sciendo

Predictive Analytics Functionalities in Supply Chain Management

Elena PUICA*
Bucharest University of Economic Studies, Bucharest, Romania[1]
**Corresponding author, elenaa.puica@gmail.com*

Abstract. *This scientific paper presents a comprehensive analysis of the capabilities of IT solutions for predictive analytics in supply chain management. The study uses a multi-method approach, including a literature review, case studies by applying a machine learning model to technology solutions currently available on the market. The study examines the various software and technology platforms available today and their key features and functionalities, focusing in particular on Scripting, Data Mining, Algorithms, Data Analysis, Modeling, Data Interaction, Data Visualization, Reporting and Data Unification. The study also assesses the potential benefits and challenges associated with implementing these solutions in supply chain management. The results of the study provide valuable information for supply chain professionals, IT managers and researchers interested in the application of predictive analytics in this field. In addition, the paper also discusses the current trend and future direction in this field.*

Keywords: Predictive Analysis in SCM1; Predictive Analytics Efficiency2; Technology efficiency3; Predictive analytics software4

Introduction

Successful and top-performing supply chains are cost-effective and time-efficient, they possess three very different qualities. First, large supply chains are agile. It reacts quickly to sudden changes in demand or supply. Second, it adapts over time as market structures and strategies evolve. Third, they align the interests of all firms in the supply network so that companies optimize chain performance when they maximize their interests. Only supply chains that are agile, adaptable and aligned provide companies with a sustainable competitive advantage.[1] Predictive Analytics includes numerous statistical and analytical methods for developing new technological approaches in supply chains. Therefore, Predictive Analytics becomes vital when a significant amount of highly sensitive data needs to be manipulated to predict future events and make recommendations called prescriptive analytics.

In this article, the results of a case study on the effectiveness of predictive analytics IT solutions in supply chain management (SCM) are presented, further complemented with the capabilities of the technology in SCM. As such, the paper provides an insight into its future potential regarding the best software technologies used in SCM in terms of Predictive Analytics. The study concludes with a conclusion on future research.

Predictive analytics in supply chain management

Predictive analytics in logistics uses both quantitative and qualitative methods to estimate the past and future behavior of material flow and storage, as well as associated costs and service levels. It is also used to improve supply chain design and competitiveness by estimating past and future levels of business process integration between functions or companies, and associated costs and service levels.

[1] *Economic Informatics Doctoral School*

DOI: 10.2478/picbe-2023-0090
© 2023 E. Puica, published by Sciendo.
This work is licensed under the Creative Commons Attribution 4.0 License.

Supply chain managers must identify techniques to gain competitive advantage to thrive, for example, the ability to predict consumer behaviors, prevent fraud, mitigate risk, identify new customers and improve operations, predict when machines will fail, increase customer satisfaction and customer loyalty in real time. Predictive Analytics software solutions in supply chain management were analyzed based on nine capabilities such as: Scripting, Data Mining, Algorithms, Data Analysis, Data Interaction Modeling, Data Visualization, Reporting, and Data Unification. Data mining applies advanced statistical techniques to uncover meaningful patterns, trends, and correlations, with the overall goal of gathering previously unknown information about data sets and transforming the data for further use.

Algorithms take the data through a sequence of steps to calculate predictive results. Effectively, predictive analytics algorithms automate the process of discovering insights in data. Predictive analytics algorithms are divided into supervised learning algorithms and unsupervised learning algorithms. Supervised learning algorithms involve an iterative process of learning from the (historical) training dataset. The training dataset contains labels and has the correct information (ie, the information the model learns to predict). The use of predictive analytics is to predict future results based on past data.

Data analytics software expresses both hopes and fears. The hopes are that it helps automate, thus speeding up and animating the coding process; to provide a more complex way of looking at relationships in data; provide a formal structure for writing and storing reports for analysis development; and helps to think more conceptually and theoretically about data.[2] The process of data analysis, or alternatively, the steps of data analysis, involves gathering all the information, processing it, exploring the data, and using it to find patterns and other insights. The process consists of gathering data requirements: the purpose of the data analysis, what type of data analysis is used, and what data is planned for analysis. Data collection: includes case studies, surveys, interviews, questionnaires, direct observation and focus groups. Data cleaning is the process by which white spaces, duplicate records and basic errors are removed. Data cleaning is mandatory before submitting the information for analysis.

Data visualization "graphically displays information in a way that people can read and understand." Charts, graphs, maps, bullet points or a number of other methods can be used. Visualization helps gain valuable insights by helping to compare sets of data and see relationships between them.

Many aspects of data unification must process complex concepts, such as understanding the original text data, creating a list of unified categorical values, and judging data quality.

Modeling is the process of taking (usually) data and constructing a simplified description of the processes that might have generated it. The description is often a computer program or mathematical formula. A model captures the knowledge presented by the data and encodes it in a language. Often the goal is to address a specific problem by modeling the world in some form and then use the model to develop a better understanding of the world.[3] (see fig. 1.)

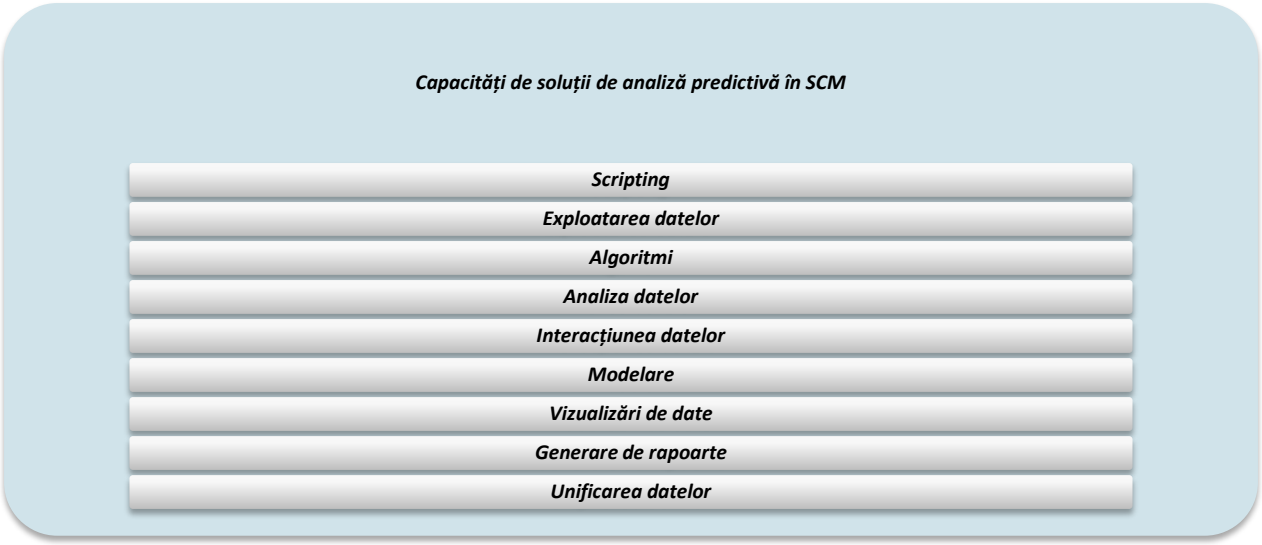


Figure 1. Predictive Analytics software capabilities

Assessing the capabilities of predictive analysis in Supply Chain Management

a. Understanding the problem

The growing number of technologies that are becoming more and more advanced make it difficult for Supply Chain Management representatives to choose the best software solution that meets the expected needs in the most efficient way. There is a risk that the adoption of certain Supply Chain Predictive Analytics IT solutions will not meet business expectations and will not result in process efficiencies. The factors that can determine this risk are multiple, and not considering several features of IT solutions before adopting them can lead to poor performance in business processes. Therefore, the objectives of the case study are to determine the most effective software technologies for predictive analytics used in Supply Chain.

b. Understanding the data

The data used in this case study was extracted from G2.com and includes indicators that measure the effectiveness of Predictive Analytics technology solutions, as well as the most important features of the software solution required for Supply Chain Management processes. The most important features of predictive analytics in Supply Chain Management are: "Scripting", "Data Mining", "Algorithms", "Data Analysis", "Data Interaction", "Modeling", "Data Visualizations", "Generating reports' and 'Data unification'. ". A score of 0 to 10 was determined from G2.com, where a score of 0 means that the capability of the respective software does not meet expectations at all, and a score of 10 means that the feature is representative of the software.

c. Data preparation

We used two data sets, the first data set includes the six indicators that analyze the best predictive analytics solutions in Supply Chain Management in terms of their effectiveness. And the second data set covers the most important features of predictive analytics solutions required for a successful Supply Chain. The target variable is predictive analytics software solutions. In this study, 28 Predictive Analytics software solutions were analyzed.

d. Modeling

Auto-ML was applied in this case study offered by the Azure Machine Learning computing solution to train and implement an analysis model of the most effective predictive analytics software solutions and the best features of the predictive analytics software solutions used in

Supply chain management. In this model, the classification algorithm technique was adopted to predict between multiple categories, a supervised learning where models learn using training data and apply those learnings to new data.

e. Evaluation

AUC is the area under the receiver operating characteristic curve. The closer it is to 1, the better the model, the range is [0, 1]. Accuracy is the ratio of predictions that exactly match the true class labels, the closer to 1, the better the model, range is [0, 1]. Accuracy is the ability of a model to avoid labeling negative samples as positive. The closer it is to 1, the better the model is, it has a range of [0, 1]. Recall is the ability of a model to detect all positive data. The closer it is to 1, the better the model is, it has the range [0, 1]. Accepted values include: macro, calculating the values for each label and finding their unweighted average; micro, calculating global values by counting the total of true positives, false negatives, and false positives; weighted mean, arithmetic mean of the score for each class, weighted by the number of true instances in each class.

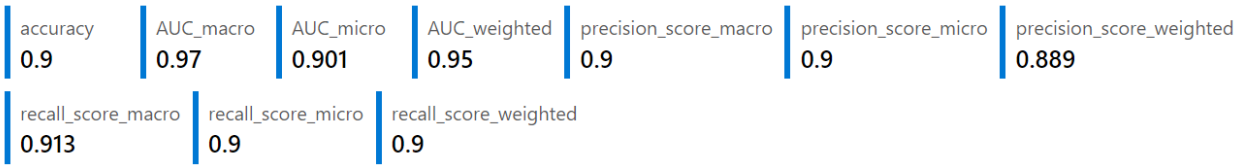


Figure 2. Model value results

Evaluation of the results for the trained model for the case study, proved to be a good model, the resulting values for metrics are as follows, the value for AUC is above 0.9, accuracy is 0.9, precision is around 0.9 and " recall' also has the value around 0.9, which indicates that it is a good model. (see Error! Reference source not found.).

f. Explanation

Global feature importance quantifies the relative importance of each feature in the test data set as a whole. It provides an overall comparison of the extent to which each feature in the data set influences the prediction. The classification model predicted the capabilities of predictive analytics software in Supply Chain Management in terms of technology capabilities. "Data Analysis" is the important feature, followed by "Data Visualization", "Data Unification" and finally "Data Mining". (see fig. 3)

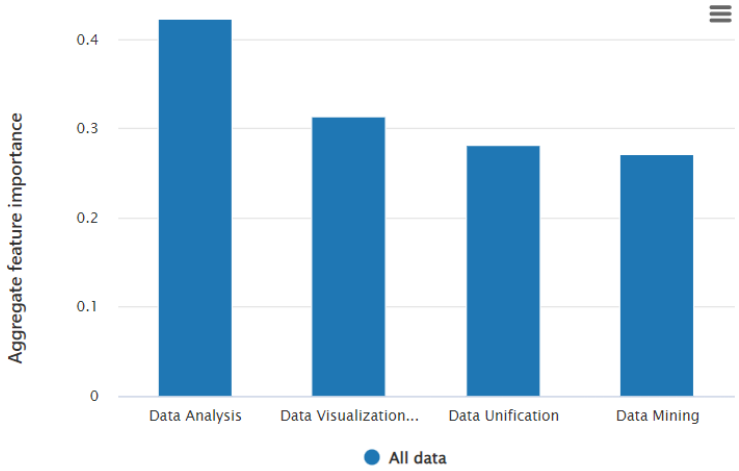


Figure 3. Global importance of features in datasets

Data analytics can be used to understand patterns and relationships in supply chain data, such as identifying which suppliers are the most reliable or which products are the most profitable. By using statistical techniques and models to analyze historical data, organizations can gain valuable insights into the supply chain process, such as identifying which products are the most profitable or which suppliers are the most reliable. For example, data analytics capabilities can be used to identify patterns in sales data to predict product demand. This can help organizations optimize their inventory management strategies and plan production schedules.

For Predictive Analytics technology, the best software solutions related to data analysis capabilities are Qlik AutoML with a score of 9.70, Dash with a score of 9.50, SAS Advanced Analytics with a score of 9.30 and the same score of 9.20 IBM SPSS Modeler and RapidMiner. (see Fig. 4)

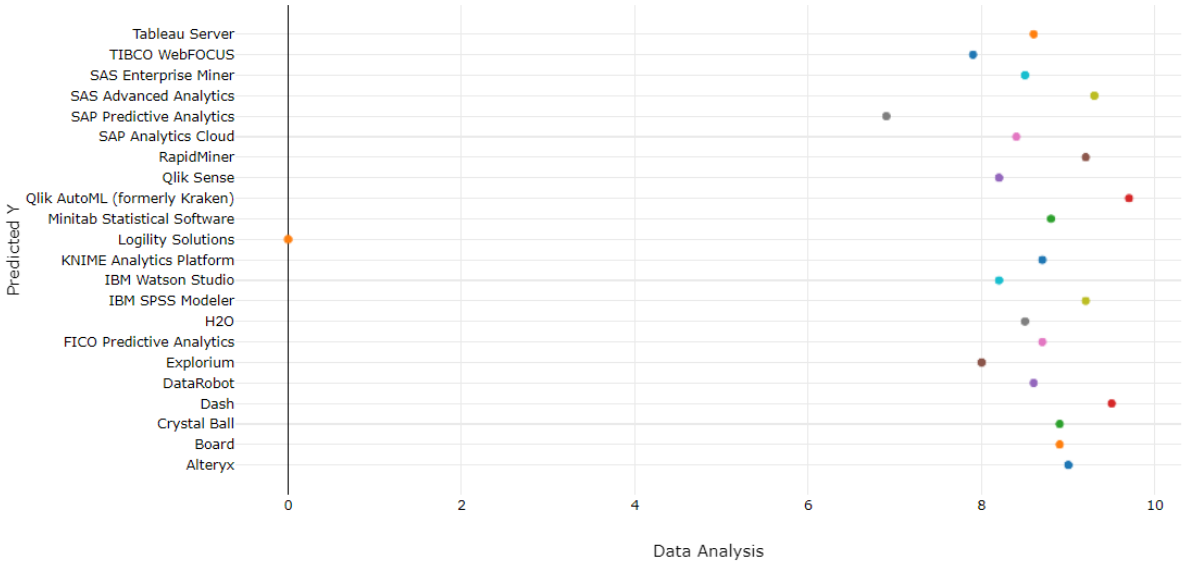


Figure 4. The likelihood of predictive analytics in SCM for data analytics capabilities

Data unification can be used to integrate data from multiple sources, such as sales data, logistics data, and supplier data, to create a single data set for analysis. This can help identify patterns and relationships that would not be visible in a single data source. For Predictive Analytics technology, the best IT solutions related to data unification capabilities are Explorium with a score of 9.20, Alteryx and SAS Advanced Analytics with the same score of 9.10, IBM SPSS Modeler with a score of 9.00 and with the same score of 8.90. Tableau Server and IBM Watson Studio. (see Fig. 5)

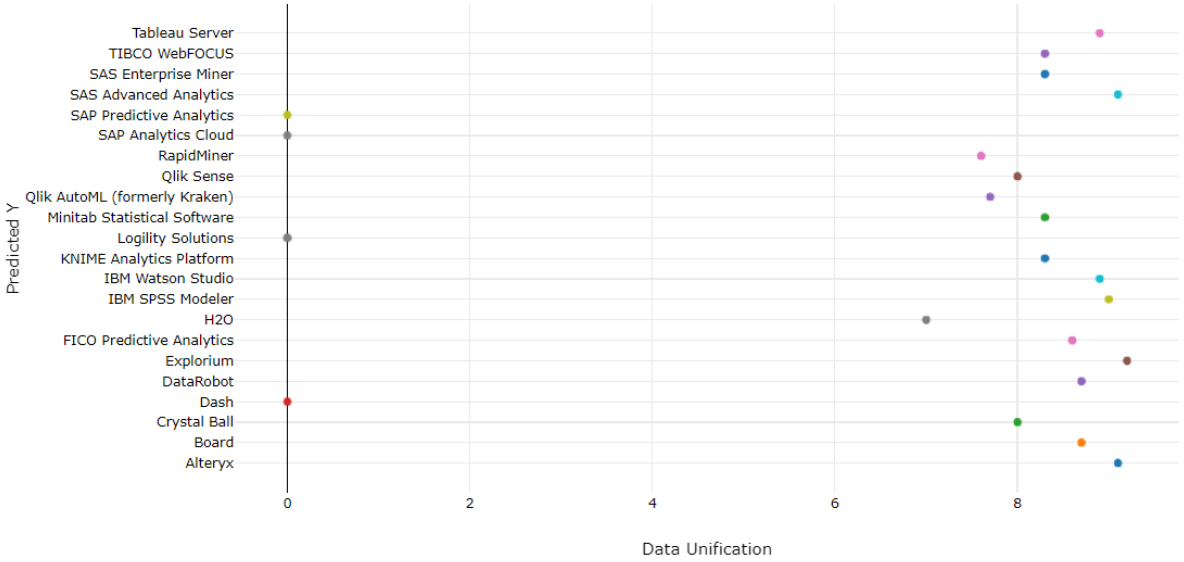


Figure 5. Probabilitatea analizei predictive în SCM pentru capabilitățile de unificare a datelor

Data visualization can be used to create visual representations of supply chain data, making it easier to understand and communicate. For example, a logistics route map can be used to visualize the flow of goods along the supply chain. For Predictive Analytics technology, the best IT solutions related to data visualization capabilities are KNIME Analytics Studio with a score of 9.80, the following software SAS Entreprise Miner, SAS Advanced Analytics and DataRobot have the same score of 9.30, IBM SPSS Modeler and IBM Watson Studio also has the same score of 9.20 and Tableau Sever with a score of 8.80. (see fig. 6)



Figure 6. Probabilitatea analizei predictive în SCM pentru capabilitățile de vizualizare a datelor

Data interaction modeling can be used to create models that describe how different variables in supply chain data interact with each other. This can help identify the factors that are most important in determining the demand for products or the success of logistics routes. For Predictive Analytics technology, the best IT solutions related to modeling capabilities are Explorium with a score of 9.20, Alteryx and SAS

Ad-Advanced Analytics with the same score of 9.10, IBM SPSS Modeler with a score of 9.00 and with the same score of 8.90 Tableau Server and IBM Watson Studio. (see Fig. 7)

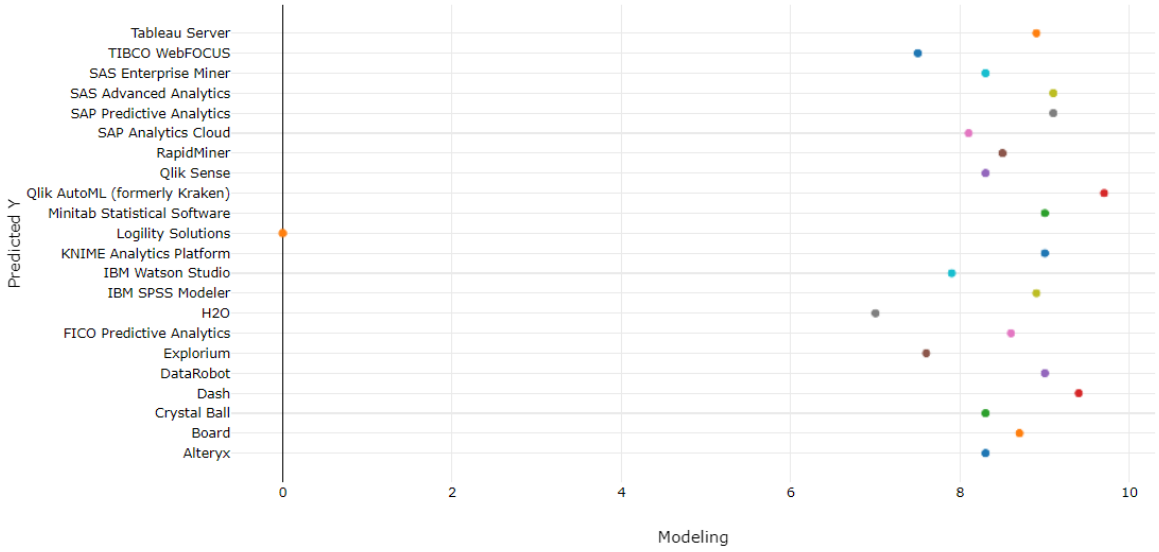


Figure 7. Probabilitatea analizei predictive în SCM pentru capabilitățile de modelare

Using data interaction modeling, organizations can create models that describe how different variables in supply chain data interact with each other. This can help identify the factors that are most important in determining product demand, logistics efficiency and overall supply chain performance. An example of this might be analyzing the relationship between inventory levels, product demand, and supplier lead times. By identifying the interactions between these variables, organizations can optimize their inventory management strategies, such as setting safety stock levels, and improve their supplier management by identifying suppliers with the most reliable lead times.

For Predictive Analytics technology, the best IT solutions related to data interaction capabilities are Qlik AutoML with a score of 9.70, Dash with a core of 9.50, followed by DataRobot with a score of 9.40. The following programs are IBM SPSS Modeler, SAS Enterprise Miner and Tableau Sever with estimated score of 9.20, 9.00 and 0.90 respectively. (see Fig. 8)

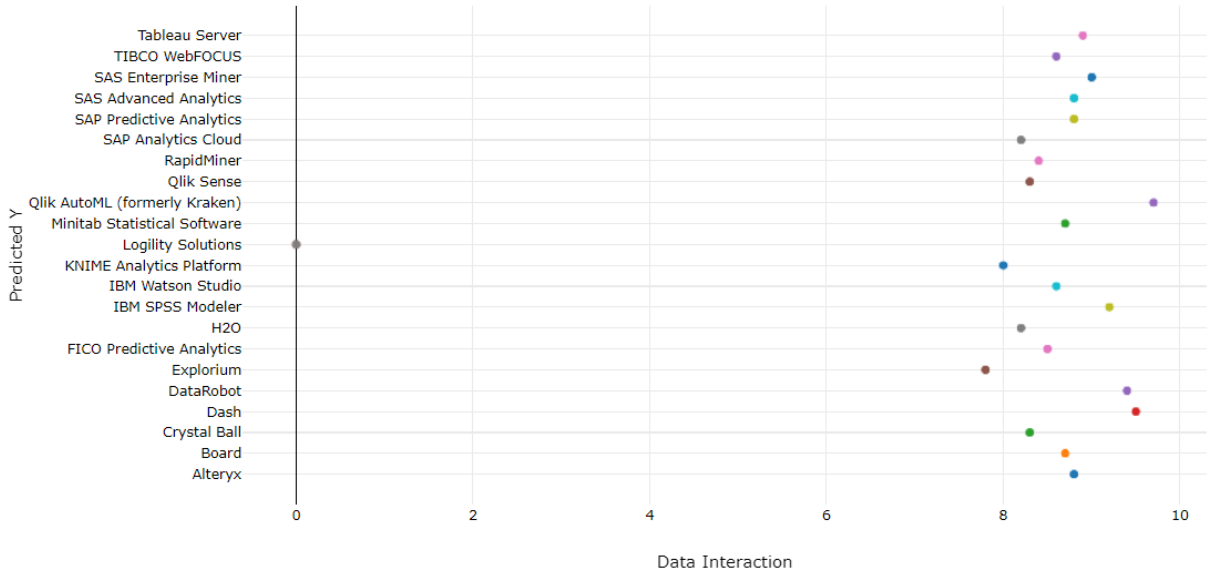


Figure 8. Probabilitatea analizei predictive în SCM pentru capabilitățile de interacțiune cu date

Algorithms such as regression, decision trees, and neural networks can be used to analyze supply chain data and make predictions. For example, regression algorithms can be used to predict product demand based on historical sales data, while decision trees can be used to optimize logistics routes. For Predictive Analytics technology, the best IT solutions related to Algorithms capabilities are Qlik AutoML with a score of 9.80, SAP Predictive Analyt-ics with a core of 9.60, followed by DataRobot with a score of 9.40. The next two programs are KNIME Analytics Platform and SAS Advanced Analytics with the same score of 9.20 and 8.90 respectively. (see Fig. 9)

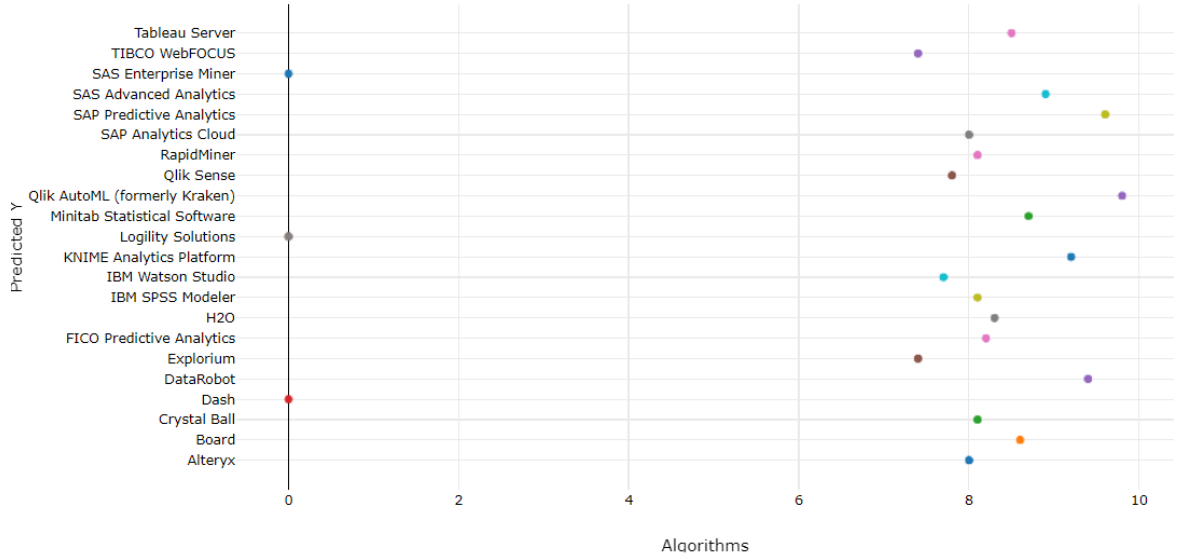


Figure 9. Probabilitatea analizei predictive în capacitățile SCM pentru algoritmi

Scripts can be used in supply chain management to automate data collection, data preparation, and model training tasks. This can help speed up the process of analyzing data and making predictions. For Predictive Analytics technology, the best IT solutions related to Scripting capabilities are Qlik AutoML

with a score of 9.50, SAS Advanced Analytics and Dash with the same score of 9.10. The next two software are IBM SPSS Mod-eler and Tableau Server which have the same score equal to 8.80. (see Fig. 10)

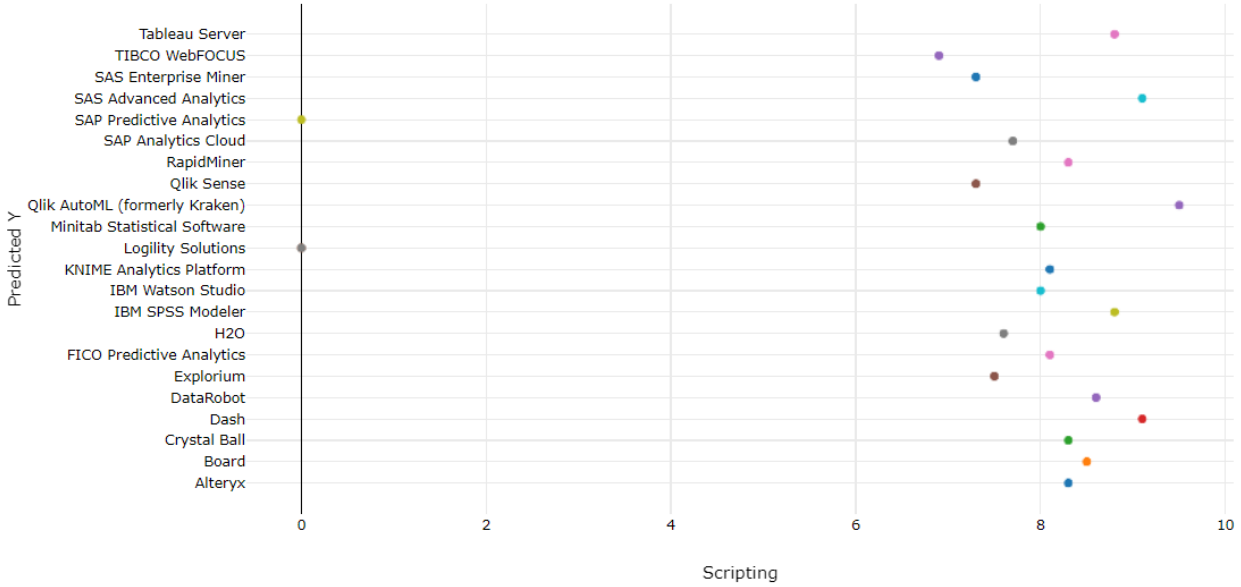


Figure 10. Probabilitatea analizei predictive în SCM pentru capabilitățile de scriptare

Data mining techniques can be used to extract useful information from large amounts of supply chain data. For example, association rule mining can be used to identify patterns in purchase data to help identify products that are often purchased together, while classification algorithms can be used to predict which products are likely to be at great search. For Predictive Analytics technology, the best IT solutions related to Data Mining capabilities are Qlik AutoML with a score of 9.80, followed by SAS Advanced Analytics, Tableau Server and Mnitab Statistical Software with the same score of 9.00. Next is SAS Enterprise Miner with a score of 8.90. (see Fig. 11)

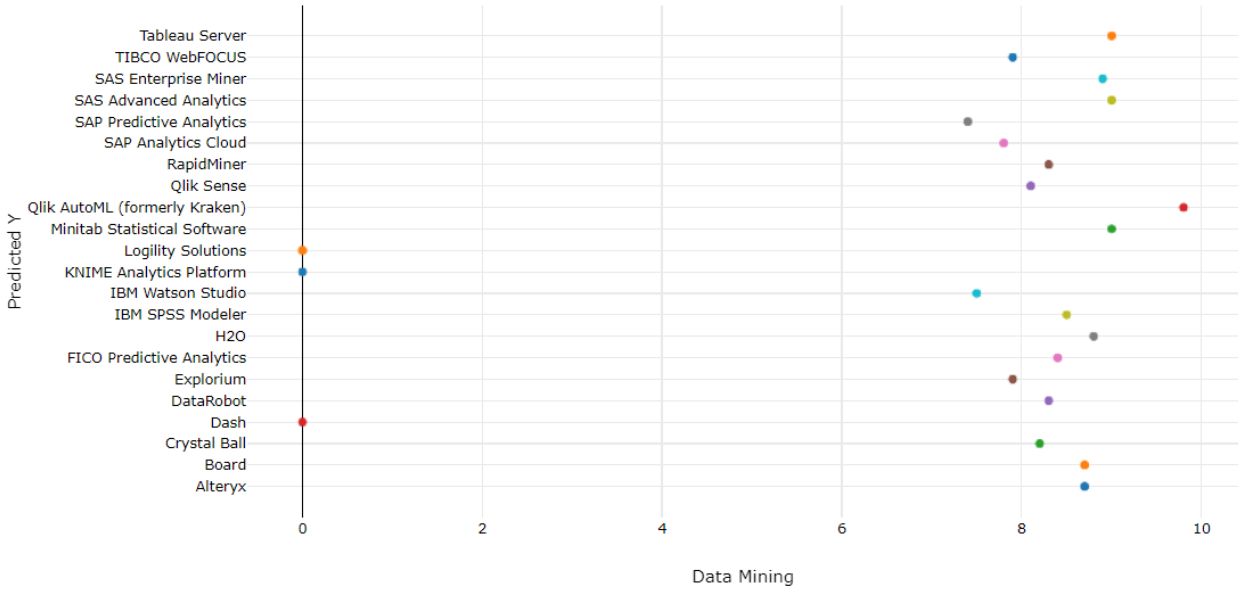


Figure11. Probabilitatea analizei predictive în SCM pentru capabilitățile de Data Mining

Data mining techniques can be used to extract useful information from large amounts of supply chain data. For example, association rule mining can be used to identify patterns in purchase data to help identify products that are often purchased together, while classification algorithms can be used to predict which products are likely to be at great search. For Predictive Analytics technology, the best IT solutions related to Data Mining capabilities are Qlik AutoML with a score of 9.80, followed by SAS Advanced Analytics, Tableau Server and Mnitab Statistical Software with the same score of 9.00. Next is SAS Enterprise Miner with a score of 8.90. (see Fig. 11)

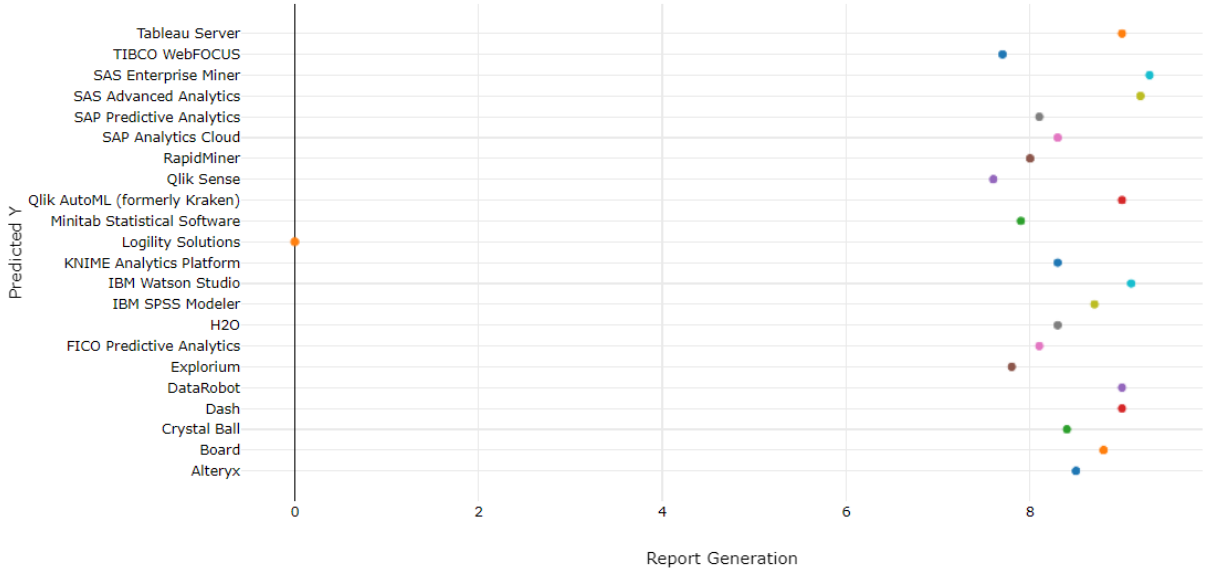


Figure 12. Probabilitatea analizei predictive în SCM pentru capabilitățile de generare a rapoartelor

The use of Predictive Analytics in supply chain management leads to for example: decision making is improved as valuable insights are provided on supply chain data, helping managers to make more informed decisions. Increased efficiency, used to optimize various aspects of the supply chain process, such as logistics, inventory management and quality control, resulting in increased efficiency and cost savings. Better forecasting, used to make more accurate predictions about product demand and other supply chain variables, enabling better forecasting and planning. Improved risk management, used to identify and assess potential supply chain risks, such as supplier disruptions, and develop strategies to mitigate those risks. Machinery maintenance predictions to predict when equipment or machinery will require maintenance, reducing downtime and costs.

Of course, the use of Predictive Analysis in supply chain management also has limitations, the implementation of predictive analysis in supply chain management can be expensive, requiring specialized software, hardware and qualified personnel. Predictive analytics relies on large amounts of accurate data, and it can be difficult or expensive to obtain the necessary data. Predictive analytics can be complex, requiring specialized knowledge and expertise to implement and interpret the results. Predictive analytics also raises some ethical concerns, such as privacy. It is important to consider these concerns and limitations when implementing predictive analytics in supply chain management.

Conclusion

The use of predictive analytics in supply chain management can greatly improve decision-making and operational efficiency within organizations. Using functionalities such as scripting, data mining, algorithms, data analysis, modeling, data interaction, data visualization, reporting and data unification, organizations can extract valuable insights from their supply chain data and implement them to optimize logistics, inventory management and quality control. These capabilities can be integrated with existing software systems such as enterprise resource planning (ERP) software, transportation management systems (TMS) and warehouse management systems (WMS) to provide a more comprehensive view of supply chain process, for supply chain optimization, demand forecasting and logistics optimization and quality control management systems. However, it is important to consider the potential limitations of using predictive analytics in supply chain management. These include the high cost of implementation, data availability and quality, complexity and ethical concerns. In addition, it is important to have the necessary technical and human resources to effectively manage and use the data and information generated by these solutions. Further research should be conducted to assess the effectiveness of using predictive analytics in different industries and supply chain contexts, as well as to address the limitations identified in this study.

References

Lee, H. L. (2004). „*The triple-a supply chain*". Harv. Bus. Rev. 82(10): 102– 113.

Christine A. Barry. (1998). „*Choosing Qualitative Data Analysis Software: Atlas/ti and Nudist Compared*". Sociological Research Online. Volume: 3 issue: 3, page(s): 16-28. 1998

Williams, G. (2011). „*Descriptive and Predictive Analytics. In: Data Mining with Rattle and R*". Use R. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9890-3_8