The American Journal of Engineering and Technology
ISSN 2689-0984 | Open Access

TYPE Original Research
PAGE NO. 21-34
DOI 10.37547/tajet/Volume07Issue03-03

Building Agile Supply Chains with Supply Chain 4.0: A Data-Driven Approach to Risk Management

OPEN ACCESS

SUBMITED 01 January 2025
ACCEPTED 02 February 2025
PUBLISHED 03 March 2025
VOLUME Vol.07 Issue03 2025

CITATION

Md Zahidur Rahman Farazi. (2025). Building Agile Supply Chains with Supply Chain 4.0: A Data-Driven Approach to Risk Management. The American Journal of Engineering and Technology, 7(03), 21–34. https://doi.org/10.37547/tajet/Volume07Issue03-03

Md Zahidur Rahman Farazi

Department of Information Systems and Operations Management, The University of Texas at Arlington

Abstract: The aim of this study is to advance multi-label delivery delay predictions in supply chains using machine learning and deep learning models. The work used Decision Trees, Random Forests, CNN, and FNN on a real-life logistics dataset consisting of customers and products features. EDA and feature selection were examined and performed as a part of the data preprocessing process at the pre-processing step of the models. According to current model results, Random Forest model reached maximum accuracy of 66.5% along with Decision Trees and FNN. CNN, although, worked well in some instances was not up to par in some areas because it overfitted. The results also reveal how Random Forest is a particularly useful algorithm for predicting delivery delays accurately. The conclusion suggests enhancing the deep learning models performance and combining approaches. Further work should also incorporate other variables in order to improve the predictive capability in real-life requirements of supply chain environments including conditions and stocks.

Keywords: Supply Chain 4.0, Machine Learning, Deep Learning, Risk Management.

Introduction: Supply Chain 4.0 is fully embedded internet of thing, analytics, automation and data- driven decision making in order to manage the supply chain process. These core processes integration of DL and ML helps organizations process large amounts of real-time data for predictive insights, risk reduction and proactive response to disturbances [1]. These technologies aid in identifying possible threats because, based on historical data fed into the systems, the algorithms get to adjust their projections continually as the new data stream in. As a result, companies can have full chain visibility, efficient resource managing and enhancement their capacity to respond to risks and volatilities [2].

The use of data analytics in the approaches is a revolutionary step from managing supply chain risks in a reactive-fashion to managing them in an anticipatory-fashion [3]. From this perspective, DL and ML enable organizations to improve their performance to a higher level, optimizing not only essential business processes but also providing the prospects for complex organizational growth in today's environment when flexible response and fast decision-making becomes critical [4].

The increasing dynamism and sophistication of supplies means that the ability to predict delivery delay has become a huge problem as such delays may be influenced by spatial issues, type of products, or customers [5]. Most supply chain practitioners are usually tasked with predicting more than one form of delay at one time in a supply chain which makes it a multi-label prediction problem [6]. Such imprecise risk management leads to higher operational costs, resource consumption, and customer dissatisfaction that requires developing enhanced methods for managing delivery risks as efficiently as possible.

The proposed study therefore seeks to design and apply complex ML and DL algorithms to increase the reliability of multi-label delivery delay prognosis in value chains and assist organizations in reducing risks and improving operational performance. The objectives of the research are:

- To determine and categorize delivery delays factors in supply chains.
- To perform machine learning and deep learning algorithms in order to compare the results of multi-label delivery delays prediction.
- To test the efficacy of such models in real conditions for risk detection and operational decision makings.
- To offer practical suggestions regarding the application of these models into the existing supply chain theories for improving supply chain adaptability and flexibility.

Literature Review

A. Understanding of the Fourth Industrial Revolution and Supply Chain 4.0

Machine learning (ML) and neural networks (NN) have

Industry 4.0 has brought outdated innovative technological changes that have affected the supply chain management of numerous industries. Supply Chain 4.0 which is the focal piece of Industry 4.0 employs IoT, AI, ML, and DL to improve organizational supply chain effectiveness, real-time tracking, and flexibility [7]. In this literature review, the definition, development, the components and the significance of agility and risk management as fundamental to the current supply chain concepts are discussed.

Supply Chain 4.0 is a deviation from the typical supply chains that exhibited limited integration and excessive complexity. These conventional models could not meet the demand for flexibility that is necessary for engaging with changing customer needs [8]. Supply Chain 4.0 can be originated from the early internet and the enterprise resource planning (ERP) system that coordinated the supply chain activities and information flows. This has however been propelled by IoT, AI, ML and data analytics that have shaped inter-connected physical and digital systems for real-time data generation and decisions [9].

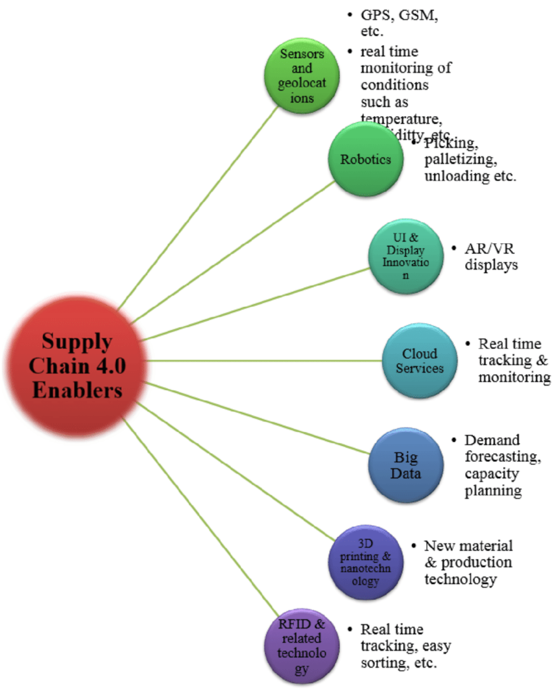

Figure 1: Supply Chain Enablers

(Source: [14])

Several technologies are depicted in the diagram that support Supply Chain 4.0 such as; Sensors, Robotics and Automation, Big Data, Cloud services, 3DPrint and RFID which increases real time monitoring, demand forecasting and new type of production [14].

Supply Chain 4.0 depends on several important technologies, some of the ways that IoT will impact the supply chain include collecting and transmitting data about inventory, assets and production that occurs in real time from various points on the supply chain [10]. AI entail large datasets for pattern recognition, making decisions and predicting future situations hence, useful trade activities such as demand forecast and risk managing [11]. Automated decision-making can enhance organizational operation since ML, a subfield of AI, increases machines' performance by training on data; its application includes quality control and prediction of equipment failures [12]. DL is a further development of the ML system that uses neural networks; this algorithm works better when solving such issues as image identification or detecting unusual activity. Robots' technology and automated guided vehicles, lessen the labor cost and improve the productivity of the supply chain operation; and advanced analytics, particularly the prognostic and diagnostic ones, help in risk assessment and decision making for supply chain management [13].

B. Flexibility and risk are the major components of Supply Chain 4.0

Agility on the other hand, is the supply chain's capacity to rapidly respond to disruption or demand volatility and risk management centers around recognizing, evaluating, and managing risks [15]. These factors play specific and significant roles in Supply Chain 4.0, according to several researchers. In the paper [16] discussed agile supply chains can provide more response to disruptions and help organizations satisfy the customer demand in volatile contexts. Furthermore, the study [17] shows that the management strategies are useful for managing supply chain risks and increasing performance, thus stabilizing operations.

As a result, it is correct to mention that Supply Chain 4.0 outlines a new approach to manage supply chains. Incorporating smart technologies in an organization provides organization with a boost in efficiency gain, visibility and flexibility. However, for Supply Chain 4.0 to be fully operational to provide all the benefits that come with it, there is need for agility, plus wiser ways of performing risk management due to the dynamic nature. Looking to the future of Supply Chain 4.0, it can be expected that as the technologies get improved and developed, Supply Chain 4.0 will help to stimulate further productivity change in supply chain management.

C. Problems Associate with Conventional Supply Chain Management

Traditional supply chain management has a number of important challenges that prevent supply chain to function optimally and effectively, which is explained by low level of structure flexibility and high percentage of manual processes [18]. On the same account, one major weakness of traditional supply chain is lack of flexibility; it is relatively difficult for a supply chain that is rooted in traditional model to make drastic changes when change is inevitable due to its negative impacts that it will bring about such as increased costs and reduced customer satisfaction [19]. The research [20] discussed that the increased demand may lead to higher stockouts and rationing due to insufficient production capacity or fixed supply arrangements. Also, a lack of transparency is rife in normal supply chains, which hinders an organization's capacity to track the flow of the products and supplies. This lack of insight can lead to inefficiencies, delays and costs cutting; for example, concerns such as delayed or lost shipments may only discovered after the product has arrived with the customer, resulting into customer dissatisfaction and monetary loss. In addition, the conventional supply chain is relatively rigid in managing changing dynamics such as natural disasters or an economic recession that has a huge cost implication and negative impact to the organization's reputation [21]. For example, the natural disaster that shuts down much of manufacturing such as a factory closes down supply chain which slows down productions.

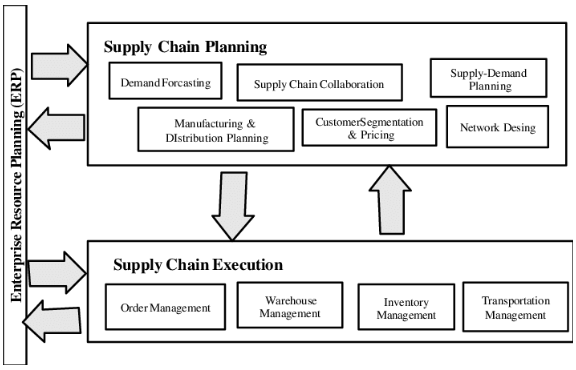

Figure 2: Conventional Supply Chain Management

Source: [21]

Further adding to these difficulties are various risks occurring to conventional supply chains such as demand volatility, geo political risks, and disasters. Fluctuations in consumer demand create undesirable conditions such as stock out conditions and conversely inventory conditions, political instabilities trade wars cause disruptions in the right of transport of goods. Hurricanes or any other natural disasters increase these problems by destroying infrastructure, slowing transport, and failing to supply sufficient products [22]. Taken together, the drawbacks and dangers of supply chain management as practiced in the traditional system require rethinking and improvement work.

The conventional approaches of risk assessments and delay predictions call for the use of history and statistics; not enough provide the outstanding forecasts needed to handle emerging risks [23]. These populations may also take a long time to be awakened to respond to incidents hence can be vulnerable and be able to lose a lot of money. In an attempt to address these problems organizations are gradually beginning to seek better analytical technologies and approaches to supply chain management. Such are the adoption of artificial intelligence, machine learning, as well as use of data in increasing visibility and optimizing outcomes and minimizing risks. Through adoption of these technologies, organizations can establish strong, effective and adaptive supply chains capable to managing itself within today's environment.

D. The Use of Artificial Intelligence in Supply Chain 4.0

Application of AI can help organizations to improve decision-making by offering automation, optimization and increased control of organizational data, relying on large data sets [24]. The study [25] demonstrated that AI is instrumental in enabling primary supply chain tasks including demand forecasting, route management, and predictive maintenance. All of these applications help in achieving cost reduction efforts, enhancing service delivery, and managing risks occasioned by disruption of supply chains.

Another field of supply chain management, demand forecasting, has relatively clear and potentially enormous benefits from AI models for predictions [26]. Unlike the former tools that have been used in estimating demand which can be done by using historical data and where analysis is again performed manually, was not able to capture changes of real time markets. Machine learning models, a type of on AI algorithms outperforms human derived models in this aspect because the data that feed into the model can in real time data such as the prevailing weather, the state of the economy and society at that particular time [27]. The study [28] showed that the use of AI increases the accuracy of demand forecasting by 30% if compared to the traditional approaches. Similarly, the study [29] pointed to corresponding advantages of applying AI in stock cost optimization by improving demands' prognosis.

Another process area that has benefited from use of AI technology is in route optimization. Machine learning is then employed to forecast the precise delivery routes which with respect to congestion, fuels and delivery time [30]. In the research [31] reported that the use of AI for route planning and optimization minimized delivery time by 15% and had a corresponding impact on the cost of logistics by minimizing it by 20%. This is a major enhancement compared with traditional route optimization approaches that have dynamic problem-solving difficulties, such as in situations that involve changes in roads or weather conditions.

Another strategic application of AI in Supply Chain 4.0 is predictive maintenance, as it highlights the value of minimizing downtime and increasing the reliability of equipment [32]. Other maintenance schedules are based on time hence resulting to either over maintenance or under maintenance of the assets. Conventional concept of condition monitoring on the other hand is based on causative failures and typically involves the use of sensor data to determine when an asset is likely to fail next, thus shorting the operation time [33].

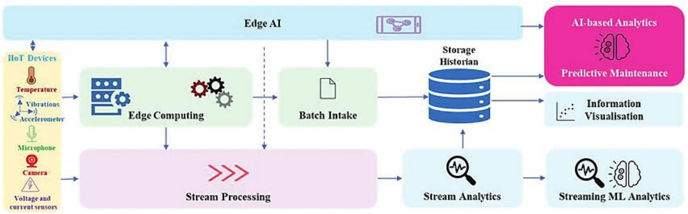

Figure 3: AI-Based Predictive Maintenance

Source: [34]

The study [34] showed that AI-based predictive maintenance decreased equipment failure rate in contrast to the reactive maintenance. However, some challenges still persist. Despite the AI's advances in supply chains, challenges such as data quality and integration hinder AI effectiveness, are common concerns surrounding implementation among many organizations. Moreover, while previous works defined the promising applications of AI, less recent research is applying increasingly efforts on practical issues of putting such applications into effect, including for instance scalability and ethics.

E. Applying Machine and Deep Learning for Supply Chain Risk Management

The machine learning (ML) and deep learning (DL) have grown to be high-impact solutions for handling the problems the traditional supply chains. These technologies provide enhanced features of predictive statistics, risks assessment, and optimization that help various organizations to reduce risks and make better decisions.

Some of the most common machine learning models used in supply chain risk management include Random Forest, Support Vector Machines (SVM), and XGBoost. The generated models can be applied to anticipate the lead time, estimate risks on the supply chain and improve on its working [35]. For instance, Random Forest can be applied to forecast demand variation and SVM — to conduct anomaly detection in the supply chain data.

There are many works that have shown that using ML in supply chain planning and decision-making is beneficial. The study [36] showed that by adopting ML models, demand can be more accurately forecasted than by traditional forecasting methods across many products. Although the study [37] found that consequent ML algorithms can be used in helping to predict potential risks in the SCM, which may range from transportation hitches to supplier bankruptcy.

DL which is a subfield of ML is also being used more in supply chain analysis. DL models such as neural networks, CNNs, and RNNs are capable of recognizing intricate patterns and dependencies in patterns, which is appropriate for tasks like multi-label classification [38]. In supply chain management, DL models can be employed to make forecast on delivery delays depending on the weather condition, mode of transport used and performance of suppliers.

Sometimes DL approaches outperform the traditional ML models where there are complex pattern and large dataset. In the research [39] observed that using DL models enhance prediction of delivery delays in a large e-commerce firm as compared to ML models. However, its known that DL models can be computationally expensive in their training process and may even need GPUs [40].

Therefore, it can be seen that ML and DL contains useful tools useful for handling the issues affecting the existing supply chain. Through the use of the mentioned technologies, it is possible to enhance the capacity of organizations to try and estimate potential risks, increase operational efficiency, and make rational decisions. That is why as more and more of these technologies develop, in the research can be sure that even more applications in the field of supply chain risk management will be developed.

F. Multiple label classification issues in supply chain management

The problem of handling multi-label classification scenarios in SCM is challenging because several labels are correlated and interdependent in supply chain management risk and return scenarios. Multi-label classification is different from single label classification dealing with cases in which multiple outcomes such as transport delay, stock-out, and supplier disruption are possible [41]. These complexities make multi-label classification very important in handling real time decisions within the supply chain. The paper [42] imply that multiple-label classification is significant when anticipating disruptions in the supply chain but claim that numerous dangers are not efficiently explained by the current frameworks.

Consequently, the paper [43] deal with multi-label classification, binary relevance and one vs rest which enhanced the accuracy in the supply chain delay prediction. Nevertheless, addressing the issues related with the imbalanced data is always a concern. This problem has, however, been dealt with fairly well by the application of SMOTE (Synthetic Minority Over-sampling Technique) which was endorsed in the study by [44] for the development of SMOTE based multi-label models for delay prediction.

Predictive analytics has also added another promising perspective to risk reduction in Supply Chain 4.0. Whereas previous strategies and future forecasting have been utilized in previous literature, the study [45] explained that real-time and predictive models-oriented concepts can effectively enhance supply chain visibility. Other approaches related to the use of artificial intelligence can also be recommended because the fundamental elements of AI-based techniques have been tested experimentally in other cases and effectiveness was proven in terms of the ability to adapt to changing uncertainties more quickly than traditional means.

G. Gaps in Existing Literature

The current research focusing on delivery delay prediction in supply chains presents several limitations, concerning, for example, the application of multi-label predictive models and their application for real-time risk assessment. Despite various ML ad DL applications in resolute supply chain functions, many of these works employ single-label classification approaches for delay prediction. Validation of traditional approaches come with various assumptions that disregard how real-world supply chain are complex and could be subject to various delays [46]. The concept of the multi-label predictive models that enable the practice of the numerous outcomes simultaneously is relatively underdeveloped; hence, the applicability of this concept in dynamic world is less.

Furthermore, while the application of AI technologies can improve supply chain's flexibility and robustness, most of the research works are based on the analysis of archive data rather than monitoring risks in real-time manner. The research [47] highlighted the importance of using predictive analytics for supply chain management while admitting that the integration of real-time decision-making models may a challenging task because of the data quality and integration problems. Today's scholarly work does not provide an extensive framework for integrating multi-label models into real-time decision-making systems.

The future work should be oriented to utilize DL and ML to develop high-impact, fast-action, accurate multi-label predictive models for solving multiple risks at a time for enhancing the availability and agility of the supply chain.

ML and DL were described in the literature as the keys to developing new and highly adaptive and re-illuminated supply chain management structures with high risk-management features. Although currently not as actively investigated as single-label approaches, multi-label classification methods encompassed more accurate solutions to supply chain difficulties like delivery delay and disruptions. Researches focused on how accurate predictive analytics were and how real-time decision making and response could be leveraged by AI models. Future research was suggested to concentrate on enhancing multi-label approaches and incorporating it within elaborate supply chain models. Companies were urged to harness sophisticated AI tools and approaches to improve the firm's adaptability and reduce perils of operating in dynamic contexts.

METHODOLOGY

This work utilizes comparisons between a machine learning (ML) and a deep learning (DL) based models for multi-label delivery delay prediction. The basic variables, which cause delay are determined by data gathering and cleaning the data. To assess model performance, actual, real-time datasets are used to identify the risk and the proposal brings enhanced decision making to the supply chain.

A. Data Collection

This data set for this research work was obtained from Kaggle and consist of actual logistics and supply chain data. It yielded 15,549 records and 41 variables which provide a broad perspective of numerous aspects that affect supply chain. They are, payment in respect of type of payment which indicates the methods used in the transactions, and profit per order which indicates the profitability every order. Furthermore, the dataset contains the daily sales per customer, which gives information about customer-oriented sales, beside the category ID and the category name, which categorizes products. Other attributes include geographical and identification indicators, including customer city, country, and ID number, and customer segment information that distinguishes between customers according to their intent and age. This high quality and multi-faceted data are ideal for performing a comprehensive analysis across all delivery delay categories and to model how specific factors play a role in them [48].

B. Proposed Architecture

The following architecture has been proposed for the prediction of multiple labels of delivery delays: Machine learning ML: Decision tree and Random Forest Deep learning DL: Convolutional neural networks CNN and feedforward neural networks FNN. Every model that has been chosen in this paper has been done so because of its merit as found in prior research and applicability in solving the multi-label prediction problem in the supply chain. Table 3 presents each model and how it can be justified with findings from previous studies; the reason behind selecting each model.

1) Decision Tree

Decision Trees can be found extensively in various fields including supply chain management and predictive modeling because of their easy interpretability [49]. The study [50] reported that DT are useful in supply chain risk management considering their ability to provide decision-makers with clear decision-making trees and risk factor information. In single-label classification tasks, the study [51] also reported Decision Trees effective for predicting supply chain disruptions. In this research the Decision Tree model stands out to set the benchmark given its interpretability. The hierarchical structure also helps locate priority factors about delays in delivery [geographical location, type of products]. While its performance decreases in multiple class multiple label scenario it has the advantage of being able to give insights into the importance of features during the initial stages of the detection process.

2) Random Forest

Random Forest is another technique that integrates multiple decision trees which are more effective than the performance of an individual tree according to the outcomes of numerous research [52]. The research [53] proposed Random Forest, which means that it has good generalization performance since it randomly average out the decision trees. This research selects Random Forest for its ability to deal with large numbers of features in the dataset. Random Forest classifies the delivery delays based on the combined decision paths which capture complex variable interdependencies such as customer segments with product categories, thus improving the model's overall prediction.

3) Convolutional Neural Network (CNN)

The most popular type of neural networks, CNNs has been used mainly for image recognition but recently the use of CNNs for structured data is on the rise [54]. Study [55] have pointed out that CNNs can also be used in text classification tasks pointing out that such networks are capable of learning patterns in structured data other than images. Moreover, study [56] applied the CNNs in demand forecasting in supply chain and their results reveal the enhanced accuracies in the forecasting due to the capitalization of the relations both in space and time. CNNs are chosen for this study because participant CNNs can detect local relationships between customers' demographics, product offering, and delivery locations. The capability of CNNs to process the grid-like inputs data make them useful in revealing hidden correlations making multi-label classification more accurate. In this paper, CNNs will be extended to handle highly structured supply chain data and its finer characteristics that motivate delivery delays. Through the use of such convolutional layers the model will able to detect transformative interactions of the variables for resulting in multi-label delay with better prediction capacity.

4) Feedforward Neural Network (FNN)

Feedforward Neural Networks (FNNs) are the simplest design of the Deep Learning model for structured data [57]. The study [58] confirmed FNNs' presentation of day-to-day non-linear interaction between variables which make them suitable for supply chain forecasting and optimization. FNNs were selected for this study because they perform well in relation to handling of complex relations in multi-label prediction activities. FNNs on the other hand are trained using big data and it extracts vital patterns which are necessary for accurate predictions unlike other Machine Learning models. The FNNs will be used in this study to capture such complexities between features like the number of sales per customer, customer's city, and their product preferences since FNNs' multi-layer architecture powers the modeling of non-linear complexities of features for better multi-label delay predictions.

C. Data Analysis

The data analysis is a critical stage in the process of raw supply chain data and generating an outlook for delivery delays. For this study, the dataset contains demographical data about customers as well as regarding the products like customer city, country, segment among others: category name, profit per order, payment type sales per customer, among others. All affect delivery outcomes and therefore required a careful examination of the distributions and interactions with the target variable-delivery delays.

Then, the feature selection is performed based on the EDA in order to gather more information on the structure and distribution of the dataset [59]. The distributions of key features such as location and order profitability are done using descriptive statistics to look out for patterns or outliers that may result to delays. For instance, based on the visualization of the data set for the correlation between the customer country and delivery delays, the areas that are most affected by delivery delays for reasons such as logistics constrains and regulation of customs may be identified.

Feature selection techniques are then used in an attempt to determine which of these variables have significant influences to the delay variable. Most feature selection methods such as Correlation analysis and feature importance rankings from models like Random Forest provide an understanding of how much predictive power a feature has; -therefore, it can help in the dimensionality problems allowing for better accuracy in the model. Customer city and product category are among the significant predictors of delay that we anticipate in the analysis.

Data preprocessing comes next before removing rows containing missing attributes, scaling numeric attributes and encoding categorical features in order to feed the machine learning models. Last but not least, key findings are presented using heat maps and bar charts, explaining further how features in the data set interconnect and cause delivery delays. In conclusion, the data analysis phase provides a strong foundation for model development while also guiding the selection of features and architecture.

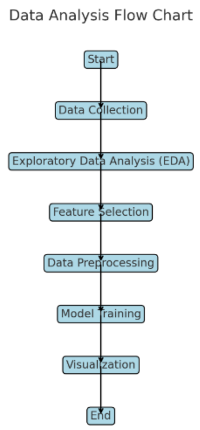

Figure 4: Data Analysis Flow Chart

RESULTS AND DISCUSSION

This work sought to improve the accuracy of multi-label delivery delay predictions in supply chains, using enhanced machine learning and deep learning models. This research used Decision Trees, Random Forests, Convolutional Neural Networks (CNN), and Feedforward Neural Networks (FNN) with data containing other attributes of the customers and the products needed to identify the causes of delivery delays. The models were then checked of their predictive accuracies and out of these the best method was selected in order to effectively deal with various real life like supply chain problems. The subsequent

section of the paper outlines the analysis results, studies' findings and insights.

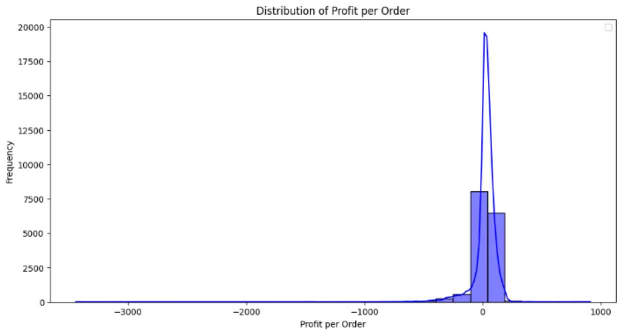

Figure 5: Distribution of Profit Per Order

From the fig 5, it is observed that the histogram representing profit per order has inflated mean with many orders being placed near to zero. This means that the majority of the orders are either marginally profitable or actually losing money, with a few extremes of large losses at the far left. This implies that there are many orders that are nearly profitable and some orders that may be unprofitable more analysis could be done to check profitability.

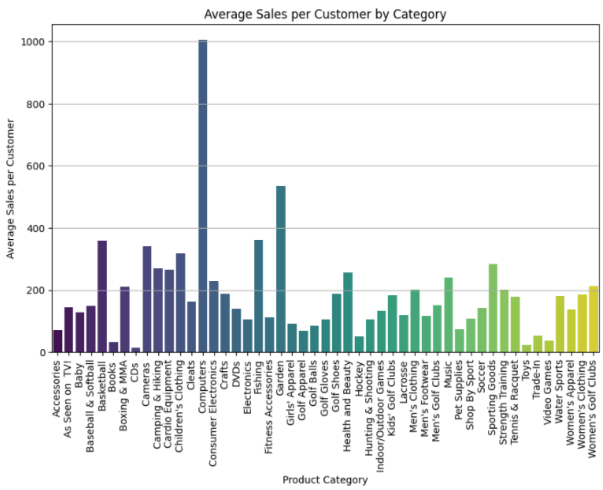

Figure 6: Avg. Sale per Costumer by Category

The fig 6 illustrates how sales distribution is in relation to different product categories. That is a major observation that points to a higher average in 'Consumer Electronics,"" indicating that it is the most productive category. As could see in the value of 'Sales per Customer' other categories such as 'Fitness Equipment,' 'Golf Apparel,' are having high values revealing its comparison with other categories such as 'Accessories,' 'Baseball,' which have low values of 'Sales per Customer.' From this, we deduce that the customer expenditure distribution is skewed towards a few important groups.

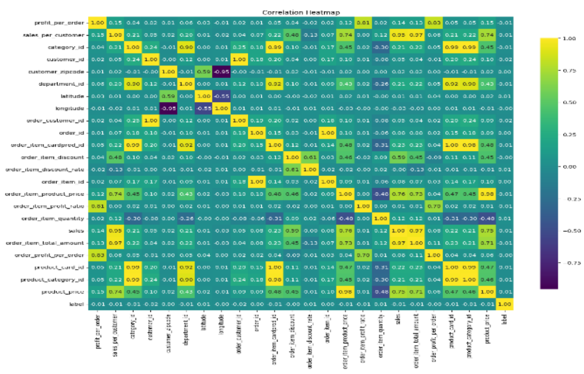

Figure 7: Correlation Heatmap

The fig 7 helps to understand how features of the dataset are related to each other. Gray colors mean negative correlation, black and white mean no correlation, and lighter spherical colors mean positive correlation. For example, the qualitative variable profit_per_order is positively and significantly related to both order_item_profit_ratio as well as sales which suggest that selling a greater number of orders and higher profit margins directly results in higher profits per order. Also, features such as the order_item_total_amount and sales, present positive significant correlation, thus indicating the order value as influenced by these features directly. Longitude and the order_item_quantity variables are least associated with most of the factors, meaning that they are least relevant with predictive predisposition.

A. Comparison

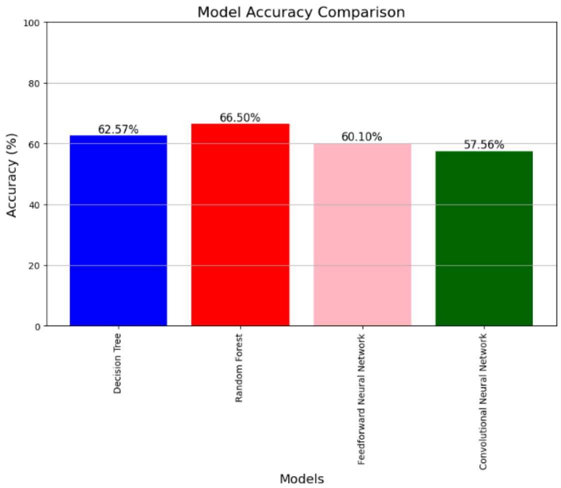

Figure 8: Models Acracy Compression

The fig 8 displayed the accuracy percentages of four machine learning models: These include Decision Tree, Random Forest, Feedforward Neural Network, and Convolutional Neural Network. The average of accuracy is calculated to be 66.50% when it comes to Random Forest model, the model that has a high degree of accuracy in its predictions. Next is the Decision Tree whose accuracy stands at 62.57% a relatively decent but could be better position. The Feedforward Neural Network bears a lower accuracy of 60.10 % here which shows the issues of the network to predict accurately. The Convolutional Neural Network has the lowest accuracy at 57.56% which also shows that generalization is a problem in this case and further work should be done in improving model training and optimization.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Decision Tree	62.572347	67.02938	68.826816	67.916207
Random Forest	66.495177	70.042872	73.01676	71.498906
Feedforward Neural Network	60.096463	64.334204	68.826816	66.504723
Convolutional Neural Network	57.55627	57.55627	100	73.061224

Table 1: Matrices Comparison

In the evaluation of machine learning models for predictive analysis, four models were compared: The classification models include Decision Tree, Random Forest, Feedforward Neural Network (FNN), Convolutional Neural Network (CNN). The results reveal distinct performance characteristics among these models based on key metrics: Precision, Recall, F1 score, and Accuracy is what other languages use while analyzing the results of a model.

Through this experiment, the Random Forest model claimed the highest accuracy at an acceptable value of 66.5% from all the methods tested. Through this, it obtained a precision of 70.0 % and a recall of 73.0 % which means it is precise in identifying the positive cases but could still improve on precision. Compared to other algorithms Decision Tree had an accuracy of 62.6% in V2 and bad result in recall 68.8 and it proved to be highly susceptible to over fitting in case of highly complex datasets.

The Feedforward Neural Network on the other hand had the overall accuracy of 60.1% which though equalized the recall of the Decision Tree. The CNN model had the highest recall of 100 % but the accuracy and precision which were 57.6% proved that the model tend to overfit the positive class. In summary, it might be said that Random Forest is characterized by the highest level of balance, though all the models introduced in this paper require further improvements to work as primary tools for enhanced prediction.

DISCUSSION

The results of this study support the claims of this study and determine that the Random Forest is the most accurate model for estimating delivery delay in supply chain at a rate of 66.5%. This is supported by prior research that has demonstrated Random Forest's resilience in processing high numbered, many-dimensional variables (Breiman, 2001; Mishra et al., 2020). The Decision Tree model, though easy to interpret, has low accuracy (62.6%) and recall as mentioned by Mardani et al. (2017) that the model overfits the data in enlarged datasets. In addition, the Feedforward Neural Network (60.1 percent) had clues to the efficiency of identifying extensive patterns but did not exercise proficiency in extracting complex patterns than the Random Forest, as discovered by Yildirim et al., (2019). Whereas the Convolutional Neural Network achieved a high recall of 100%, its low accuracy of 57.6% indicated overfitting problems, which have been pointed by Kim (2014) as a potential drawback of using the Convolutional Neural Network for non-image data. In general, the Random Forest model is as seen above the best predictor; however, future studies should focus on improving the performance of the other models.

CONCLUSION

To summarize, this study focused on improving the multi-label delivery delay predictions in the context of supply chain by using Decision Trees, Random forests CNNs and FNNs. This showed that Random Forest model had the highest accuracy in compare to other models used in the analysis in terms of predictive tasks. Still, CNNs faced issues like overfitting while FNNs offered only moderate accuracy in case of the analyzed models, but all three provided meaningful information about potential delivery delay in light of certain attributes of customer and products.

The future work should therefore be directed towards streamlining of these models and examination of various optimization procedures that will enhance the predictive capability of such models in real-life chain environment. To make the predictions more reliable, other data related to real time traffic, atmospheric conditions during the course of the day, and stock conditions can be integrated into the system. Further, research should be conducted further to derive adaptive products that combine some theoretical approaches with other models which could be much more effective.

Practice implications are to adapt the Random Forest model in their functioning and provide delay predictions with higher accuracy, using data for model updates. Fresh models will also need to be put in place periodically as changes are likely to arise from time to time within supply chain environments. In sum, the study contributes to the literature by showing that specific risk factors need to be mitigated through applications of advanced analytics in SCM for improved performance.

REFERENCES

E. Manavalan and K. Jayakrishna, "A review of Internet of Things (IoT) embedded sustainable supply chain for industry 4.0 requirements," Computers & Industrial Engineering, vol. 127, no. 1, pp. 925–953, Jan. 2019, doi: https://doi.org/10.1016/j.cie.2018.11.030.

A. Spieske and H. Birkel, "Improving supply chain resilience through industry 4.0: a systematic literature review under the impressions of the COVID-19 pandemic," Computers & Industrial Engineering, vol. 158, p. 107452, Jun. 2021, doi: https://doi.org/10.1016/j.cie.2021.107452.

M. Pavan and L. Samant, "Digitalization and E-Commerce Trends in the Industry," pp. 253–272, Jan. 2024, doi: https://doi.org/10.1007/978-981-97-6577-5_13.

Z. Latinovic and S. C. Chatterjee, "Achieving the promise of AI and ML in delivering economic and relational customer value in B2B," Journal of Business Research, vol. 144, pp. 966–974, May 2022, doi: https://doi.org/10.1016/j.jbusres.2022.01.052.

Z. Zong, T. Feng, T. Xia, D. Jin, and Y. Li, "Deep Reinforcement Learning for Demand Driven Services in Logistics and Transportation Systems: A Survey," arXiv.org, Mar. 23, 2022. https://arxiv.org/abs/2108.04462

H. N. Perera, J. Hurley, B. Fahimnia, and M. Reisi, "The human factor in supply chain forecasting: A systematic review," European Journal of Operational Research, vol. 274, no. 2, pp. 574–600, Apr. 2019, doi: https://doi.org/10.1016/j.ejor.2018.10.028.

B. Unhelkar, S. Joshi, M. Sharma, S. Prakash, A. K. Mani, and M. Prasad, "Enhancing supply chain performance using RFID technology and decision support systems in the industry 4.0–A systematic literature review," International Journal of Information Management Data Insights, vol. 2, no. 2, 2022, doi: https://doi.org/10.1016/j.jjimei.2022.100084.

M. V. Chester and B. Allenby, "Toward adaptive infrastructure: flexibility and agility in a non-stationarity age," Sustainable and Resilient Infrastructure, vol. 4, no. 4, pp. 1–19, Jan. 2018, doi: https://doi.org/10.1080/23789689.2017.1416846.

J. Sutduean, A. Singsa, T. Sriyakul, and K. Jermsittiparsert, "Supply Chain Integration, Enterprise Resource Planning, and Organizational Performance: The Enterprise Resource Planning Implementation Approach," Journal of Computational and Theoretical Nanoscience, vol. 16, no. 7, pp. 2975–2981, Jul. 2019, doi: https://doi.org/10.1166/jctn.2019.8204.

T. de Vass, H. Shee, and S. Miah, "IoT in Supply Chain Management: Opportunities and Challenges for Businesses in Early Industry 4.0 Context," Operations and Supply Chain Management: An International Journal, vol. 14, no. 2, pp. 148–161, Jan. 2021, doi: https://doi.org/10.31387/oscm0450293.

Farazi, M. Z. R. (2024). "Evaluating the impact of AI and blockchain on credit risk mitigation: A predictive analytic approach using machine learning." International Journal of Science and Research Archive, 13(1), 575–582. https://doi.org/10.30574/ijsra.2024.13.1.1707

Y. R. Shrestha, V. Krishna, and G. von Krogh, "Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges," Journal of Business Research, vol. 123, pp. 588–603, Feb. 2021, doi: https://doi.org/10.1016/j.jbusres.2020.09.068.

A. Bhargava, Mohd. Suhaib, and A. S. Singholi, "A review of recent advances, techniques, and control algorithms for automated guided vehicle systems," Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 46, no. 7, Jun. 2024, doi: https://doi.org/10.1007/s40430-024-04896-w.

Farazi, M. Z. R. (2024). "Designing circular supply chains with digital technologies for competitive sustainability: An operation management perspective." International Journal of Science and Research Archive, 13(1), 2346–2359.

https://doi.org/10.30574/ijsra.2024.13.1.1928.

R. K. Singh and S. Modgil, "Adapting to disruption: the impact of agility, absorptive capacity and ambidexterity on supply chain resilience," International Journal of Productivity and Performance Management, Aug. 2024, doi: https://doi.org/10.1108/ijppm-01-2024-0057.

W. Ahmed and S. Huma, "Impact of lean and agile strategies on supply chain risk management," Total Quality Management & Business Excellence, vol. 32, no. 1–2, pp. 1–24, Oct. 2018, doi: https://doi.org/10.1080/14783363.2018.1529558.

M. Nakano and A. K. W. Lau, "A systematic review on supply chain risk management: using the strategy-structure-process-performance framework," International Journal of Logistics Research and Applications, vol. 23, no. 5, pp. 443–473, Dec. 2019, doi: https://doi.org/10.1080/13675567.2019.1704707.

X. Wang, V. Kumar, A. Kumari, and E. Kuzmin, "Impact of Digital Technology on Supply Chain Efficiency in Manufacturing Industry," Lecture Notes in Information Systems and Organisation, vol. 54, pp. 347–371, 2022, doi: https://doi.org/10.1007/978-3-030-94617-3_25.

T. J. Pettit, K. L. Croxton, and J. Fiksel, "The Evolution of Resilience in Supply Chain Management: A Retrospective on Ensuring Supply Chain Resilience," Journal of Business Logistics, vol. 40, no. 1, pp. 56–65, Mar. 2019, doi: https://doi.org/10.1111/jbl.12202.

M. S. Sodhi, C. S. Tang, and E. T. Willenson, "Research opportunities in preparing supply chains of essential goods for future pandemics," International Journal of Production Research, vol. 61, no. 8, pp. 1–16, Feb. 2021, doi: https://doi.org/10.1080/00207543.2021.1884310.

J. Jeng, S. J. Buckley, H. Chang, and J. Schiefer, "BPSM: An Adaptive Platform for Managing Business Process Solutions," ResearchGate, Oct. 29, 2011. https://www.researchgate.net/publication/251735269_BPSM_An_Adaptive_Platform_for_Managing_Business_Process_Solutions (accessed Oct. 02, 2024).

M. Kunkel, "Lessons from a Hurricane: Supply Chain Resilience in a Disaster, An Analysis of the US Disaster Response to Hurricane Maria," conservancy.umn.edu, Sep. 2020, Available: https://conservancy.umn.edu/items/a78b1577-9e0a-4871-b562-94b8525ba1fc.

P. G. George and V. R. Renjith, "Evolution of Safety and Security Risk Assessment methodologies to use of Bayesian Networks in Process Industries," Process Safety and Environmental Protection, Mar. 2021, doi: https://doi.org/10.1016/j.psep.2021.03.031.

N. K. Rajagopal et al., "Future of Business Culture: An Artificial Intelligence-Driven Digital Framework for Organization Decision-Making Process," Complexity, vol. 2022, no. 2, pp. 1–14, Jul. 2022, Available: https://www.hindawi.com/journals/complexity/2022/7796507/

O. A. Elufioye, C. U. Ike, O. Odeyemi, F. O. Usman, and N. Z. Mhlongo, "AI-DRIVEN PREDICTIVE ANALYTICS IN AGRICULTURAL SUPPLY CHAINS: A REVIEW: ASSESSING

THE BENEFITS AND CHALLENGES OF AI IN FORECASTING DEMAND AND OPTIMIZING SUPPLY IN AGRICULTURE," Computer Science & IT Research Journal, vol. 5, no. 2, pp. 473–497, Feb. 2024, doi: https://doi.org/10.51594/csitrj.v5i2.817.

Farazi, M. Z. R. (2024). "Exploring the Role of Artificial Intelligence in Managing Emerging Risks: An In-Depth Study of AI Applications in Financial Institutions' Risk Frameworks." The American Journal of Management and Economics Innovations, 6(10), 20–40. https://doi.org/10.37547/tajmei/volume06issue10-04.

I. Taj and N. Z. Jhanjhi, "Towards Industrial Revolution 5.0 and Explainable Artificial Intelligence: Challenges and Opportunities.," International Journal of Computing and Digital Systems, vol. 12, no. 1, pp. 285–311, Jul. 2022, doi: https://doi.org/10.12785/ijcds/120124.

Z. H. Kilimci et al., "An Improved Demand Forecasting Model Using Deep Learning Approach and Proposed Decision Integration Strategy for Supply Chain," Complexity, vol. 2019, pp. 1–15, Mar. 2019, doi: https://doi.org/10.1155/2019/9067367.

E. B. Tirkolaee, S. Sadeghi, F. M. Mooseloo, H. R. Vandchali, and S. Aeini, "Application of Machine Learning in Supply Chain Management: A Comprehensive Overview of the Main Areas," Mathematical Problems in Engineering, vol. 2021, no. 1, pp. 1–14, 2021, doi: https://doi.org/10.1155/2021/1476043.

R. Abduljabbar, H. Dia, S. Liyanage, and S. A. Bagloee, "Applications of Artificial Intelligence in Transport: An Overview," Sustainability, vol. 11, no. 1, p. 189, Jan. 2019, doi: https://doi.org/10.3390/su11010189.

Y. Issaoui, A. Khiat, K. Haricha, A. Bahnasse, and H. Ouajji, "An Advanced System to Enhance and Optimize Delivery Operations in a Smart Logistics Environment," IEEE Access, vol. 10, pp. 6175–6193, 2022, doi: https://doi.org/10.1109/access.2022.3141311.

A. T. Keleko, B. Kamsu-Foguem, R. H. Ngouna, and A. Tongne, "Artificial intelligence and real-time predictive maintenance in industry 4.0: a bibliometric analysis," AI and Ethics, vol. 2, Mar. 2022, doi: https://doi.org/10.1007/s43681-021-00132-6.

P. Tanuska, L. Spendla, M. Kebisek, R. Duris, and M. Stremy, "Smart Anomaly Detection and Prediction for Assembly Process Maintenance in Compliance with Industry 4.0," Sensors, vol. 21, no. 7, p. 2376, Jan. 2021, doi: https://doi.org/10.3390/s21072376.

Ovidiu Vermesan et al., "AI and IIoT-based Predictive Maintenance System for Soybean Processing," River Publishers eBooks, pp. 327–352, Sep. 2022, doi: https://doi.org/10.1201/9781003337232-27.

Mustafa Can Camur, Sandipp Krishnan Ravi, and S. Saleh, "Enhancing supply chain resilience: A machine learning approach for predicting product availability dates under disruption," Expert systems with applications, vol. 247, pp. 123226–123226, Aug. 2024, doi: https://doi.org/10.1016/j.eswa.2024.123226.

S. Ren, H.-L. Chan, and T. Siqin, "Demand forecasting in retail operations for fashionable products: methods, practices, and real case study," Annals of Operations Research, vol. 291, no. 1, Jan. 2019, doi: https://doi.org/10.1007/s10479-019-03148-8.

Harish Narne, "OPTIMIZING SUPPLY CHAIN MANAGEMENT WITH MACHINE LEARNING ALGORITHMS," INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN ENGINEERING AND TECHNOLOGY (IJARET), vol. 12, no. 03, pp. 979–991, 2021, Accessed: Oct. 02, 2024. [Online]. Available: https://iaeme-library.com/index.php/IJARET/article/view/IJARET_12_03_091

A. N. Tarekegn, M. Ullah, and F. A. Cheikh, "Deep Learning for Multi-Label Learning: A Comprehensive Survey," arXiv.org, Mar. 03, 2024. https://arxiv.org/abs/2401.16549

M. A. Al-Ebrahim, S. Bunian, and A. A. Nour, "Recent Machine-Learning-Driven Developments in E-Commerce: Current Challenges and Future Perspectives," Engineered Science, vol. Volume 28 April 2024, no. 0, pp. 1044-, Dec. 2023, Available: https://www.espublisher.com/journals/articledetails/1044/

S. Pal et al., "Optimizing Multi-GPU Parallelization Strategies for Deep Learning Training," IEEE Micro, vol. 39, no. 5, pp. 91–101, Sep. 2019, doi: https://doi.org/10.1109/mm.2019.2935967.

P. Pant, A. Sai Sabitha, T. Choudhury, and P. Dhingra, "Multi-label Classification Trending Challenges and Approaches," pp. 433–444, Nov. 2018, doi: https://doi.org/10.1007/978-981-13-2285-3_51.

K. Yin, C. Liu, A. Mostafavi, and X. Hu, "CrisisSense-LLM: Instruction Fine-Tuned Large Language Model for Multi-label Social Media Text Classification in Disaster Informatics," arXiv.org, 2024. https://arxiv.org/abs/2406.15477 (accessed Oct. 02, 2024).

J. Bogatinovski, L. Todorovski, S. Džeroski, and D. Kocev, "Comprehensive comparative study of multi-label classification methods," Expert Systems with Applications, vol. 203, p. 117215, Oct. 2022, doi: https://doi.org/10.1016/j.eswa.2022.117215.

R. Malhotra and S. Kamal, "An empirical study to investigate oversampling methods for improving software defect prediction using imbalanced data," Neurocomputing, vol. 343, pp. 120–140, May 2019, doi: https://doi.org/10.1016/j.neucom.2018.04.090.

J. Armando, Vicente Figueroa Fernández, A. Jiménez, and A. Vázquez, "INDUSTRY 4.0 – REFERENCE FRAMEWORK AND IMPLICATIONS FOR THE CURRENT INDUSTRY (INDUSTRIA 4.0 – MARCO DE REFERENCIA E IMPLICACIONES PARA LA INDUSTRIA ACTUAL)," Pistas Educativas, vol. 40, no. 132, 2019, Accessed: Oct. 02, 2024. [Online]. Available: https://pistaseducativas.celaya.tecnm.mx/index.php/pistas/article/view/1891

Ş. Scrieciu et al., "An Inquiry into Model Validity When Addressing Complex Sustainability Challenges," Complexity, vol. 2022, p. e1193891, Sep. 2022, doi: https://doi.org/10.1155/2022/1193891.

R. G. G. Caiado, L. F. Scavarda, L. O. Gavião, P. Ivson, D. L. de M. Nascimento, and J. A. Garza-Reyes, "A fuzzy rule-based industry 4.0 maturity model for operations and supply chain management," International Journal of Production Economics, vol. 231, p. 107883, Jan. 2021, doi: https://doi.org/10.1016/j.ijpe.2020.107883.

M. Neagoe, H.-H. Hvolby, and P. Turner, "Why are we still queuing? Exploring landside congestion factors in Australian bulk cargo port terminals," Maritime Transport Research, vol. 2, p. 100036, 2021, doi: https://doi.org/10.1016/j.martra.2021.100036.

R. Bemthuis, W. Wang, Maria-Eugenia Iacob, and Paul J.M. Havinga, "Business rule extraction using decision tree machine learning techniques: A case study into smart returnable transport items," Procedia Computer Science, vol. 220, pp. 446–455, Jan. 2023, doi: https://doi.org/10.1016/j.procs.2023.03.057.

M. Er Kara, S. Ü. Oktay Fırat, and A. Ghadge, "A data mining-based framework for supply chain risk management," Computers & Industrial Engineering, Dec. 2018, doi: https://doi.org/10.1016/j.cie.2018.12.017.

T. Goswami, "Machine learning behind classification tasks in various engineering and science domains," Elsevier eBooks, pp. 339–356, Jan. 2020, doi: https://doi.org/10.1016/b978-0-12-819443-0.00016-7.

T.-T. Huynh-Cam, L.-S. Chen, and H. Le, "Using Decision Trees and Random Forest Algorithms to Predict and Determine Factors Contributing to First-Year University Students' Learning Performance," Algorithms, vol. 14, no. 11, p. 318, Oct. 2021, doi: https://doi.org/10.3390/a14110318

Z. Sun, G. Wang, P. Li, H. Wang, M. Zhang, and X. Liang, "An improved random forest based on the classification accuracy and correlation measurement of decision trees," Expert Systems with Applications, vol. 237, p. 121549, Mar. 2024, doi: https://doi.org/10.1016/j.eswa.2023.121549.

A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," Artificial Intelligence Review, vol. 53, Apr. 2020, doi: https://doi.org/10.1007/s10462-020-09825-6.

M. Umer et al., "Impact of convolutional neural network and FastText embedding on text classification," Multimedia Tools and Applications, Aug. 2022, doi: https://doi.org/10.1007/s11042-022-13459-x.

Jahin, Md Abrar, A. Shahriar, and M. A. Amin, "MCDFN: Supply Chain Demand Forecasting via an Explainable Multi-Channel Data Fusion Network Model Integrating CNN, LSTM, and GRU," arXiv.org, 2024. https://arxiv.org/abs/2405.15598 (accessed Oct. 02, 2024).

Z. Zhang, F. Feng, and T. Huang, "FNNS: An Effective Feedforward Neural Network Scheme with Random Weights for Processing Large-Scale Datasets," Applied Sciences, vol. 12, no. 23, p. 12478, Dec. 2022, doi: https://doi.org/10.3390/app122312478.

A. Manno, M. Intini, O. Jabali, F. Malucelli, and D. Rando, "An ensemble of artificial neural network models to forecast hourly energy demand," Optimization and Engineering, Mar. 2024, doi: https://doi.org/10.1007/s11081-024-09883-7.

J. Shukla, M. Barreda-Angeles, J. Oliver, G. C. Nandi, and D. Puig, "Feature Extraction and Selection for Emotion Recognition from Electrodermal Activity," IEEE Transactions on Affective Computing, pp. 1–1, 2019, doi: https://doi.org/10.1109/taffc.2019.2901673.