Review Paper

A review of data mining in knowledge management: applications/findings for transportation of small and medium enterprises

Siti Aishah Mohd Selamat[1] · Simant Prakoonwit[1] · Wajid Khan[1]

Received: 15 October 2019 / Accepted: 24 March 2020

Abstract
A core subfield of knowledge management (KM) and data mining (DM) constitutes an integral part of the knowledge discovery in database process. With the explosion of information in the new digital age, research studies in the DM and KM continue to heighten up in the business organisations, especially so, for the small and medium enterprises (SMEs). DM is crucial in supporting the KM application as it processes the data to useful knowledge and KM role next, is to manage these knowledge assets within the organisation systematically. At the comprehensive appraisal of the large enterprise in the transportation sector and the SMEs across various industries—it was gathered that there is limited research case study conducted on the application of DM–KM on the transportation SMEs in specific. From the extensive review of the case studies, it was uncovered that majority of the organisations are not leveraging on the use of tacit knowledge and that the SMEs are adopting a more traditional use of ICTs to its KM approach. In addition, despite DM–KM is being widely implemented—the case studies analysis reveals that there is a limitation in the presence of an integrated DM–KM assessment to evaluate the outcome of the DM–KM application. This paper concludes that there is a critical need for a novel DM–KM assessment plan template to evaluate and ensure that the knowledge created and implemented are usable and relevant, specifically for the SMEs in the transportation sector. Therefore, this research paper aims to carry out an in-depth review of data mining in knowledge management for SMEs in the transportation industry.

Keywords Knowledge management · Data mining · SMEs · Transportation · Data mining models · Knowledge

1 Introduction

In an information era, knowledge is deemed as the lifeblood of an organisation and that its survival is relatively dependent on it. Arising as a crucial business element, the utilisation of knowledge aids organisation to remain competitive in today's volatile business environment—notably in the services industries [4]. The differentiating competitive factors of an organisation are in its intangible data assets. Therefore, managing knowledge is as critical in contrast to administering any other assets of the organisation. Knowledge management (KM) implementation creates effective and efficient pathways for organisation to employ its intellectual assets [2]. Several KM approaches have been advocated in depositing information at work for organisations [27]. These approaches include instruments and methods that require an eminent level of resources in order to deliver the KM goals. The vast amount of data have been collected by many organisations. The organisational challenge is discovering valuable information in the large pool of data collected and translating these data into useful and actionable insights [44]. To facilitate the creation, sharing, integration and distribution of knowledge, these organisations are employing the use of information technology in its KM initiatives [44]. The created knowledge that is not shared, integrated and distributed

✉ Siti Aishah Mohd Selamat, aishah@bournemouth.ac.uk | [1]Faculty of Science and Technology, Bournemouth University, Dorset, England, UK.

SN Applied Sciences (2020) 2:818 | https://doi.org/10.1007/s42452-020-2589-3

Published online: 05 April 2020

precisely will erode easily causing wastage of organisational resources and time.

The complexity and ambiguity of the business environment are presenting the small and medium enterprises (SMEs) with an unparalleled set of challenges brought upon the knowledge economy. In consequence, it is, therefore, vital that SMEs remain flexible and continue to innovate momentously. The competitive dilemmas faced by the SMEs can be curbed with the implementation of KM by elevating the organisation's innovation ability in leading a sustainable performance. According to Wang and Yang [49], the deployment of KM can elevate the SMEs organisational performance [49]. Evangelista et al. [19] concurred that the application of KM in an organisation (regardless of its geographical locations and size) has indeed become the key source for organisational success [19]. SMEs, in turn, need to riposte to the emerging business changes to keep abreast with the ever-changing customer demands. KM is no longer a primary concern just for the large enterprises; it has become a key agenda for the SMEs in its business strategies too [49]. Nevertheless, the studies on KM implementation have mainly been concentrated on for the large enterprises. In resultant, the current research findings relate more to the large enterprises setting in comparison with the SMEs [44]. The challenges posed by the SMEs are distinctively different from the larger companies. SMEs, for example, have limited financial and human resources to work within the first instance [49]. Therefore, the research studies and result findings derived from the large enterprise case studies may not be suitable and applicable to the SMEs context [16]. A study on SMEs KM efforts is a key to identify the implementation pitfalls and successes—charting a point of reference for other SMEs alike to refer and emulate. The insurgency need for KM research on SMEs has seen a rise of development studies in recent years [28, 48].

In an in-depth report shared by Eurostat, the transportation and storage comprise 5% of the 22.3 million non-financial business economies in 2014 [18]. Further to this, in comparison with the large enterprises, it was recorded that the SMEs contributed a sum of 390 billion euros to the economy in 2015 [11]. A concise report by IDC further reveals that a total number of 48,500 SMEs have yet to implement any data analytic capabilities in its businesses [10]. The European Commission has established that large enterprises were able to escalate its business productivity by up to 10% with the implementation of data analytic [10]. Even with the substantial data on the positive implementation outcome of data analytics for a business—the transport and storage SMEs are not cordially embracing vigorously. Within the United Kingdom business sphere alone, the large enterprises have recorded a zealous adoption rate of 98.8% in comparison with the SMEs in 2012 [9]. Data analytics is essential because for new knowledge to be extracted, SMEs need to data mine its data asset using a data mining process model [33]. Therefore, within the knowledge management tools—data mining is an essential component of it.

In summation, there is an insurgent requisite for SMEs to delve into the application of both DM and KM in order to optimise these advanced technology benefits. It is further crucial for SMEs in the transportation sector to code attentive to this in order to remain competitive and relevant in the industry.

With the identification of the key research gap, this paper aims to carry out an in-depth review of data mining in knowledge management for SMEs in the transportation industry. Several research questions are established to unravel the research aim and these includes—(1) Reviewing the existing literature on DM and KM (2) Examining DM–KM case studies in the transportation sector and SMEs context and last but not least, (3) Analysing the DM–KM case studies results. The novelty of this research paper would create a substantial contribution to knowledge in the field of transportation and SME studies.

To establish the relevancy of this research, the articles used are kept within the timeframe of the last 10 years only. The literature collection is divided into four parts. The first and second part contain literature articles on KM and DM in transportation SMEs. The third and fourth part compose of literature papers on the DM–KM application in the transportation sector from the large organisation and SMEs, respectively. In order to authenticate the suitability and relevance of these articles, each full-text article is screened individually. Only relevant articles in relation to the key topic of DM–KM in the transportation and SME literature will be included for this research upon screening.

As a whole, this paper presents a critical study of DM–KM in the area of transportation sector with emphasise on the SMEs in particular. The literature collection was obtained from various online databases like Science Direct, EBSCOhost, ACM Digital Library, Springer, Google Scholar, IGI Global and Semantic Scholar. The research is targeted to provide fellow researchers, policymakers' and industrial business leaders with prominent research discovery in the study of DM–KM. In addition, this paper envisages to raise the understanding of DM and KM for SME's capitalisation to reap business advancement. In Sect. 2, the paper presents information on DM in transportation SMEs to provide the reader the characteristic of DM in the transportation of the transportation industry and why DM is crucial for the SMEs with the transportation sector. Section 3 covers the topic of KM, the various knowledge types and KM method available and KM application barriers. Section 4 covers the different types of DM models for the readers' comprehension before delving into the case studies of the paper. Sections 5 and 6 present the DM–KM application case studies findings for the transportation sectors, followed by findings in the SMEs' context. Section 7 provides a comprehensive analysis based on the case study findings and results of DM–KM application from both the transportation sector and SMEs' context. Lastly, in Sect. 8, the paper concludes with a summation of the research key findings and discussion of future research work to embark on.

2 Background: data mining in transportation SMEs

DM is the art of science in analysing and examining valuable information from a large pool of datasets or database to bring about unique knowledge [53]. The mining process involves deep computing of data analysis from a broad range of computational approaches such as decision trees, statistical analysis, rule induction and refinement, neural networks and graphics visualisation [45]. There are two primary goals in carrying out DM—(1) to spawn predictions and (2) to unearth new and distinct insights [20]. As the global data volume is predicted to scale up to 50 billion terabytes by 2020 [55], the significance of DM to the transportation SMEs expects to have an affluent impact to the sector. The growth in data collection by the transportation SMEs is critical during the DM course of actions as it creates and translates data into knowledge for business applications [35]. The implementation of DM can revolutionise a conventional transportation SME into an organisation with a competitive edge. In the mining of customer data for an example, the transportation SMEs would be able to understand its customers' behavioural habits, patterns or extract any other unique insights. This set of new information would be very useful to the sales and marketing division to devise new cross-selling opportunities, new product development initiatives and customers' retention strategies [53]. The mining of operational data, on the other hand, would facilitate transportation SMEs (in the private hiring industry) in overseeing its daily operational activities effectively and efficiently. For an instance, the operations division would be able to identify an effective and low-cost transport vehicle deployment for group pick-ups and drops off [26]. From the above examples, it is evident that DM can bring significant improvement to the SMEs business performance.

3 Knowledge management

KM can be described as the governance and optimisation of intellectual assets of both or either, the explicit knowledge (that are found in objects) or tacit knowledge (which are possessed by an individual or a group of communities) [48]. A further definition comparison between knowledge and KM are as per outlined below:

Knowledge	Knowledge can be defined as a mix of contextual information, expert insight embodying personal experiences and values to evaluate or incorporate new information or experiences [28]. For instance, in an organisation, these set of knowledge are not only set the organisation's repositories or documents—but also in its work routines, processes and practices
Knowledge management	KM, on the other hand, refers to the systematic management of knowledge asset, which involves the process of creation, sharing and application—to achieve business objectives [28]

Several pieces of research addressed that the majority applied KM projects involve one of the listed three aims below [6, 17]:

1. To create visibility in the usage of knowledge within the organisation; using visualisation platforms such as maps, hypertext tools and yellow pages.
2. To institute a pro-knowledge culture by spurring knowledge sharing initiatives throughout the various divisions in the organisation.
3. To build a knowledge-based infrastructure to stimulate knowledge interaction and collaboration.

3.1 Knowledge types

There are two types of knowledge within the bounds of KM and business, which is explicit and tacit knowledge. Explicit knowledge refers to a codified and formalised type of knowledge [15]. Unlike tacit knowledge, explicit knowledge is not difficult to discern, store and reclaim. An explicit knowledge in a business can be found namely in databases, documents, notes, memos, etc. On the other hand, tacit knowledge refers to a type of knowledge that is mainly intuition or experienced base—which is a difficult knowledge to define [15]. Because tacit knowledge is majorly personal in nature, it is tough to give an interpretation of. Nonetheless, amongst the two type of knowledge, the value of tacit knowledge outweighs that of explicit knowledge as it can aid organisation with breakthrough discovery [51]. According to Gamble and Blackwell [23], the lack of emphasis on tacit knowledge can nibble an organisation's innovation capabilities [23].

3.2 Knowledge management method

KM implementations are driven by the use of information and communication technologies (ICT) that assist with knowledge acquisition/creation, dissemination, conversion and utilisation [46]. The KM method can be classified into two method—(1) Traditional and (2) Advanced. Traditional method refers to traditional ICT tools like websites, portal, electronic databases (like video and audio recordings), multimedia presentation, discussion forms and email that are being used by an organisation for information dissemination and sharing [46]. Advanced method on the hand refers to new ICT tool like social networking tools, web 2.0/3.0, collaborative technologies 2.0, internal blogging, wikis and many more [46]. It is characterised as the next generation of community-driven web platform in which the internal stakeholders of the organisation is able to connect and collaborate in a conducive and productive information space [34]. Unlike the traditional method, the adoption of the advanced method allows a greater platform of knowledge sharing and promotes collaboration that may spark innovation and creativity with the organisation.

3.3 Knowledge management application barriers in SMEs

In an extensive research by Cerchione et al. [6], it was concluded that there are three areas that are indirectly or directly impeding the adoption of KM of the SMEs [6]. In the SME's business sphere, the KM of the organisation is usually human embedded—a common knowledge possesses and shared amongst the entire associate within the organisation. In other words, the KM culture is rather humanly managed and shared, without any technological intervention. The remaining factors include an immedicable deficiency in the financial and human resources as characterised by the SMEs. Although the three areas seem to delineate the major hindrance factors of KM implementation by the SMEs, it should be highlighted that the advancement of technology has enabled that the Information and Communication Technologies (ICTs) to offer SMEs with KM tools are low cost and ease-of-use [46]. As such, SME's that intends to apply KM to its organisation does not require high financial investment. In addition, with the ease-of-use, the SMEs do not need to get tormented by the need to get specialist skills in its workforce. Newer tools are embedded to provide and support the socialisation between the team members of the group [46]. In conclusion, it is worthy to note that the key hindrance factors as identified in the literature are weakening by technological intervention—majorly decreasing the financial and human resource barriers.

According to Cerchione et al. [6], SMEs denote a positive perception of the strategic significance of KM to the organisation; in consequence, using Information and Communication Technologies (ICT) in implementing [6]. Nonetheless, it has been found that the SMEs usually adopt a more traditional method in comparison with the large enterprise—whereby majorities are adopting the latest KM tools [6]. These leading to three inherent issues relating to the relevancy, usability and quality of the knowledge engendered [22]. As stated by Frost [22], knowledge is deemed as vital when it can be utilised to meet the organisation's business objectives [22]. A successful KM project that is being run would be made redundant if it is unable to generate the relevant knowledge as per the criteria sets out. This would in turns classify the KM project as a plausible failure. In all, the usability and quality of a set of knowledge are intertwined with its relevancy. The repercussion of an irrelevant knowledge would result in an outdated content with little or zero quality for the organisation's usage. In summary, the literature has pointed out that the significance of KM application for SME's is the key major concern that needs to be looked at. Practising KM without any distinctive benefits would cause wastage of resources.

4 Data mining process models

For knowledge to be created, a DM process model is need to be applied in order to extract and mine large dataset into useful and unique insights [33]. Based on an industrial poll organised by KdNuggets.com, the three most commonly used DM process models are Knowledge Discovery in Database (KDD), Cross-Industry Standard Process for Data Mining (CRISP-DM) and Sample, Explore, Modify, Model and Assess (SEMMA) model [29]. The definition and function of each model will be elaborated in the next few subsections.

4.1 KDD

KDD refers to the extensive process of discovering knowledge in data with the application of DM methods such as statistics, machine learning, artificial intelligence, and data visualisation [14]. In overall, the KDD consists a total of nine steps, which are both iterative and interactive in essence [42]. Iterative meaning the end user may require revisiting the earlier steps if necessary. Interactive, on the other hand, refers to the end user being able to be involved in each step. The KDD steps involved include (1) Establishing the application domain (2) Building a target data set (3) Data cleaning and pre-processing (4) Data transformation (5) Selecting the suitable DM task [50] Selecting the

DM algorithm (7) Applying the DM algorithm (8) Evaluation and interpretation of mined data (9) Consolidation of newly created knowledge [14].

4.2 CRISP-DM

A concept developed in the mid 1990s, CRISP-DM was developed by four renowned organisations—Daimler–Chrysler, SPSS, Teradata and OHRA. The DM model is recognised as a standard model by industrial practitioners because of its applicability across various industries. In total, the CRISP-DM model comprises of six steps. Likewise the KDD model, the CRISP-DM model is also iterative, allowing the end user to move backward and forward flexibly when required. The six steps of CRISP-DM includes (1) Business understanding (2) Data understanding (3) Data preparation (4) Modelling (5) Evaluation (6) Deployment [7].

4.3 SEMMA

SEMMA is a systematic tool developed by Suite of Analytics (SAS) Institute that functions specifically with its in-house enterprise miner tool. Similar to KDD and CRISP-DM, the SEMMA model is also iterative in essence. The overall step is derived from the acronym SEMMA itself—which refers to Sample, Explore, Modify, Model and Assess [21]. The SEMMA model primary focus is on the model development aspects of DM.

5 DM–KM case studies in the transportation sectors

This section comprises the case studies of DM–KM application in the transportation sectors for large enterprises. According to the Organisation for Economic Co-operation and Development (OECD), large enterprise refers to organisations that have more than 250 employees [37]. Ten large enterprises from the transportation sector will be identified. The selections of large enterprises are not limited to any geographical location, sector and transportation type. The reason being, any limited set will result in a further limited research resource for analysis. Based on the case studies research, it is noted that DM–KM is extensively implemented in all the transportation modes that comprise of sea, land and air as depicted in Table 1. For instance, in the sea transportation mode, a research conducted by Greis and Nogueira [25] identify that KM is used to facilitate the discovery of potential high-risk shipment [25]. Mirabadi and Sharifian [36] research from the land transportation mode on the hand also utilises KM to discover new information on the contributing factors of unsafe condition using the past accidental data [36]. In another land transportation mode research by Shin et al. [43], KM is used to facilitate in the forecasting of taxi pick-ups points using passengers' data [43]. Last but the least, a research by Wong et al. [52] in the area of air transportation employs KM to discover customer retention strategies using the passengers' travel behaviour and feedbacks data. Based on the compilation of case studies in the transportation sector as reflected in Table 1, KM is majorly used for predictions, generating new insights and to facilitate decision-making. The type of knowledge commonly used in all case studies is explicit knowledge. And the common KM approach used by the private and public transportation sector is the advanced method. The most commonly used DM model applied is CRISP-DM. In an extensive study, it was uncovered that the DM models applied are embedded in an assessment phase to validate the quality and usability of the DM application. Unlike the DM application, the KM application does not embed with a relevancy, usability and quality assessment. Further to this, it was also identified that collectively there is no DM–KM assessment given that there is no integration between the DM model and KM approach applied. Last but not least, one distinctive finding derived from Table 1 is that there is little research studies being carried out on SMEs in the transportation sector.

6 DM–KM case studies in the SMEs' context

From the earlier section, it was established that there is limited study being done for SMEs in the transportation sector. Therefore, to ensure that this paper is complete, this section intends to cover the use of DM–KM within the SMEs' context in the different industries. Like the DM–KM case studies carried out for transportation sectors, ten SMEs will be selected to study its DM–KM applications. Given the limited study in the area of SME transportation, the SME identified is open across various. According to OECD, what makes SMEs vary across the different region [37]. For the European Union (EU), an SME is classified with employing 250 members of staff, whereby in the USA, consider an SME with less than 500 employees. The SME classification and geographical location is non-restrictive for this particular case study as the study is primarily focusing on the DM–KM application and not related to any culture, behavioural or economic research. Further to this, there are no limitations on the type of industry the SMEs are in.

The ten case studies compiled as displayed in Table 2 include SMEs from the technology, food and beverage, tourism, finance, trading, manufacturing and aviation industry. Likewise in the case studies of DM–KM application in the transportation sector, the compilation of the

Table 1 DM–KM application in the transportation sector

Case studies	Knowledge resource	Knowledge type	KM method	KM usage	DM model applied	DM model integration with KM	DM model assessment	KM—relevancy, usability, quality assessment	DM–KM assessment
[47]	Private sector; Land transportation; Train industry	Explicit knowledge	Advanced	To enable railroad demand predictions To facilitate the operations department in manpower and operational activities planning	CRISP-DM	No	Yes	No	No
[52]	Private sector; Air transportation; Airflight industry	Explicit knowledge	Advanced	To mine passengers travel behaviour, demographics and service feedbacks To facilitate and plan customer retention activities	KDD	No	No	No	No
[26]	Public sector; Land transportation; Bus industry	Explicit knowledge	Advanced	To identify accident connections between the tram and car on the electric tramway To facilitate accident preventive measures	CRISP-DM	No	Yes	No	No
[43]	Private sector; Land transportation; Taxi industry	Explicit knowledge	Traditional	To analyse passenger's taxi pick-up point To facilitate in predicting prospective pick-up point for uncatered taxis	SEMMA	No	Yes	No	No

Table 1 (continued)

Case studies	Knowledge resource	Knowledge type	KM method	KM usage	DM model applied	DM model integration with KM	DM model assessment	KM—relevancy, usability, quality assessment	DM–KM assessment
[36]	Public sector; Land transportation; Train industry	Explicit knowledge	Advanced	To analyse past accident data To uncover new insights on the contributing factors of unsafe condition	CRISP-DM	No	Yes	No	No
[56]	Public sector; Land transportation; Train industry	Explicit knowledge	Advanced	To mine the various accident databases To facilitate intelligent accident treatments decision-makings	KDD	No	No	No	No
[25]	Public sector; Sea transportation; Shipping Cargo industry	Explicit knowledge	Advanced	To mine real-time data from wireless sensor networks, electronic "smart" tags and radio frequency identification (RFID) To facilitate in discovering potential high-risk shipments	CRISP-DM	No	Yes	No	No
[3]	Public sector; Land transportation; Bus industry	Explicit knowledge	Advanced	To mine both the past meteorological data and event driven totalising database To determine the best fuel-efficient resources for route planning	SEMMA	No	Yes	No	No

Table 1 (continued)

Case studies	Knowledge resource	Knowledge type	KM method	KM usage	DM model applied	DM model integration with KM	DM model assessment	KM—relevancy, usability, quality assessment	DM–KM assessment
[31]	Public sector; Air transportation; Airflight industry	Explicit knowledge	Advanced	To analyse the past aviation incident data To predict prospective incident and ramification	CRISP-DM	No	Yes	No	No
[12]	Public sector; Land transportation; Bus industry	Explicit knowledge	Advanced	To mine the collected vehicle fleet data To facilitate resource planning by predicting passenger demands	CRISP-DM	No	Yes	No	No

case studies of DM–KM application in the SMEs' context similarly reflects that KM is commonly used for predictions, generating new insights and to facilitate decision-making. For an example, in the finance industry, both Mandala et al. [32] and Koyuncugil and Ozgulbas [30] research involve deploying KM to aid in financial risk prediction and detection [30, 32]. In the technology industry, Bozdogan and Zincir-Heywood [5] and Pytel et al. [39] employ KM to facilitate decision-making to enhance organisational effectiveness and efficiency [5, 39]. In another example from the manufacturing industry, Packianather et al. [38] research employs KM to uncover the firm's consumer unique purchasing behavioural—forecasting potential purchasing behaviour to facilitate the SME's sales and marketing strategies [38]. Based on the compilation of case studies in the SMEs' context as reflected in Table 2, the common knowledge type used by the SMEs is explicit knowledge. And the majority of the SMEs are adopting a traditional approach to its KM application. A distinctive similarly with the transportation sector findings, the most commonly used DM model applied by the SMEs is CRISP-DM and that the model is embedded with an assessment phase to validate the quality and usability of the DM application. Likewise, the findings found in the transportation sectors case studies, the SMEs KM application are not embedded with a relevancy, usability and quality assessment. In addition, there is no DM–KM assessment available because the DM model and KM approach application and not collectively integrated together.

7 Case studies analysis and discussion

In this section, the case studies findings of Sects. 5 and 6 as discussed earlier are compiled in Table 3 for analysis and further discussion. The analysis of Table 3 findings will be dissected in two parts. The first part will cover the analysis of the DM–KM application trends. Next, in the second part, the analysis of DM–KM application issues will be presented. In the last part, a critical discussion focusing on the implication of DM–KM application issue in the SMEs' context will be addressed.

7.1 Analysis I: DM–KM application trends

From Table 3, it is gathered that collectively a total number of 20 case studies has been extracted. Of which there was an equal amount of ten case studies participation belonging to the transportation sectors and SMEs' context accordingly—providing evenness in the research study. The case studies industry application of the transportation sector (in Table 1) has identified that majority of the research findings were based on the private and public sectors and that there were limited studies to be found on the SMEs in the transportation sector in specific. And in order to succumb

Table 2 DM–KM application in the SMEs context

Case studies	Knowledge resource	Knowledge type	KM approach	KM usage	DM model applied	DM model integration with KM	DM model assessment	KM—relevancy, usability, quality assessment	DM–KM assessment
[40]	UK based SME wholesaler; Food and beverage industry	Explicit knowledge	Traditional	To analyse the various factors that affect the consumer demand (e.g. past sales record, product price, events, weather, holidays, etc.) To facilitate product demand predictions	CRISP-DM	No	Yes	No	No
[41]	Tourism SMEs; Service industry	Explicit knowledge	Traditional	To strengthen the analysis technique To enhance credit fraud detections	KDD	No	No	No	No
[39]	Project planning for SME; information technology industry	Explicit knowledge	Traditional	To analyse the historical data with the 23 pre-defined cost drivers To facilitate effort and cost approximation for small range software assignments	CRISP-DM	No	Yes	No	No
[32]	Rural bank; Financial industry	Explicit knowledge	Traditional	To analyse lender's character, capital, capacity, collateral, constraints and condition of the economy To establish credit assessment in classifying potential lenders as low risk or high risk	CRISP-DM	No	Yes	No	No
[30]	SMEs; financial industry	Explicit knowledge	Traditional	To analyse the financial data obtained from the balance sheet To facilitate in potential financial risk detection for SMEs	KDD	No	No	No	No

Table 2 (continued)

Case studies	Knowledge resource	Knowledge type	KM approach	KM usage	DM model applied	DM model integration with KM	DM model assessment	KM—relevancy, usability, quality assessment	DM–KM assessment
[5]	SMEs; IT management	Explicit knowledge	Traditional	To create a knowledge base on IT support management using the public sources automatically To facilitate IT support request	SEMMA	No	Yes	No	No
[8]	SMEs; trading	Explicit knowledge	Traditional	To discern the customer's potential purchasing behaviour to facilitate the SMEs' sales and marketing strategies To design contrive data mining application to analyse periodic sales order patterns for SME's usage specifically	CRISP-DM	No	Yes	No	No
[38]	SMEs; Manufacturing	Explicit knowledge	Traditional	To create new knowledge for consumer forecasting and organisational decision-making To uncover the firm's consumer unique purchasing behavioural To create new knowledge on untapped financial data of the firm such as payment orders, invoices, etc.	KDD	No	No	No	No

Table 2 (continued)

Case studies	Knowledge resource	Knowledge type	KM approach	KM usage	DM model applied	DM model integration with KM	DM model assessment	KM—relevancy, usability, quality assessment	DM–KM assessment
[54]	SMEs; Aviation	Explicit knowledge	Traditional	To create new knowledge in aircraft maintenance management to improve the aircraft launching and landing safeness To uncover the best maintenance process using the firm's existing historical data To create new knowledge using the system performance, supply records and maintenance data to uncover diagnosis and repair escalation	CRISP-DM	No	Yes	No	No
[1]	SMEs; Information technology	Explicit knowledge	Traditional	To create new knowledge of the firm's customer segmentation for targeted and strategic marketing decision-makings To drive down the firm's marketing cost driving out from the targeted marketing actions and deliverables	CRISP-DM	No	Yes	No	No

Table 3 Case studies analysis: DM–KM application in the transportation sector and SMEs context

	DM–KM application in the transportation sector	DM–KM application in the SMEs context
Total case studies participation	10	10
Case studies industry application	Private and public sectors; Air, sea and land transportation	SMEs sectors
Knowledge resources industry	Rail, aviation, public transportation, taxi, shipping, private transportation	Information technology, aviation, manufacturing, trading, financial, tourism, food and beverages
KM key usages	Predictions, planning, pattern recognition, decision-making, preventive measures	Segmentation, prediction, preventive measures, planning, cost control, pattern recognition, decision-making
Knowledge type application	Explicit knowledge—100% Tacit knowledge—0%	Explicit knowledge—100% Tacit knowledge—0%
KM approach	By traditional method—10% By advanced method—90%	By traditional method—100% By advanced method—0%
DM model applied	CRISP-DM model—60% SEMMA—20% KDD—20%	CRISP-DM model—60% SEMMA—10% KDD—30%
DM model assessments	Yes—80% No—20%	Yes—70% No—30%
KM—relevancy, usability, quality assessment	Yes—0% No—100%	Yes—0% No—100%
DM model integration with KM	Integrated—0% Un-integrated—100%	Integrated—0% Un-integrated—100%
DM–KM assessment	Yes—0% No—100%	Yes—0% No—100%

this gap, case studies industry applications from the SMEs' context (in Table 2) were collected to provide an overview of the DM–KM application from the SMEs' perspective in the various industries. Based on Table 3, an intriguing finding highlights a synonymous finding that both the transport sector and SMEs' of various industries use KM typically to derive unique or new insights, decision-making, planning and prediction. These suggest that despite the variance of organisational size and industry the core KM usage and synonymously similar.

7.2 Analysis II: DM–KM application issues

In this section, the DM–KM application issues will be discussed. One distinctive finding from Table 3 reflects that both the large enterprise in the transportation sector and SMEs' are conventionally taking advantage of its explicit knowledge only—without tapping on the tacit knowledge. This key finding correlates with the literature finding as stated in Sect. 3.1 that there is lack emphasise on the usage of tacit knowledge despite its greater advantage over explicit knowledge in spurring organisational innovation. Another distinctive finding highlights that unlike the large organisations in the transportation sector, the SMEs are heavily employing a traditional method in the KM approach instead of the advanced method. Out of the ten case studies from the transportation sector, a total of 90% of the large enterprise are using the advanced method of ICT in its KM approach. Contradictory to the SMEs', which accounted for 0% as in all of the SMEs case studies, a traditional approach to ICT in its KM approach was being applied. This in turns affirms the literature finding in Sect. 3.1 that SMEs are usually applying a traditional approach to ICT in its KM approach. On the DM model application, both the transportation sector and SMEs case study group record 60% usage of the CRISP-DM model. KDD and SEMMA make up the remaining 40% of the DM model application of both case study groups. This finding provides a correlation to the literature in Sect. 4 that CRISP-DM, SEMMA and KDD are the commonly applied model in the real-life industrial application. In the area of DM model assessment, it is critical to highlight that only the CRISP-DM and SEMMA are embedded with an assessment phase in its overall model processing to assess the DM model outcome. The assessment is carried out in terms of usability and quality model outcome. From Table 3, a total number of 80% of the transportation sectors case studies have adopted a DM model (either CRISP-DM or SEMMA) that is embedded with an assessment phase as part of its modelling phase. The remaining 20% of the case studies have adopted the KDD model which does not come with an assessment component to validate the outcome of the DM model. The case studies on the SMEs on the other hand reflected that a total number of 70% of the organisations have adopted a DM model (either CRISP-DM or SEMMA) that has an assessment component. The remaining 30% of the case studies have adopted the KDD model which do not have the assessment component. In contra to KM, Table 3 reflects that there is 0% KM assessment adopted in both the transportation sector and in the SMEs context case studies. In addition, seeing that the DM model is not integrated and incorporated with KM, a collective DM–KM assessment is not present neither in both the transportation sector and SMEs case studies—with each recorded 0%, respectively.

7.3 Discussion

Based on the analysis conducted in Sect. 7.1, it is positive to note that the DM–KM application is widely applied in both the large enterprise and SMEs of the various industries. Additionally, the DM–KM applications are synonymously used by both the large enterprise and SMEs with the core aim of business improvement. The analysis in Sect. 7.2 has however uncovered several numbers of key gaps that may potentially have an implication if not addressed promptly—more so for the SMEs' in the transportation sector. First and foremost, it is uncovered in Sect. 7.2 that explicit knowledge and traditional method were prominently being used more so by the SMEs of the various industries. To reap the benefits of tacit knowledge, SMEs will need to upgrade the adoption of traditional KM method to an advanced method using new ICT tools to elevate KM sharing, participation and collaboration. The traditional method of KM would not be able to support the use of tacit knowledge in a KM application [46]. Further to this, the larger enterprise is ahead of the SMEs in its KM approach by leveraging primarily on advanced method—and it will not be long before these large enterprises start exploiting its tacit knowledge to increase its competitive advantage. The correlation of explicit knowledge with traditional method and tacit knowledge with the advanced method is evidence from the literature review study carried out in Sects. 3.1 and 3.2. As iterated in Sect. 3.1, explicit knowledge can be found namely in databases [15] which can be classified as a traditional method where traditional ICT tools include websites, portal and databases [46]. Tacit knowledge, on the one hand, is experience-based knowledge [15], which can be collected through advanced method ICT tool like social networking tools [46].

Secondly, in Sect. 7.2, it was revealed that in reference to the component of assessment—it was highlighted that there is a gap in the presence of KM assessment and DM–KM assessment. These correspond with the literature finding in Sect. 3.3 that the leading KM issue relates to the relevancy, usability and quality of knowledge being adopted. Despite the presence of a DM assessment (more so for CRISP-DM and SEMMA model), it does not continue on to assess the KM application. In other words, the DM assessment functions in silos. This is a critical gap as in order to fully evaluate the effectiveness and efficiency of the DM–KM output, an integrated assessment needs to be present—to ensure that the DM and KM implementation are carried out seamlessly. As DM processes the data to become knowledge [53], KM role next is to manage these knowledge assets within the organisation [13]. Therefore, DM and KM need to be integrated and to work in conjunction with one another.

8 Conclusions and future work

Data are deemed as the lifeblood of an organisation in this twenty-first century. And it is only useful when the data are being processed, structured, organised, interpreted and presented to become meaningful information and knowledge [44]. SMEs and especially, transportation SMEs in particular, can no longer neglect the need to adopt advanced technology in its business in order to remain relevant and competitive against the large enterprises. As affirmed by the European Commission, the practice of data analytic can raise business productivity by up to 10% [10]. In altogether, there is an insurgent requisite for SMEs to delve into the application of DM and KM in order to optimise these advanced technology benefits. The SMEs' organisational concerns in the application of KM are related to the relevancy, usability and quality of the knowledge being produced [6]—that requires addressing.

In this paper, the DM–KM application in the transportation and SMEs context are being examined. To facilitate with an in-depth analysis and discussion, the case studies findings of both the transportation sector and SMEs context are compiled as presented in Table 3. The table uncovered several imperative findings. Firstly, this paper reveals that there is limited case study research on SMEs in the transportation sector in regard to DM–KM application. Secondly, an intriguing finding highlights a synonymous finding that both the transport sector and SMEs' of the various industries use KM typically for deriving unique or new insights, decision-making, planning and prediction. These suggest that despite the variance of organisational size and industry, the core KM usage and synonymously similar. Thirdly, the SMEs are lagging behind in its KM practice with the use of explicit knowledge and traditional method. SMEs will need to upgrade its adoption of explicit knowledge and traditional KM method to tacit knowledge with the usage of advanced KM method to elevate KM sharing, participation and collaboration—which can spur organisational innovation greatly. The core limitations deduced from the case studies are in the area of assessment. It was identified in Sect. 7.3 that there is a gap in the presence of an integrated DM–KM assessment to evaluate the outcome of the DM–KM application accordingly. These strikingly correspond with the literature finding in Sect. 3.3 that the leading KM issue faced by the SMEs relates to the relevancy, usability and quality of knowledge being adopted. Despite the research deriving an insightful conclusion, as highlighted by Geletkanycz and Tepper [24], the research limitation of each research study should be factored for scientific reasoning [24]. The key research limitation of this research study would be the small sample size of ten companies used to carry out the case study of DM–KM in transportation sector and SMEs context. The implications of this may potential derive a different research result. However, the research limitations persist in view of the limited literature resources in the area SMEs transportation studies.

Following up from this paper, the future research work aims to develop a DM–KM Assessment Plan Template (DM–KM APT) for the SMEs in the transportation sector. The DM–KM APT aims to evaluate both the DM and KM application outcome with the objective of ensuring that the knowledge created and implemented are usable and relevant. This would resolve the existing limitation uncovered in Sect. 7.3. The DM–KM APT will be constructed and applied on three transportation SMEs that will be identified. Three transportation SMEs identified are UK SME coach operator that has been operating a minimum of 3 years. The DM–KM APT will be evaluated by the three UK SME coach operators using an USE (usefulness, satisfaction and ease-of-use) survey method. The outcome of the USE survey method will evaluate the validity of the DM–KM APT application.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Afolabi IT, Worlu RE, Uwadia OC (2016) Data mining approach for target marketing SMEs in Nigeria. Covenant J Inform Commun Technol 4
2. Akhavan P, Pezeshkan A (2014) Knowledge management critical failure factors: a multi-case study. VINE 44:22–41
3. Almeida J, Ferreira J (2013) BUS public transportation system fuel efficiency patterns. In: 2nd international conference on machine learning and computer science (IMLCS'2013)
4. Asrar-Ul-haq M, Anwar S (2016) A systematic review of knowledge management and knowledge sharing: trends, issues, and challenges. Cogent Bus Manag 3:1127744
5. Bozdogan C, Zincir-Heywood N (2012) Data mining for supporting it management. In: Network operations and management symposium (NOMS), 2012 IEEE. IEEE, pp 1378–1385
6. Cerchione R, Esposito E, Spadaro MR (2016) A literature review on knowledge management in SMEs. Knowl Manag Res Pract 14:169–177
7. Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C, Wirth R (2000) CRISP-DM 1.0 step-by-step data mining guide. SPSS Inc, Chicago
8. Cheung C, Li F (2012) A quantitative correlation coefficient mining method for business intelligence in small and medium enterprises of trading business. Exp Syst Appl 39:6279–6291
9. Commision E (2017) Annual report on European SMEs 2015/2016—SME recovery continues
10. Commission E (2013) Business opportunities: big data
11. Commission E (2017) Annual report on European SMEs 2015/2016—SME recovery continues
12. Cristóbal T, Lorenzo JJ, García CR (2015) Using data mining to improve the public transport in Gran Canaria Island. In: International conference on computer aided systems theory. Springer, pp 781–788
13. Dalkir K, Liebowitz J (2011) Knowledge management in theory and practice. MIT Press, Cambridge
14. Dash P, Pattnaik S, Rath B (2016) Knowledge discovery in databases (KDD) as tools for developing customer relationship management as external uncertain environment: a case study with reference to state Bank of India. Indian J Sci Technol 9
15. Davenport TH, de Long DW, Beers MC (1998) Successful knowledge management projects. Sloan Manag Rev 39:43
16. Durst S, Edvardsson IR (2012) Knowledge management in SMEs: a literature review. J Knowl Manag 16:879–903
17. Enshassi A, Falouji I, Alkilani S, Sundermeieri M (2016) Knowledge management critical success factors in construction projects. Int J Sustain Constr Eng Technol 7:69–84
18. EUROSTAT (2014) Breakdown of number of enterprises within the non-financial business economy, EU-28
19. Evangelista P, Esposito E, Lauro V, Raffa M (2010) The adoption of knowledge management systems in small firms. Electron J Knowl Manag 8:33–42
20. Fayyad U (1997) Knowledge discovery in databases: an overview. In: Lavrač N, Džeroski S (eds) Inductive logic programming. ILP 1997. Lecture notes in computer science, vol 1297. Springer, Berlin, pp 1–16. https://doi.org/10.1007/3540635149_30
21. Fernandez G (2010) Data mining using SAS applications. CRC Press, Boca Raton
22. Frost A (2014) A synthesis of knowledge management failure factors. Recuperado el 22:1–22
23. Gamble PR, Blackwell J (2001) Knowledge management: a state of the art guide. Kogan Page Publishers, London
24. Geletkanycz M, Tepper BJ (2012) Publishing in AMJ–part 6: discussing the implications. Academy of Management Briarcliff Manor, NY
25. Greis NP, Nogueira ML (2011) Use of data mining for validation and verification of maritime Cargo movement. University of North Carolina, Chapel Hill
26. Haluzová P (2008) Effective data mining for a transportation information system. Acta Polytech 48:24–29
27. Jafari M, Rezaeenour J, Akhavan P, Fesharaki MN (2010) Strategic knowledge management in aerospace industries: a case study. Aircr Eng Aerosp Technol 82:60–74
28. Jennex ME, Smolnik S, Croasdell D (2016) The search for knowledge management success. In: 2016 49th Hawaii international conference on system sciences (HICSS). IEEE, pp 4202–4211

29. KDnuggets (2014) What main methodology are you using for your analytics, data mining, or data science projects? Poll. http://www.kdnuggets.com/polls/2014/analytics-data-mining-data-science-methodology.html. Accessed 25 Sept 2016
30. Koyuncugil AS, Ozgulbas N (2012) Financial early warning system model and data mining application for risk detection. Exp Syst Appl 39:6238–6253
31. Lukacova A, Babic F, Paralic J (2014) Building the prediction model from the aviation incident data. In: 2014 IEEE 12th international symposium on applied machine intelligence and informatics (SAMI). IEEE, pp 365–369
32. Mandala IGNN, Nawangpalupi CB, Praktikto FR (2012) Assessing credit risk: an application of data mining in a rural bank. Procedia Econ Finance 4:406–412
33. Mariscal G, Marbán Ó, Fernández C (2010) A survey of data mining and knowledge discovery process models and methodologies. Knowl Eng Rev 25:137–166
34. Martín-de Castro G (2015) Knowledge management and innovation in knowledge-based and high-tech industrial markets: the role of openness and absorptive capacity. Ind Market Manag 47:143–146
35. McAfee A, Brynjolfsson E (2012) Big data: the management revolution. Harvard Bus Rev 90:60–68
36. Mirabadi A, Sharifian S (2010) Application of association rules in Iranian Railways (RAI) accident data analysis. Saf Sci 48:1427–1435
37. OECD (2005) OECD SME and entrepreneurship outlook: 2005. OECD, Paris, p 17
38. Packianather MS, Davies A, Harraden S, Soman S, White J (2017) Data mining techniques applied to a manufacturing SME. Procedia CIRP 62:123–128
39. Pytel P, Britos P, García-Martínez R (2013) A proposal of effort estimation method for information mining projects oriented to SMEs. In: Poels G (ed) Enterprise information systems of the future. Springer, Berlin
40. Raju Y, Kang PS, Moroz A, Clement R, Hopwell A, Duffy A (2016) Investigating the demand for short-shelf life food products for SME wholesalers. Int J Soc Behav Educ Econ Manag Eng 9(6):1976–1980
41. Rebón F, Castander I, Argandoña J, Gerrikagoitia JK, Alzua-Sorzabal A (2015) An antifraud system for tourism SMEs in the context of electronic operations with credit cards. Am J Intell Syst 5:27–33
42. Shafique U, Qaiser H (2014) A comparative study of data mining process models (KDD, CRISP-DM and SEMMA). Int J Innov Sci Res 12:217–222
43. Shin I-H, Park G-L, Saha A, Kwak H-Y, Kim H (2009) Analysis of moving patterns of moving objects with the proposed framework. In: Computational science and its applications–ICCSA 2009, pp 443–452
44. Silwattananusarn T, Tuamsuk K (2012) Data mining and its applications for knowledge management: a literature review from 2007 to 2012. arXiv preprint arXiv:1210.2872
45. Smyth UF, Gregory P-S, Padhraic S (2008) From data mining to knowledge discovery in databases. AI Mag 17:37
46. Soto-Acosta P, Soto-Acosta P, Cegarra-Navarro J-G, Cegarra-Navarro J-G (2016) New ICTs for knowledge management in organizations. J Knowl Manag 20:417–422
47. Viglioni G, Cury MVQ, da Silva PAL (2007) Methodology for railway demand forecasting using data mining. In: Proceedings of the SAS global forum, Brazil
48. Wang H (2010) Knowledge management in small and medium sized enterprises. In: 2010 The 2nd IEEE international conference on information management and engineering (ICIME. IEEE, pp 420–424
49. Wang M-H, Yang T-Y (2016) Investigating the success of knowledge management: an empirical study of small-and medium-sized enterprises. Asia Pac Manag Rev 21:79–91
50. Wehner P, Piberger C, Göhringer D (2014) Using JSON to manage communication between services in the Internet of Things, pp 1–4
51. Wellman J (2009) Organizational learning: how companies and institutions manage and apply knowledge. Springer, Berlin
52. Wong J-Y, Chung P-H (2007) Managing valuable Taiwanese airline passengers using knowledge discovery in database techniques. J Air Transp Manag 13:362–370
53. Xindong W, Xingquan Z, Gong-Qing W, Wei D (2014) Data mining with big data. IEEE Trans Knowl Data Eng 26:97–107
54. Young T, Fehskens M, Pujara P, Burger M, Edwards G (2010) Utilizing data mining to influence maintenance actions. In: AUTOTESTCON, 2010 IEEE. IEEE, pp 1–5
55. Zeng D, Lusch R (2013) Big data analytics: perspective shifting from transactions to ecosystems. IEEE Access 28:2–5
56. Zhang C, Huang Y, Zong G (2010) Study on the application of knowledge discovery in data bases to the decision making of railway traffic safety in China. In: 2010 international conference on management and service science (MASS). IEEE, pp 1–10

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.