REVIEW ARTICLE

RSP Science Hub

International Research Journal on Advanced Science Hub
2582-4376
Vol. 05, Issue 07 July
www.rspsciencehub.com

Check for updates

http://dx.doi.org/10.47392/irjash.2023.046

Big Data Analytics for Supply Chain Optimization: A Review of Methodologies and Applications

Shubham Agarwal[1], *Nishant Moghe*[1], *Vaishali Wadhe*[1]

[1]*Department of Artificial Intelligence and Data Science (AI & DS), K J Somaiya Institute of Technology, Mumbai - 400022, India*

Emails: 2nishantmoghe20@gmail.com, 3vwadhe@somaiya.edu

Article History

Received: 5 June 2023
Accepted: 12 July 2023
Published: 28 July 2023

Keywords:

Big data;
Supply Chain;
Optimization;
Analytics;
Demand Forecasting;
Optimized Logistics

Abstract

Supply chain and its management play a critical role in determining how the business or the organisation performs in the mentioned criteria. With the growth of Big Data Analytics (BDA), most institutions have started to utilise large amounts of data in the form of datasets, databases, and other sources from which data can be collected. This paper focuses on the topic of Supply Chain Optimization and the detailed process of how it can be achieved. The paper first traverses through the topic of Big Data and how it is utilised in Supply Chains and therefore looks at various ways in which data affects Supply Chain Management. Further in the paper we address the topic of how Big Data benefits the optimization of Supply Chains and what limitations exist in the implementation. Lastly, the paper covers the improvements that could be brought into the field, as well as the future scope of Big Data Analytics in Supply Chain Optimization.

1. Introduction

Traditional supply chain management approaches often relied on historical data and manual decision-making processes, which were prone to inefficiencies and suboptimal outcomes. As supply chains become increasingly complex and interconnected, organizations face the challenge of optimizing their operations to meet customer expectations while minimizing costs and risks. Supply chain optimization is crucial for organizations seeking to gain a competitive edge in today's complex and dynamic business environment (Ahmad and Kamruzzaman). The abundance of data generated by different supply chain components offers an opportunity to leverage big data analytics for enhanced decision-making and performance improvement. By exploring various analytical techniques and real-world examples, the paper seeks to highlight the potential benefits and challenges of leveraging big data analytics in supply chain management. This review paper critically examines the methodologies and applications of big data analytics in supply chain optimization, shedding light on the ways in which organizations can harness data-driven insights to gain a competitive edge.

2. Overview of Big Data

In the contemporary world, valuable data which is not apparent to many, exists in various formats including structured, semi-structured, and unstructured data. This data is generated from various sources (social media, satellites, scanners, sensors, public repositories, etc.) and is generated in colossal amounts - known as Big Data. Generally, Big Data is characterized by its Volume, Velocity, Variety, and Veracity. Management and efficient utilization of this data is not straightforward and requires

the utilization of techniques such as distributed storage systems and parallel processing.

Characteristically, Big Data Analytics is being widely adopted across a diverse range of industries to help businesses become more profitable, sustainable and substantial. (Roy and A) Big Data Analytics comprises various steps which, if and when performed duly, formulate substantial results and help uncover invaluable insights. These steps include identifying the problem statement, identifying the suitable sources for data acquisition, data cleaning, visualization, reporting and dashboarding. There are other tasks that may or may not be required based on the particular scenario. In summary, Big Data involves various processes which require expertise of personnel involved in the same (Shadi et al.)

3. Outline of Supply Chain

The management of the supply chain has gone through a transformation, progressing from traditional, linear processes to complicated and interconnected systems. The former supply chain process was sequential, starting with procurement and culminating with product delivery, but had its limitations, including deficient visibility, obsolete data, and disjointed communication. Problems arose, such as inaccurate prediction of demand, surplus inventory, long lead times, and fragmented communication. However, with the emergence of globalization, advancements in technology, and evolving customer expectations, modern-day supply chains have become increasingly intricate. They include numerous levels of suppliers, dynamic demand trends, diverse products, and complex distribution networks. This complexity creates greater uncertainty, heightened risk exposure, and obstacles in managing the end-to-end supply chain efficiently. (Lee and Mangalaraj) Data is produced throughout the supply chain, providing valuable insights for optimization and decision-making.

Procurement generates data on supplier performance, pricing, and contract terms, while production generates data on machine performance and quality control. Transportation and logistics generate data on delivery times and carrier performance, and customer-relationship management generates data on order history and customer preferences. Using this data for supply chain optimization is made possible through advanced technologies such as IoT devices and sensors, allowing for real-time data capture and analysis. Big data analytics offers an opportunity to study vast amounts of structured and unstructured data, (Pan, Luo, and Fu) providing insights for demand forecasting, inventory management, transportation optimization, and overall performance improvement in the competitive business landscape of today.

4. Methodology

Big data analytics techniques in supply chain optimization include different phases and techniques aimed at harnessing the power of data to improve supply chain operations. These methods can be categorized into data acquisition and integration, data preprocessing and cleansing, descriptive analytics, predictive analytics, and prescriptive analytics.

Data collection and integration collects data from various sources within the supply chain, such as procurement, production, logistics, and customer interactions. This data may include information about supplier performance, production rates, inventory levels, transportation routes and customer preferences. Integrating this data into a central system gives you a complete picture of your supply chain, enabling better decision-making and analysis. Data preprocessing and cleaning are important steps to ensure the accuracy and reliability of the data collected. These include techniques such as data normalization, outlier detection, and missing data handling. Eliminating discrepancies and errors makes the data suitable for further analysis and modelling.

Descriptive analytics focuses on analyzing historical data to gain insight into how the supply chain has performed in the past. It uses techniques such as data visualization, statistical analysis, and data mining to uncover patterns, trends, and relationships. Descriptive analytics provide a comprehensive understanding of the current state of your supply chain and help identify areas for improvement.

Predictive analytics uses historical data to build models and make predictions about future events and outcomes in the supply chain. We use techniques such as time series analysis, regression and machine learning algorithms to forecast demand, identify potential disruptions and optimize inventory. Predictive analytics enable proactive decision-making and improve supply chain responsiveness.

Prescriptive analytics go beyond prediction to

provide recommendations for optimal decision-making. Determine the best course of action in various supply chain scenarios using advanced optimization techniques, simulation models and algorithms. Prescriptive analytics enable companies to optimize factors such as production planning, inventory allocation, transportation routes, and supplier selection.

By using these techniques for big data analytics in supply chain optimization, companies can gain a competitive advantage by increasing efficiency, reducing costs, increasing customer satisfaction and reducing risk. . Data integration, preprocessing and cleansing, descriptive analytics, predictive analytics, and prescriptive analytics work together to provide a comprehensive approach to optimizing supply chain operations in the era of big data.Working of a Supply Chain shown in fig 1.

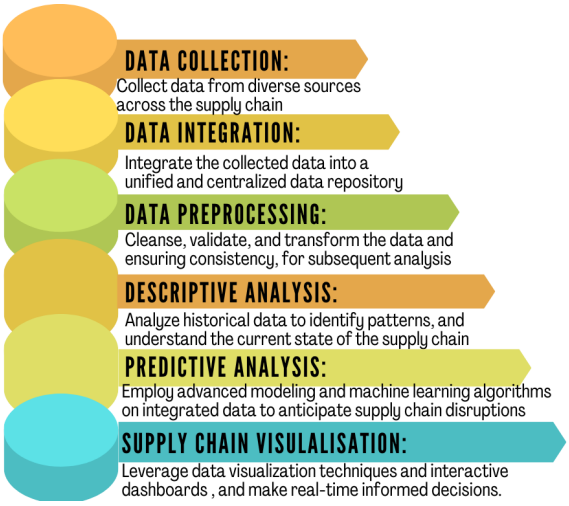

FIGURE 1. Working of a Supply Chain

5. Application of Big Data Analytics in Supply Chain Management

Supply Chain Management has been a mammoth task for many companies over the years, especially ones at the inception of their journeys. Big Data Analytics is used to make the process more streamlined (Thakur) It is used to specifically address various aspects of a supply chain, including demand forecasting, inventory optimization, supplier relationship management, route optimization, risk management, and Real-time Tracking of resources or products.

5.1. Demand Forecasting

Demand forecasting is the process of accurately predicting and reporting the forthcoming demand of products that the business or organization provides. This is done in various ways and numerous factors are considered for the same which may include historical data of the market trends, factual performance, social media sentiment, overall sales data and consumer feedback.

5.2. Inventory Optimization

Inventory optimization is one of the most crucial aspects of a supply chain which is attended to using Big Data. It enables organizations to analyze humongous volumes of empirical data. This data can be used to gather insights on sales of various products, customer behavior, performance of the supply chain, etc. Predictive analysis is used to identify any trends existing in the data which are invisible to the naked eye. (Wang et al.)

5.3. Supplier Relationship Management

It is vital for businesses and organizations to keep track of the performance of the suppliers of their products and resources. This tracking enables organizations to assess potential risks and take measures to avoid these risks from transforming into catastrophes. Integration of data from various sources done after the proper cleaning and validation of the same is critical. This data can be used for business activities such as supplier selection and other types of negotiations. (Shafique et al.)

5.4. Optimized Logistics

Transportation and logistics are critical aspects and factors which heavily influence a product's and subsequently, the business's success. To optimize a process that takes place on such a large scale, it is important to consider factors such as weather conditions and patterns, target locations and destinations, optimal vehicles, traffic conditions, and the most package weight. This can help the organization minimize delivery costs and improve customer service.

5.5. Real-Time Tracking of Resources and Products

Characteristically, logistics require tracking of packages, especially before a professional relationship is established between the multiple entities involved.

Big Data helps businesses monitor their operations in real-time. By integrating data from sources such as scanners, RFID tags and other IoT devices, organisations can collect data regarding the real-time temperatures, location, and overall condition of the package. Such data can be analyzed to gain the trust of consumers and make this data visible to them. (Juma and Kilani)

5.6. Market Analysis

Big Data helps businesses gain valuable information about the real-time market trends, geopolitical situations, and other factors in order to avoid risks related to phenomena such as natural disasters, and political unrest.

6. Analyzing Amazon's Supply Chain Management

6.1. Before using big data analytics:

Before implementing big data analytics, Amazon's supply chain faced many challenges. The company struggled with inventory management, and limited visibility into demand patterns often resulted in shortages and overstocks. Inefficient transportation routes and logistics led to longer delivery times and higher costs. Additionally, the lack of real-time data integration hindered decision-making and collaboration among stakeholders.

6.2. After using big data analytics:

By implementing big data analytics, Amazon transformed its supply chain, achieving significant increases in efficiency, cost savings, and customer satisfaction (Hegde and Gray). The use of advanced analytics techniques and real-time data integration enabled the following improvements:

6.2.1. Demand forecasting and inventory management:

By leveraging big data analytics, Amazon was able to improve the accuracy of its demand forecasts. Insights into historical sales data, market trends, and customer behavior enabled more accurate inventory. Study of Amazon's Supply Chain Optimization planning. This resulted in a 30% reduction in out-of-stocks and a 20% reduction in overstock. This optimization reduced storage costs by 15

6.2.2. Optimizing transportation and logistics:

Big data analytics played a key role in optimizing Amazon's transportation and logistics operations. By analyzing real-time data from GPS, sensors and traffic information, the company cut delivery times by 25%. Transport costs increased by 15% through dynamic route decisions and optimized cargo consolidation based on traffic conditions.

6.2.3. Supplier management and collaboration:

Amazon uses big data analytics to improve supplier management and collaboration efforts. By analyzing supplier performance data, quality metrics and customer feedback, we identified high-performing suppliers and implemented effective risk mitigation strategies. Supplier performance increased by 20% and supplier disruption decreased by 25%. Optimizing warehouse operations and operations: Amazon was able to use big data analytics to streamline warehouse operations and labor management. By analyzing data from its warehouse management system, the company identified bottlenecks, optimized picking and packing processes, and improved warehouse layouts. This optimization increased warehouse productivity by 30% and reduced labor costs by 10%. (Ghosh) . Study of Amazon's Supply Chain Optimization shown in fig 2.

7. Limitations of this Approach

Data has numerous applications in supply chain management which can prove to be major advantages for the same. However, it is crucial to consider some of the limitations, complexities and potential disadvantages of this approach of managing the supply chain. The limitations can arise due to various characteristics of BDA, like its cost, complexity of implementation, data quality, integration, privacy risks, and overreliance on data and technology.

The implementation of BDA in supply chain management is a tedious task. It has two major obstacles – cost and complexity. BDA has a variety of preconditions, some of them include technology infrastructure, analytical tools, personnel for implementation and analysis and medium for data storage (databases). Big Data cannot be analyzed out of the box by just anyone, it requires skilled personnel specialised in building statistical and analytical models using various Machine Learning techniques. The complexity arises due to the data's velocity, veracity, volume and value. Big data is often generated at large speeds and volumes, it becomes increasingly complex to deal with as the amount of data increases. Veracity refers to the readability of the data. Big Data is generated and collected in various

Supply Chain Aspect	Improvement
Demand Forecasting	30% reduction in out-of-stocks and 20% reduction in excess inventory
Inventory Management	Decreased Carrying costs by 15%
Transportation	Reduce delivery times by 25% and transportation costs by 15%
Supplier Management	Improved supplier performance by 20% and reduced supplier disruptions by 25%
Warehouse Operations	Warehouse productivity increased by 30% and labor costs were reduced by 10%

FIGURE 2. Study of Amazon's Supply Chain Optimization

forms which are not always readable or ready for analysis. The complexity is, therefore, also commensurate to the level of analysis required.

Data exists in all forms across a multitude of platforms. As a result, the quality of data is not always satisfactory. The data needs to be cleaned and integrated properly to ensure that the data is clean, consistent and devoid of inaccuracies. The task of cleaning has a low margin of error, if done incorrectly it can lead to loss of important data and insights. Furthermore, the integration of data requires personnel skilled in technologies like SQL server, MongoDB and other database services. Poor data can lead to incorrect insights and inaccurate decision-making, making the implementation of BDA complex. (Khattab et al.)

Security risks are a big factor to consider while implementing BDA for a supply chain. The data associated with a supply chain is often sensitive, as it may include customer details, contract information, cell-phone numbers and other intellectual property. To protect this data, proper measures need to be established and implemented. Inadequate measures may lead to damage to the company's reputation and monetary resources caused by regulatory non-compliance and data breaches. (Neustel)

The technology infrastructure used for the implementation of BDA needs to be robust to handle all types of data available. This includes various aspects of it including servers, computational power, and storage capacity. BDA is heavily reliant on the available technology and data. Any type of failure or disruption in the architecture may lead to incorrect analysis, loss of data and inefficient management of the supply chain. The crux of BDA lies in data-driven decision-making. This is a highly efficient method of making decisions as it removes the possibility of personal bias. However, it is difficult for statistical models to compute the context of the data. They are unable to use intuition. It is crucial to attain a balance between data-driven decision-making and judgment of human professionals to obtain optimized results.

8. Prospects for innovation and growth

Emerging trends and future research directions in the field of big data analytics for supply chain optimization include several exciting areas with the potential to shape the future of supply chain management. A key trend is the integration of artificial intelligence (AI) and machine learning (ML) techniques. Researchers can focus on developing advanced ML algorithms that can process rich supply chain data and leverage AI technologies like natural language processing and computer vision to extract meaningful insights from unstructured data sources. Additionally, exploring the potential of reinforcement learning and autonomous decision-making systems may further improve the supply chain optimization process. (Shi et al.)

Another emerging trend is the widespread adoption of Internet of Things (IoT) devices and sen-

sor technology. Future research may address the integration of IoT-enabled devices and sensor networks to enable end-to-end visibility and traceability across supply chains. This includes exploring data fusion techniques, examining data security and privacy issues, and developing analytical models to harness the potential of IoT-generated data for predictive and prescriptive analytics.

Blockchain technology also holds promise when it comes to optimizing supply chains. Researchers can explore the application of blockchain to ensure data integrity, traceability and reliability in supply chain transactions. Research areas include supplier verification, commodity traceability, smart contracts for automated trading, and improving supply chain resilience. Additionally, exploring the integration of blockchain with other emerging technologies such as IoT and AI could open up new opportunities for data sharing and collaboration. (Seyedan and Mafakheri)

Focusing on sustainable and resilient supply chains is another important direction for future research. As sustainability and resilience become key factors in supply chain management, it is important to develop frameworks and models that consider environmental, social and economic dimensions. This includes integrating sustainability metrics into optimization algorithms, analyzing the impact of disruptions, developing strategies to build resilient supply chains, and supporting sustainable and ethical supply chain practices. Includes consideration of the role of big data analytics.

Ethical and legal implications are also gaining importance in the area of big data analytics for supply chain optimization. Researchers can explore frameworks for responsible data use, develop privacy-preserving technologies, and ensure compliance with data protection regulations. Additionally, considering ethical considerations when using AI and ML algorithms in decision-making processes can contribute to responsible supply chain practices.

By embracing these emerging trends and conducting further research in these directions, supply chain professionals and researchers can unlock new possibilities, improve decision-making processes, and drive innovation in supply chain optimization. can be promoted. These areas offer exciting prospects for advancing the potential of big data analytics and reshaping the future of supply chain management. (Sun and Huo)

9. Conclusion

In conclusion, the use of big data analytics to supply chain optimization offers the possibility of fundamentally altering current supply chain procedures. Organizations may make wise decisions to improve operational efficiency and promote supply chain optimization by utilising cutting-edge technologies. Additionally, big data analytics offers solutions through real-time data integration, predictive analytics, and enhanced visibility to address issues such as limited visibility, inaccurate demand forecasts, inefficient inventory management, and disconnected collaboration between stakeholders. The report also highlighted emerging trends and future research directions in this area. Integrating artificial intelligence and machine learning, embracing the Internet of Things and sensor technology, exploring the potential of blockchain, focusing on sustainable and resilient supply chains, and tackling ethical and legal implications, was identified as a key area for further research. Advancing research in these directions will open up new possibilities in the field of big data analytics for supply chain optimization and may drive future innovations. In summary, big data analytics has the potential to reshape the landscape of supply chain management. Applications in demand forecasting, inventory management, supplier management, transportation optimization, warehousing, and supply chain visualization have demonstrated significant business benefits. By harnessing the power of big data analytics, companies can optimize their supply chains, improve operational efficiency and gain a competitive advantage in today's dynamic business environment. Insights gleaned from this review provide a foundation for further research and development in the area of big data analytics for supply chain optimization, paving the way for a more efficient and responsive supply chain ecosystem.

References

Ahmad, S and M Kamruzzaman. "Optimization in apparel supply chain using Artificial Neural Network". *European Journal of logistics, Purchasing and Supply Chain Management* 10.1 (2022): 1–14.

Hegde, Chiranth and K E Gray. "Use of machine learning and data analytics to increase drilling efficiency for nearby wells". *Journal of Natural Gas Science and Engineering* 40 (2017): 327–335.

Juma, Luay and Sona Kilani. "Adoption enablers of big data analytics in supply chain management practices: the moderating role of innovation culture". *Uncertain Supply Chain Management* 10.3 (2022): 711–720.

Khattab, Shadi, et al. "The relationship between big data analytics and green supply chain management by looking at the role of environmental orientation: Evidence from emerging economy". *Uncertain Supply Chain Management* 10.2 (2022): 303–314.

Lee, In and George Mangalaraj. "Big Data Analytics in Supply Chain Management: A Systematic Literature Review and Research Directions". *Big Data and Cognitive Computing* 6.1 (2022): 17–17.

Pan, Qiaohong, Wenping Luo, and Yi Fu. "A csQCA study of value creation in logistics collaboration by big data: A perspective from companies in China". *Technology in Society* 71 (2022): 102114–102114.

Roy, A and A. "The emerging era of Big Data Analytics". *Big Data Analytics* 1 (2016): 1–1.

Seyedan, Mahya and Fereshteh Mafakheri. "Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities". *Journal of Big Data* 7.1 (2020): 53–53.

Shadi, Khattab, et al. "The relationship between big data analytics and green supply chain management by looking at the role of environmental orientation: Evidence from emerging economy". *Uncertain Supply Chain Management* 10.2 (2022): 303–314.

Shafique, M N, et al. "The Role of Big Data Predictive Analytics and Radio Frequency Identification in the Pharmaceutical Industry". *IEEE Access* 7 (2019): 9013–9021.

Shi, C, et al. "Swarm intelligence in big data analytics". *Intelligent Data Engineering and Automated Learning* (2013): 417–426.

Sun, Zhaohao and Yanxia Huo. "The Spectrum of Big Data Analytics". *Journal of Computer Information Systems* 61.2 (2021): 154–162.

Thakur, V K. "Review Paper on Big Data Analytics". *International Journal for Research in Applied Science and Engineering Technology* 8.6 (2020): 785–788.

Wang, Gang, et al. "Big data analytics in logistics and supply chain management: Certain investigations for research and applications". *International Journal of Production Economics* 176 (2016): 98–110.

To cite this Article: Agarwal, Shubham, Nishant Moghe, and Vaishali Wadhe. "**Big Data Analytics for Supply Chain Optimization: A Review of Methodologies and Applications** ." International Research Journal on Advanced Science Hub 05.07 July (2023): 215–221. http://dx.doi.org/10.47392/irjash.2023.046