Electron Commer Res (2017) 17:169–183
DOI 10.1007/s10660-016-9242-7

Does big data mean big knowledge? Integration of big data analysis and conceptual model for social commerce research

Xuemei Tian[1] · **Libo Liu**[1]

Published online: 1 October 2016

Abstract The Big Data era has descended on many communities, from governments and e-commerce to health organizations. Information systems designers face great opportunities and challenges in developing a holistic big data research approach for the new analytics savvy generation. In addition business intelligence is largely utilized in the business community and thus can leverage the opportunities from the abundant data and domain-specific analytics in many critical areas. The aim of this paper is to assess the relevance of these trends in the current business context through evidence-based documentation of current and emerging applications as well as their wider business implications. In this paper, we use BigML to examine how the two social information channels (i.e., friends-based opinion leaders-based social information) influence consumer purchase decisions on social commerce sites. We undertake an empirical study in which we integrate a framework and a theoretical model for big data analysis. We conduct an empirical study to demonstrate that big data analytics can be successfully combined with a theoretical model to produce more robust and effective consumer purchase decisions. The results offer important and interesting insights into IS research and practice.

Keywords Big data · Business intelligence · Customer knowledge management · Customer purchase behavior

✉ Xuemei Tian
stian@swin.edu.au

Libo Liu
liboliu@swin.edu.au

[1] Department of Business Technology and Entrepreneurship, Faculty of Business and Law, Swinburne University of Technology, Melbourne, Australia

1 Introduction

Since the 1960s, the interplay between hardware, software, and communications has led to previously unforeseen advances in information systems. These developments have been accompanied by shifts in the relationship between 'carrier and content', and between manifestations of the latter as various data, information, knowledge, and indeed wisdom [1]. During this time, concerns over data and information overload increased, along with those focused on related organizational and managerial challenges [2].

One response to these concerns saw the emergence of sub-disciplines, such as competitive and business intelligence, and data, information, and knowledge management [3–5]. Others came from management and organizational theory, notably the resource-based theory, which emphasized the importance of internal capabilities and competencies [6, 7], and later, knowledge-based views [8, 9], which focused on the heterogeneous bundles of intangible resources (i.e., imperfectly mobile, imperfectly imitable and non-substitutable) as a basis for competitive advantage and organizational success.

As understanding of the nexus between information technology, information systems, and business has increased, another core lesson has been learned—that technology on its own is unlikely to provide the answer to the questions continually posed by the management [10]. Whether it is relational databases, data mining, or enterprise systems, the latest 'big thing' usually turns out to be much less of a transformative force when introduced to the market. This is as likely to be true for the latest such breakthrough, namely Big Data [11]. How to better utilize big data assets, in addition to business assets, and human capital, to create value has become a fertile ground for corporate competitive advantage. As big data analysis becomes the next frontier for advancement of knowledge, innovation, and enhanced decision-making process, the significance of its impact on society as a whole should not be underestimated [12].

The aim of this paper is to assess the relevance of these trends in the current business context through evidence-based documentation of current and emerging applications as well as their wider business implications. In this paper, we use the BigML software package to investigate how two social information channels (i.e., friends-based and opinion leaders-based social information) influence consumer purchase decisions on social commerce sites. We undertake an empirical study by combining a theoretical model with the analysis of big data using BigML. This methodology demonstrates that big data analytics can be successfully combined with a theoretical model to produce more robust and effective consumer purchase decisions. However, for this to happen, big data technologies must be further developed to cope with the proposed big data analysis framework.

The rest of the paper is organized as follows. In Sect. 2, we provide a review of the literature related to big data, business intelligence, social information processing theory and online consumer reviews. In Sect. 3, we identify the research gap based on the literature review. In Sect, 4, we conduct an empirical study to develop and explain our methodology and compare the results from different models. In Sect. 5, we provide robust checks for our final model and conclude by discussing the wider implications of our study.

2 Literature review

2.1 Big data and business intelligence

Big data has been labelled as the fifth wave in the technology revolution, after the mainframe, the PC, the Internet, and Web 1.0 eras, and more recently, mobile and Web 2.0 eras [13]. It combines an architectural paradigm shift in data movement in which instead of bringing data for centralized computation, the aim is to push the computation to distributed locations [14]. Big data analytics have been used to describe data sets and analytical techniques in applications that are so large (from terabytes to exabytes) and complex (from sensor to social media data) that they require advanced and unique data storage, management, analysis, and visualization technologies [12]. Considering the seemingly limitless spectrum, big data can utilized for social networks, web server logs, traffic flow sensors, satellite imagery, broadcast audio streams, banking transactions, MP3 s of rock music, the content of web pages, scans of government documents, GPS trails, telemetry from automobiles, financial markets and so on [15]. Dumbill (2012) explains that although any data can be regarded as 'Big' when size is too large to be handled by conventional systems, it is also the case that the dramatically increased *volume* (there is too much of it), *velocity* (it is moving too fast), and *variety* (it is not structured in a useable way) do not longer fit in the structures of database architectures [12, 15]. More recently, three more characteristics have been mentioned by other researchers, such as Daniel [16]: (1) *value* (a source of competitive advantage), (2) *veracity* (the biases, noise and abnormality in data), and (3) *verification* (refers to data verification and security). Daniel [16] argues that these three attributes also represent Big Data's fundamental properties as they are linked to data accuracy, a concept associated with the longevity of data and their relevance to analysis outcomes, as well as the length required to store data in a useful form for appropriate value-added analysis.

Clearly, today more than ever, data is a key resource and corporations are striving to generate it at innumerable data points to create value and obtain competitive advantage [17]. As will be seen below, further iterations of the concept are ongoing, as through a combination of technological change, regulatory processes, and managerial decision making, businesses are collecting and storing data at an astonishing rate. In turn this is enabling organizations to employ business intelligence to make faster and more reliable information-based business decisions [18].

It is claimed that the advent of big-data technologies has rendered obsolete the separation of information-centric systems, such as data warehouses (once reserved for strategic decision support) and those, like enterprise resource systems (supporting daily operations) [19]. The integrated management of transactional and operational data, especially in new post-bureaucratic structures promises improved business value through better-informed decisions, the discovery of hidden insights and automation of business processes [20]. In addition, the resultant decisions can be made on the basis of data and rigor rather than gut-feeling [12], albeit increasingly in situations where big data is placing transformative data discovery and advanced analytics tools into the hands of customers [21] and where concerns have already been expressed about risks to the security of proprietary knowledge [22].

Furthermore, while there is plenty of enthusiasm in business circles for big data, there are those who argue that there is little special about it, seeing it as merely a magnification of the features found in knowledge management, and others predicting that its popularity will eventually wane owing to its complexity and a shortage of qualified workers [23]. It is also important to balance matters of data size against the context of its application. A survey conducted jointly by IBM and Oxford University Business School revealed that 30 % of respondents did not know what 'big' meant for their organizations [24]. Elsewhere it has been observed that considerations of adequacy and relevance might require 'not so big data' [25], something that might apply to smaller businesses, who are also expected to adopt these technologies, for example through SaaS or as a component of ERP systems [26].

An important component of the ongoing case for big data is calls for education and training of a new breed of statisticians and analysts for the proper use of BI applications [27], big data scientists operating as multifunctional problem solvers communicating between different departments [28, 29]. Likewise, the professional development and career progression of in-house analysts—already familiar with the organization's unique business processes and challenges—has also been identified as a top priority for business executives [24]. This might well extend to the operational personnel, to reflect their needs for analytical and decision making competencies consequent upon ongoing systems integration and the breaking down of traditional organizational hierarchies [19].

As with other such concepts, the meaning and interpretation of business intelligence (BI) has shifted over the years. At its simplest, BI refers to the ability to use information to gain a competitive advantage. This includes data on education, skills, and the past performance of employees to help businesses identify the critical talent within their organizations and ensure its development and retention [29]. BI is often referred to as the techniques, technologies, systems, practices, methodologies, and applications that analyze critical business data to help an organization better understand its business and respond to the market through timely decisions [30]. Moving beyond BI to the development of more advanced apps and an enhanced ability to drill down and answer the questions asked by BI engines [31], BI and Big Data initiatives must advance together, rather than separately on parallel tracks. The real value lies in integrating the existing BI and analytics capabilities with the new big data technologies and techniques to focus on how these new capabilities can augment and extend the existing environment [31].

It is clear that much work remains to be done for the anticipated benefits of big data to be realized. The remaining issues of clarification around the specific content of these large datasets (primitive data elements, information, knowledge or meta-knowledge) and the relationship between terms like data, information, and knowledge [25]. There are also risks to do with the misuse of data or with ‘false discoveries’ in trying to find meaningful needles in massive haystacks of data [11]. Hence, this calls for improvement in such areas as data analysis, acquisition, extraction and cleaning, integration, representation, and interpretation thorough integration of systems. More generally, there are matters of the benefits and risks of big data, both in organizational terms and as regards society as a whole, whether these relate to potential problems in the areas of privacy, security and ethics, or to emergence of a new version of the Digital Divide, a digital network divide [1, 32]. Critically, this ‘softer’ dimension to developments in big data extends to behavioral and cultural issues at organizational level, and would extend to a much wider set of norms and behaviors.

The promises of Big Data are real, however, there is currently a wide gap between its potential and realization. Many leading researchers, such as Agrawal et al. [33] and Zicari [34], emphasize the opportunities and challenges in big data research and argue that the challenges include not just the obvious issues of scale, but also heterogeneity, lack of structure, error-handling privacy, timeliness, provenance, and visualization. These problems impede the progress at all phases in the pipeline (acquisition/recording, extraction/cleaning annotation, integration/aggregation/representation, analysis/modeling, and interpretation). Better understanding and management of these problems is the key of creating business value from data [33]. The task of creating real value from data cannot be achieved by simply by focusing on development of new technologies; it also requires us to fundamentally rethink how we manage data analysis. The purpose of this paper is to introduce a concept on how to manage data analysis.

2.2 Social information processing theory and online consumer reviews

The growing availability and popularity of social commerce sites have increased the importance of online social information as a market force [35]. Previous studies have revealed an association between eWOM (electronic word-of-mouth) and product sales/revenue, which is mostly explained through either the awareness or persuasive effects [36]. Alshibly and Chiong [37] showed that customer empowerment can increase e-government success. Opinion-based social information in the form of eWOM communication can take place through several channels. For example, consumers can post their opinions, and reviews of products on weblogs, review websites, and social networking sites. Other studies have found that the volume of consumer reviews is significantly related with product sales [38, 39].

Nowadays, online consumer reviews are becoming increasingly prevalent in the majority of online shopping sites. Consumers use them either to find products that match their preferences or to find information useful for offline purchases. Typically, online consumer reviews can be deconstructed into several dimensions. Scholars focusing on these have reported the different influences of such reviews on the choice and sale of products. For instance, Kostyra et al. [40], focusing on valence (i.e., average rating), volume (i.e., number of customer ratings) and variance (i.e., variation in ratings) of online consumer reviews (also see [41]), found that volume and variance only serve as moderators of the effect of valence. As well it was revealed that they do not have a direct effect on consumers' product choice. Focusing on the volume and valence of online movie reviews, Liu [42] documented that review volume, rather than review valence, increases both aggregate and weekly box office revenue. By taking this finding a step further, Duan et al. [43] proposed a dynamic simultaneous equation system in which the two dimensions could be considered as both a precursor to, and an outcome of, retail sales. That is, both a movie's box office revenue and review valence significantly increase review volume, which in turn leads to better box office performance. Moreover, as noted earlier, previous studies concerning the effects of online consumer reviews have investigated the issues either from the perspective retailers or consumers. In our study, we focus on the effect valence of review on consumer purchase decision.

3 Research framework

In the foregoing literature review a number of commonalities emerged between sources. Firstly, researchers examined knowledge contribution in general, with a special focus on knowledge sharing. Although some studies noted that social relational activities influence knowledge sharing [44], few studies explicitly investigated social relational activities. Secondly, most studies relied on subjective data collected through surveys, case studies, and focus groups to explore how and why customers participate in online social communities. Straub et al. [45] argued that actual usage and perceived usage are not always congruent. Instead of using self-reported contributions, we used field data to measure the actual contributions. Thirdly, we worked on the premise that extracting useful knowledge from big data requires not only scalable analytics services, but also theoretical support.

Boyer et al. [46] identified the obstacles to successful BI projects as not as much technological, as lying within the organization itself. They thus advocated strategic enterprise BI programs that would enable business users to make informed decisions based on the collaborative leveraging of enterprise-wide information. This prediction seems relevant today and to the circumstances of big data, whose analytics-driven insights must be closely linked to business strategy, easy for end users to understand, and embedded into organizational processes to enable action at the right time [47]. The strategic significance of such activities is emphasized by Elbashir et al. [48] who recognize the critical importance of collaborative knowledge synergies between senior management, CIOs and IT managers in facilitating improved decision making across the organization's value chain. For present purposes this is a reminder of the importance of the knowledge component, both in relation to current expectations of big data, and in acknowledgement of the fact that many organizations are still failing to use available data and knowledge effectively, let alone being in a position to accommodate the demands of big data.

Admittedly big data really could turn out to be different, but another lesson from experience is that the more data and information are available, the greater the need for human judgement in decision making. Even in reporting big data-related advances in cognitive augmentation and job automation, McGovern [49] concedes that there are still many tasks where the combination of humans and machines

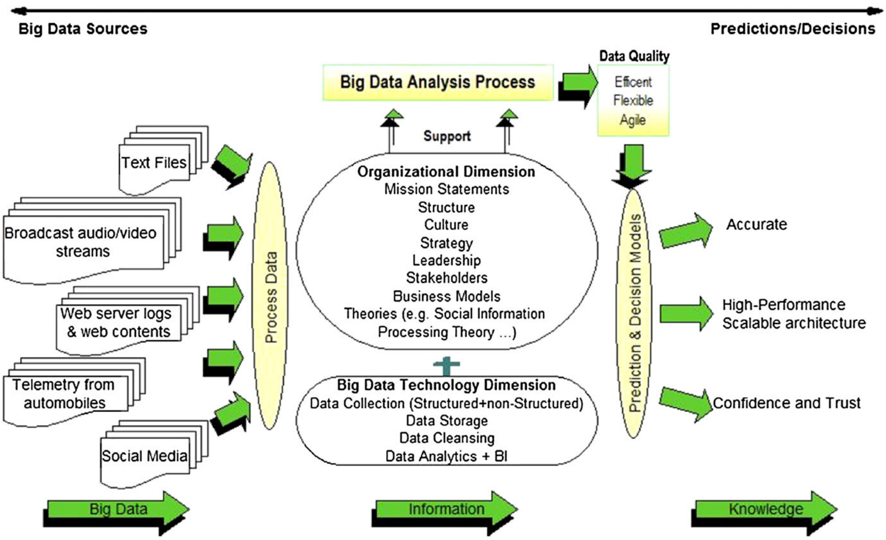

Fig. 1 Big data analysis framework

produce superior results. As shown in Fig. 1, the availability of data alone does not equate to value creation [25], and what is often missing is the knowledge to extract wisdom from it, and to take decisions about what data to keep and what to discard, and how to store what is kept reliably with the right metadata [50]. Processing data requires not only the right technologies and analytic tools/techniques, but also consideration of the wider organizational dimension. This will mark the transformation from big/massive data to knowledge which can be used for making business predictions and decisions. This raises questions not only as to the relative value of big data predictions and knowledge-based decisions, but also as to future relationships between big data and knowledge management.

In this paper, to support the above arguments, we use two empirical studies to investigate the influence of social relational activities on customer purchase decisions. We also compare the findings from two studies which indicate that big data analysis may not provide the best results without consideration of the organizational dimension.

4 Empirical study

We conducted an empirical study by combining a big data tool (BigML) and a conceptual model. We used BigML to examine how the two social information channels (i.e., friends-based social information and opinion leaders-based social information) influence consumer purchase decisions on social commerce sites.

4.1 Data collection

The data for this study was collected from a popular social commerce site in Asia, which provides a platform for consumers to share their experience related to the purchased cosmetics, and to interact with other consumers. A consumer in the social commerce site can rate, for example, a cosmetics product (of a particular brand) while sharing her or his experience about this product. The consumer can also follow other consumers whose posts or ratings are useful. Figure 2 shows the social network structure of ego consumer network in this community. As the provision of ratings is optional, some consumers choose not to provide ratings about their products or a particular brand.

In total, the website includes 10,097 products from 60 different brands. All of brands were used for analysis in this study. Moreover, there are 163,845 consumers within sixty brands. We crawled panel data from this site in Nov 2012. Based on the customers ID, we crawled the network data for each customer from the community. Specifically, we collected the "following" and "follower" of customers' lists and built an egocentric network for each customer. Then we crawled friends-based social information (i.e., friends' review valence and friends' purchase) and opinion leaders-based social information (i.e., opinion leaders' review valence and opinion leaders' purchase). Table 1 presents the summary statistics of the variables.

4.2 Operationalization of variables

Opinion leaders' review valence is operationalized as the total score of ratings (on products in a particular brand) provided by individuals who are the members' reference persons.

Friends' review valence is operationalized as the total score of ratings provided by individuals who are friend of a consumer in the social commerce site.

Opinion leaders' purchase is operationalized as the total number of products in the buy-lists of a consumer's reference persons in the social commerce site.

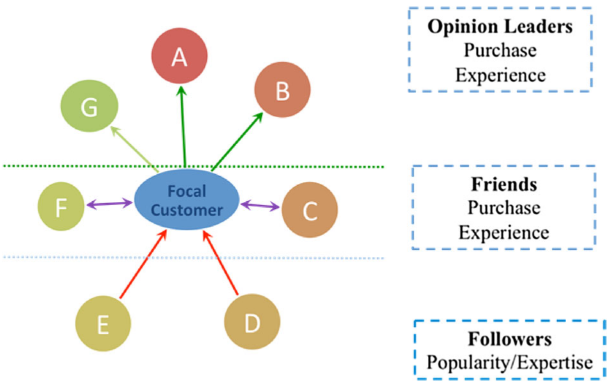

Fig. 2 Social network of ego customers

Table 1 Descriptive statistics

	Mean	SD
Friends	6.92	8.37
Opinion leaders	4.07	17.66
Friends' review valence	5.99	16.50
Friends' purchase	1.53	8.05
Opinion leaders' review valence	5.05	29.56
Opinion leaders' purchase	1.07	14.61

Friends' purchase is operationalized as the total number of products in the buy-lists of a consumer's friends in the social commerce site.

Consumer purchase behavior is operationalized as the number of products in a consumer's buy-list. Such operationalization was considered because consumers can add products to their buy-lists to indicate that they have already bought specific products in the social commerce site.

4.3 Data analysis tool: BigML

BigML (https://bigml.com) is an approach to machine learning. Users can set up data sources, create, visualize and share prediction models, and use models to generate predictions. It implements a set of basic data mining procedures. BigML includes several advanced procedures that support interactive analysis of big data. For example, data histograms are first rendered on a data sample and then updated when the entire data set is loaded and processed. Decision trees provide a supervised machine learning technique that starts with the root classification node, which is then iteratively refined. Decision trees in BigML are rendered on the client application dynamically, and evolve as their nodes are induced on the computational server. More important, the process of preparing a model in BigML can be divided into several consecutive steps: (1) data preparation, (2) association model learning, (3) model selection, (4) model testing, and (5) development of model.

4.4 Phase 1: BigML analysis: base model

To analyze our data in BigML, we set consumer purchase decisions as our objective. We first used BigML to run the selected dataset with the *Base Model* (1 Click Model). The results of the decision tree are shown in Fig. 3. This indicates how important each factor is in the prediction of consumer purchase decisions. Friends' purchase is the most important factor, which is followed by opinion leaders' review and opinion leaders' purchase. The result of the BigML analysis ran with the Default Model is not consistent with the regression analysis. This means that the Default Model does not fit the business rules well, which are artifacts of the data.

BigML's Default Model does not fit the business rules, and therefore we should adjust the training model by integrating a conceptual framework from the existing literature.

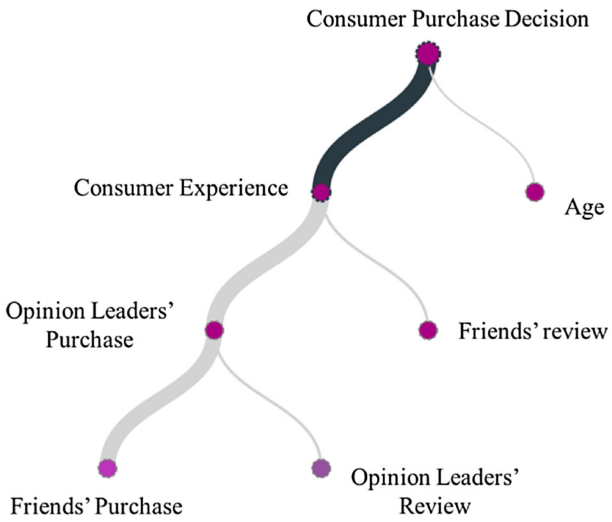

Fig. 3 Results from training data with the default model

4.5 Phase 2 BigML analysis: integrating the conceptual model into the base model

4.5.1 Conceptual research model

Based on social information processing theory, we propose a conceptual research model shown in Fig. 4, which should be integrated with the default model of BigML.

Consequently we developed the following hypotheses:

H1a Friends' review valence will positively influence ego-consumer purchase behavior.

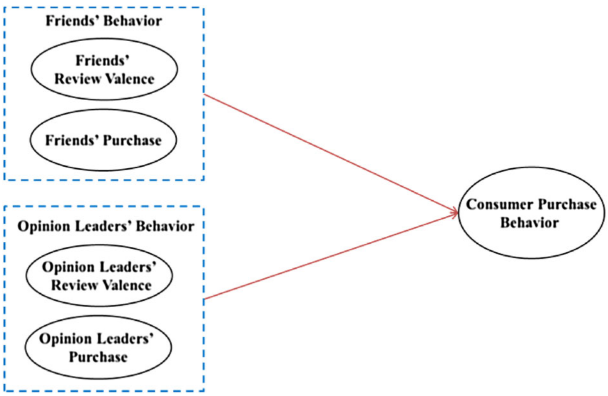

Fig. 4 Research model

H1b Friends' purchase will positively influence ego-consumer purchase behavior.

H2a Opinion leaders' review valence will positively influence ego-consumer purchase behavior.

H2b Opinion leaders' purchase will positively influence ego-consumer purchase behavior.

Based on the conceptual framework, we conducted the following steps to adjust the model in BigML: (1) association model learning; (2) model testing; and (3) development of model of purchase decision marking. The results from running the data set with the adjusted model are shown in Fig. 5. The results indicate that the adjusted model fits the theoretical model better and therefore provides more accurate results concerning the antecedents of consumers' purchase decisions.

4.6 Robust check

Consumer purchase decision is represented by using a count variable summing up all the products that are added to a buy-list. Negative binomial regression is often used to analyze count data. Thus we use negative binomial regression to analyze the dataset. To test the robustness of the adjusted model from BigML, we ran a negative binomial regression analysis for the model.

$$\begin{aligned}\text{Consumer\_Purchase\_Decision} = {} & \beta_0 + \beta_1 \text{Opinionleader\_Review} \\ & + \beta_2 \text{Friend\_Review} \\ & + \beta_3 \text{Opinionleader\_Purchase} \\ & + \beta_4 \text{Friend\_Purchase} + \text{w} \qquad (1)\end{aligned}$$

The regression analysis results are summarized in Table 2. The results indicate that friends' reviews valence have a significant positive effect ($\beta = 0.012$, $p < 0.001$) on consumers' purchase decisions. The regression results suggest that

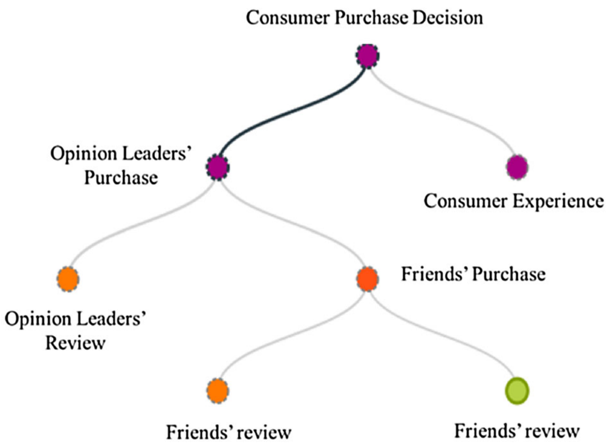

Fig. 5 BigML analysis model 2

Table 2 Results of regression analysis

Model	Unstandardized coefficients		Sig.
	B	Std. error	
(Constant)	.364	.004	.000
Friends' review valence	.012	.000	.000
Friends' purchase	.051	.001	.000
Opinion leaders' review valence	.006	.000	.000
Opinion leaders' purchase	.035	.001	.000

opinion review also has a significant positive effect ($\beta = 0.051$, $p < 0.001$) on consumers' purchase decisions. Table 2 reveals that opinion leaders' review valence exerts a significant positive moderating effect on consumers' purchase decisions ($\beta = 0.006$, $p < 0.001$). The positive coefficient ($\beta = 0.035$, $p < 0.001$) indicates that opinion leaders' purchase exerts a significant positive effect on consumers' purchase decisions. Furthermore, the results are consistent with the adjusted model of BigML. Therefore, our model developed in BigML is robust.

5 Discussion and conclusion

5.1 Research implications

Recently, the Big Data era has descended on many communities, from governments to e-commerce to health organizations. To avoid being data driven, the big data analysis process should be adopted to leverage the opportunities presented by the abundant data and domain-specific analytics. In addition, while machine learning algorithms exist, most of them produce "black box" models, which are difficult to understand. This study presents a complete implementation of a big data analysis process by integrating machine learning with a theory based on business rules from the existing literature. Instead of developing machine learning models solely on specific training datasets, researchers should integrate theory based business rules to adjust the machine learning models and use test datasets to test the adjusted machine learning models in order to attain comprehensive models that are more relevant and better reflect the reality (see Fig. 6).

The literature of big data emphasises the application of algorithms to pattern analysis and prediction. We have created a big data processing model to emphasise that processing data requires not only the right technologies and analytic tools/ techniques, but also consideration of the wider organizational dimension. This will drive the transformation from big/massive data to knowledge to business value processes (as shown in Fig. 5 below). This raises questions not only as to the relative value of big data predictions and knowledge-based decisions, but also as to future relationships between big data and knowledge management. Big Data is still in its infancy and is quintessentially a technology or set of still developing

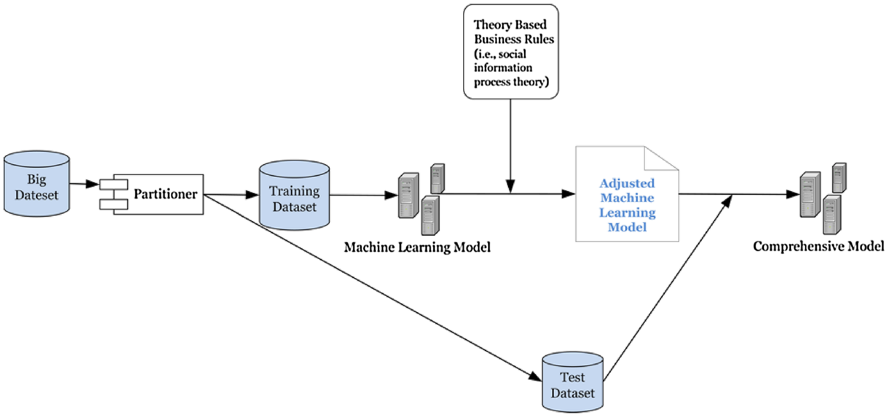

Fig. 6 Big data analysis process

technologies. Experience has shown time and again that the contribution of new technologies is dependant very much on organizational and especially, human factors.

5.2 Conclusion

Making sound business decisions based on accurate and current information and knowledge requires more than simple intuition, and BI has become indispensable to organizational success in the global economy. We need not only to be skilled engineers, but also domain experts. Big data will become big knowledge through the combination of data mining of association rules and the conceptual model of business rules. Data mining and domain experts will be able to use this for extending the possibilities of the model not only for user-defined business rules, but also for models generated from association rules gained from machine learning.

This paper started from the perception that big data and business intelligence, while facing challenges in developing systems for a new data savvy generation, had the potential to leverage opportunities presented by the abundant data and domain-specific analytics needed in many critical areas.

This study assessed the relevance of these trends in the current business context through evidence-based documentation of current and emerging applications as well as their wider business implications. We used the comprehensive model after verifying that the machine learning model was inaccurate and thus unsuitable to the task in hand. The proposed framework was tested using in an empirical study the results of which suggest that big data can become big knowledge by the combination of data mining of association rules and a conceptual model of business rules. This big data analysis framework and theoretical model contributes to the existing literature and offers important and interesting insights to IS research and practice.

References

1. Buchholtz, S., Bukowski, M., & Sniegocki, A. (2014). Big and open data in Europe: A growth engine or a missed opportunity. Warsaw, demosEuropa und Warsaw Institute for Economic Studies. Retrieved from http://www.bigopendata.eu/wp-content/uploads/2014/01/bodeurope2020fullreportsinglepage.pdf.
2. Reich, B. H., & Benbasat, I. (2000). Factors that influence the social dimension of alignment between business and information technology objectives. *MIS Quarterly, 24*(1), 81–113.
3. Biere, M. (2003). *Business intelligence for the enterprise*. Upper Saddle River, NJ: Prentice Hall Professional.
4. Williams, S., & Williams, N. (2010). *The profit impact of business intelligence*. Burlington, NJ: Morgan Kaufmann.
5. Wixom, B. H., Watson, H. J., & Werner, T. (2011). Developing an enterprise business intelligence capability: The Norfolk Southern Journey. *MIS Quarterly Executive, 10*(2), 61–71.
6. Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal, 5*(2), 171–180.
7. Barney, J. B. (2001). Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of Management, 27*(6), 643–650.
8. Eisenhardt, K. M., & Santos, F. M. (2002). *Knowledge-based view: A new theory of strategy. Handbook of strategy and management* (Vol. 1, pp. 139–164). London: Sage.
9. Spender, J. C., & Grant, R. M. (1996). Knowledge and the firm: Overview. *Strategic Management Journal, 17*(S2), 5–9.
10. Reynolds, P., & Yetton, P. (2015). Aligning business and IT strategies in multi-business organizations. *Journal of Information Technology, 30*, 101–118.
11. Lohr, S. (2012). *The age of big data*. New York: New York Times.
12. McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data. *The Management Revolution. Harvard Business Review, 90*(10), 61–67.
13. Freeland, C. (2012). *Big data's big impact*. Accessed January 12.
14. Wixom, B., Ariyachandra, T., Douglas, D., Goul, M., Gupta, B., Iyer, L., et al. (2014). The current state of business intelligence in academia: The arrival of big data. *Communications of the Association for Information Systems, 34*(1), 1.
15. Dumbill, E. (2012). *What is big data? An introduction to the big data landscape*. Sebastopol, CA: O'Reilly Media Inc.
16. Daniel, B. (2014). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*. doi:10.1111/bjet.12230.
17. Batra, S. (2014). Big data analytics and its reflections on the DIKW hierarchy. *Review of Management, 4*(1/2), 5–17.
18. Kapoor, B. (2010). Business intelligence and its use for humans. *The Journal of Human Resource and Adult Learning, 6*(2), 21–30.
19. Berner, M., Graupner, E., & Maedche, A. (2014). The information panopticon in the big data era. *Journal of Organization Design, 3*(1), 14–19.
20. Laney, D., Hung, L. H., & Lapkin, A. (2013). Big data means big business. *Financial Times*.
21. Fast Company (2014). The world's top 10 most innovative companies in Big Data. https://www.fastcompany.com/most-innovative-companies/2014/industry/big-data.
22. Wagner, D., Vollmar, G., & Wagner, H.-T. (2014). The impact of information technology on knowledge creation: An affordance approach to social media. *Journal of Enterprise Information Management, 27*(1), 31–44.
23. Boulton, C. (2013). Much Hadoop about nothing. *Wall Street Journal, 27*(1), 31–44.
24. Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). *Analytics: The real-world use of big data*. Somers, NY: IBM Global Business Services.
25. Zhang, D. (2013). Granularities and inconsistencies in big data analysis. *International Journal of Software Engineering and Knowledge Engineering, 23*(06), 887–893.
26. Economist. (2010, Febraury 25). A different game: Information is transforming traditional businesses. *The Economist*.
27. Lund, S. (2013). *Game changers: Five opportunities for US growth and renewal*. New York: McKinsey.

28. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., et al. (2011). *Big data: The next frontier for innovation, competition, and productivity*. New York: McKinsey Global Institute.
29. Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review, 84*(1), 98.
30. Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165–1188.
31. Bennett, M. (2014). *Leverage the convergence of BI and Big Data in 2015 or miss out*. Edmonds, WA: Martha Bennett's Blog.
32. McNeely, C. L., & Hahm, J. O. (2014). The big (data) bang: Policy, prospects, and challenges. *Review of Policy Research, 31*(4), 304–310.
33. Agrawal, D., Bernstein, P., Bertino, E., Davidson, S., Dayal, U., Franklin, M., et al. (2012). *Challenges and opportunities with big data A community white paper developed by leading researchers across the United States*. Washington, D.C.: Computing Research Association.
34. Zicari, R. V. (2014). Big Data: Challenges and opportunities. In R. Akerkar (Ed.), *Big data computing* (pp. 103–128). New York: CRC Press, Taylor & Francis Group.
35. Godes, D., Mayzlin, D., Chen, Y., Das, S., Dellarocas, C., Pfeiffer, B., et al. (2005). The firm's management of social interactions. *Marketing Letters, 16*(3–4), 415–428.
36. Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive marketing, 21*(4), 23–45.
37. Alshibly, H., & Chiong, R. (2015). Customer empowerment: Does it influence electronic government success? A citizen-centric perspective. *Electronic Commerce Research and Applications, 14*, 393–404.
38. Davis, A., & Khazanchi, D. (2008). An empirical study of online word of mouth as a predictor for multi-product category e-commerce sales. *Electronic Markets, 18*(2), 130–141.
39. Duan, W., Gu, B., & Whinston, A. B. (2009). Informational cascades and software adoption on the internet: An empirical investigation. *MIS Quarterly, 33*(1), 23–48.
40. Kostyra, D. S., Reiner, J., Natter, M., & Klapper, D. (2015). Decomposing the effects of online customer reviews on brand, price, and product attributes. *International Journal of Research in Marketing, 33*, 11–26.
41. Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science, 29*(5), 944–957.
42. Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing, 70*(3), 74–89.
43. Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter?—An empirical investigation of panel data. *Decision Support Systems, 45*(4), 1007–1016.
44. Liang, T.-P., Ho, Y.-T., Li, Y.-W., & Turban, E. (2011). What drives social commerce: The role of social support and relationship quality. *International Journal of Electronic Commerce, 16*(2), 69–90.
45. Straub, D., Limayem, M., & Karahanna-Evaristo, E. (1995). Measuring system usage: Implications for IS theory testing. *Management Science, 41*(8), 1328–1342.
46. Boyer, J., Frank, B., Green, B., Harris, T., & Van De Vanter, K. (2010). *A practical guide for achieving BI excellence*. New York: IBM Corporation.
47. Balboni, F., & Cook, S. (2012). *Analytics in the boardroom: Accelerating competitive advantage*. Somers, NY: IBM Institute for Business Value.
48. Elbashir, M. Z., Collier, P. A., Sutton, S. G., Davern, M. J., & Leech, S. A. (2013). Enhancing the business value of business intelligence: The role of shared knowledge and assimilation. *Journal of Information Systems, 27*(2), 87–105.
49. McGovern, T. (2015). Big Data Now.
50. Association, C. R. (2012). Challenges and opportunities with big data. http://www.cra.org/ccc/files/docs/init/bigdatawhitepaper.pdf.