CAPSI 2019 Proceedings

Portugal (CAPSI)

10-2019

Unlocking Machine Learning Business Value

Carolina Reis

Pedro Ruivo

Tiago Oliveira

Paulo Faroleiro

Unlocking Machine Learning Business Value

Carolina Reis, NOVA School of Business and Economics, Portugal, 23959@novasbe.pt

Pedro Ruivo, NOVA Information Management School, Portugal, pruivo@novaims.unl.pt

Tiago Oliveira, NOVA Information Management School, Portugal, toliveira@novaims.unl.pt

Paulo Faroleiro, NOVA School of Business and Economics, Portugal, paulo.faroleiro@novasbe.pt

Abstract

Machine learning (ML) stands out as one of the most successful advanced analytics for dealing with big data. However, as a quite recent tool amongst organizations, there are some doubts hanging over this technology. Through an original lens, we expect to substantiate how organizations can sustained ML business value. We developed a conceptual model, grounded on the resource-based view, that aims to validate key antecedents of ML business value. Through a positivist approach, we imply ML use, big data analytics maturity, top management support and process complexity enhance ML business value, in terms of firm performance. Due to the pioneering nature of our research model, we expect to support our data analysis with the partial least squares. To the authors' best knowledge, this represents the first study aiming such findings on the ML discipline.

Keywords: machine learning; business value; resource-based view

1. INTRODUCTION

The big data movement has the power to completely transform businesses and their current position amongst competition (Côrte-Real, Oliveira, & Ruivo, 2017; McAfee & Brynjolfsson, 2012). The contemporaneous technologies within organizations are already collecting more data than ever before (H. Chen, Chiang, & Storey, 2012; LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). However, the ultimate step, transforming all these data into actionable insights, has not been reported with the same success level (Ferraioli & Burke, 2018; LaValle et al., 2011). In fact, leaders around the world are still searching for the best advanced analytics to retrieve data-driven knowledge (LaValle et al., 2011).

Executives from the largest corporations in America seems to have found the answer: machine learning (ML) (Jarrahi, 2018). ML stands out as one of the best-performing advanced analytics for dealing with big data (Dean, 2014; Jimenez-Marquez, Gonzalez-Carrasco, Lopez-Cuadrado, & Ruiz-Mezcua, 2019). This branch of artificial intelligence (Kuzey, Uyar, & Delen, 2014; Michalski, R. S., Carbonell, J. G., & Mitchell, 1983) has emerged due to recent advances in computational aptitudes that dramatically decreased the costs of its algorithms (Agrawal, Gans, & Goldfarb, 2017; Kwak, Lee, Park, & Lee, 2017). However, as a quite recent tool within organizations, ML still raises several questions (Pyle & José, 2015).

Regarding companies' opinion about the ML business value (BV), there can be identified three different clusters. The first group is composed by firms that are either unaware of ML or are still researching about it (Lismont, Vanthienen, Baesens, & Lemahieu, 2017). The second set comprises organizations that already decided to implement ML systems, but until now, it has been a gradual process with low positive outcomes (LaValle et al., 2011; Lismont et al., 2017). The last cluster includes organizations who have fully embraced ML and is already paying off (Economist Intelligent Unit & SAP, 2018; Hernández, Perez, Gupta, & Muntés-Mulero, 2018).

The purpose of this study is to provide organizations with clear evidence on how to sustain ML BV. Through an original lens, we aim to present a theoretical framework that provides relevant insights to attain ML BV. The overall question under examination is: "Which are key antecedents of machine learning business value?".

The remainder paper is organized as follows. In section two, we specify "machine learning", "business value" and "machine learning business value" terminologies alongside our structured literature review. In section three, we deliver an overview of the resource-based view (RBV) - the solid theoretical ground of our model - as well as our proposed research model. Our work culminates with future steps' definition.

2. THEORETICAL BACKGROUND

2.1. Machine Learning

In this paper, ML is defined as a system fuelled by data (Jimenez-Marquez et al., 2019; Thomas, Abraham, & Dapeng, 2018) that uses algorithms but do not rely on rule-based programming (Kaplan & Haenlein, 2019). Its ability to learn and improve (Gollapudi & Phillips, 2016; Jarrahi, 2018) with minimal human intervention (Ben-David & Shalev-Shwartz, 2014; Kwak et al., 2017) enhances pattern recognition (Antons & Breidbach, 2018; Bose & Mahapatra, 2001) which, ultimately, leads to data-driven knowledge that are critical for businesses (Bohanec, Borštnar, & Robnik-Šikonja, 2017; Jarrahi, 2018).

ML can be performed through three types of learning approaches (Jimenez-Marquez et al., 2019; Kaplan & Haenlein, 2019): supervised learning, unsupervised learning and reinforcement learning. Each learning approach may be further divided into distinct ML algorithms. In this particular study, we will focus on the broader concept of the ML discipline.

2.2. Business Value – Firm Performance

The information technology BV has been widely connected to firm's performance (Côrte-real, Ruivo, Oliveira, & Popovič, 2019; Ruivo, Oliveira, & Neto, 2012). In this study, firm performance relates to the increase of sales, profitability and return on investment as well as to growth of customer acquisition and retention (Ren, Wamba, Akter, Dubey, & Childe, 2017; Ruivo, Oliveira, & Neto, 2014).

In the current fast pace environment, it seems exploring and exploiting vast amounts of data flows can radically improve a firm's performance (McAfee & Brynjolfsson, 2012; Vidgen, Shaw, & Grant, 2017). However, in order to accomplish such benefits, organizations must abandon rigid capabilities, that do not deliver a satisfactory return (Grant, 2016), and shift to the innovative technologies that are much more likely to reap data-driven information (Nwankpa & Datta, 2017).

2.3. Machine Learning Business Value

ML BV refers to the organizational performance impact created and sustained by the ML solution. Through pattern recognition, ML seems to provide relevant data-driven insights that are crucial for business differentiation (Bohanec et al., 2017). Not only can ML improve internal processes (Bose & Mahapatra, 2001), such as machinery maintenance (D. Q. Chen, Preston, & Swink, 2015), but it can also redesign products and services to fully match customer's expectations (H. Chen et al., 2012; Singh et al., 2017). Ultimately, this technology seems to drive sustained BV across industries and domains (Jarrahi, 2018; Schreck, Kanter, Veeramachaneni, Vohra, & Prasad, 2018).

To better understand the extent of the ML effect on firm's performance, we developed a structured literature review. All articles from journals amongst "Information Management" and "General Management, Ethics and Social Responsibility" sets within the ABS List 2015 across the spectrum between 2000 and the current time were included. We used the web of science search engine with the feature topic as "Machine Learning". From our initial pool of 1448 potential candidates, we excluded papers not related with ML BV or organizational performance (exclusion based on title, keywords, abstract and full-text content). Also, papers too technical or too focused on the application of ML to a specific case were discarded.

We encountered six articles inside the described scope. Bose's & Mahapatra's work (2001) presented the strengths and weaknesses of ML techniques in the context of data mining. The best outcomes were identified in the finance business area for prediction problems. Cavalcante, Brasileiro, Souza, Nobrega, & Oliveira (2016) provided an in-depth literature review of ML within the financial industry. The goal was to showcase which ML techniques had the greater potential for each financial forecasting job. In turn, Jimenez-Marquez et al. (2019) focus on the ML potential for data mining in the social media context. On a different perspective, we found Parnas's view (2019), where he presented ML has having the same potential as any conventional software. Savage (2012) unveiled the potential of ML image and text recognition within the medicine industry. In Lismont's et al. descriptive survey (2017), ML stood out from all big data analytics in terms of potential to find actionable data-driven insights.

The presented literature review uncovered some gaps amongst literature. To the authors' best knowledge, there are no studies substantiating the important antecedents of ML BV. At the same time, we understand the existing papers tackling the ML BV are limited to an industry, area or task. To bridge the identified gaps, this research intends to provide a conceptual model that exposes the antecedents of ML BV for companies in general.

3. INTEGRATIVE MODEL OF MACHINE LEARNING BUSINESS VALUE

3.1. Resource-Based View

The RBV describes an organization as a bundle of resources and capabilities (R&C), where these R&C determine the firm's ability to achieve and sustain BV (Barney, 1991). If the R&C are valuable, rare, imperfectly imitable and non-substitutable, they will most certainly detect market opportunities and effectively respond to potential threats (Barua, Konana, Whinston, & Yin, 2004).

The RBV is regarded as an important theory to understand the big data and analytics' practices (Vidgen et al., 2017). Advanced analytics, such as ML, represent R&C that, when developed closely to the business processes, can become superior assets (Vidgen et al., 2017) that empower companies with knowledge to surpass competition (Shanks, Gloet, Asadi Someh, Frampton, & Tamm, 2018).

Within past literature, we encountered several research models assessing the process of BV creation of information technologies and information systems that elected the RBV as their theoretical paradigm (Melville, Kraemer, & Gurbaxani, 2004; Ruivo, Oliveira, & Neto, 2015). Following the same line of though, we selected this theory as our baseline.

3.2. Integrative Model and Hypotheses

Following the RBV rationales, we propose the following research model (Figure 1) to investigate potential antecedents of the ML effect on firm performance.

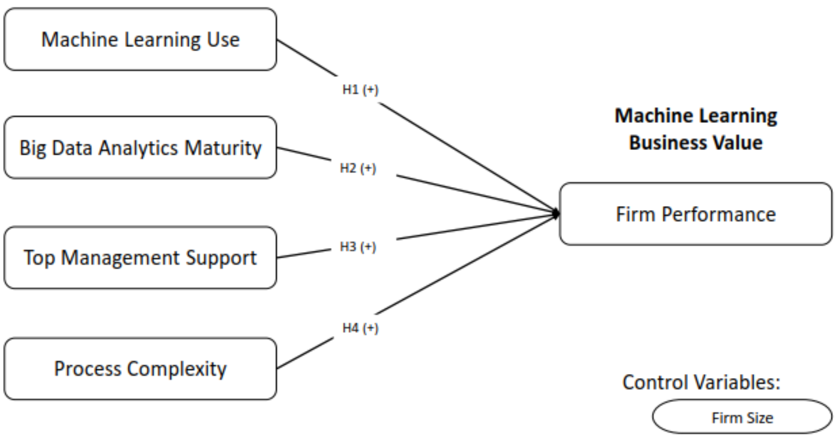

Figure 1 **-** Conceptual model: machine learning business value

Through a positivist approach, we expect four constructs will have a significant effect on the ML BV process. The next postulated hypotheses were based on past literature insights.

H1: ML use has a positive influence on firm performance.

The variable in question gained relevance when past studies disclosed that firms are not retrieving value from the ML solution because they are actually not deploying ML within their business processes (Schreck et al., 2018).

H2: Big data analytics maturity has a positive influence on firm performance.

Researchers have alert that the earlier organizations mature their expertise on big data analytics (Côrte-Real, Ruivo, & Oliveira, 2019), the earlier they will advance to more proficient techniques, such as ML (Lismont et al., 2017).

H3: Top management support has a positive effect on firm performance.

The endorsement of new technologies by senior management in public is presaged to diminish the resistance by workers to new changes in the business processes (Rai, Brown, & Tang, 2009; Ramamurthy, Sen, & Sinha, 2008).

H4: Process complexity has a positive effect on firm performance.

Complexity and uncertainty across business processes open opportunities from ML techniques to identify patterns and associations that may have been overlooked by humans (Antons & Breidbach, 2018; Bohanec et al., 2017).

Regarding our control variable we used firm size, as it is equally handled on previous studies (Rai et al., 2009; Ruivo, Oliveira, Johansson, & Neto, 2013) to capture the variations in data that are not explained by constructs.

4. RESEARCH METHODOLOGY (FUTURE STEPS)

Our logical next step will comprise the development of a questionnaire, where constructs will be based on past literature aiming greater results. Moreover, we expect an expert panel of 5 individuals, with extensive experience in survey development on the information systems area, will review the elaborated content. To assess constructs reliability, we will perform a pilot test with 30 firms and their feedbacks will be taken into consideration. We plan to measure the constructs with reflective items on a seven-Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (7).

With the assistance of SAP, a German software company and one of the world's largest corporations, a large-scale online survey will target corporations with a ML solution in place based in America or Europe in 2019. Due to the nature of the research model and the fact that it has not been yet tested in the past, data will be analysed through partial least squares (Ruivo, Oliveira, & Mestre, 2017; Zhu & Kraemer, 2005).

5. CONCLUDING REMARKS

In this paper, we suggest a theoretical model with grounds on the resource-based view to exploit the integrative value creation process of Machine Learning (ML). With the research question: "Which are key antecedents of machine learning business value?" at hands, we intend to provide relevant guidelines for companies to enhance their ML BV. To the author's best knowledge, this study comprises the first attempt to uncover the antecedents of ML business value. With our goal in mind, we are focus on continuing this course with the development of our questionnaire and then the analysis of the collected data using partial least squares. The research project aims at producing contributions both to theory and practice by introducing a ground-breaking work that provides clear evidence that ML is playing an ever-increasing role within firms' performance.

REFERENCES

Agrawal, A., Gans, J. S., & Goldfarb, A. (2017). What to expect from artificial intelligence. *MIT Sloan Management Review, 58*(3), 23–26.

Antons, D., & Breidbach, C. F. (2018). Big data, big insights? Advancing service innovation and design with machine learning. *Journal of Service Research*, *21*(1), 17–39.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Barua, A., Konana, P., Whinston, A. B., & Yin, F. (2004). An empirical investigation of net-enabled business value. *MIS Quarterly*, *28*(4), 585–620.

Ben-David, S., & Shalev-Shwartz, S. (2014). *Understanding machine learning: From theory to algorithms. Cambridge University Press*.

Bohanec, M., Borštnar, M. K., & Robnik-Šikonja, M. (2017). Explaining machine learning models in sales predictions. *Expert Systems with Applications*, *71*, 416–428.

Bose, I., & Mahapatra, R. K. (2001). Business data mining - a machine learning perspective. *Information & Management, 39*(3), 211–225.

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L. F., Nobrega, J. P., & Oliveira, A. L. I. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, *55*, 194–211.

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems, 32*(4), 4–39.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, *36*(4), 1165–1188.

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of big data analytics in european firms. *Journal of Business Research, 70*, 379–390.

Côrte-Real, N., Ruivo, P., & Oliveira, T. (2019). Leveraging internet of things and big data analytics initiatives in European and American firms: is data quality a way to extract business value? *Information and Management*, *[in press]*.

Côrte-real, N., Ruivo, P., Oliveira, T., & Popovič, A. (2019). Unlocking the drivers of big data analytics value in firms. *Journal of Business Research, 97*, 160–173.

Dean, J. (2014). *Big data, data mining and machine learning. Value creation for business leaders and practitioners* (1st ed.). Wiley Publishing.

Economist Intelligent Unit, & SAP. (2018). *Lessons from fast learners*.

Ferraioli, J., & Burke, R. (2018). Drowning in data, but starving for insights. *Deloitte Insights*.

Gollapudi, S., & Phillips, D. (2016). *Practical machine learning. Tackle the real-world complexities of modern machine learning with innovative and cutting-edge techniques. Packt Publishing.*

Grant, R. M. (2016). *Contemporary strategy analysis: Text and cases.* Chichester, West Sussex, United Kingdom: John Wiley & Sons Ltd.

Hernández, Á. B., Perez, M. S., Gupta, S., & Muntés-Mulero, V. (2018). Using machine learning to optimize parallelism in big data applications. *Future Generation Computer Systems, 86*, 1076–1092.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: human-AI symbiosis in organizational decision making. *Business Horizons*, *61*(4), 577–586.

Jimenez-Marquez, J. L., Gonzalez-Carrasco, I., Lopez-Cuadrado, J. L., & Ruiz-Mezcua, B. (2019). Towards a big data framework for analyzing social media content. *International Journal of Information Management*, *44*, 1–12.

Kaplan, A., & Haenlein, M. (2019). Siri, siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25.

Kuzey, C., Uyar, A., & Delen, D. (2014). The impact ofmultinationality on firm value: A comparative analysis of machine learning techniques. *Decision Support Systems*, *59*, 127–142.

Kwak, C., Lee, J., Park, K., & Lee, H. (2017). Let machines unlearn – Machine unlearning and the right to be forgotten. In *Twenty-third Americas Conference on Information Systems*. Boston.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management, 52*(2), 13.

Lismont, J., Vanthienen, J., Baesens, B., & Lemahieu, W. (2017). Defining analytics maturity indicators: a survey

approach. *International Journal of Information Management*, *37*(3), 114–124.
McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard Business Review*.
Melville, B. N., Kraemer, K., & Gurbaxani, V. (2004). Review: information technology and organizational performance: an integrative model of IT business value, *28*(2), 283–322.
Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (1983). *Machine learning*. *Springer-Verlag Berlin Heidelberg*.
Nwankpa, J. K., & Datta, P. (2017). Balancing exploration and exploitation of IT resources: The influence of digital business intensity on perceived organizational performance. *European Journal of Information Systems*, *26*(5), 469–488.
Parnas, D. L. (2019). More to learn about machine learning. *Communications of the ACM*, *62*(2), 6–7.
Pyle, D., & José, C. S. (2015). An executive's guide to machine learning. *McKinsey Quarterly*.
Rai, A., Brown, P., & Tang, X. (2009). Organizational assimilation of electronic procurement innovations. *Journal of Management Information Systems*, *26*(1), 257–296.
Ramamurthy, K. (Ram), Sen, A., & Sinha, A. P. (2008). An empirical investigation of the key determinants of data warehouse adoption. *Decision Support Systems*, *44*(4), 817–841.
Ren, S. J., Wamba, S. F., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics , business value and firm performance in a big data analytics environment. *International Journal of Production Research*, *55*(17), 5011–5026.
Ruivo, P., Oliveira, T., Johansson, B., & Neto, M. (2013). Differential effects on ERP post-adoption stages across Scandinavian and Iberian SMEs. *Journal of Global Information Management*, *21*(3), 1–20.
Ruivo, P., Oliveira, T., & Mestre, A. (2017). Enterprise resource planning and customer relationship management value. *Industrial Management and Data Systems*, *117*(8), 1612–1631.
Ruivo, P., Oliveira, T., & Neto, M. (2012). ERP use and value: Portuguese and Spanish SMEs. *Industrial Management and Data Systems*, *112*(7), 1008–1025.
Ruivo, P., Oliveira, T., & Neto, M. (2014). Examine ERP post-implementation stages of use and value: empirical evidence from Portuguese SMEs. *International Journal of Accounting Information Systems*, *15*(2), 166–184.
Ruivo, P., Oliveira, T., & Neto, M. (2015). Using resource-based view theory to assess the value of ERP commercial-packages in SMEs. *Computers in Industry*, *73*, 105–116.
Savage, N. (2012). Better medicine through machine learning. *Communications of the ACM*, *55*(1), 17–19.
Schreck, B., Kanter, M., Veeramachaneni, K., Vohra, S., & Prasad, R. (2018). Getting value from machine learning isn' t about fancier algorithms — it's about making it easier to use. *Harvard Business Review*.
Shanks, G., Gloet, M., Asadi Someh, I., Frampton, K., & Tamm, T. (2018). Achieving benefits with enterprise architecture. *Journal of Strategic Information Systems*, *27*(2), 139–156.
Singh, J. P., Irani, S., Rana, N. P., Dwivedi, Y. K., Saumya, S., & Roy, P. K. (2017). Predicting the "helpfulness" of online consumer reviews. *Journal of Business Research*, *70*, 346–355.
Thomas, M. A., Abraham, D. S., & Dapeng, L. (2018). Federated machine learning for translational research. In *Twenty-fourth Americas Conference on Information Systems*. New Orleans.
Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, *261*(2), 626–639.
Zhu, K., & Kraemer, K. L. (2005). Post-adoption variations in usage and value of e-business by organizations: Cross-country evidence from the retail industry. *Information Systems Research*, *16*(1), 61–84.