Chapter 93

Integration of Data Mining and Business Intelligence in Big Data Analytics:
A Research Agenda on Scholarly Publications

Atik Kulakli

https://orcid.org/0000-0002-2368-3225
American University of the Middle East, Kuwait

ABSTRACT

The purpose of this chapter is to analyze and explore the research studies for scholarly publication trends and patterns related to the integration of data mining in particular business intelligence in big data analytics domains published in the period of 2010-2019. Research patterns explore in highly prestigious sources that have high impact factors and citations counted in the ISI Web of Science Core Collection database (indexes included SCI-Exp and SSCI). Bibliometric analysis methods applied for this study under the research limitations. Research questions formed based on bibliometric principles concentrating fields such as descriptive of publication, author productivity, country-regions distribution, keyword analysis with contribution among researchers, citation analysis, co-citation patterns searched. Findings showed strong relations and patterns on these important research domains. Besides this chapter would useful for researchers to obtain an overview of publication trends on research domains to be concerned for further studies and shows the potential gaps in those fields.

INTRODUCTION

Data management becomes the center of the decision-making process; Business Intelligence and Analytics (BI&A) has emerged as accessible (Buhl *et al*., 2013) research area to be studied in academia and industry. With the movement of Big Data (BD) initiatives, these popular fields become more attractive

for researchers (Maté *et al*., 2015). Big data could be defined as the most unstructured and extremely massive data sets to be analyzed to reveal trends, associations, and patterns, related to human interactions (e.g., Social media data). In comparison, data mining is structured data in the databases to generate new information by mining the relations in data patterns. The data warehousing platforms should be established incorporation level with necessary data marts to analyze the large data sets. Moreover, big data analytics (as proper business decision tool); is a complex process of examining large and various data sets to uncover information in unstructured platforms with a variety of data sources and systems to discover hidden patterns, correlations, customer insights, and decisions as well as behaviors.

Data Mining (DM) with Big Data relations studied by Wu *et al.* (2014), it has outlined that Big Data concentrates large-volume, multiple sourced, complex, unstructured and autonomous (various) sources from the entire data platforms. In the research, the HACE theorem studied and proposed the Big Data processing model with mining techniques. While Big Data and traditional Business Intelligence discussed in different platforms, researchers studied the skill set and knowledge requirements for Big Data and Business Intelligence (Debortoli, Müller & vom Brocke, 2014) and conducted latent semantic analysis (LSA). According to findings, business knowledge is a crucial indicator of successful development and execution. Followed by Business Intelligence needs bigger than Big Data competencies, whereas Big Data is much more "human-capital-intensive" than Business Intelligence projects. Sun, Sun & Strang (2018) examined to use of Big Data Analytic services (service-oriented architecture-BASOA) to enhance Business Intelligence. Big Data ontology has three-layer analyses, such as descriptive, predictive, and prescriptive.

Since the number of academic publications has been increased and become very complex to follow, the trends would be a critical point for researchers and audiences. The impact of outcomes has resulted in directing research efforts to be concentrated on the most recent development fields in parallel. Scholars conduct different qualitative and quantitative studies to explore, discover, analyze, and organize the information to provide through wider scientific communities. Therefore, bibliometric methods as a usual proven approach provide systematic, transparent, process-based on the statistical measurement of science, and scientific activities (Aria & Cuccurullo, 2017).

This chapter aims to analyze and explore the research studies on big data, data mining in particular business intelligence, and big data analytics domains published in a recent research period (a decade). The study uses the Web of Science Core Collection database as a primary research platform and analysis the data for the period of 2010 to 2019. In this chapter, the author used a science mapping method with open-source Bibliometrix R-package that conducting bibliometric studies on scholarly high ranked sources. The chapter organized based on the IGI Global's structured template which devoted in seven main sections as Introduction, Background (literature review), Main focus (issues, controversies, problem statement, research methodology, bibliometric study definitions, and data collection strategy), Results and Discussions (analysis, findings, addressing the research questions), Solutions and Recommendations followed by Future research directions (Theoretical implications and research needs, Practical implications, and Limitations and future research), and Conclusion. The study also shows the research patterns and concentration areas in scholarly publications to find out the gap to highlight future research direction needs.

Related Content

SBASH Stack Based Allocation of Sheer Window Architecture for Real Time Stream Data Processing
Devesh Kumar Lal and Ugrasen Suman (2020). *International Journal of Data Analytics (pp. 1-21).*
www.irma-international.org/article/sbash-stack-based-allocation-of-sheer-window-architecture-for-real-time-stream-data-processing/244166

Wearable Devices Data for Activity Prediction Using Machine Learning Algorithms
Lakshmi Prayaga, Krishna Devulapalli and Chandra Prayaga (2019). *International Journal of Big Data and Analytics in Healthcare (pp. 32-46).*
www.irma-international.org/article/wearable-devices-data-for-activity-prediction-using-machine-learning-algorithms/232334

Prediction of the Quality of Fresh Water in a Basin
Sira M. Allende, Daniel C. Chen, Carlos N. Bouza-Herrera, Agustin Santiago and Jose Maclovio Sautto (2017). *Handbook of Research on Intelligent Techniques and Modeling Applications in Marketing Analytics (pp. 351-367).*
www.irma-international.org/chapter/prediction-of-the-quality-of-fresh-water-in-a-basin/170357

Ontology-Based IoT Healthcare Systems (IHS) for Senior Citizens
Sakshi Gupta and Umang Singh (2021). *International Journal of Big Data and Analytics in Healthcare (pp. 1-17).*
www.irma-international.org/article/ontology-based-iot-healthcare-systems-ihs-for-senior-citizens/287604

A Machine Learning-Based Exploration of Relationship Between Security Vulnerabilities of IoT Devices and Manufacturers
Ritu Chauhan and Gatha Varma (2020). *International Journal of Data Analytics (pp. 1-12).*
www.irma-international.org/article/a-machine-learning-based-exploration-of-relationship-between-security-vulnerabilities-of-iot-devices-and-manufacturers/258917