Role of Big Data and Predictive Analytics in Business organizations: A Review of Current Status and Research Issues

Pavithra K
Department of Computer Science, MGM College, Udupi, Karnataka, India

Abstract

The role of Big Data and predictive analytics in business organizations is set to rise in importance. In today's globalized competitive business environment, companies are facing challenges in dealing with big data issues of rapid decision-making for improved productivity. This paper overviews and consolidates the efforts of various researchers in designing strategies, frameworks and component functionalities of big data analytics for overall firm performance. In this paper, the impact of Big Data and Predictive Analytics on supply chain management, customer relationship management and organizational performance are also discussed. The limitations and scope for future work in this field are also outlined.

Keywords - *Supply Chain Management, Customer Relationship, Organizational Growth, Data Scientist, IoT.*

I. INTRODUCTION

From the academic researchers to corporate leaders, big data are the subject of attention. Big data analytics, which is evolved from business intelligence and decision support systems, enable business organizations to analyze an immense volume, variety and velocity of data from a wide range of customers, suppliers and business partners to support decision making and to create new, data-driven business models. Since big data analytics helps to get insights into market trends, customer buying patterns, product quality assessment, maintenance process, manufacturers, retailers, software companies, and consultants are effectively using predictive analytics with

Big data in supply chain management, customer relationship management and organizational growth.

This paper presents a review of literature on the role and applications of Big Data, Predictive analytics for decision making and success of any organization. In addition, this paper identifies research issues to be tackled in this area.

II. BIG DATA AND PREDICTIVE ANALYTICS

In digital era, enormous amount of data are generated from various sources and processing and inferring from these data have led to growth of big data. Hence big data is characterized by its high volume, velocity and variety. "Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making." (Gartner IT Glossary, n.d.) Similarly, Tech America Foundation defines big data as follows: "Big data is a term that describes large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information." (Tech America Foundation's Federal Big Data Commission, 2012)

The taxonomy of analytics categorizes analytics into three types: descriptive, predictive and prescriptive analytics.

Descriptive analytics takes place either at standardized periods or whenever needed using techniques such as OLAP, and aims at identifying problems and opportunities within existing processes and functions.

Predictive analytics involves the use of mathematical algorithms and programming to discover explanatory and predictive patterns within data, which guides to predict what will happen in the future and also provide reasons as to why it may happen.

Prescriptive analytics involves the use of data and mathematical algorithms to determine and assess alternative decisions that involve objectives and requirements characterized by high volume and complexity, with the aim to improve business performance.

All of the above three categories of analytics can be applied to rapid decision making in achieving organizational growth. But, this paper focuses solely on applications of predictive analytics.

Predictive analytics comprise a variety of techniques that predict future outcomes based on historical and current data. At its core, predictive analytics seek to uncover patterns and capture relationships in data. Predictive analytics techniques are subdivided into two groups. Some techniques attempt to discover the historical patterns in the outcome variable(s) and extrapolate them to the future. Others aim to capture the interdependencies between outcome variable(s) and explanatory variables, and exploit them to make predictions.

Based on the underlying methodology, techniques can also be categorized into two groups: regression techniques (e.g., multinomial logit models) and machine learning techniques (e.g., neural networks). Another classification is based on the type of outcome variables: techniques such as linear regression address continuous outcome variables (e.g., sale price of houses), while others such as Random Forests are applied to discrete outcome variables (e.g., credit status). Figure.1 shows the typical steps involved in Predictive Analytics Process.

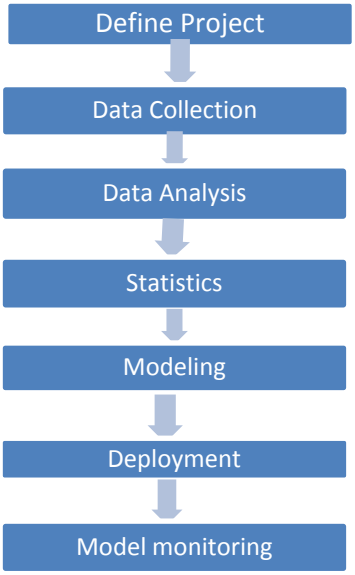

Fig.1: Predictive Analytics Process

III. LITERATURE REVIEW

In the current competitive world, companies should have good measures to handle unpredictable customer demands, varied performance of suppliers, logistics management and even collaboration with competitors. [1] In the rise of maker movement trend, consumer behavior has become an integral part of both production and demand. Internet-based business generates lump sum of data pertain to marketing trends, collaboration with suppliers, collaboration with business partners and competitors, customer buying patterns, reviews from customers etc. Applying data analytics on the collection of these terabytes of data can generate strategies to improve business performance. Resource-based view is followed by [2] to conceptualize Big Data and Predictive Analytics assimilation as a capability that impacts on Supply Chain Management and Organizational Performance. The authors suggested that connectivity and information sharing under the mediation effect of top management commitment are positively related to Big Data and Predictive Analytics acceptance, and positively related to Supply Chain Management and Organizational Performance. This study relied on homogeneous data following survey-based approach. It is clear that a mixed research approach with heterogeneous big data may improve the proposed model and may also offer better insights into Big Data and Predictive Analytics comprehension.

We need strong predictive analytics applications and theory, to predict consumer behavior has implications for product innovation, production, distribution, and demand. But, there is lack of academic research into Big data and predictive analytics in business organizations as well as trained Data Scientists.

Table 1 shows the summary of the survey conducted by [4] with expert interviews from companies promoting the use of predictive analytics companies, which lists out the required skill sets for Data scientists in the field of Supply Chain Management. They also provided an insight into training the next-generation data scientists.

Table I: Expert interviews on desired skill sets

Expert and title	Company	Desired skill sets
Tim Rey, Director of Advanced Analytics	Steelcase	Being able to convert data into business gain; being inquisitive about problems; creativity, having a mathematical slant; statistics; machine learning; operations research
Philip Lear, Manager of Trade Analytics	Kellogg's	Critical thinking; mathematics; programming; however, it is not really only about crunching numbers and getting statistics, but to develop insights from numbers; understanding the business behind it is thus important; passion
David Dorleans, Manager, Advanced Risk and Compliance Analytics	PriceWaterhouseCoopers	Ability to analyze the data, but then also to convey useable results and implications to executives (communication skills)
Mike Marshall, Director of Marketing and Statistical Science	J.D. Power & Associates	Quantitative skill sets, ability to find and see patterns, passion for discovering things; inquisitive mindset; technical capabilities and skill sets
Richard Rodts, Manager of Global Academic Programs for Data Analytics	IBM	Understanding what questions to ask (not necessarily with a big technology background); being able to address business needs; leverage technology to look further into data to facilitate better decisions; communication skills (need to tell a story about why the data matter); mathematics; sociology
Jeremie Juban, Chief Data Scientist, Statistics, Data Mining, Machine Learning	The Weather Company	Being able to spend time with the data, coupled with the desire to understand what is behind the data

Source: https://accounting.broad.msu.edu/welcome/ms-business-analytics/events/

In order to adopt Big Data Analytics, the companies not only need to update technological infrastructure, but also the top management need to constantly improve and update the methodologies and techniques for improved firm performance and growth. The recently developing trend of Internet of Things (IOT) framework tightly connects systems and humans together, which further assists companies to utilize Big Data Analytics more effectively and to improve their productivity. Marketing data sources with the new technologies such as IP address tracking, cookie tracking, and loyalty card usage provide an increased ability to predict customer behavior and the implications of marketing on it.

Predictive Analytics deals mainly with structured data, but unstructured data are the major contents of big data. Therefore, the distinct characteristics of big data such as heterogeneity, noise accumulation and spurious correlations call for developing new statistical methods to comprehend predictive models.

IV. LIMITATIONS AND FURTHER RESEARCH DIRECTIONS

From the study of the existing works, I observed the following limitations and scope for future works: In this review, I focused on literature from only academic journals and it is observed that there is plenty of academic survey-based researches which lack practical implications from practitioners. Another drawback is that these works considered homogeneous data rather than actual heterogeneous big data. Moreover, there is lack of well-trained practitioners working with Big Data Analytics in business organizations; hence more researches form practitioner's point of view is expected.

Researcher can focus on developing a better business model that exploits Big Data Analytics. More importantly, there is a need for introducing methodologies and techniques for robust data collection and data cleansing depending on the stakeholder groups in a business.

V. CONCLUSION

In this paper, I reviewed the literature on Big Data and Predictive Analytics in Business Organizations and explored the applications of Big Data and Predictive Analytics at various levels of business such as supply chain management, retailing, customer retention and management, firm performance. It is observed the need for training data scientists in business organizations to effectively make use of Big Data Analytics.

Further, researchers may consider the differing impacts of having a centralized, distributed, or hybrid structure for successfully promoting Big Data Analytics' use within the organization.

In addition, legal and ethical issues in the use of predictive analytics, as it pertains to consumer data, need to be dealt with.

REFERENCES

[1] S.C. Lenny Koh, Mehmet Demirbag, Erkan Bayraktar, Ekrem Tatoglu, Selim Zaim, The impact of supply chain management practices on performance of SMEs, Industrial Management & Data Systems, Vol. 107(1), 2007, 103-124.

[2] Angappa Gunasekaran, Thanos Papadopoulos , Rameshwar Dubey, Samuel Fosso Wamba, Stephen J. Childe , Benjamin Hazenf, Shahriar Akter, Big data and predictive analytics for supply chain and organizational performance, Journal of Business Research 70 (2017) 308–317.

[3] Matthew A. Waller, Stanley E. Fawcett, Click Here for a Data Scientist: Big Data, Predictive Analytics, and Theory Development in the Era of a Maker Movement Supply Chain, Journal of Business Logistics, 2013, 34(4): 249–252.

[4] Obias Schoenherr, Cheri Speier-Pero, Data Science, Predictive Analytics, and Big Data in Supply Chain Management: Current State and Future Potential, Journal of Business Logistics, 2015, 36(1): 120–132.

[5] Ren, S. Ji-Fan., Fosso Wamba, S., 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), 1-16.

[6] Wang, Gang and Gunasekaran, Angappa and Ngai, Eric W.T. and Papadopoulos, Thanos (2016), Big data analytics in logistics and supply chain management: Certain investigations for research and applications. International Journal of Production Economics,176. pp. 98-110.

[7] D. P. Acharjya, Kauser Ahmed P, Survey on Big Data Analytics: Challenges, Open Research Issues and Tools, International Journal of Advanced Computer Science and Applications, Vol. 7, No. 2, 2016, 511-518.

[8] Amir Gandomi, Murtaza Haider, Beyond the hype: Big data concepts, methods, and analytics, International Journal of Information Management 35 (2015) 137–144.

[9] Jay Lee*, Hung-An Kao, Shanhu Yang, Service innovation and smart analytics for Industry 4.0 and big data Environment, Proceedings of the 6th CIRP Conference on Industrial Product-Service Systems, Procedia CIRP 16 (2014) 3 – 8.

[10] Uthayasankar Sivarajah, Muhammad Mustafa K, Zahir Irani, Vishanth Weerakkody, Critical analysis of Big Data challenges and analytical methods Journal of Business Research 70 (2017) 263–286.