Advanced Data Analytics in Logistics Demand Forecasting

A. Agatić *, E. Tijan **, S. Hess *** and T. Poletan Jugović ****

* University of Rijeka, Faculty of Maritime Studies/Department of Maritime and Transportation Technology, Rijeka, Croatia
agatic@pfri.hr

** University of Rijeka, Faculty of Maritime Studies/Department of Maritime Logistics and Management, Rijeka, Croatia
etijan@pfri.hr

*** University of Rijeka, Faculty of Maritime Studies/Department of Maritime and Transportation Technology, Rijeka, Croatia
shess@pfri.hr

**** University of Rijeka, Faculty of Maritime Studies/Department of Maritime and Transportation Technology, Rijeka, Croatia
poletan@pfri.hr

Abstract – The logistics demand forecasting is increasingly influenced by digitalization processes in logistics business. Traditional approach to logistics demand forecasting based on human expertise and statistical assessment is still very present, but the use of Big Data, Artificial Intelligence and Machine Learning becomes more prominent. By using these technologies, logistics demand forecasting becomes not only more reliable but also more agile and self-adjusting, with better insight into changing market conditions in the real-time perspective. In this paper, the Authors research the evolution of Data Analytics in logistics demand forecasting, and provide an insight to the features of Big Data, Artificial Intelligence and Machine Learning used for Advanced Data Analytics in logistics demand forecasting.

Keywords – logistics demand forecasting; Advanced Data Analytics; Big Data; Artificial Intelligence; Machine Learning

I. INTRODUCTION

Logistics companies are operating in a very challenging logistics market. It is expected of logistics companies to easily adapt to shipment patterns and customers’ behavior, to provide on-time deliveries through efficient routes and reduce the risks of cargo inventory errors and miscalculations [1]. Furthermore, the growing focus in the logistics industry is placed on the use of digital technologies and digital transformation. These changes increase the necessity of logistics companies to accurately forecast possible future demand scenarios and to make plans and decisions by using the advantages of digital technologies.

Logistics demand forecasting is the process of accurately anticipating the demand for products, services, and shipments throughout the supply chain. The logistics demand forecasting enables the creation of plan for the unexpected scenarios, costs cutting and streamlining of existing and future manufacturing logistics planning efforts including: load distribution, flexibility in case of disruptions, seasonality, inventory costs and inbound logistics planning [2]. The ultimate goal in logistics demand forecasting is to minimize forecasting errors.

The Council of Supply Chain Management conducted a study of Data Analytics in business of shippers and third-party logistics companies. The results have shown that 93% of shippers and 98% of third-party logistics companies consider data-driven decision-making as crucial in supply chain activities [1]. Technologies such as Big Data, Artificial Intelligence and Machine Learning are becoming ever-more important, allowing logistics companies to process vast amounts of information more quickly, and therefore make accurate demand forecasting [3].

In this paper, the Authors have analyzed the evolution of Data Analytics in logistics demand forecasting. Furthermore, the Authors compared the use of Big Data, Artificial Intelligence and Machine Learning technologies to traditional Data Analytics in logistics demand forecasting. This paper aims at analysis of advantages and challenges of the aforementioned technologies used for Advanced Data Analytics in logistics demand forecasting.

II. EVOLUTION OF DATA ANALYTICS IN LOGISTIC DEMAND FORECASTING

Data Analytics solutions in logistics demand forecasting have evolved over the years. Changes in Data Analytics are caused by advancement in technologies and methods used. The five types of Data Analytics used in logistics demand forecasting are [4]:

1. Descriptive Analytics: the process of collecting and interpreting data to describe what has occurred. The Descriptive Analytics detect that something is wrong or right in logistics demand,
without analyzing the reasons. Raw data is used to provide insights into past events which affect the logistics demand. It uses two primary techniques: data aggregation and data mining to analyze past events. In Descriptive Analytics data is analyzed by statistical methods: averages, standard deviation, variance, instead of the complex calculation [5]. Furthermore, data is shown by visual tools: line graphs, pie and bar charts, etc.

2. Diagnostic Analytics: the process of collecting and interpreting different data sets to identify discontinuity in logistics demand. In the Diagnostic Analytics several methods are used: regression analysis, ratios, likelihoods and the distribution of outcomes [6]. Supervised Machine Learning algorithms for classification and regression are also used in Diagnostic Analytics. The Diagnostic Analytics detect not only changes in logistics demand, but also patterns and connection between events and demand forecasting results.

3. Predictive Analytics: uses descriptive and predictive variables from the past to analyze and identify the possible future outcomes in logistics demand. In the Predictive Analytics, various different methods are used: data mining, statistical modelling (mathematical relations between variables) and Machine Learning algorithms (classification, regression and clustering techniques). Successful use of traditional forecasting techniques with more advanced predictive algorithms enables effective interpretation of large amounts of data and provides competitive advantages. The Predictive Analytics determine demand drivers and use of the determined drivers to proactively respond to the market. It’s expected that global Predictive Analytics market will grow up to USD 21.5 billion by 2025 [7].

4. Prescriptive Analytics: a combination of data, mathematical models and business strategies to infer actions which influence future outcomes. Prescriptive Analytics provide insight on how to make desired outcomes happen, not only which outcomes may happen. Furthermore, Prescriptive Analytics include simulation, probability maximization and optimization. Large companies such as Amazon, Target and McDonald’s are already using Prescriptive Analytics in demand forecasting. In the report published by Allied Market Research, the global Prescriptive Analytics market is expected to reach the value of USD 12.35 billion by 2026 [8].

5. Cognitive Analytics: integration of digital technologies: semantics, Artificial Intelligence algorithms and a number of learning techniques such as Deep Learning and Machine Learning. By using the set of digital technologies, Cognitive Analytics can provide smarter and self-adjusting logistics demand forecasting and become more effective over time by learning from interactions with data and forecasting experts. The Cognitive Analytics almost removes the boundary between the physical and the virtual and automates processes to bring new capabilities to logistics demand forecasting. Cognitive Analytics imitate human thinking to make cognitive applications smarter and effective over time [9]. The Cognitive Analytics market is expected to reach the value of USD 14.95 billion by 2026 [10].

All the described Data Analytics types are used by logistics companies for demand forecasting. However, the MHI Annual Industry survey from 2020 has shown that 57% of the supply chain companies are not using Predictive Analytics, but they plan to use it by 2025 [11]. Logistics companies can choose which Data Analytics type to use according to available data sources, company size, data infrastructure, forecasting period, demand patterns, stability of market, market size, digital culture, etc. [12].

III. IMPROVING THE LOGISTICS DEMAND FORECASTING BY ADVANCED DATA ANALYTICS

Data Analytics for logistics demand forecasting has evolved from manual to automated. In the past few years, digital technologies are changing the logistic demand forecasting, especially in the case of large logistics companies. The Advanced Data Analytics is autonomous or semi-autonomous examination of data or content using sophisticated techniques and tools to discover deeper insights, make predictions or generate recommendations [13]. Advanced Data Analytics includes various methods and technologies: data/text mining simulations, complex event processing, Big Data, Machine learning, Artificial Intelligence, Neural Networks, etc. [13]. Key advantages of Advanced Data Analytics compared to traditional Data Analytics are [14]:

- Detailed analytics: large amount of data from different sources provides more detailed insight into the factors which influence the logistic demand.
- Forecasting accuracy: use of advanced techniques, e.g. Artificial Intelligence, Machine Learning, Deep Learning algorithms, etc., enable faster, accurate and self-adjusting logistics demand forecasting.
- Confidence level: real time and accurate data with future-oriented insights provide highly reliable logistic demand forecasting.

The Authors choose to analyze the main advantages and challenges of Big Data, Artificial Intelligence and Machine Learning in logistics demand forecasting. These technologies are analyzed because of their increasing importance in Advanced Data Analytics for logistics demand forecasting and it is expected that their use will continue to increase in the future.
A. Big Data

Big Data is important in logistics demand forecasting in order to make a large amount of data useful and meaningful, particularly in the case of large logistics companies. Big Data has two key points—to develop effective techniques (e.g., data exploration, interactive analysis and planning, embedded analytics, stream analytics, etc., [15]) able to make accurate predictions and to gain insights into the relations between factors influencing the logistics demand [16]. The four steps in Big Data Analytics are: data sources identification, data integration, data analytics and interpretation of results—concrete actions [15].

Big Data provide the different level of data structures from many different resources, while traditional data provide only structured data from the organizational databases and trade patterns. Furthermore, Big Data has a very high velocity, unlike the traditional data which depend on the business volume. Data flow is fixed in traditional logistics demand forecasting, while Big Data circulates continuously. In the logistics demand forecasting Big Data enable real time analytics of logistics demand forecasting, reports and planning of logistics demand with future predictability and perspective broader than internal business decisions.

Table I. shows the comparison of Big Data to traditional data in demand forecasting [17].

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Traditional Data</th>
<th>Big Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure of data</td>
<td>Structures are defined</td>
<td>Structured, semi-structured and unstructured data</td>
</tr>
<tr>
<td>Data volume</td>
<td>Based on business volumes and extend of digitalization.</td>
<td>Very high. In petabytes and even more.</td>
</tr>
<tr>
<td>Variety of data sources</td>
<td>Database systems and trading partners data</td>
<td>Business information systems, text (email, documents), weblogs, sensors, RFID, social networks etc.</td>
</tr>
<tr>
<td>Velocity</td>
<td>Low to moderate-based on volume of business</td>
<td>High</td>
</tr>
<tr>
<td>Flow</td>
<td>Fixed</td>
<td>Continuous accumulation of data</td>
</tr>
<tr>
<td>Analytics</td>
<td>Hystorical view, Status reports</td>
<td>Real-time, direct feedback from the consumer, sentiment analysis, opinions</td>
</tr>
<tr>
<td>Functions</td>
<td>Decision support to senior executives on internal business decisions.</td>
<td>Direct market feedback from the consumer, which can be used for planning market strategies, planning, etc.</td>
</tr>
</tbody>
</table>

Despite numerous advantages of using Big Data, there are also some challenges in using Big Data in logistics demand forecasting. For example, the amount of data which needs to be processed is increasing rapidly. To use in demand forecasting, data should be sorted and managed with specialized software solutions such as Talend, Skytree, Splice Machine, Spark, Hadoop, etc. Furthermore, data quality depends on the sophisticated statistical and computational methods for analytics. Use of Big Data in organizations requires infrastructure, changes in business processes and experienced data scientists [17].

The Supply Chain Big Data Analytics market was valued at USD 3.55 billion in 2020. It is expected to reach USD 9.28 billion by 2026 [18]. Furthermore, the overall Big Data Analytics market is estimated to grow 4.5 times, gaining a revenue of USD 68.09 billion by 2025 (from 14.85 billion in 2019) [19]. In spite of the COVID-19 uncertainty, Big Data continues to be a top priority for enterprises as its use helps to adjust demand forecasting and to remain competitive [19].

B. Artificial Intelligence

Using Artificial Intelligence in logistics demand forecasting has several advantages. Traditional demand forecasting methods mostly consider only linear factors such as seasonality. Using Artificial Intelligence on larger datasets from different sources may provide a better insight into current market conditions i.e. factors influencing the demand forecasting.

Artificial Intelligence can utilize Big Data and analyze any number of factors that might impact logistics demand (e.g., real-time weather conditions, holidays, fuel prices, traffic etc.). As a result, Artificial Intelligence can provide better insights into logistics demand and become more efficient, resulting in faster processes and cost savings [20]. With the ability to predict errors before they occur, Artificial Intelligence provides logistics companies with increased ability to improve business profitability [21].

Algorithms used in Artificial Intelligence almost always outperform expert’s judgment. In the few cases where Artificial Intelligence didn’t outperform experts, the results were usually concluded in a tie. Algorithms used in Artificial Intelligence are not completely replacing human intelligence, but analyzing the data they provide can improve accuracy of forecasting [22].

The most challenging issue when using Artificial Intelligence in logistics demand forecasting is the availability and accuracy of data. In order to make a correct demand forecasting, events which affect the demand have to be identified and accounted for. For an Artificial Intelligence application to learn from these events, they need to be fully understood and coded [23]. Furthermore, many logistics companies struggle with the implementation of digital culture in their business. The use of Artificial Intelligence in logistics demand forecasting models brings out what is termed the ‘explainability’ problem. This term describes the managers who should use Artificial Intelligence applications, in which logic of obtaining the results is difficult to explain [23].
According to McKinsey Digital, Artificial Intelligence can reduce errors of forecasting by 30 to 50% in supply chain networks. Improved accuracy enables up to a 65% reduction in lost sales due to inventory out-of-stock situations and warehousing costs decrease around 10% to 40% [24].

The Artificial Intelligence market was valued at USD 21.46 billion in 2018, and is estimated to reach the value of USD 190.61 billion by 2025 [25].

C. Machine Learning

Machine Learning Forecasting achieves high accuracy in logistics demand forecasting. With Machine Learning, processors learn from mining Big Data without human interventions to make accurate insights in logistics demand. Machine Learning Forecasting is self-correcting and powerful when compared to traditional logistics demand forecasting methods such as average, moving average, trend, multiple linear regression, etc. [26]. Steps in Machine Learning demand forecasting are: data collecting, data preprocessing (fill missing values, removing outliers, and exploratory data analysis), application of to the machine Learning model, evaluation of the model, results interpretation [27].

Table II. shows the comparison of Machine Learning demand forecasting and traditional demand forecasting [28]. In traditional logistics demand forecasting, data is stored in a computer and then analyzed by predetermined set of rules to generate a result. Machine Learning uses large amounts of data, and unlike in traditional demand forecasting, Machine Learning forecasting can easy incorporate multiple variables and data sources with regard to automation.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Traditional forecasting</th>
<th>Machine Learning forecasting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data volume</td>
<td>Small</td>
<td>Large</td>
</tr>
<tr>
<td>Ability to consider numerous variables and data sources</td>
<td>Adding extra variables and sources require substantial effort</td>
<td>Multiple variables and data sources can be smoothly incorporated thanks to the high level of automation</td>
</tr>
<tr>
<td>Manual work</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Technology requirements</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Maintenance complexity</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Forecasting period</td>
<td>Mid/long-term planning Established products Stable demand</td>
<td>Short/medium-term planning New products Volatile demand scenarios</td>
</tr>
</tbody>
</table>

By applying Machine Learning, the computer becomes adaptive, dynamically responding to changes in the data and update of logistics demand forecasting accordingly [29]. Furthermore, manual work is low in Machine Learning, but maintenance is high. Technology requirements to use Machine Learning are high, while in traditional forecasting are low.

In traditional Data Analytics, statistical programs (e.g., SPSS, Statistica, R, etc.) and/or basic Excel calculations are sufficient to analyze data and to make demand forecasts. Demand forecasting based on Machine Learning implies more technology requirements to enable data analysis and demand forecasting. Data storage capacities i.e. memory should be enough to handle large amounts of data. Data selection, collection and preprocessing (filtering, categorization and feature extraction) are the key factors for a model’s accuracy and predictive value [30]. Therefore, data aggregation and storage are significant elements that influence hardware features. For data intensive tasks such as data preparation or disk-enabled analytics software, the hardware acceleration is needed. Furthermore, for intensive tasks that can be run in parallel, such as matrix algebra, graphical processing units (GPUs) are necessary [30].

The large amount of data and more technology used makes maintenance of all elements included in the Machine Learning based forecasting more complex. Traditional forecasting is suitable for Mid/long-term planning with established products and stable demand. The advantage of the Machine Learning is the ability to make “ad-hoc” demand forecasts i.e. adaptable logistics demand scenarios.

One of the challenges of Machine Learning in logistics demand forecasting is the time necessary for algorithms to learn and develop enough to fulfill their purpose with a considerable accuracy and relevancy [31]. Another challenge is the accurate interpretation of results generated by the algorithms and use of right algorithms for the purpose of demand forecasting [31].

In the situation of COVID-19 uncertainty, companies experienced in using Machine Learning tend to continue to gain maximum advantage from their demand forecasting even as the pandemic rages. These companies continue to build their data infrastructure and have adapted their models to Covid-19’s uncertainties [32].

The Global Machine Learning market is expected to reach the value of USD 96.7 billion by 2025, with growth of 43.8% CAGR (compound annual growth rate) from 2019 to 2025 [33].

The Institute of Business Forecasting and Planning (IBF) provided a survey on the essential changes in future demand forecasting. The survey results have shown that Artificial Intelligence and Machine Learning will be leading technological advancements in the logistics companies by 2025 [34]. The 70% of the respondents identified Artificial Intelligence as the most prominent technology [34]. Furthermore, the survey results have shown that advanced decision making is going to be a core competence in demand forecasting by 2025 [34].

Logistics companies are aware of the Big Data, Artificial Intelligence and Machine Learning in logistic demand forecasting and decision making. For example, a leading logistics provider DHL uses Big Data, Artificial Intelligence and Machine Learning in their business i.e.
demand forecasting [35]. With modern forecasting tools, logistics companies can shift to a strategy based on accurate demand forecasting and achieve greater operational efficiency.

IV. CONCLUSION

Logistics companies use various Data Analytics methods and tools to make accurate demand forecasting. In different types of Data Analytics, different methods are available to forecast logistics demand. Logistics companies are free to choose the Data Analytics type according to their goals in demand forecasting and their abilities (infrastructure, financial sources, digital culture, etc.).

Recently, digital technologies such as Big Data, Artificial Intelligence and Machine Learning have increasingly been used for logistics forecast demand. These technologies enable Advanced Data Analytics in demand forecasting. It is possible to analyze large data volume with different structures. Analyzed data can originate from different sources, which is important to observe demand comprehensively i.e. to include all factors relevant for forecasting. Furthermore, analyzed data may be used for much longer forecasting horizon. In traditional methods, data is historical and forecasts are tightly coupled to past events. Data is accumulated continuously, which is important for the real-time insight in all factors which influence demand forecasting. Although these technologies have many advantages, some challenges exist: complexity, technology requirements, digital company culture, specific skills and software required to analyze large amount of data, etc.

Large logistics companies such as DHL use Advanced Data Analytics i.e. Big Data, Artificial Intelligence and Machine Learning technologies in logistics demand forecasting. Use of Big Data, Artificial Intelligence and Machine Learning can make the process of logistics demand forecasting self-corrective and adaptive – capable to respond to data changes and update the forecasting automatically (e.g., in case of changes caused by COVID-19).

REFERENCES