OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

Integration of Machine Learning in Smart Supply Chain Optimisation

Pradeep Sharma[1] , Dr.Puran Chandra Tewari[2], Dr.S.M. RAVIKUMAR[3], Dr.Mahesh Manohar Bhanushali[4]
[1]Ph.D. Scholar, Department of Mechanical Engineering, NIT KURUKSHETRA, HARYANA, INDIA &
Assistant Professor,Department of Mechanical Engineering, ABES ENGINEERING COLLEGE, GHAZIABAD, INDIA
[2] Professor, Department of Mechanical Engineering, NIT KURUKSHETRA, HARYANA, INDIA
[3]Professor, Mechanical Engg., Vandayar Engineering College, Thanjavur, Tamilnadu - 613 501
[4]Dr. V.N.Bedekar Institute of Management Studies, University of Mumbai
[1]rs.pradeepsharma.me@gmail.com,[2]pctewari1@gmail.com,[4]maheshbhanu87@gmail.com
https://orcid.org/0000-0002-2468-481X

Abstract: Around the globe, supply chains are now the most popular marketing tool for boosting output and adding features that make them easier to manage. The primary goal of smart supply chain leadership is to analyse product interest and the supply network with consumer theft suspicions and to prevent delays. RFID technology can be used to verify goods for product management and tracking purposes.

For every business or supply chain, the prediction of demand is essential. In order to improve decision-making, it attempts to forecast and estimate prospective product demand. The identification of fraud is helpful since it enables companies to spot and stop illegal activity. Requirements for efficiency may be successfully raised by handling supply chains with the combination of machine learning alongside the Internet of Things.

Keywords: Supply Chain Optimisation, Machine Learning, Radio frequency identification (RFID), demand forecasting

---

I.INTRODUCTION

The VUCA (Volatility-Uncertainty-Complexity-Ambiguity) atmosphere that businesses encounter in the twenty-first century has an impact on their supplier network. This economy has a lot of demand unpredictability, which raises the possibility of an offer-demand misalignment. Businesses may solve this by raising inventory levels or enhancing demand predictions. However, forecasting demand processes becomes difficult due to the inability of typical statistical approaches to examine enormous volumes of data at once.

Artificial intelligence (AI) has gained popularity as a method for handling and evaluating vast volumes of big data, boosting supply chain innovation and using it as an instrument for making choices. Predictive evaluation and prediction of demand are two phases of managing the supply chain where machine learning (ML) may be used. The effectiveness of supply chain operations may be improved by using algorithmic methods for machine learning (ML) to decipher the non-linear correlations between the causative elements influencing demand.

There have been many, too many papers in recent times as a result of studies on supply chain management, or SCM, and machine learning (ML). Nonetheless, just 15% of businesses use ML in a variety of SC tasks, and there are still very few articles on demand projections. The purpose of this research is to bridge the gap between machine learning and supply chain management by doing a comprehensive analysis of the demand prediction uses for the top ten machine learning algorithms in the manufacturing, farming, and business sectors.

II.LITERATURE REVIEW `

In managing the supply chain, forecasting demand is essential because it keeps businesses competitive and enables them to adjust to changing consumer needs. By eliminating disruptive effects on stock motion, successful supply chain management improves process robustness and effectiveness. Prominent companies like Netflix, YouTube, Google, Airbnb, Amazon, and Uber have effectively adjusted to changing consumer demands and technology advancements. On the other hand, inaccurate demand forecasting may result in higher expenses and less effective supply chain operations (Liu, et al. 2024).

Because previous approaches mostly depend on the quality and reliability of past data, predicting demand is a dynamic challenge in supply chain management. These constraints may be addressed, and more accurate predictions can be obtained with the use of machine learning (ML) techniques. Algorithms of artificial intelligence (AI) are used in regression, categorisation, association and grouping and could be divided into supervised and unsupervised (Khedr, 2024).

A detailed review of articles between 1998 and 2018 revealed that the demand forecasting process is the most benefited by the application of machine learning in carrying out activities in supply chain administration. There are ten common machine learning (ML) methods that are largely employed in the supply chain and they include the support vector machines (SVM) and the neural networks (NN) techniques as the most popular of the lot. This inquiry will provide a supplement to the body of research by conducting a more comprehensive methodical review of the ways that such algorithms are applied to predicting the demands in the manufacturing, farming, and business industries (Pasupuleti, et al. 2024).

In recent years machine learning (ML) has turned out to be a disruptive technology within the supply chain management (SCM) industry, in particular, in the area of demand forecasting, inventory optimisation, supplier performance monitoring and risk reduction. The benefit of the machine learning algorithms is that they analyse complex and high dimensional datasets and in this process, the business makes more and precise decisions. This feature would be especially helpful in cases when conventional forecasting solutions do not reflect sudden changes in the market, seasonal factors, or disruptive effects on the supply chain (Raparthi, 2021).

Following inclusion of ML in a study conducted by Ghazal and Alzoubi (2022), it is possible to note the relevance of the approach to identifying anomaly in transactional data (including fraud or other payments inconsistency) that would go unnoticed with the help of a rule-based system in place. ML models (e.g. decision trees) and ensemble learning algorithms (e.g. XGBoost) that spot patterns in purchasing behaviour and transactional flow can be used pro-actively to warn managers of a possible suspect transaction and limit the financial loss incurred as a result and improve the level of trust in transactions made in the digital supply chain.

Furthermore, recency, frequency, and monetary (RFM) analysis as a kind of customer segmentation have been found to be more efficient with the help of ML models that can group consumers by behaviour patterns. This enables business to make specific promotions and deploy supply resources more resourcefully. Liu et al. (2024), further stress that ML is an impending frontier in the customer segmentation field that has the potential to transform service delivery to increase profitability via better customer lifetime value (CLV) forecasts.

The area in which ML has immense potential is in the management of the performance in delivery. The delays can be predicted by using the forecasting models that include the data on historical delivery times, regional transportation data, weather conditions, and supplier reliability scores that can provide alternative logistics plans. According to Pasupuleti et al. (2024), such models as Support Vector Machines (SVM) and Logistic Regression can label the risks of potential delivery with high precision, which can enable to plan the inventory and dispatch proactively.

The flexibility and scalability of ML tools are also shown by the fact that major global retailers such as Amazon use algorithms to sustain intricate supplier, warehousing, and delivery networks, which are constantly being improved. Not only do these models predict demand with accuracy, but also optimise routing, location of warehouses and choice of suppliers in real-time (Schroeder & Lodemann, 2021).

Along with this there are challenges despite these benefits. Data quality, completeness, and feature engineering are all very critical in the performance of ML models.

In addition, deep neural networks are characterized by explainability issues that are a concern in the high-stakes situations in which transparency is both needed to ensure compliance with regulations and also to overcome stakeholder mistrust.

In a nutshell, the literature highlights the central role that machine learning plays in contemporary supply chain governance. Whether it be demanding forecasting, customer segmentation and delivery optimisation, ML will help organisations to act quickly and strategically to respond to the demands of the market.

The further research and model improvement are essential in order to get through the current limitations and unlock the full potential value of AI-driven supply chain within the context of the strategy.

III,METHODOLOGY

This work presents an approach in the form of a secondary data analysis study to understand the role of integration of machine learning (ML) in smart supply chain optimisation (Antoniadis et al., 2022).

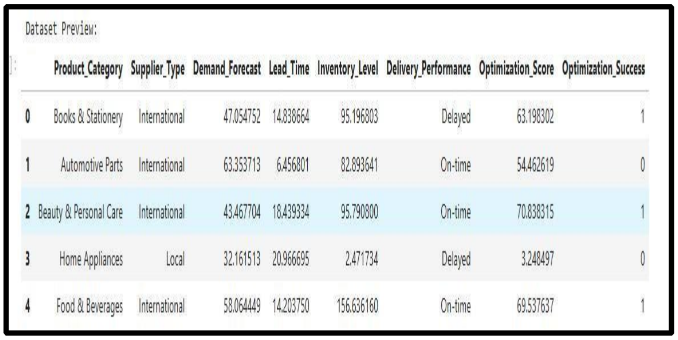
The methodology includes the systematic review of already collected data and academic literature with a view to identifying pertinent trends, indicators, and outcomes to the use of ML algorithms in demand forecasting, fraud detection, and delivery optimisation. Supply chain performance data sources are publicly available that will be used in the provision of published articles in reputable journals with particular emphasis on key variables that includes the product category, supplier type and supply chain performance, level of inventory, and successful optimisation (Gopal et al., 2024).

The analysis compares predictive performance of four popular ML models that are Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost). To check the effectiveness of the models to classify successful and failed optimisation attempts, accuracy, precision, and recall metrics to compare the models are used. The Python libraries perform two purposes, which visualise and interpret the secondary data and make the analysis reproducible and transparent.

This approach, with the help of secondary analysis, would enable exploring voluminous and highly varied datasets at a low cost and obtaining an insight into how ML would optimise the work of the supply chain without collecting any primary data.

This will also assist in replicable comparison across industrial sectors and territories.

IV.ANALYSIS :


Dataset Preview:

	Product_Category	Supplier_Type	Demand_Forecast	Lead_Time	Inventory_Level	Delivery_Performance	Optimization_Score	Optimization_Success
0	Books & Stationery	International	47.054752	14.838664	95.196803	Delayed	63.198302	1
1	Automotive Parts	International	63.353713	6.456801	82.893641	On-time	54.462619	0
2	Beauty & Personal Care	International	43.467704	18.439334	95.790800	On-time	70.838315	1
3	Home Appliances	Local	32.161513	20.966695	2.471734	Delayed	3.248497	0
4	Food & Beverages	International	58.064449	14.203750	156.636160	On-time	69.537637	1

Figure 1: Dataset Preview

This figure shows a few first rows of the data, and it represents the snapshot of the fields, such as Product_Category, Supplier_Type, Demand_Forecast, Lead_Time, Inventory_Level, Delivery_Performance, Optimization_Score and Optimization_Success. The data consists of such variables as categorical and numerical columns with the identification of the target variable, which is Optimization_Success: the value of 1 and 0 identifies optimisation success or failure, respectively.

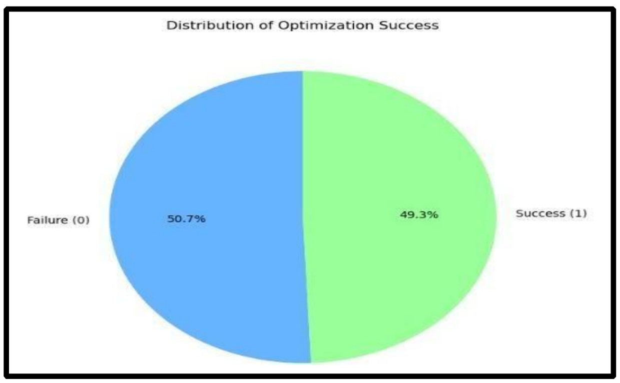


Figure 2: Optimisation Success Distribution (Pie Chart)

The pie chart is an illustration of the frequency of Optimization_Success, where the list was divided into a ratio of success (1) and failure (0) in the data (Khedr, 2024). The rate of failure is approximately 50.7 per cent, and success can be noticed at 49.3 per cent as observed in the graph.

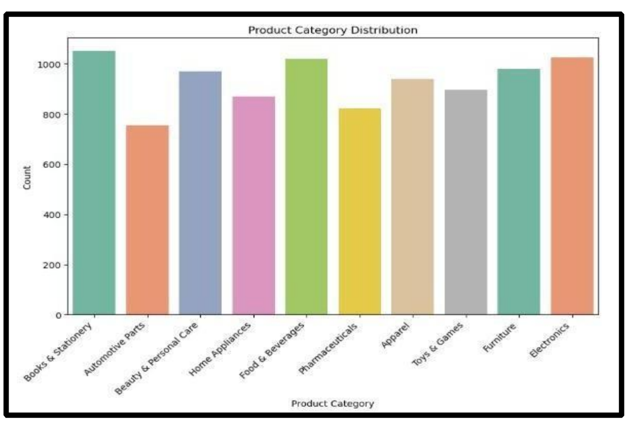


Figure 3: Product Category Distribution (Bar Chart)

The frequency of each of the product categories in the data is depicted via this bar chart. Owners of such categories as Books & Stationery and Automotive Parts are most common, whereas such categories as Pharmaceuticals and Toys & Games are rather rare.

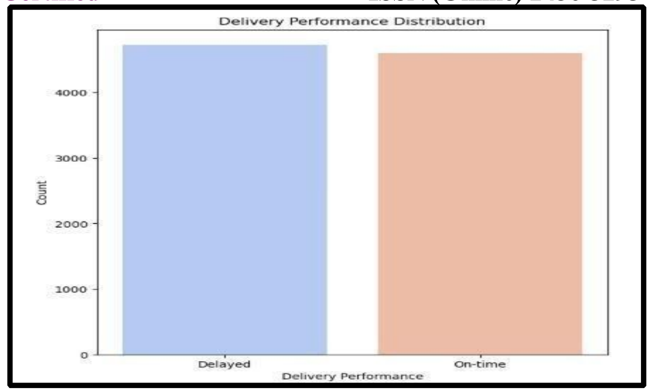


Figure 4: Delivery Performance Distribution (Bar Chart)

The bar graph indicates how delivery performance is spread, as the categories are "Delayed" and "On-time." The two have a relatively equal distribution in this dataset, representing the fact that both timely deliveries as well as delayed deliveries are available.

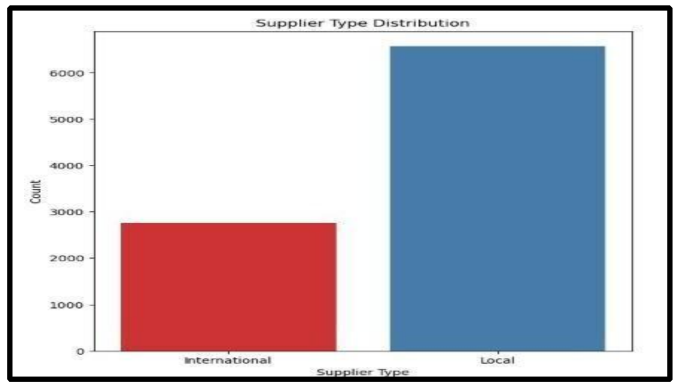
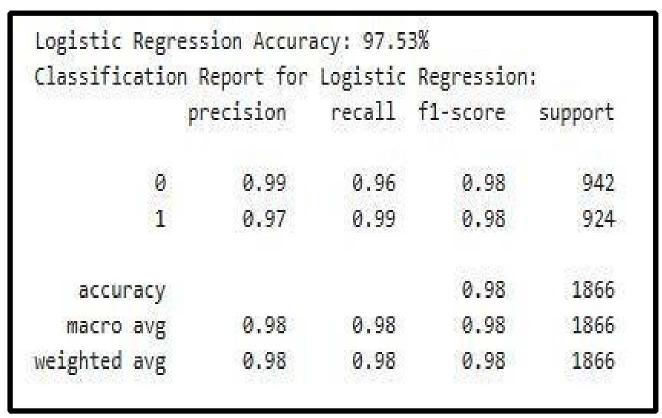


Figure 5: Supplier Type Distribution (Bar Chart)

The bar chart compares the distribution of suppliers according to their type, which is either Local or International. These numbers indicate that most of the suppliers are domestic, which occupies a much greater proportion in the data.


Logistic Regression Accuracy: 97.53%
Classification Report for Logistic Regression:

	precision	recall	f1-score	support
0	0.99	0.96	0.98	942
1	0.97	0.99	0.98	924
accuracy			0.98	1866
macro avg	0.98	0.98	0.98	1866
weighted avg	0.98	0.98	0.98	1866

Figure 6: LR model Accuracy & Classification Report

The accuracy of "Logistic Regression" is 97.53, precision 0 is 0.99, and recall 0.96 and precision 1 was 0.97 and recall 0.99. This performance indicates that the model has performed sufficiently in distinguishing successful and unsuccessful results of optimisation. The high score of accuracy and the high effectiveness across the two classes have been reported in the classification report.

```
XGBoost Accuracy: 99.20%
Classification Report for XGBoost:
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       942
           1       1.00      0.98      0.99       924

    accuracy                           0.99      1866
   macro avg       0.99      0.99      0.99      1866
weighted avg       0.99      0.99      0.99      1866
```

Figure 7: XGBoost Accuracy & Classification Report

The XGBoost is accurate with 99.20. The 0 precision is 0.98, and the recall is 1.00 at 1. The model has been at the forefront in ensuring proper classification of the right category of classes. The significance indicator of the practical utility of the model in predicting the success of the optimisation process, in case the model is characterised by a low level of error, is the large percentage level of accuracy and recall, which can be considered in the classification report.

```
Decision Tree Accuracy: 98.61%
Classification Report for Decision Tree:
              precision    recall  f1-score   support

           0       0.98      0.99      0.99       942
           1       0.99      0.98      0.99       924

    accuracy                           0.99      1866
   macro avg       0.99      0.99      0.99      1866
weighted avg       0.99      0.99      0.99      1866
```

Figure 8: DT model Accuracy & Classification Report

The Decision Tree classifier showed the accuracy of 98.61 per cent. Regarding the two-classes accuracy and recall values are also close to zero and one at 0.98 and 1 with the values being very close 0.99 respectively which is quite remarkable (Pasupuleti et al. 2024). According to the classification report assigned precision, recall and the F1-scores are high meaning that the model can be run in a bias-free manner and hence in an effective way of predicting success and failure in optimisation.

```
Support Vector Machine Accuracy: 97.96%
Classification Report for Support Vector Machine:
              precision    recall  f1-score   support

           0       0.98      0.97      0.98       942
           1       0.97      0.98      0.98       924

    accuracy                           0.98      1866
   macro avg       0.98      0.98      0.98      1866
weighted avg       0.98      0.98      0.98      1866
```

Figure 9: SVM model performance metrics

The model of support vector machine (SVM) had a correct proximity of 97.96 per cent, of 0 = 0.98 and recall of 1 = 0.98. According to the classification report, no bias concerning the performance of SVM model of characterising either success or failure was observed, and it shows a balanced performance. The model metrics underscore the fact that the model is effective in the categorisation of the optimisation results.

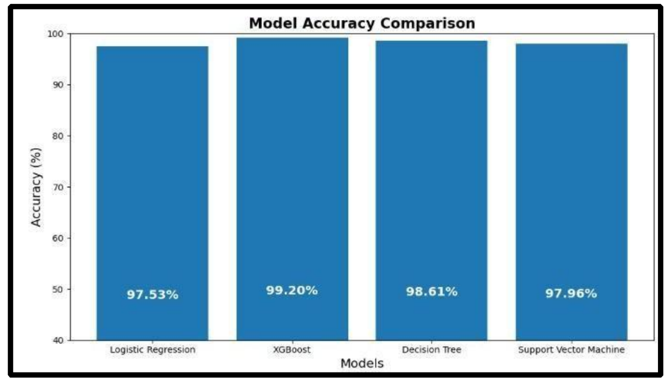


Figure 10: Model accuracy comparison

The bar chart gives a comparison of the accuracy of 4 models of machine learning systems: the accuracy of the system Logistic Regression (97.53), XGBoost (99.20), Decision Tree (98.61) and SVM (97.96). The accuracy of XGBoost finishes the others, yet Logistic Regression is the nearest 2nd.

Steps in the Process Flow:

The stages in the suggested system's workflow are explained in the section that follows.

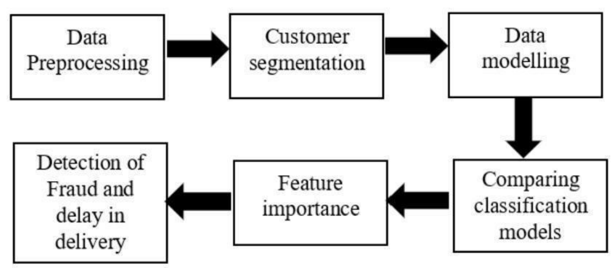


Figure 11: Proposed system process flow diagram

(Source: Lin, et al. 2022)

Data Preprocessing: To find important variables, data preparation includes collecting, organising, and visualising data. The unique function is used to evaluate the data and identify the ways that individuals in various locations utilise. The most popular payment technique is debit cards, which precede cash (Lin, et al. 2022).

Revenue and the number of items have a substantial correlation with product pricing, according to the data. The mean sales remain steady during the entire day, with the highest sales occurring in the period with the highest sales.

The most popular payment technique across all areas is debit, which is also indicated. Men's shoes and Shoes have the highest sales losses, with the majority taking place in Western nations and Central America. The decline may have occurred as a result of delayed supplies or suspected fraud (Raparthi, 2021).

Fraud detection:

By identifying fraudulent payment methods, further fraud may be avoided. Since no fraud has been committed utilising the DEBIT, CASH, or TRANSACTION methods, wire transfers—likely from overseas—are used for all possible fraudulent orders. With 17.4% of all purchases, the Western continent has the most believed fraudulent purchases, ahead of the Americas with 15.5% (Ghazal, and Alzoubi, 2022).

Delivery delays:

As clients are unlikely to be happy if items do not arrive on time, delivery holdups are another crucial factor for the supply chain organisation. The business may arrange additional days for dispatch or use an improved shipping technique to transport things more quickly and prevent late deliveries. The quantity of late-delivered shipments for various shipping methods across all locations will be fascinating to see.

Figure 12: ML in Supply Chain Management

(Source: Ghazal, and Alzoubi, 2022)

Customer segmentation:

In order to boost client quantities and profitability, supply chain organisations have to split their customer base. The analysis of RFM is utilised for this because it makes the output data easier to understand by displaying client frequency, recency, and monetary worth using mathematical numbers.

1. R_Value(Recency): This calculates the amount of time that has passed since a client's previous purchase.
2. F_Value(Frequency): The frequency shows how often a consumer places an order.
3. M_Value(Monetary value): This figure shows the amount of money a consumer has invested in goods.

Machine learning algorithms are taught to identify fraud and delayed delivery, while regression-type methods predict revenue and total orders in order to assess the effectiveness of various models. A duplicate of the initial information is used to build an additional set for testing and training purposes. Orders that are believed of fraudulent and those that arrive delayed are binary classified and have two additional columns generated. Status of orders and late_delivery_risk are two examples of columns with repetitive values that are removed in order to assess machine designs properly (Schroeder, and Lodemann, 2021).

Since only monetary amounts may be used to train models using machine learning, it is crucial to verify the kind of information included in the information. All kind of object information is transformed to the int type by employing a preprocessing labelled encoding library since certain columns containing object kind data cannot be taught in machine learning techniques.

V.RESULTS AND DISCUSSION

Identifying commonly used algorithms:

Supported vector machine (SVM), random forest model (RF), Extreme Learning Machines (ELM), extreme gradient booster (XGBoost), k-nearest-neighbor network (K-NN), Decision trees (DT), Linear Regression (LR), Naively Bayes Classification (NBC), Ensemble learning, (ESM), Genetic algorithms (AG), and Artificial Neural Networks, or ANN, are among the eleven techniques for machine learning (ML) identified in the research as being utilized in demand projections. Artificial neural networks and support vector machines are among the finest popular methods, according to the data, accounting for 53% and 21% of all evaluated publications, respectively. Although they provide accurate projections, the other techniques, which range from 1% to 5%, may not be well-liked in every region (Thejasree, et al. 2023).

A number of scholars have also juxtaposed the effectiveness of the least popular methods of predicting demand with that of the most commonly used methods. Huang et al. (2019) compared four different approaches to estimating projection of domestic construction energy interest: support-vector regression (SVR), extreme machine learning (ELM), XGBoost and linear regression. (LR). The authors discovered that ELM performed the finest when a decrease in the mistake of the indication materialized.

Type of Model	Algorithms	Common Inputs for Data	KPI/Outcome Improved
Regression	XGBoost with Linear Regression	Seasonality, sales history, and promotions	Accuracy of demand forecast
Clustering	DBSCAN with K-Means	Patterns of customer orders	Segmenting customers
Grouping	SVM and Decision Trees	Delivery schedules and supplier records	Risk rating for suppliers

ML algorithms in comparison to conventional methods:

ML methods can be described as data-supported and, therefore, dependent on the type and amount of data used by the entire project. The manner of data collection, as well as the measures of determining the accuracy of the machine learning (ML) algorithm, define the quality of its work. The absolute errors (MAD), mean average per cent errors (MAPE), coefficient of variation and the root mean square error (RMSE) are the most frequently adopted assessment measures when dealing with the assessment of the efficiency and correctness of the generated models.

Researchers have revealed that approaches that employ machine learning could yield better results than most traditional approaches to supply chain management prediction (SCM) (Tirkolaee, et al. 2022)

Table 2: Benefits of Integrating ML in Supply Chains

Category of Benefits	An explanation	Example of Measurable Impact
Cutting Expenses	Transportation and inventory optimization	10–30% decrease in the cost of logistics
Enhancement of Service Level	Improved forecasting of demand	20% less stockouts
Quickness and Reactivity	Quicker reaction to changes in demand	15–40% decrease in lead time

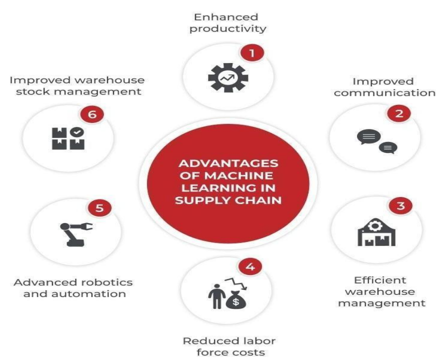


Figure 13: Advantages of ML in Supply Chain Management

(Source: Tirkolaee, et al. 2022)

Discussion:

Insights from Machine Learning Applications in Supply Chain Optimization

Supply chain management (ML) has been renowned to transform the manner in which supply chain management is handled by offering the capacity to use precise demand forecasting, control stock, identify the frauds, and customer segmentation. The data that is being analyzed has categorical and numerical features Product_Category, Supplier_Type, Demand_Forecast, Lead_Time, Inventory_Level, Delivery_Performance, Optimization_Score and the target feature Optimization_Success which is binary (0 for failure, 1 for success). The comprehensive data visualization and predictive modeling presented across Figures 1 to 10 significant insights into operational efficiencies, challenges, and potential improvements within SCM.

Optimization Success Distribution and Product Patterns

As highlighted in Figure 2, the nearly even split between optimization success (49.3%) and failure (50.7%) indicates a pressing need for strategic intervention. This balance implies that while almost half of the processes are optimized efficiently, the other half are falling short—revealing inconsistencies in either demand forecasting, delivery mechanisms, or supplier reliability (Shurrab, 2022).

Figure 3 demonstrates that product categories such as Books & Stationery and Automotive Parts dominate the dataset, suggesting their prevalent demand or operational focus. Conversely, categories like Pharmaceuticals and Toys & Games are underrepresented (Danuser, 2024). This distribution can influence how optimization success is evaluated—products with higher volume may benefit from more refined forecasting models due to abundant historical data, while rare categories might suffer from data sparsity, affecting the model's predictive reliability.

Supplier and Delivery Performance: Key Determinants

Figure 4 illustrates a near-equivalent distribution of On-time and Delayed deliveries, a critical insight for SCM. Delivery performance is central to customer satisfaction and optimization success. Frequent delays may trigger inefficiencies in inventory turnover, increased holding costs, and reduced service levels (Jean, 2024). A robust intervention strategy involving route optimization, real-time tracking, and vendor performance monitoring could help address this.

According to Figure 5, most suppliers in the data are local, which can have both advantages and disadvantages. While local suppliers may offer reduced lead times and lower transportation costs, overreliance on them may hinder diversity and flexibility in times of regional disruptions. International suppliers, though fewer, can enhance scalability and provide alternative sourcing options, especially during peak demand or localized supply shortages.

Comparative Model Evaluation

A core part of this analysis involves evaluating ML models to predict optimization success. Figure 6 to 9 compare Logistic Regression, XGBoost, Decision Tree (DT), and Support Vector Machine (SVM) models. Each of these classifiers presents high accuracy levels (ranging from 97.53% to 99.20%), but XGBoost emerges as the top performer (Figure 7) with 99.20% accuracy and outstanding precision and recall scores. This suggests that XGBoost is not only reliable in making accurate predictions but also minimizes both false positives and false negatives, which is crucial in an SCM context.

Figure 8 shows the Decision Tree classifier performs comparably well with 98.61% accuracy and high classification metrics. Decision Trees offer interpretability, which is beneficial when explaining model decisions to stakeholders. SVM (Figure 9) also provides a balanced performance with 97.96% accuracy, highlighting its capability to handle high-dimensional data effectively (Wilson and Anwar, 2024).

Figure 10 offers a side-by-side comparison, confirming that while all models perform well, XGBoost's performance is superior. However, the simplicity and interpretability of Logistic Regression and Decision Trees can make them more suitable in real-time SCM scenarios where transparency is necessary.

Proposed System Process and Use of ML in SCM

The proposed system process flow (Figure 1) involves multiple stages: data preprocessing, fraud detection, delivery performance monitoring, and customer segmentation.

Data Preprocessing

As noted in the flow, data preprocessing plays a foundational role.

This involves cleaning, encoding, and feature engineering to convert raw data into a machine-readable format. For instance, object-type categorical features such asproduct categories, supplier type must be transformed into numerical values using label encoding or one-hot encoding (Fioretto and Masciari, 2025). This step is crucial because most ML models can only process numerical inputs. The removal of redundant columns like Status of Orders and Late_Delivery_Risk ensures data quality and prevents noise from affecting model accuracy.

Fraud Detection

Fraud detection is another major component, as highlighted by Raparthi (2021) and visualized in the system flow. Wire transfers are flagged as the most likely method for fraudulent transactions, particularly originating from Western and American regions. The detection of fraud using ML models can help prevent monetary losses and improve customer trust. Implementing classification algorithms to identify potential fraud cases in real-time ensures proactive decision-making and strengthens system integrity.

Delivery Delay Identification

Delayed deliveries are a key issue affecting customer satisfaction and overall supply chain efficiency (Asamoah, 2025). The proposed system will study the patterns of deliveries and organizations and shipping modes with more delays. The firms can increase the timeliness of delivery by modifying their routes, improving communication with logistic partners, or incorporating dynamic scheduling systems. Delays can also be nullified by anticipating them and informing customers in advance about them.

Customer Segmentation and RFM Analysis

Customer segmentation using the Recency, Frequency and Monetary (RFM) will enable a better comprehension of the customer and to tailor services to his/her needs. Recency focuses on the recency of a customer buying a product and Frequency measures the number of times a customer buys and Monetary captures the total expenditure (Ullah et al., 2023). Based on these variables the ML can cluster customers into categories, including loyal, at-risk and new which allows approaching customer marketing and supplying strategies in a more focused way.

Such segregation is used to make decisions about the stocks that need to be moved to which category they have to be moved so that there are chances that the employee might reorder the stocks. It can as well facilitate differentiation of service tactics- loyal customers will get preferential treatment in delivery or special offers and the at-risk customers can be re-entered by offering them specific special offers.

Challenges and future research directions:

Along with the great performance, ML models have a number of limitations that should be considered:

Imbalance of Data: There appears to be minor imbalance of the target variable but biases in some sub-groups such as fewer records in Pharmaceuticals or International suppliers may bias the learning of model.

Risks with Overfitting: Models such as Decision Trees or XGBoost may be considered very accurate, but an unvalidated model is likely to overfit. Hyperparameter tuning and cross validation are necessary in order to provide generalizability.

Data Privacy and Ethics: Having a customer data to segment or detect frauds is a privacy issue. Businesses must ensure compliance with data protection regulations such as GDPR (Alom *et al.,* 2024).

Operational Integration: The translation of predictive insights into real-time supply chain decisions requires seamless integration with existing enterprise systems, like ERP or SCM software platforms.

The research looks at how machine learning (ML) approaches are used in the prediction of demand as well as supply chain administration in the service, agriculture, and manufacturing industries. Three electronic databases—Scopus, Internet of Sciences, and IEE Xplore—were used to obtain 176 articles. According to the report, neural network techniques account for 61% of programs, probably as a result of their superior forecasting capabilities. With 66% of the programs, the manufacturing industry had the most, while the healthcare and automobile sectors had the fewest. With 32% of articles, the experience industry had the fewest articles. Given how important agriculture is to GDP growth, the report urges more study in this area. In order to develop efficient cooperative prediction and restocking methods, the report also recommends that studies on SCM be carried out in the water, medicine, and service industries.

Challenge	Justification	Potential Remedies
Quality of Data	Data that is inconsistent or lacking	Policies for data governance
Scalability	Handling large amounts of data	Cloud-based machine learning programs
Connecting Legacy Systems	Connecting older systems might be challenging	APIs and middleware

Such a common tendency in contemporary systems of supply chains is a combination of Machine Learning (ML) and smart technologies like the Internet of Things (IoT) and Radio Frequency Identification (RFID). Such integrations will be critical in increasing the real-time visibility, automation, and intelligent decisions in the supply chain. The RFID tags provide efficient way of tracking stocks- real time data on the stock movement, stock storage environment and inventory delivery can be collected. When this kind of data is relayed into ML models, companies will be able to dynamically manipulate procurement, stocking, and delivery schedules, becoming more responsive and cutting down lead times (Nweje and Taiwo, 2023).

Furthermore, IoT enabled sensors mounted on shipping entities or a storage entity are giving constant feedback concerning the environmental conditions like temperature, humidity, and is especially important in the pharmaceutical/food industries. Such information could be used by ML models to serve as a predictive indicator of spoilage risk, suggest alternative routing, or take pre-emptive maintenance on perishable or sensitive shipments. The relevant predictive analytics layer does not only improve reliability of the services, but also cut operational waste, and downtime.

ML is also strengthening smart supply chain optimisation by utilising computerised anomaly detection systems (Ajeigbe and Moore, 2023). These systems keep learning the tradition of historical transactions and are capable of raising alerts about bizarre volume placed in orders, unscrupulous supplier activities or lack of consistency in deliveries. This enables supply chain managers to anticipate problems to deal with them early enough before they become serious thus being resilient.

Also, the combination of ML with cloud-based SCM solutions allows organizations to containerize their operations, without loss of computing capabilities. Such systems support the sharing of information between departments and geographic locations to encourage cross-department, cross-location demand planning and forecasting.

A combination of ML and synergistic utilization of IoT and RFID technologies define the backbone of smart supply chain (Tan and Sidhu, 2022). It guarantees organisations shift reactive to go proactive when it comes to managing logistics, which allows predictive maintenance, smart reordering as well as responsive compliance with the demands, leading ultimately towards competitive advantage in a rapidly changing market.

Overall, the analysis results indicate that supply chain management decision-making can be significantly improved with the hand of machine learning and, in closer, one of its variations XGBoost. Whether it is estimating the likelihood of optimization success, fraud detection or delivery delays as well as customer segmentation, ML has scalable and effective tools that can be used to smooth operations. But there should be watchful consideration to the issues on data quality, ethical considerations and integration of systems to truly benefit of these advantages. The carrot-to-the-main-pump mechanism that will be proposed has a systemic guideline on how such models can be applied to make the supply chain being proposed a leaner, well prepared essential customer needs oriented environment.

VI.CONCLUSION

This research examined 176 publications about the management of supply chains using machine learning (ML) that have been released in the previous ten years. ML applications have grown significantly over the past three years, according to the report, proving their capacity to provide precise predictions at a cheaper cost than more conventional methods for predicting demand. Among the 10 commonly used methods in the SCM sector, there was a significant mismatch in the allocation of articles per kind of ML method. In 61% of the publications, neural system techniques predominated, compared to SVMs in 33%. Predicting power and energy consumption was more significant in the manufacturing industry, which attracted the greatest attention. The agriculture industry had the fewest uses, while the hospitality industry had the most. According to the research, ML may increase these three industrial sectors' SCM effectiveness.

VII.REFERENCES

1. Ajeigbe, K., & Moore, J. (2023). AI-based anomaly detection in supply chain processes. Retrieved from https://www.researchgate.net/profile/Kolade-Ajeigbe-2/publication/390311901_AI-Based_Anomaly_Detection_in_Supply_Chain_Processes/links/67e90f6c76d4923a1ae2dad8/AI-Based-Anomaly-Detection-in-Supply-Chain-Processes.pdf
2. Alom, N. B., Zakaria, A. F., Rahmat, S. B., Abdullah, S. N. B., Yusof, Z. B., Jamaludin, H. B., & Razak, A. B. (2024). An examination of customer behavior analytics, privacy issues, and data protection laws in the age of big data and machine learning. Retrieved from https://www.researchgate.net/profile/Zainuddin-Bin-Yusof/publication/386507641
3. Antoniadis, J., Arzoumanian, Z., Babak, S., Bailes, M., Bak Nielsen, A. S., Baker, P., ... & Brazier, A. (2022). The International Pulsar Timing Array second data release: Search for an isotropic gravitational wave background. Monthly Notices of the Royal Astronomical Society, 510(4), 4873–4887. https://academic.oup.com/mnras/article-abstract/510/4/4873/6503453
4. Asamoah, J. Y. S. (2025). Evaluating key determinants of customer satisfaction: A focus on customer relations, service quality, product quality, and supply chain management. Retrieved from https://www.theseus.fi/handle/10024/883142
5. Danuser, D. J. (2024). Evoking & embodying science: Representing science and social differences in early 21st century television commercials (Doctoral dissertation, University of Pittsburgh). https://d-scholarship.pitt.edu/46826/
6. Fioretto, S., & Masciari, E. (2025). A comparative analysis of predictive process monitoring: Object-centric versus classical event logs. Knowledge and Information Systems, 1–44. https://link.springer.com/article/10.1007/s10115-025-02461-y
7. Ghazal, T. M., & Alzoubi, H. M. (2022). Fusion-based supply chain collaboration using machine learning techniques. Intelligent Automation & Soft Computing, 31(3), 1671–1687. Retrieved from https://scholar.google.com
8. Gopal, P. R. C., Rana, N. P., Krishna, T. V., & Ramkumar, M. (2024). Impact of big data analytics on supply chain performance: An analysis of influencing factors. Annals of Operations Research,333(2), 769–797. https://link.springer.com/article/10.1007/s10479-022-04749-6
9. Jean, G. (2024). Inventory management strategies: Balancing cost, efficiency, and customer satisfaction. Retrieved from https://www.researchgate.net/profile/Guillaume-Jean-6
10. Khedr, A. M. (2024). Enhancing supply chain management with deep learning and machine learning techniques: A review. Journal of Open Innovation: Technology, Market, and Complexity, 100379. https://www.sciencedirect.com/science/article/pii/S2199853124001732

11. Lin, H., Lin, J., & Wang, F. (2022). An innovative machine learning model for supply chain management. Journal of Innovation & Knowledge, 7(4), 100276. https://www.sciencedirect.com/science/article/pii/S2444569X22001111
12. Liu, T., Guan, X., Wang, Z., Qin, T., Sun, R., & Wang, Y. (2024). Optimizing green supply chain circular economy in smart cities with integrated machine learning technology. Heliyon,10(9). https://www.cell.com/heliyon/fulltext/S2405-8440(24)05856-0
13. Nweje, U., & Taiwo, M. (2025). Leveraging artificial intelligence for predictive supply chain management. International Journal of Science and Research Archive, 14(1),230–250. https://eprint.scholarsrepository.com/id/eprint/58/
14. Pasupuleti, V., Thuraka, B., Kodete, C. S., & Malisetty, S. (2024). Enhancing supply chain agility and sustainability through machine learning: Optimization techniques for logistics and inventory management. Logistics, 8(3), 73. https://www.mdpi.com/2305-6290/8/3/73
15. Raparthi, M. (2021). Blockchain-based supply chain management using machine learning: Analyzing decentralized traceability and transparency solutions. Blockchain Technology and Distributed Systems, 1(2), 1–9. Retrieved from https://www.researchgate.net
16. Schroeder, M., & Lodemann, S. (2021). A systematic investigation of the integration of machine learning into supply chain risk management. Logistics, 5(3), 62. https://www.mdpi.com/2305-6290/5/3/62
17. Shurrab, H. (2022). Balancing demand and supply in complex manufacturing operations: Tactical-level planning processes.Chalmers Tekniska Hogskola. https://core.ac.uk/download/pdf/521281873.pdf
18. Tan, W. C., & Sidhu, M. S. (2022). Review of RFID and IoT integration in supply chain management.Operations Research Perspectives,9,100229. https://www.sciencedirect.com/science/article/pii/S2214716022000070
19. Thejasree, P., Manikandan, N., Vimal, K. E. K., Sivakumar, K., & Krishnamachary, P. C. (2023). Applications of machine learning in supply chain management: A review. In Industry 4.0 Technologies: Sustainable Manufacturing Supply Chains(pp. 73–82). https://link.springer.com/chapter/10.1007/978-981-99-4819-2_6
20. Tirkolaee, E. B., Sadeghi, S., Mooseloo, F. M., Vandchali, H. R., & Aeini, S. (2021). Application of machine learning in supply chain management: A comprehensive overview. Mathematical Problems in Engineering, 2021, 1476043. https://onlinelibrary.wiley.com/doi/full/10.1155/2021/1476043
21. Ullah, A., Mohmand, M. I., Hussain, H., Johar, S., Khan, I., Ahmad, S., Mahmoud, H. A., & Huda, S. (2023). Customer analysis using machine learning-based classification algorithms for effective segmentation. Sensors, 23(6), 3180. https://www.mdpi.com/1424-8220/23/6/3180
22. Wilson, A., & Anwar, M. R. (2024). The future of adaptive machine learning algorithms in high-dimensional data processing. International Transactions on Artificial Intelligence, 3(1), 97–107. https://journal.pandawan.id/italic/article/view/656