A TEACHING AND LEARNING CASE STUDY ON DATA MINING USING ASSOCIATION RULES FOR SMES

M. Willetts, A.S. Atkins, C. Stanier

Staffordshire University (UNITED KINGDOM)

Abstract

Big Data Analytics is widely adopted by large companies but to a lesser extent by small to medium-sized enterprises (SMEs). SMEs comprise 99% of all businesses in the UK (6 million), employ 61% of the country’s workforce and generate over half of the turnover of the UK’s private sector (£2.2 trillion). SMEs represent 99% of all businesses in Europe and 90% of all businesses worldwide. Therefore, assisting them to gain competitive advantage by the adoption of technology is important. One of the key barriers to adoption is the shortage of case studies. This paper documents the process in which a positioning tool has been developed to help SMEs analyse their readiness to adopt Big Data Analytics. The positioning tool has been applied to a medium-sized logistics company who are currently analysing Big Data captured through the telematic sensors on their fleet of trucks. The case study proposes how the company can enhance their analytics capability further by undertaking data mining through the form of applying association rule mining to gain competitive advantage. This paper outlines how the positioning scoring tool was used in the case study, and how association rule mining was undertaken and the type of rules which may be identified. The development of this case study provides an approach which could be replicated by educators to develop case studies in other sectors such as manufacturing, retail and the service industry.

Keywords: Big Data Analytics, Case Study, SMEs, Competitive Advantage, Data Mining, Association Rule Mining

# INTRODUCTION

Big Data Analytics refers to the variety of software tools and techniques such as data mining and social media analytics which are utilised to extract insights from Big Data sources. Big Data is defined as: *‘an umbrella term used to describe a wide range of technologies that capture, store, transform and analyse complex data sets which can be of a high volume, generated at a high velocity in a variety of formats*’ [1, p. 3034]. Mikalef et al. [2, p. 1] state that awidely used definition of Big Data Analytics is*: ‘a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high velocity capture, discovery and/or analysis’*. Big Data Analytics is widely adopted by large companies but to a lesser extent by SMEs [3]–[5]. One form of Big Data Analytics is data mining.

In previous research, 21 barriers to the adoption of Big Data Analytics were identified and a strategic framework has been developed [1], [6]. The framework has been validated through both quantitative [7] and qualitative studies. A software tool has been developed to assist SMEs in assessing their current level of Big Data Analytics readiness on a scale of 1 being very low to 5 being very high [8]. The framework and the scoring tool are shown in Fig 1. As described in [8], the framework is derived from an extensive literature search which identified 69 barriers from publications. A thematic analysis was performed on the barriers and the refined barriers were grouped into pillars based on theoretical frameworks [1], [6]. The barriers were validated quantitatively through the utilisation of an online questionnaire which was fully completed by 102 SMEs and qualitatively through focus group interviews with 8 SME practitioners. The framework was updated to reflect the feedback received and one of the original pillars was removed due to the results of the Cronbach’s Alpha statistical analysis. The weightings applied to the barriers are based on the number of citations in the literature and feedback from the SME practitioners from the focus group interviews as the richer feedback they provided could not be acquired through a questionnaire. A significant barrier identified from the literature and interviews is the lack of case studies [3], [4] of SMEs adoption of Big Data Analytics, which this paper addresses. The weightings were tested on a position study before being applied to a case study demonstrating how an SME successfully adopted social media analytics [8]. This approach was repeated with the case study outlined in this paper.

 Figure 1 - Calculation Sheet showing an example of the scoring tool and its relationship to the framework

# METHODOLOGY

As documented in [8], the validation of the Big Data Analytics framework was achieved by assessing two case studies of SMEs who have adopted Big Data Analytics. Both case studies show the application of the tool across three stages: the first stage when the business does not utilise data analytics, the second phase when the business is undertaking some analysis of Big Data using Business Intelligence approaches but is not fully utilising Big Data Analytics Data and a proposed third stage when Big Data Analytics could be utilised. Two SMEs of different sizes working in two different sectors were approached to provide case studies. The first case study [8] presents a small recruitment company, who utilise social media analytics and are actively attempting to improve their usage through the form of sentiment analysis. The second case study, outlined in this paper, presents a medium-sized logistics company, Company A who analyse large volumes of telematics sensor Big Data. Due to commercial confidentiality of the very sensitive information shared by the participants, both of the case studies have been anonymised including the software utilised.

The stages performed to document the case studies were as follows:

1. A standard presentation was presented to the representative of the SME assisting with the case study. The presentation documented the framework and the tool to show the SME how to apply it to their business.
2. A workshop was held with the SME representative. They provided information about each phase of their business’ analytics maturity journey including background information about the company, the data captured and the software utilised.
3. The SME scored themselves using the tool from the guidance provided.
4. The output of the session was documented in the relevant stages of the case study.

## Selection of Data and Software

To demonstrate how association rule mining could be adopted by Company A, some examples were created using the Weka software tool and both the Apriori and Predictive Apriori algorithms were utilised as they evaluate the rules based on different metrics. The first challenge was to obtain a suitable dataset. Company A were not willing to share a dataset due to commercial confidentiality, however they did describe the data they were able to retrieve from their system and gave a demonstration of their telematics software. An open source dataset was found published by Scania [9] for Air Pressure System failure, however this dataset was limited and it was not clear how to apply this in relation to Company A. A more comprehensive dataset was found on Kaggle [10] showing delivery data for a logistics company in India.

This dataset has similarities to Company A’s data, therefore it was decided to adopt this format and create a similar dataset with input from Company A. Data creation packages such as Mockaroo [11] were evaluated to automate the generation of example datasets, however for the purposes of this study it was decided to manually create the datasets using a spreadsheet algorithm. With support and input from Company A, an example dataset was created as a CSV file in Excel for import into Weka.

# RESULTS

## Pre-Data Analytics Stage

Company A is a medium-sized logistics business which was founded in the early 1980s and operates from three offices in the UK. The business predominantly transports goods across England, Scotland and Wales. However, trailers are also sent to an EU country but not the tractor units (the vehicle which pulls the trailers). The business transports goods for clients of all sizes, including large high street retailers and construction firms.

The business employed over 100 staff distributed across five departments including Transport. There are teams within the Transport department dedicated to clients in specific markets including steel and pallets. The business utilises a fleet of 50 leased trucks. The truck leases last three years and at the end of the three-year period the vehicles are replaced with new trucks and a new lease commences. Twenty trucks are utilised for lighter loads while the other 30 haul heavier loads. The business owns around 400 trailers and over 100 of these may be in the EU country at any one time. The business initially operated in a 1980s operational style, heavily relying on paper records and manually updating a magnetic board to show the location of the drivers. The business was very reluctant to change unless senior management was confident that a return on investment would be achieved. Therefore, a major change such as the adoption of Big Data Analytics would have been very difficult to implement without the senior management’s support, despite the business having a dedicated IT Director at this stage.

The business previously had a dedicated IT department, managed by the IT director. The IT department provided IT support and maintained the applications utilised by the business. The IT Director developed a database system to record all of the orders the business received, which recorded fields including the customer details, the pickup location, delivery location and the value of the order. Data relating to the vehicles were stored on ad-hoc spreadsheets and Microsoft Word documents. However, the majority of financial records were stored in paper format and retained in large filing cabinets. A large magnetic board was utilised to map the location of the 50 vehicles, with each driver represented by their own magnet. This required to be manually updated throughout the day. The business did not utilise any dedicated analytics software other than Microsoft Excel which was utilised to produce financial and sales reports, in addition to summarising data from the bespoke database system. The business was compliant with the data protection legislation and aware of the ethical issues of storing and processing data. The information in the scenario was applied to the framework and weightings shown in Fig. 1 and the score for this stage of the case study was calculated as 1.7 which indicates a low level of readiness for the adoption of Big Data Analytics.

Table 1 - Results from Stage 1 of the Scenario

|  |  |
| --- | --- |
| **Overall Score** | **1.7** |
| **Stage** | * Stage 1 – Basic Analytics
 |
| **Software** | * Microsoft Excel
 |
| **Cost** | * £7 per month per user (Office 365) or £249 for an Office Home & Business 2019 licence [12]
 |
| **Expertise** | * Basic spreadsheet skills
* Many written tutorials are available online or videos such as on YouTube
* Training courses are available
 |

## Business Intelligence

Following the appointment of a new senior manager who was experienced in the logistics sector, Company A changed their operating model. Company A identified a number of areas where the business could make changes to reduce costs and increase efficiency. The first major change was that the internal IT staff were replaced by an outsourced IT support provider. The IT infrastructure, including file storage and applications were mostly moved from on premise to a cloud infrastructure. The legacy database system was retained on-site at the headquarters to ensure that staff retained access to historic records.

Company A also decided to review its truck leases to ensure that the vehicles leased provided value for money. The new trucks were fitted with telematic sensors. Although cost was the primary reason for selecting the chosen manufacturer, the inclusion of sensors and a cloud-based telematics software package for analysing the sensor data, referred to here as Package A, was an additional factor in the decision. Telematic sensors generate very large volumes of Big Data, which Company A would not have been able to analyse with their existing applications prior to the adoption of Package A. An example of this type of telematics software is Geotab which collects data from 2 million devices, resulting in 40 billion data points per day (2020). Therefore, as Company A increases its fleet and the number of sensors, the volumes of Big Data collected will grow at a very high velocity.

Package A is a software as a service application (SaaS), which is web browser based and provides its users with a wide range of features including:

* Positioning: allows users to track the location of all vehicles in the fleet. Real-time information is captured from the telematics sensors on the trucks and is displayed on a map within the application.
* Driver activity: Package A allows tachograph activity to be downloaded remotely. Tachographs are a legal requirement for all vehicles if the vehicle being driven is covered under European Union or European Agreement Concerning the Work of Crews of Vehicles Engaged in International Road Transport (AETR) rules [13]. Tachographs record information regarding driving time, speed and distance for the purpose of ensuring that drivers and their employers follow the rules on drivers’ hours.
* Driver performance: Package A produces a variety of reports relating to drivers’ performance including miles per gallon, braking, standstill times, average speed, engine and gear usage, fuel efficiency. Company A heavily utilises the daily reports on drivers, as they are used for performance management, for example if a driver has been consistently breaking the speed limit.
* Messaging: the application allows SMS (short message service) text messages to be sent to the drivers’ cab.
* Fuel and the Environment: the application produces reports regarding fuel consumption, CO2 and nitrous oxide emission. Package A allows carbon footprint reports to be produced per driver.

Applying the framework factors and weightings to the information in the Business Intelligence stage resulted in a score of 3.2 as shown in Table 2. Although this indicates an awareness of the benefits of Big Data Analytics, Company A are limited in the type of analysis which they can conduct.

Table 2 - Results from Stage 2 of the Scenario

|  |  |
| --- | --- |
| **Overall Score** | **3.2** |
| **Stage** | * Stage 2 – Business Intelligence
 |
| **Software** | * Telematics software with inbuilt analytics such as Geotab [14], MAN TeleMatics [15] or Microlise [16]
 |
| **Cost** | * Either available as an extra with the leased trucks or available to buy the sensors and software separately
 |
| **Expertise** | * Support available from the vendors
* Simple analytics which do not require training to interrogate
 |

## Big Data Analytics

Package A includes a variety of inbuilt standard reports at both driver and vehicle level. Data can be drilled down into, for example from the map displaying the location of the fleet, a vehicle can be selected and the user can drill down into the history of where the vehicle has been. Automated reports can be scheduled and sent to specific users by email and alerts can be setup. The reports produced by Package A would be classified as Business Intelligence, as they are performing descriptive analytics to show summaries of historical data.

The business does not currently have sensors on any of the 400 trailers, therefore Company A is not able to accurately track the position of trailers. The business also has a transport information system, Package B, which records details relating to orders, staff and finance. Trailer details are recorded as part of the job but there is currently no central repository containing information on all of the trailer details. Package B has the ability to integrate with Package A and other systems, including capturing data from sensors on trailers. Company A are regarded as being currently at Stage 2, the Business Intelligence stage of Big Data Analytics adoption.

Although Company A is capturing large volumes of Big Data from the telematics sensors on their truck tractors but none on their trailers, they are limited to performing descriptive analytics on historical data because of the limitations of Package A. Company A could adopt Big Data Analytics to achieve competitive advantage through the form of data mining. A widely used approach in data mining is Association Rule mining. Association Rule Mining is utilised to identify associations between items or itemsets, which has extended to Big Data [17]. Two examples of Association Rule Mining algorithms which could be applied to Company A are Apriori and Predictive Apriori. Apriori is a seminal algorithm for mining frequent itemsets for Boolean Association and utilises an iterative approach for finding rules, known as a level-wise search [18]. Witten et al. [19, p. 216] state: ‘*Apriori follows a generate-and test methodology for finding frequent item sets, generating successively longer candidate item sets from shorter ones that are known to be frequent. Each different size of candidate item set requires a scan through the dataset to determine whether its frequency exceeds the minimum support threshold’*. García et al [20] state that in the association rule X⇒ Y, in a transaction where X occurs, the probability of Y also occurring is high. X is known as the antecedent and Y is known as the consequent [20]. Association rules are measured by support and confidence as the criteria to identify the most important relationships, with support being the number of transactions which contain both X and Y, and confidence being the number of transactions which contain both X and Y divided by the number of transactions containing X [20], [21]. Witten et al. [19] discuss four metrics for ranking association rules: Confidence, Lift, Leverage and Conviction. Oweis *et al.*[22] state that Lift can be used to identify interesting patterns in the data. Lift > 1 signifies a positive correlation, Lift < 1 indicates that there is a negative correlation. For a rule to be considered useful, its Lift value must be greater than 1 and the larger this is, the stronger the association. Leverage indicates the frequency of X and Y appearing together when they are independently distributed [23]. When Leverage is equal to zero, both X and Y are independent, however the greater the leverage is, the closer the relationship between X and Y. Li *et al.* [23]suggest that conviction is used to measure the independenceof variables and similarly to the lift, the greater the value of confidence, the greater the correlation between the elements. Neelima, Satyanarayana and Murthy [21] summarise the advantages and disadvantages of utilising the Apriori algorithm. The advantages include: it is easy to implement; it utilises a large itemset; is easily parallelised; and it allows a user-defined minimum support threshold. The main disadvantage is that it is not efficient in large databases as it requires a large volume of repeated dataset scans.

Predictive Apriori is a version of Apriori which combines support and confidence into a single measure, predictive accuracy [24]. Predictive Apriori differs to Apriori as rules are ranked by “*expected predictive accuracy*” as it attempts to maximise the expected accuracy of an association rule instead of confidence [25]. However, Akosa [26] states that predictive accuracy can be misleading when imbalanced datasets are used.

Company A could utilise a data mining application to perform data mining on the sensors currently on the truck tractors (the front of the lorry with the cab), including mileage, position, speed, braking and fuel, utilisation, in addition to the data captured by Package A including driver, journey and vehicle information. There are many data mining software packages available, including free and open-source applications. Weka [27] is an open-source application which allows data mining to be performed on smaller datasets, providing a starting point for businesses which want to investigate data mining and the Weka package has been used for illustrative purposes in this paper. Commercial products are available including KNIME [28], IBM SPSS and RapidMiner.

Although the sales of commercial vehicles have fallen, the telematics market is expected to continue to grow, projecting to be a $17.1 billion market [29]. Singh [29, p. 1] states: ’*while driver wages account for about 30% of a fleet’s expenses, fuel costs make up about 26% of fleet expenditure. The big cost buster here is telematics which has the capacity to increase productivity by 10-15% and reduce overtime by another 10-15%. It also enables fleets to save about 20-25% on fuel expenses, on average, by promoting better driving practices. And, it helps fleets to shave 20-30 minutes on daily driving time’*. Therefore, further efficiencies could be achieved by utilising data mining. There have been attempts to future proof telematics solutions, for example the Open Telematics API [30] which attempts to standardise the interfaces between telematic sensors and software. Previously, if a telematics provider ceased trading, their customers would have been forced to purchase new telematics sensors for their replacement software. However, the Open Telematics API [30] will allow the interface between the vehicles to remain the same between solutions, giving further assurance to adopters of telematic sensors.

One example of how association rule mining could help Company A is to identify problematic routes, for example routes or destinations which frequently result in late deliveries. Associations may be identified with routes in city centres or on busy roads which result in frequent stopping and starting. To mitigate this, the routes may be changed or destinations may be avoided to prevent unprofitable jobs being agreed. Similarly, combinations of drivers and routes which are unprofitable may be identified, which can allow the business to intervene.

Another problem which Company A frequently encounter in the winter is battery failure which can be very costly. Despite replacing the trucks every five years, a common problem is that the batteries fail during the winter. Drivers often use appliances such as kettles and heated seats, particularly when they are parked overnight, consuming the battery. Company A have reported that they have noticed that this issue occurs when the battery charge is low and the temperature is below zero degrees Celsius. Association rule mining would be able to identify patterns such as this to allow Company A to intervene. Fig. 2 shows how these rules would appear in Weka. One of the association rules found in the example provided is that when drivers are travelling to Telford with a low battery indicator, it frequently results in battery failure as the batteries have not been given time to charge. A solution to this problem could be to ensure drivers are allocated several long-distance journeys a week to charge the battery, rather than only short journeys which do not allow the battery to charge. Other factors may be identified from association rule mining, such as particular drivers who are utilising unusually high volumes of power.



Figure 2 - Screenshot from Weka showing how associations may be found between temperature, the battery indicator and failure

Company A is not able to track the location of its trailers. Although the trailers allocated to each job can be recorded manually in Package B and its current location, no other information is recorded such as when a trailer was last used and the mileage. One of the biggest issues currently encountered by Company A is Anti-lock Braking System (ABS) failure when trailers which are either old or have not been used for several weeks are selected by a driver for a job. The drivers do not currently have a way of identifying which trailers they should use (trailers which have been recently used and are known to be reliable); when an unsuitable trailer is selected it, it can often lead to ABS failure. Company A have previously contacted telematics suppliers for quotes for sensor solutions for the trailers but deemed the cost to be too expensive. A solution could be to install Radio-Frequency Identification (RFID) tags on the trailers as RFID is a reliable technology, widely used in many industries including mining, healthcare and construction [31]–[33]. The tags are relatively inexpensive with costs being as low a few pence each [33], with heavy duty RFID tags which can be applied to both metal and non-metal surfaces for less than £1 each [34]. The tags could be scanned at the start and end of every journey by the driver using their mobile phone. The data could be transmitted to Package B, which has integrations with many systems and the developers can add bespoke functionality, which Company A has previously purchased. The data from the tags could be mined to find associations with the trailers and drivers, routes and other attributes. Fig. 3 shows the process of how this solution could be implemented. Alternatively, QR codes or barcodes may also be used.



Figure 3 - Example of how low-cost RFID tags could be installed on trailers

It should be noted that the individual conducting the association rule mining will need domain knowledge, in addition to the expertise required to undertake the data mining and understand the techniques being performed. A rule which does not return information of value is described as a ‘trivial rule’ and domain expertise may be required to differentiate which rules are trivial from rules which may provide value to the business. An individual with domain knowledge will be more likely to be able to identify anomalies from relevant patterns in the data which they can analyse further. The information in the scenario was applied to the framework and weightings shown in Fig. 1 and the score for this stage of the case study was calculated as 4.0 as shown in Table 3. This indicates a high level of expertise in Big Data Analytics with scope to expand further.

Table 3 - Results from Stage 3 of the Scenario

|  |  |
| --- | --- |
| **Overall Score** | **4.0** |
| **Stage** | * Stage 3 – Big Data Analytics – Association Rule Mining
 |
| **Software** | * A data mining tool which supports Weka for illustrative purposes
 |
| **Cost** | * Free tools are available such as Weka which could be used as a teaching or learning tool, as illustrated in this section
* Paid software is available such as RapidMiner
* Most providers reviewed offer free trials or free tiers
 |
| **Expertise** | * Online tutorials and user guides are available
* Some software providers offer online training such as RapidMiner
* Most solutions reviewed offer support
* Consultants available from £500-£1,000 per day
 |

# CONCLUSIONS

This paper has described the process of developing a case study designed for use with the Big Data Analytics framework to support the adoption of Big Data Analytics by SMEs. This scenario outlined a case study of how Big Data Analytics can be adopted and illustrated this through the use of Association Rule Mining. The development of the case study followed the same process as a previous case study on an SME’s adoption of social media analytics [8]. The process of developing the case study included the development of a dataset and the selection of appropriate association rule mining techniques, with examples of rules which may be found. The case study documented three stages of Company A’s data analytics maturity journey: pre-Business Intelligence, Business Intelligence and Big Data Analytics. The case study outlined in the paper will be useful as a teaching resource. Future publications will document the application of the scoring tool to further case studies of SMEs in different sectors such as healthcare and manufacturing.

REFERENCES

[1] M. Willetts, A. S. Atkins, and C. Stanier, “A Strategic Big Data Analytics Framework To Provide Opportunities for SMEs,” in *INTED2020 Proceedings*, 2020, vol. 1, pp. 3033–3042.

[2] P. Mikalef, M. Boura, G. Lekakos, and J. Krogstie, “Big data analytics and firm performance: Findings from a mixed-method approach,” *J. Bus. Res.*, vol. 98, pp. 261–276, 2019.

[3] S. Coleman, R. Göb, G. Manco, A. Pievatolo, X. Tort-Martorell, and M. S. Reis, “How Can SMEs Benefit from Big Data? Challenges and a Path Forward,” *Qual. Reliab. Eng. Int.*, vol. 32, no. 6, pp. 2151–2164, Oct. 2016.

[4] M. Iqbal, S. H. A. Kazmi, A. Manzoor, A. R. Soomrani, S. H. Butt, and K. A. Shaikh, “A study of big data for business growth in SMEs: Opportunities & challenges,” in *2018 International Conference on Computing, Mathematics and Engineering Technologies: Invent, Innovate and Integrate for Socioeconomic Development, iCoMET 2018 - Proceedings*, 2018, vol. 2018-Janua, pp. 1–7.

[5] M. Bianchini and V. Michalkova, “OECD SME and Entrepreneurship Papers No. 15 Data Analytics in SMEs: Trends and Policies,” 2019.

[6] M. Willetts, A. S. Atkins, and C. Stanier, “Barriers to SMEs Adoption of Big Data Analytics for Competitive Advantage,” in *4th International Conference on Intelligent Computing in Data Sciences, ICDS 2020*, 2020, pp. 1–8.

[7] M. Willetts, A. S. Atkins, and C. Stanier, “Quantitative Study on Barriers of Adopting Big Data Analytics for UK and Eire SMEs,” in *Data Management, Analytics and Innovation*, 2022, pp. 349–373.

[8] M. Willetts, A. S. Atkins, and C. Stanier, “Teaching and Learning Case Study on Social Media Analytics for Small And Medium-Sized Enterprises,” in *ICERI2021 Proceedings*, 2021, pp. 3158–3167.

[9] Scania, “UCI Machine Learning Repository: APS Failure at Scania Trucks Data Set,” 2017. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/APS+Failure+at+Scania+Trucks. [Accessed: 19-Sep-2021].

[10] Kaggle, “Delivery truck trips data | Kaggle,” 2021. [Online]. Available: https://www.kaggle.com/ramakrishnanthiyagu/delivery-truck-trips-data. [Accessed: 19-Sep-2021].

[11] Mockaroo, “Mockaroo - Random Data Generator and API Mocking Tool | JSON / CSV / SQL / Excel,” 2021. [Online]. Available: https://www.mockaroo.com/. [Accessed: 19-Sep-2021].

[12] Microsoft, “Microsoft Office 2019 for Business | Microsoft 365,” 2021. [Online]. Available: https://www.microsoft.com/en-GB/microsoft-365/get-started-with-office-2019. [Accessed: 19-Jul-2021].

[13] GOV.UK, “Tachographs: rules for drivers and operators,” 2020. [Online]. Available: https://www.gov.uk/tachographs. [Accessed: 21-Dec-2020].

[14] Geotab, “Fleet Tracking and Telematics Solutions | Geotab UK,” 2021. [Online]. Available: https://www.geotab.com/uk/. [Accessed: 23-Aug-2021].

[15] MAN, “MAN TeleMatics | MAN DE,” 2021. [Online]. Available: https://www.man.eu/de/en/service/man-telematics-37198.html. [Accessed: 23-Aug-2021].

[16] Microlise, “Our IIoT & OEM Offer | IIOT,” 2021. [Online]. Available: https://www.microlise.com/iiot/. [Accessed: 23-Aug-2021].

[17] S. Bagui and P. C. Dhar, “Positive and negative association rule mining in Hadoop’s MapReduce environment,” *J. Big Data*, vol. 6, no. 1, 2019.

[18] J. Han, M. Kamber, and J. Pei, *Data Mining : Concepts and Techniques : Concepts and Techniques (3rd Edition)*, Third edit. 2012.

[19] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, Third Edit. 2011.

[20] E. García, C. Romero, S. Ventura, and T. Calders, “Drawbacks and solutions of applying association rule mining in learning management systems,” in *CEUR Workshop Proceedings*, 2007, vol. 305, pp. 13–22.

[21] S. Neelima, N. Satyanarayana, and P. K. Murthy, “Enhanced Apriori Algorithm Using Top down Approach,” vol. 5, no. 6, 2015.

[22] N. E. Oweis, M. M. Fouad, S. R. Oweis, S. S. Owais, and V. Snasel, “A Novel Mapreduce Lift Association Rule Mining Algorithm (MRLAR) for Big Data and Maritime Transport Cairo, Egypt,” *IJACSA) Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 3, 2016.

[23] Z. Li, X. Li, R. Tang, and L. Zhang, “Apriori Algorithm for the Data Mining of Global Cyberspace Security Issues for Human Participatory Based on Association Rules,” *Front. Psychol.*, vol. 0, p. 2680, Feb. 2021.

[24] E. García, C. Romero, S. Ventura, and C. De Castro, “A collaborative educational association rule mining tool,” *Internet High. Educ.*, vol. 14, no. 2, pp. 77–88, Mar. 2011.

[25] M. Sharma, J. Choudhary, and G. Sharma, “Evaluating the performance of apriori and predictive apriori algorithm to find new association rules based on the statistical measures of datasets,” *Int. J. Eng. Res. Technol.*, vol. 1, no. 6, pp. 1–5, 2012.

[26] J. S. Akosa, “Predictive accuracy: A misleading performance measure for highly imbalanced data,” *SAS Glob. Forum*, vol. 942, pp. 1–12, 2017.

[27] The University of Waikato, “Weka 3 - Data Mining with Open Source Machine Learning Software in Java,” *The University of Waikato*. 2016.

[28] KNIME, “KNIME | Open for Innovation,” 2021. [Online]. Available: https://www.knime.com/. [Accessed: 29-Aug-2021].

[29] S. Singh, “Commercial Vehicle Sales Are Ailing, So Why Is The Connected Truck Telematics Market Still In Good Health?,” 03-Jun-2020. [Online]. Available: https://www.forbes.com/sites/sarwantsingh/2020/06/03/commercial-vehicle-sales-are-ailing-so-why-is-the-connected-truck-telematics-market-still-in-good-health/?sh=3d1360c7e126. [Accessed: 23-Jan-2021].

[30] The Open Telematics API, “Open Telematics API · Apiary,” 2019. [Online]. Available: https://opentelematicsapi.docs.apiary.io/#. [Accessed: 24-Aug-2021].

[31] L. Zhang and A. S. Atkins, “A Decision Support Application in Tracking Construction Waste Using Rule-based Reasoning and RFID Technology,” *Int. J. Comput. Intell. Syst.*, vol. 8, no. 1, p. 128, 2015.

[32] A. A. AlRasheed, A. Atkins, and R. Campion, “Using Knowledge Management and Visualisation Concepts to Improve Patients and Hospitals Staff Workflow,” Aug. 2017.

[33] A. Atkins, L. Zhang, and H. Yu, “Application of RFID and Mobile technology in Tracking of Equipment for Maintenance in the Mining Industry,” in *10th Underground Coal Operators’ Conference, University of Wollongong & the Australasian Institute of Mining and Metallurgy*, 2010, pp. 350–358.

[34] Seritag, “Warehouse Tag - Seritag,” 2021. [Online]. Available: https://seritag.com/nfc-tags/warehouse-tag?vat=true&gclid=Cj0KCQjwpreJBhDvARIsAF1\_BU3yK5VStc7Dg9GRDgoL5aKD9CNmFxSExhThFs\_01C1U-o5DixKVaK8aAsFnEALw\_wcB. [Accessed: 31-Aug-2021].