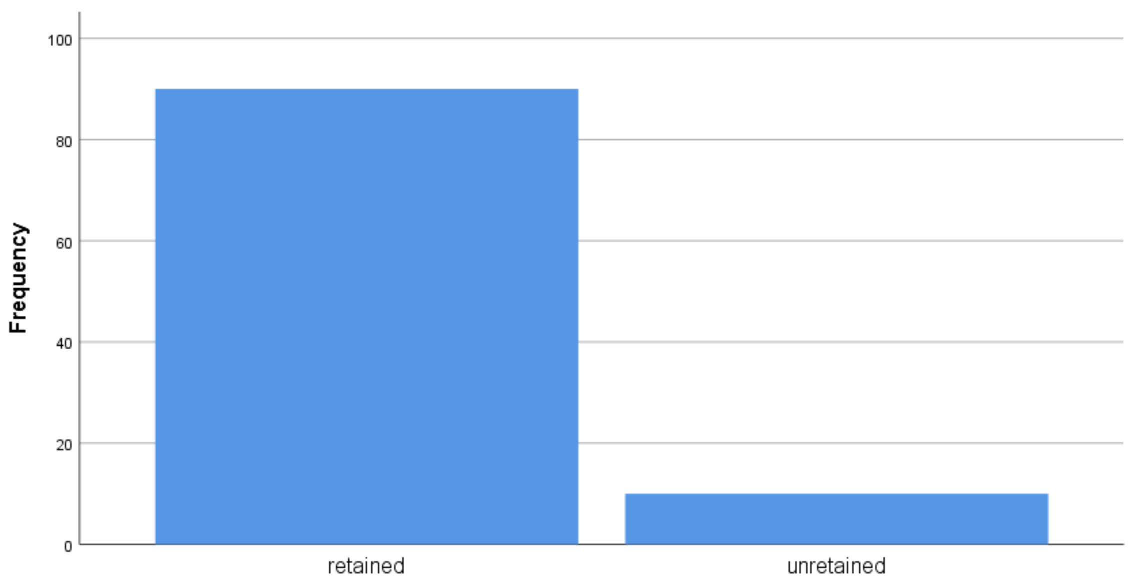


Figure 50: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is good,Academic engagement is engaged, Academic progress is normal and Attendance is below 75%.

Figure 5-50: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘DE’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘normal’ and Attendance is below 75%. His or her chance of being retained is 90%. The Figure 5-50 is a visual representation of the Table 5-50 above.

Table 5-51: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘DE’, Academic skills is good, Academic engagement is ‘engaged’, Academic progress is ‘slow’ and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retained	45	45.0	45.0	45.0
	unretained	55	55.0	55.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-51 above. Results in Table 5-51 show that this student has a 45% (0.45) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 55% (0.55) chance of being dropped out of the university. The Figure 5-51 below is a visual representation of the Table 5-51 above.

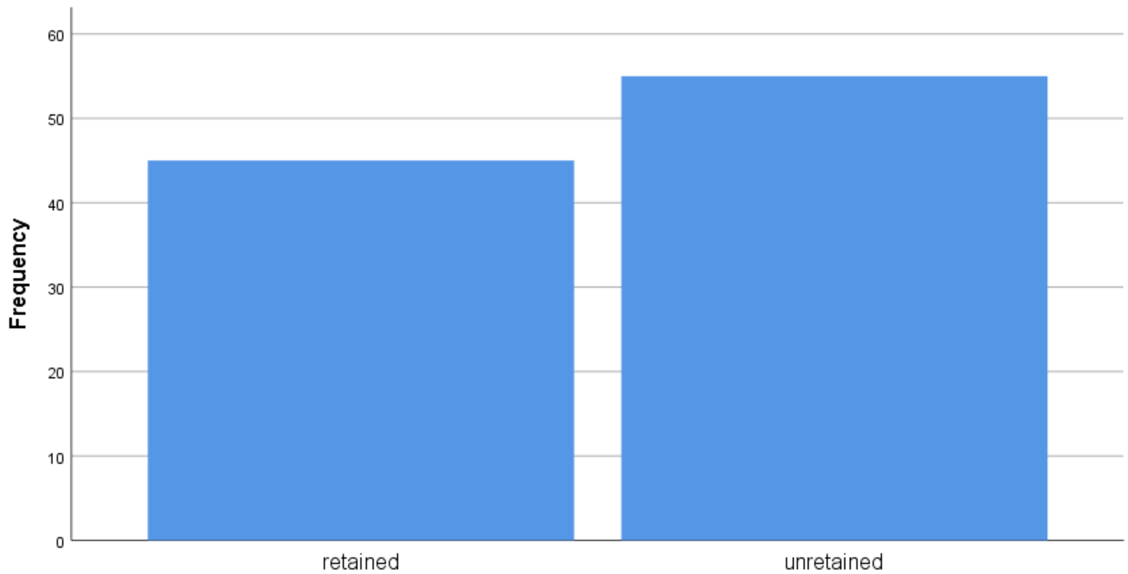


Figure 51: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is good,Academic engagement is engaged, Academic progress is slow and Attendance is 75% or above.

Figure 5-51: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75%. or above. His or her chance of being retained is 45%. The Figure 5-51 is a visual representation of the Table 5-51 above.

Table 5-52: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	40	40.0	40.0	40.0
	unretained	60	60.0	60.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-52 above. Results

in Table 5-52 show that this student has a 40% (0.4) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 60% (0.56) chance of being dropped out of the university. The Figure 5-52 below is a visual representation of the Table 5-52 above.

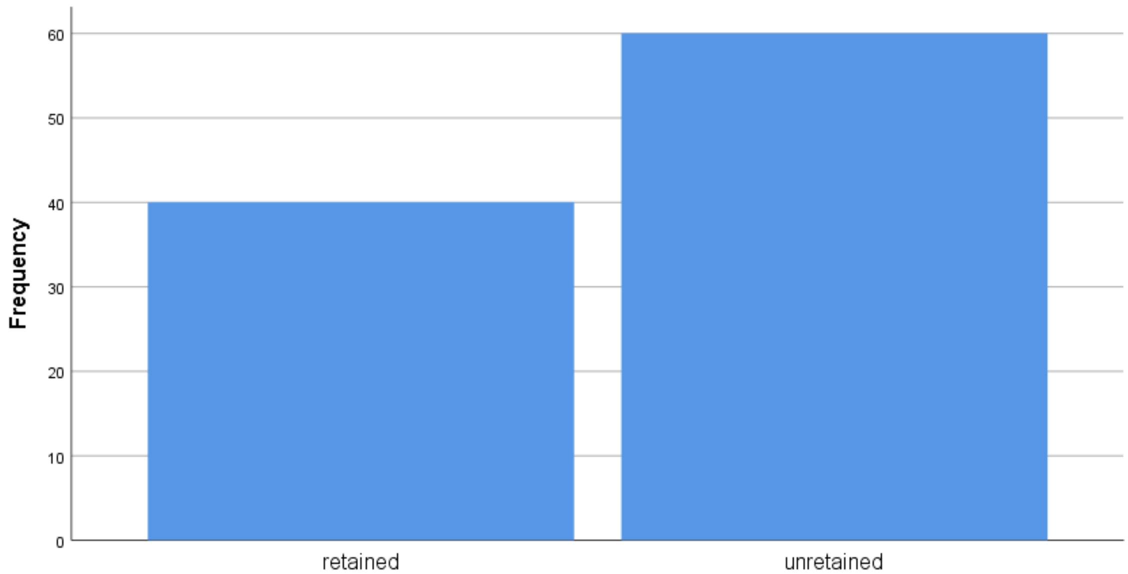


Figure 52: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is good,Academic engagement is engaged, Academic progress is slow and Attendance is below 75%.

Figure 5-52: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is below 75%. His or her chance of being retained is 40%. The Figure 5-52 is a visual representation of the Table 5-52 above.

Table 5-53: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	25	25.0	25.0	25.0
	unretained	75	75.0	75.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-3 above. Results in Table 5-53 show that this student has a 25% (0.25) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 75% (0.75) chance of being dropped out of the university. The Figure 5-53 below is a visual representation of the Table 5-53 above.

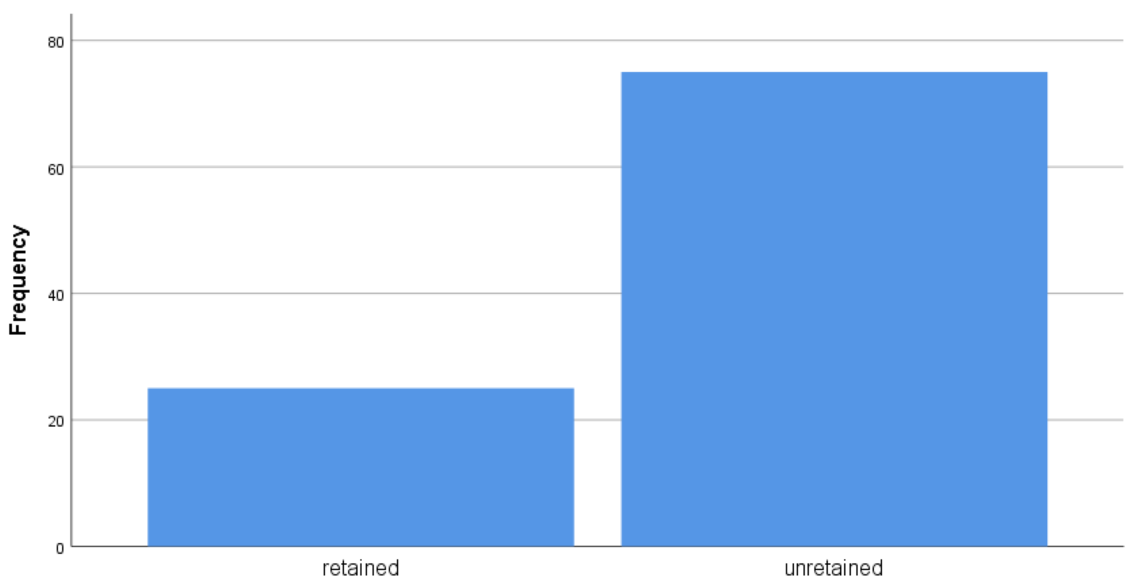


Figure 53: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is good,Academic engagement is not engaged, Academic progress is normal and Attendance is 75% or above.

Figure 5-53: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above. His or her chance of being retained is 25%. The Figure 5-53 is a visual representation of the Table 5-53 above.

Table 5-54: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	20	20.0	20.0	20.0
	unretained	80	80.0	80.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-54 above. Results in Table 5-54 show that this student has a 20% (0.2) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 80% (0.8) chance of being dropped out of the university. The Figure 5-54 below is a visual representation of the Table 5-54 above.

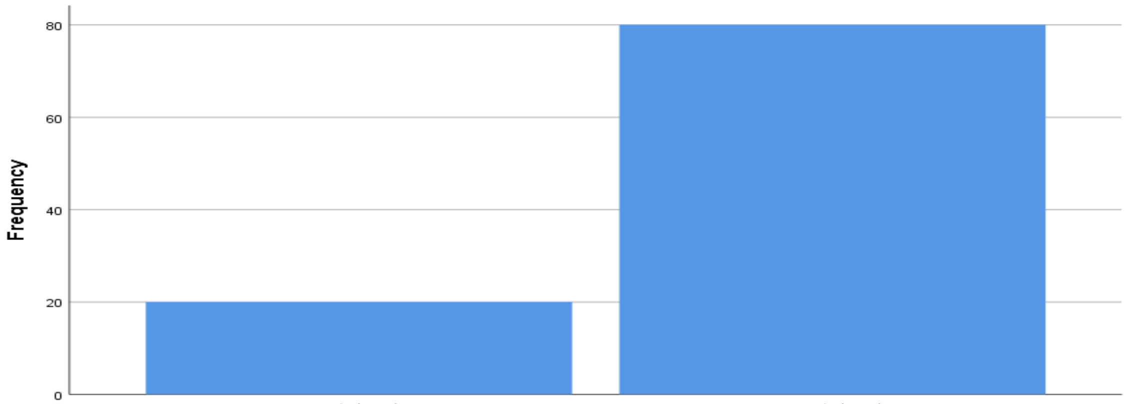


Figure 54: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is good,Academic engagement is not engaged, Academic progress is normal and Attendance is less than 75%.

Figure 5-54: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. His or her chance of being retained is 20%. The Figure 5-54 is a visual representation of the Table 5-54 above.

Table 5-55: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-55 above. Results in Table 5-55 show that this student has a 5% (0.05) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-55 below is a visual representation of the Table 5-55 above.

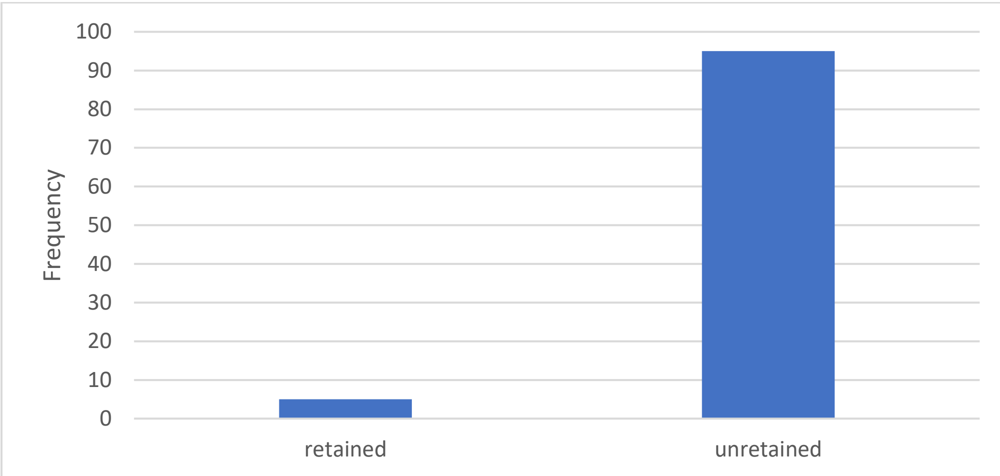


Figure 5-55: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘DE’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is 75% or above. His or her chance of being retained is 5%. The Figure 5-55 is a visual representation of the Table 5-55 above.

Table 5-56: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘DE’, Academic skills is good, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-56 above. Results

in Table 5-56 show that this student has a 5% (0.05) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-56 below is a visual representation of the Table 5-56 above.

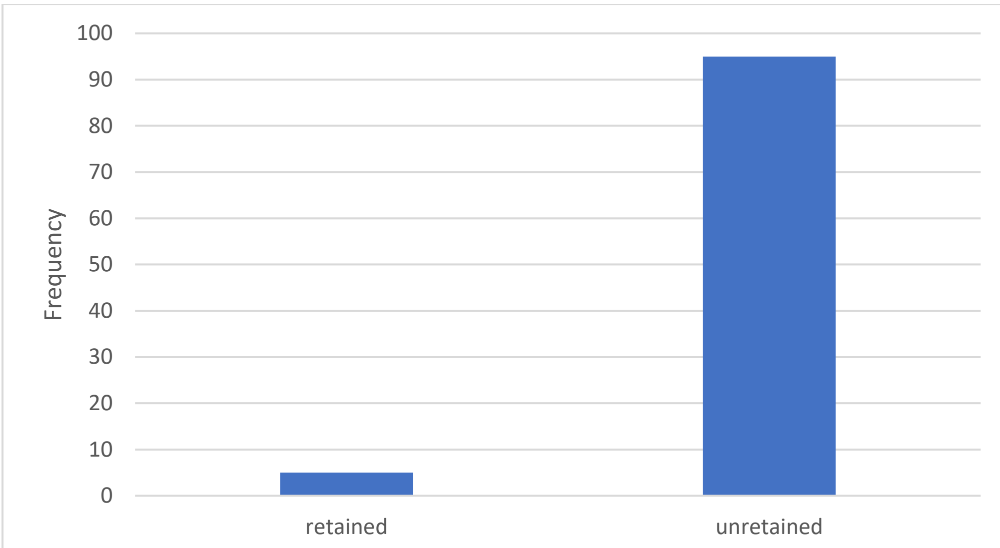


Figure 5-56: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%. His or her chance of being retained is 5%. The Figure 5-56 is a visual representation of the Table 5-56 above.

Table 5-57: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is 75% or above.

Valid	retained	75	75.0	75.0	75.0
	unretained	25	25.0	25.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-57 above. Results in Table 5-57 show that this student has a 75% (0.75) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 25% (0.25) chance of being dropped out of the university. The Figure 5-57 below is a visual representation of the Table 5-57 above.

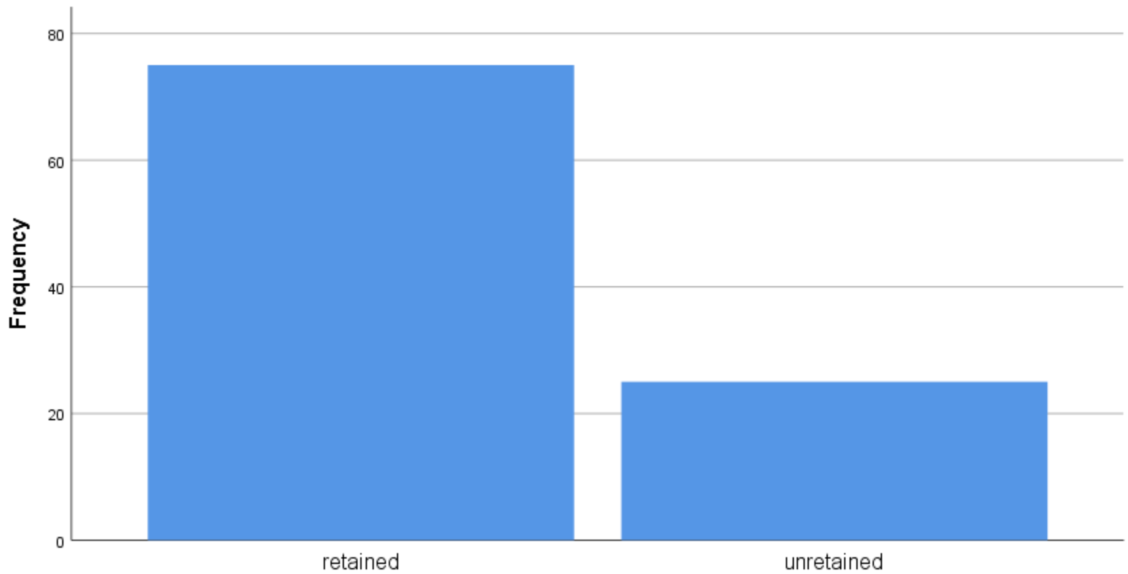


Figure 57: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is poor,Academic engagement is engaged, Academic progress is normal and Attendance is 75% above.

Figure 5-57: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is good, Academic engagement is 'engaged', Academic progress is 'normal and Attendance is 75% or above. His or her chance of being retained is 75%. The Figure 5-57 is a visual representation of the Table 5-57 above.

Table 5-58: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	65	65.0	65.0	65.0
	unretained	35	35.0	35.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-58 above. Results in Table 5-58 show that this student has a 65% (0.65) chance of being retained therefore any student in this category is not at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 35% (0.35) chance of being dropped out of the university. The Figure 5-58 below is a visual representation of the Table 5-58 above.

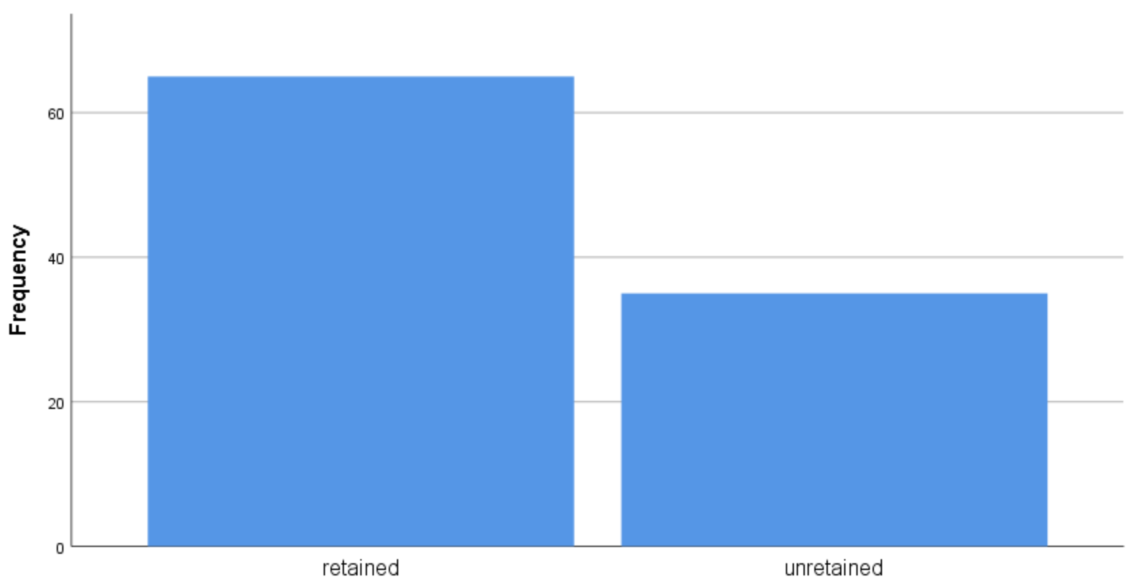


Figure 58: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is poor,Academic engagement is engaged, Academic progress is normal and Attendance is less than 75%.

Figure 5-58: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'normal' and Attendance is less than 75%. His or her chance of being retained is 65%. The Figure 5-58 is a visual representation of the Table 5-58 above.

Table 5-59: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	35	35.0	35.0	35.0
	unretained	65	65.0	65.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-59 above. Results in Table 5-59 show that this student has a 35% (0.35) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 65% (0.65) chance of being dropped out of the university. The Figure 5-59 below is a visual representation of the Table 5-59 above.

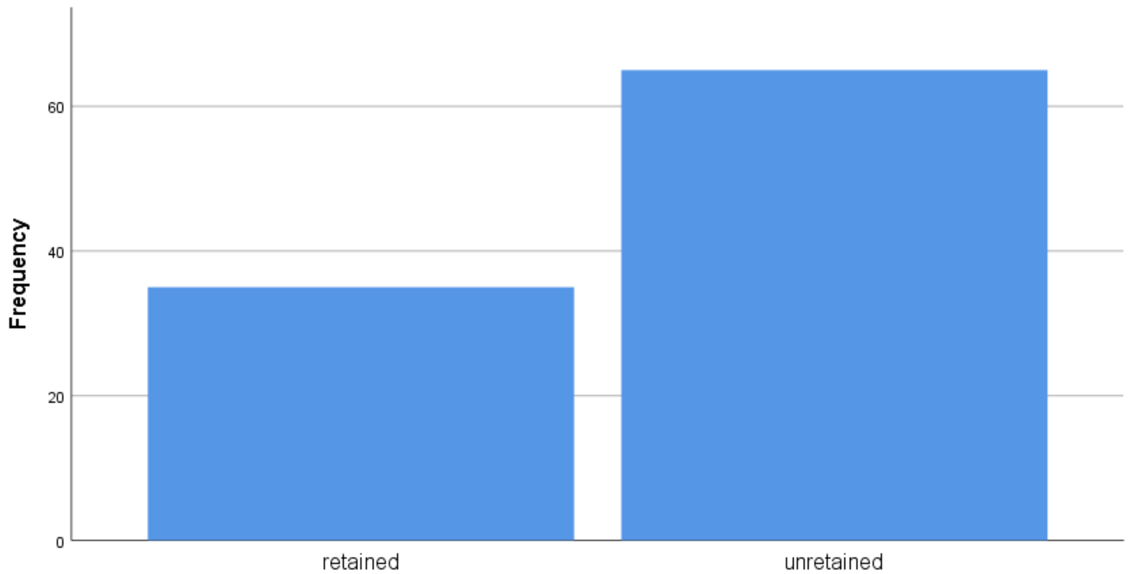


Figure 59: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is poor,Academic engagement is engaged, Academic progress is slow and Attendance is 75% or above.

Figure 5-59: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 35%. The Figure 5-59 is a visual representation of the Table 5-59 above.

Table 5-60: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retained	30	30.0	30.0	30.0
	unretained	70	70.0	70.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-60 above. Results

in Table 5-60 show that this student has a 30% (0.3) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 70% (0.7) chance of being dropped out of the university. The Figure 5-60 below is a visual representation of the Table 5-60 above.

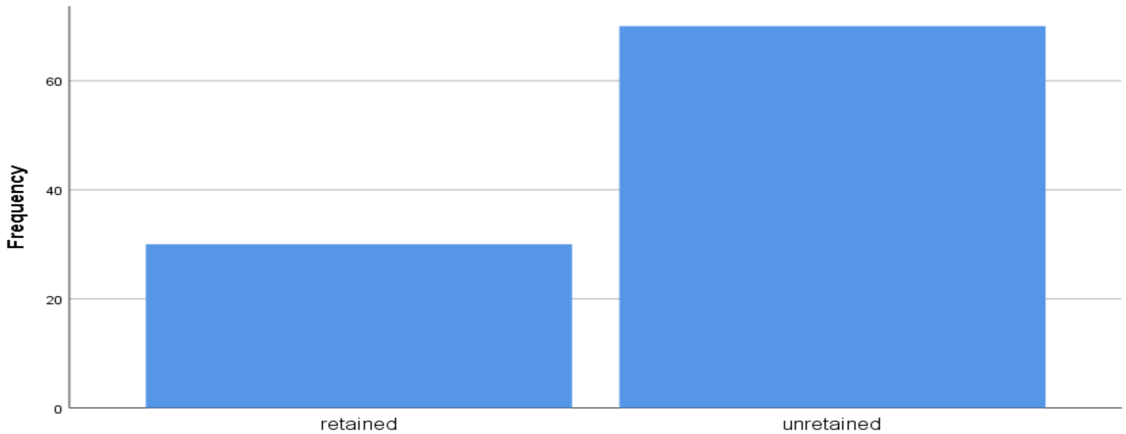


Figure 60: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is poor,Academic engagement is engaged, Academic progress is slow and Attendance is less than 75%.

Figure 5-60: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'engaged', Academic progress is 'slow' and Attendance is less than 75%. His or her chance of being retained is 30%. The Figure 5-60 is a visual representation of the Table 5-60 above.

Table 5-61: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Retained	20	20.0	20.0	20.0
	unretained	80	80.0	80.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-61 above. Results in Table 5-61 show that this student has a 20% (0.2) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 80% (0.8) chance of being dropped out of the university. The Figure 5-61 below is a visual representation of the Table 5-61 above.

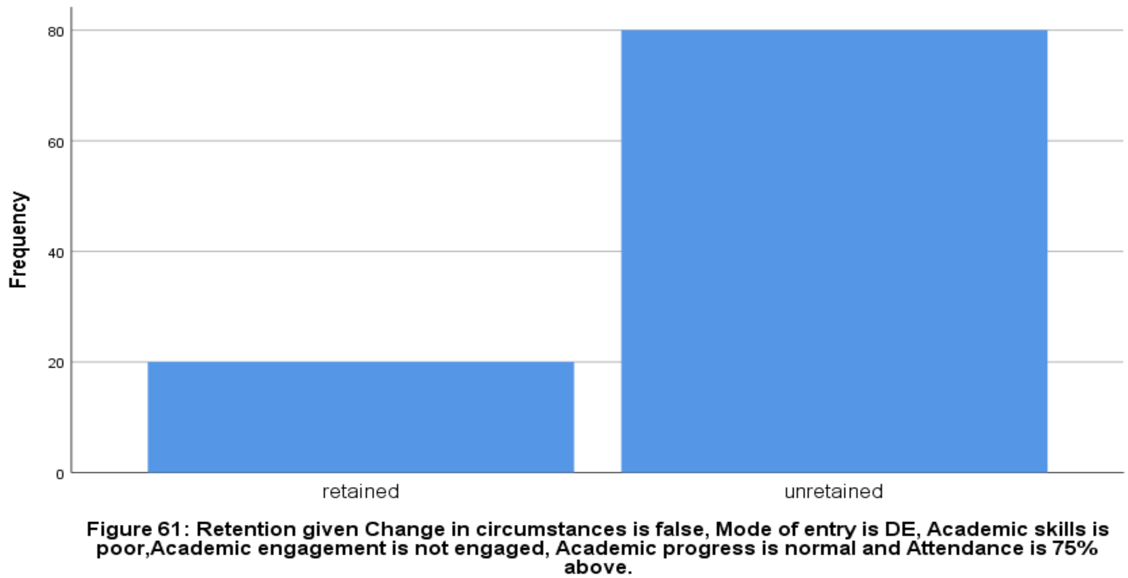


Figure 5-61: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘DE’, Academic skills is poor, Academic engagement is ‘unengaged’, Academic progress is ‘normal’ and Attendance is 75% or above. His or her chance of being retained is 20%. The Figure 5-61 is a visual representation of the Table 5-61 above.

Table 5-62: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	15	15.0	15.0	15.0
	unretained	85	85.0	85.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-62 above. Results in Table 5-62 show that this student has a 15% (0.15) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a 85% (0.85) chance of being dropped out of the university. The Figure 5-62 below is a visual representation of the Table 5-62 above.

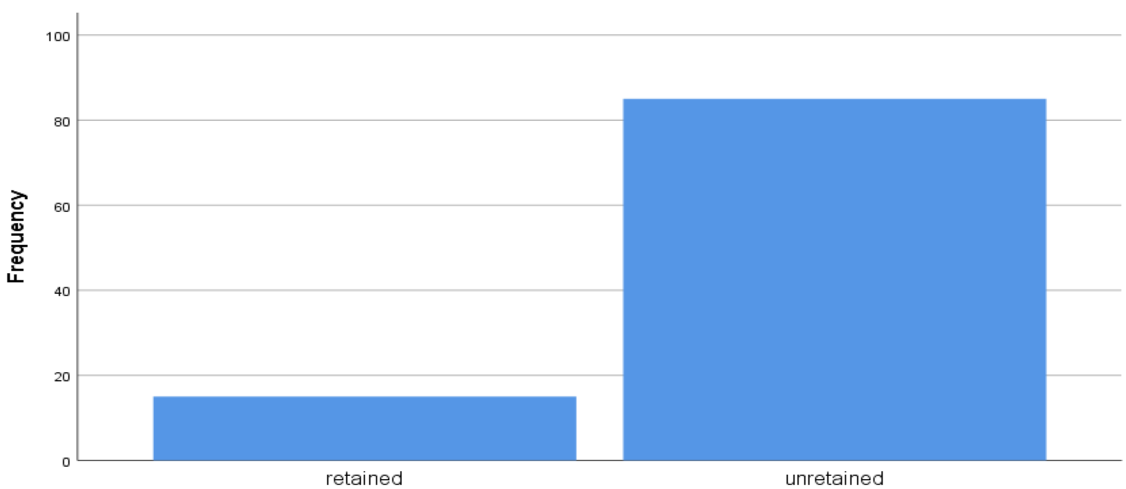


Figure 62: Retention given Change in circumstances is false, Mode of entry is DE, Academic skills is poor,Academic engagement is not engaged, Academic progress is normal and Attendance is below 75%.

Figure 5-62: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'normal' and Attendance is below 75%. His or her chance of being retained is 15%. The Figure 5-62 is a visual representation of the Table 5-62 above.

Table 5-63: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-63 above. Results in Table 5-63 show that this student has a 5% (0.05) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-63 below is a visual representation of the Table 5-63 above.

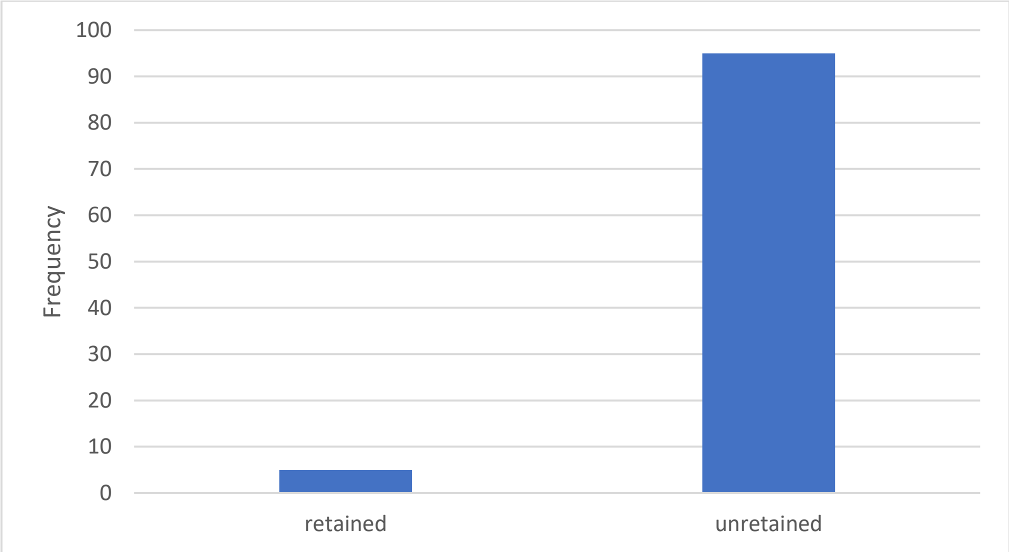


Figure 5-63: Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is 75% or above. His or her chance of being retained is 5%. The Figure 5-63 is a visual representation of the Table 5-63 above.

Table 5-64: Frequency table for the chance of a student being retained given: Change in circumstances is false, Mode of entry is through 'DE', Academic skills is poor, Academic engagement is 'unengaged', Academic progress is 'slow' and Attendance is below 75%.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	retained	5	5.0	5.0	5.0
	unretained	95	95.0	95.0	100.0
	Total	100	100.0	100.0	

The frequency table for the estimates of probabilities for the above conditional probability as elicited from the domain experts in this study is displayed in Table 5-64 above. Results in Table 5-64 show that this student has a 5% (0.05) chance of being retained therefore any student in this category is at risk of being dropped out of university. Since total probability is 100% (1.0), therefore any student in this category has a has a 95% (0.95) chance of being dropped out of the university. The Figure 5-64 below is a visual representation of the Table 5-64 above.

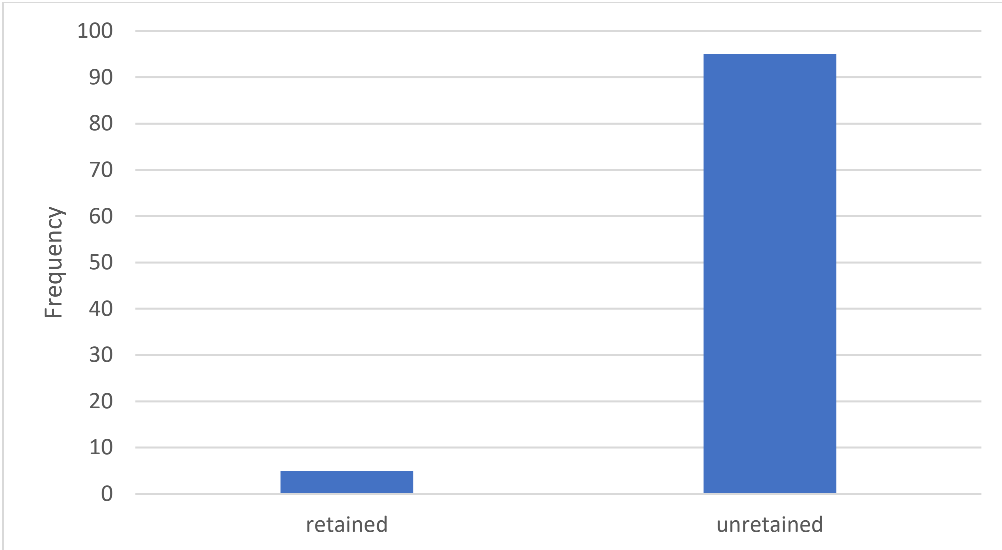


Figure 5-64: A Bar chart representing the chance of a student being retained given: Change in circumstances is false, Mode of entry is through ‘DE’, Academic skills is poor, Academic engagement is ‘unengaged’, Academic progress is ‘slow’ and Attendance is below 75%. His or her chance of being retained is 5%. The Figure 5-64 is a visual representation of the Table 5-64 above.