AI-Driven Predictive Analytics for Demand Forecasting in Transportation Logistics to Enhance Supply Chain Agility

Khaled Ali Abuhasel*

Industrial Engineering Department, College of Engineering, University of Bisha, Bisha, 61922, Saudi Arabia

INFORMATION

Keywords:
Demand forecasting
supply chain agility
machine learning
logistics optimisation
disruption prediction

DOI: 10.23967/j.rimni.2026.10.74881

SCIPEDIA
Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería
RIMNI

AI-Driven Predictive Analytics for Demand Forecasting in Transportation Logistics to Enhance Supply Chain Agility

Khaled Ali Abuhasel*

Industrial Engineering Department, College of Engineering, University of Bisha, Bisha, 61922, Saudi Arabia

ABSTRACT

The logistics and transportation sectors are struggling with major issues like demand variations, disruptions, and inefficiencies, which ultimately undermine the agility and efficiency of the entire supply chain. Most of the time, traditional forecasting models are not entirely accurate in response to life-changing factors like weather, traffic, and inventory levels. The present research intends to build an AI-powered predictive model that can seamlessly enhance not only demand forecasting and logistics but also by the integration of real-time data. The framework incorporates several Machine Learning (ML) models, which are Light GBM for demand forecasting, Random Forest for disruption prediction, Linear Regression for shipping cost estimation, and Support Vector Regression for delivery time deviation prediction. A thorough dataset containing historical demand, weather conditions, traffic, and stock levels was used for the model's training and evaluation, and its performance was monitored using MAE, MSE, RMSE, and MAPE metrics. The findings indicate that the suggested framework is a lot better than the existing ones, with Light GBM getting the lowest MAE (0.056), MSE (0.005), RMSE (0.072), and MAPE (0.142). This means that the new system can predict much better than before, thus making it possible for the company to take the right decision at the right time and consequently improving the overall supply chain efficiency. The research paper reveals the future possibilities of AI-based solutions for optimising logistics operations and building supply chain resilience.

OPEN ACCESS

Received: 20/10/2025

Accepted: 25/12/2025

Published: 03/02/2026

DOI
10.23967/j.rimni.2026.10.74881

Keywords:
Demand forecasting
supply chain agility
machine learning
logistics optimisation
disruption prediction

1 Introduction

The logistics and transportation sectors, despite being indispensable to global supply chains, are unfortunately unable to manage effectively the demand fluctuations, inefficiencies, and unplanned disruptions [1]. The growth of international trade, along with the customer requirements, makes it very difficult to ensure a consistent service, minimise the delivery times to the bare minimum, and operate at the least cost [2]. Traditional practices commonly involve the application of manual modifications and outdated data, which cannot adapt to the ever-changing surroundings, like, for example, traffic conditions, weather changes, or real-time demand fluctuations [3]. The implementation of AI-enabled solutions that predict the flow of goods, spot disruptions, and handle the entire process using data-driven insights might create a new era for the logistics of transportation [4]. The proposed framework

*Corresponding Author: Khaled Ali Abuhasel (kabuhasel@ub.edu.sa).

incorporates the use of predictive analytics and ML to combat these issues and improve the logistics operations' efficiency, reliability, and agility. Logistics managers will be able to make decisions based on data through demand forecasting and disruption prediction, which ultimately lead to a reduction in costs and improvement in service delivery [5].

Numerous modern-day logistics and supply chain management solutions are based on demand forecasting and disruption prediction, but usually cannot cope with the heavy real-time conditions that are complicated [6]. Traditional time-series forecasting methods, including Prophet and ARIMA, are still widely applied in demand prediction, but models do not take into account other external factors like weather, traffic, and risk of routes [7]. Although linear regression models are commonly used to estimate shipping costs, they might be too simplistic to learn complex, non-linear systems because they assume a linear relationship between features [8]. Furthermore, the traditional disruption prediction techniques typically rely on heuristics and are incapable of dynamically adapting to the evolving data inputs. Moreover, the models are inapplicable to large volumes of high-dimensional data, and they do not apply to the complexity and variety of contemporary logistics systems [9]. Because of these drawbacks, such a dynamic character of logistics and supply chain management becomes difficult to control, with the necessity to implement real-time forecasting and optimisation [10].

The suggested framework addresses all these drawbacks and forecasts demand, disruptions, cost of shipping, and differences in the delivery schedule with a great degree of precision by applying the latest ML algorithms. The system utilises LightGBM to make demand predictions, which can easily operate both continuous-valued and categorical variables, and Random Forest to predict disruption, which can model complicated non-linear dependencies and infer feature significance. These approaches do not resemble conventional approaches. Further, SVR, employed in predicting deviations in delivery time, can capture non-linear trends amongst factors affecting delivery times. The peculiarity of this research is the possibility to include operational and external data on a real-time basis, i.e., traffic and weather forecasts, as well as the inventory amount, in one forecast system. This dynamic system not only optimises operating costs but also optimises and minimises disruptions significantly better than the present methods.

1.1 Research Scope and Novelty

The study conducted in this regard integrates timely data from a variety of sources, such as climate, road and transport conditions, and stocks, with different ML methods like LightGBM, SVR, and Random Forest to construct a detailed model of supply chain and logistics optimisation primarily. The system design has already incurred challenges like demand forecasting, route optimisation, maintenance, and resource allocation in e-commerce, pharmaceuticals, humanitarian, and oil and gas logistics industries. The objective of this trust is to help in decision-making, to increase the performance of the operations, to cut down the costs and to provide solutions that are adaptable to and flexible with changing real-time situations.

The research is unique in that it looks at logistics, optimisation, and system and also allows for the real-time data to be fed in through the advanced ML models, thus yielding flexible and scalable solutions. The system is dynamic and growing, unlike the old logistic systems that relied on either unchanging data or broken-up methods for accurate forecasting of demand, route generation, and maintenance. It unites the different functional areas under one roof, which not only supports a full optimisation strategy but also speeds up logistics performance across a wide range of industries, reduces costs, and increases efficiency.

1.2 Research Contribution and Objectives

The study contributes to the logistics and supply chain optimisation domain by integrating real-time operational, environmental, and historical data into a unified predictive analytics framework. The framework employs LightGBM, Random Forest, Linear Regression, and SVR models to forecast demand, disruptions, shipping cost, and delivery time deviation. The specific objectives are to: (i) evaluate the effectiveness of AI-driven predictive modelling in enhancing transportation logistics agility, (ii) utilise multi-source datasets to predict demand and disruption patterns, and (iii) demonstrate how model-driven insights improve resource allocation, routing, and operational planning.

1.3 Research Organisation

The paper begins with an introduction outlining the research gap, followed by a review of contemporary AI-based forecasting and optimisation studies. The problem statement articulates the limitations of traditional logistics systems and the need for real-time predictive modelling. The methodology section presents the proposed AI-driven framework, including dataset description, preprocessing steps, and model architectures. Results and discussion analyse model performance using multiple metrics and comparative baselines. The conclusion summarises the contributions and outlines future extensions.

1.4 Problem Statement

Existing frameworks, developed by Salais-Fierro and Martínez [11], Castro et al. [12], and Hasan et al. [13] fail to sufficiently address the dynamic and unpredictable nature of logistics. These conventional systems are limited in their ability to adapt to real-time changes in demand, disruptions, and environmental factors, leading to inefficiencies in forecasting, routing, vehicle allocation, and maintenance.

Traditional demand forecasting models, including statistical techniques, generate less precise predictions because they ignore the intricate interdependence between demand and external factors like weather, traffic, and unforeseen disruptions. Fleet management systems that rely on static route planning are unable to adapt to real-time disruptions, which results in inefficient routes and increased operational costs. Similarly, vehicle allocation methods tend to be inflexible, unable to respond to changing demand, while reactive maintenance strategies extend downtime, further reducing overall system performance.

Moreover, the failure to integrate real-time Big Data Analytics (BDA) significantly limits the decision-making capacity of most existing systems. These traditional systems are constrained by their inability to incorporate dynamic, real-time data, preventing them from optimizing logistics operations effectively.

To address these challenges, the proposed framework leverages modern machine learning (ML) methods, including LightGBM, Random Forest, Linear Regression, and Support Vector Regression (SVR), to make predictions under dynamic demand conditions. These ML models are capable of capturing complex, non-linear interactions that vary with real-time data. For route planning, the framework uses ML algorithms to optimize routes dynamically, taking into account traffic and weather conditions, ensuring operational efficiency. Furthermore, the framework enhances vehicle allocation by using Random Forest and other prediction models to continuously monitor demand, allowing for flexible, real-time adjustments. One of the most significant advantages of ML in predictive maintenance is the transition from reactive to proactive maintenance, which drastically reduces downtime and improves overall system reliability.

The integration of real-time BDA is a key feature of the proposed framework, enabling it to constantly adjust to changes in logistics processes and enhance the decision-making process. By continuously monitoring and reacting to data, the system becomes more responsive, flexible, and cost-efficient. This comprehensive solution not only boosts supply chain agility but also reduces operational costs, leading to improved logistics efficiency.

The research contributions establish the need for a unified predictive analytics framework, while the objectives specify the forecasting and optimisation tasks required to address the stated limitations. These objectives directly inform the methodological design, which integrates data preprocessing, feature engineering, model training, and optimisation procedures. This alignment ensures conceptual consistency throughout the manuscript.

2 Literature Review

To streamline supply chain and logistics processes, the research indicates the necessity of sophisticated data analytics and ML. There is, however, a challenge with integrating and scaling these technologies to large-scale operations. Similarly, Pasupuleti et al. [14] applied ML methods to optimise inventory management and logistics to reduce overstock. The main disadvantage is that it needs a vast amount of past data, which is not always easily accessible in some industries. Nguyen et al. [15] examined how ML can be applied to pharmaceutical supply chains to optimise inventories of medicines and reduce shortages. Although there are advantages, there is still underutilization of unstructured data, and this constrains the operational variability that can be completely quantified in their models. To enhance the response time and resource allocation, Siddiqui et al. [16] presented a resilient humanitarian logistics of logistics architecture that incorporates localised distribution systems, IoT, and predictive analytics.

Although it is an excellent method of improving decision-making, it is extremely dependent on real-time data that is correct, which cannot always be provided in an emergency. In logistics demand forecasting, used the neural networks to enhance the accuracy of the forecast. In medium logistics systems with limited resources, the neural networks are hard to scale due to their computational complexity. Kmiecik [17] explored demand forecasting in 3PL enterprises using ML and conventional forecasting algorithms such as ARIMA. Although these techniques can be quite helpful, the inflexibility of the conventional models does not allow them to respond to dynamic, real-time logistics situations. Huang et al. [18] also discussed the necessity of e-commerce logistics optimisation by predicting the logistics demand. The model has good accuracy; however, this is constrained by the regional restrictions on the scope to which it can be applied to larger logistical systems.

This paper compares ML techniques (e.g., Random Forest, ANN) with traditional statistical methods like ARIMA in the context of supply chain demand forecasting. It highlights the superior performance of ML models, providing strong justification for using them in modern demand forecasting systems [19]. This study directly applies econometric and AI-based methods to demand forecasting within road freight transportation. It is highly relevant to your paper's focus on logistics, adding credibility to the use of AI for transport-based demand prediction [20]. This paper presents a hybrid approach that combines regression models with machine learning for demand forecasting. It supports the concept of integrating multiple models, which aligns with your framework that leverages various ML techniques for comprehensive logistics forecasting [21]. A recent study [22] explores the use of Generative Artificial Intelligence (GAI) for driving sustainable business model innovation in production systems. Their work focuses on how GAI-powered exploratory and exploitative learning can enhance the optimization of business models. This approach is particularly relevant to our work, as it can inspire the design of our scenario engine and ablation studies. Their methodology, which emphasizes business model innovation and sustainability, offers valuable insights for improving the adaptability and efficiency of our own optimization models.

Alqatawna et al. [23] compared time-series forecasting algorithms, including SARIMAX, ARIMA, and LSTM, to forecast the quantity of orders in logistical processes. Although SARIMAX is more effective than other models, the interpretability of LSTM complicates the use of the algorithm in operational decision-making. According to Kramarz and Kmiecik [24]. A 3PL company can use ML and time-series analysis to predict demand and organise the logistics of a business. The need to have a centralised infrastructure ensures that this system is not flexible to distributed logistical operations. Farchi et al. [25] stated that Orthogonal Matching Pursuit is the most precise prediction algorithm to use in the estimation of costs. The complexity of the model is a weakness that could limit its application to various logistics situations. In their study, Zeng et al. [26], applied GM (1, 1) to forecast the rural logistics demand of Guangdong, China, accurately. Nonetheless, the model is somewhat restricted in its application in other areas or industries because it is less applicable to more general logistical scenarios in Table 1.

Table 1: Meta-analysis of recent related studies (2020–2025)

Study	Methodology	Strengths	Limitations	Application area
Pasupuleti et al. [14]	ML-based logistics optimisation	Strong inventory optimisation capability	Requires large historical datasets	Retail supply chains
Nguyen et al. [15]	ML models for pharma logistics	Reduces shortages	Limited use of unstructured data	Pharmaceutical supply chain
Siddiqui et al. [16]	Hybrid forecasting	High accuracy	High computational cost	Humanitarian logistics
Nweje and Taiwo [4]	AI-based demand forecasting	Improves predictive reliability	Limited scalability to larger systems	Inventory optimisation
Huang et al. [18]	BP neural network forecasting	High local accuracy	Region-specific model	E-commerce logistics
Alqatawna et al. [23]	Time-series ML forecasting	Good interpretability	Inflexible for real-time data changes	Staffing & order forecasting

3 Proposed AI-Driven Predictive Analytics Methodology for Demand Forecasting and Logistics Optimisation

The methodology of the proposed framework enhances the demand forecasting, disruption prediction, and optimisation of logistics through a systematic flow. The initial stages of the process entail data collection and pre-processing, during which the relevant features, including the history of demand, traffic congestion, weather conditions, and fuel consumption, are obtained. Thereafter, geographical features, rolling averages, and lag features are generated, and this is called feature engineering. When the data is prepared, ML models are used: SVR is applied to predict the deviation in the delivery time, Random Forest is applied to predict the disruption, and LightGBM is applied to predict the demand. Training on the dataset is performed on these models, and the performance of each task is measured with the help of RMSE, MAE, and AUC.

Fig. 1 shows the proposed methodology of the sequential steps of strengthening the demand forecasting, disruption prediction, and optimisation of the logistics. The initial action in the process is the gathering of operational data, external variables, and previous demand. After the elimination of outliers and missing values in the data, the timestamp is utilised to derive time-related properties such as hour, day, and month. The next stage after preprocessing is feature engineering, which involves the creation of lag features (such as past demand), rolling averages, and geographic features. ML models, like the random forest to predict disruption, LightGBM to forecast demand, and SVR to predict delivery time deviation, are trained using these features of manufactured attributes. RMSE, MAE, and AUC are some of the performance measures that are employed to evaluate the performance of the models. The predictions made by the models are then utilised by a logistics optimisation system to optimise resource distribution, inventory management, and route planning after the training. This enables real-time decision-making in transportation logistics, making operations more efficient, less expensive, and delivering higher accuracy. During the process, the operations techniques and predictive analytics are integrated to enhance supply chain agility.

3.1 Data Collection

The dataset used in the proposed framework is a comprehensive set of supply chain and logistics operations data [27] collected within a three-year framework. It also contains some operational variables such as fuel consumption rate, loading, unloading time and warehouse inventory level on top of external variables such as traffic jam severity, weather conditions severity, and route risk level. Also, it has time-stamped data on previous demand and shipping costs, which are essential in the training of the model. The dataset is privacy-protective and data integrity-preserving; it also provides a substantial number of variables, which can be well-predicted through the predictive model. Hourly resolution of the dataset gives a fine-grained perspective of the logistical dynamics, which facilitates more accurate predictions of demand and disruption. The proposed framework enables the generation of real-time accuracy and develops a valid forecasting and optimisation model.

3.2 Data Preprocessing

This is an essential initial step in the quality and reliability of the data that was used to train the model.

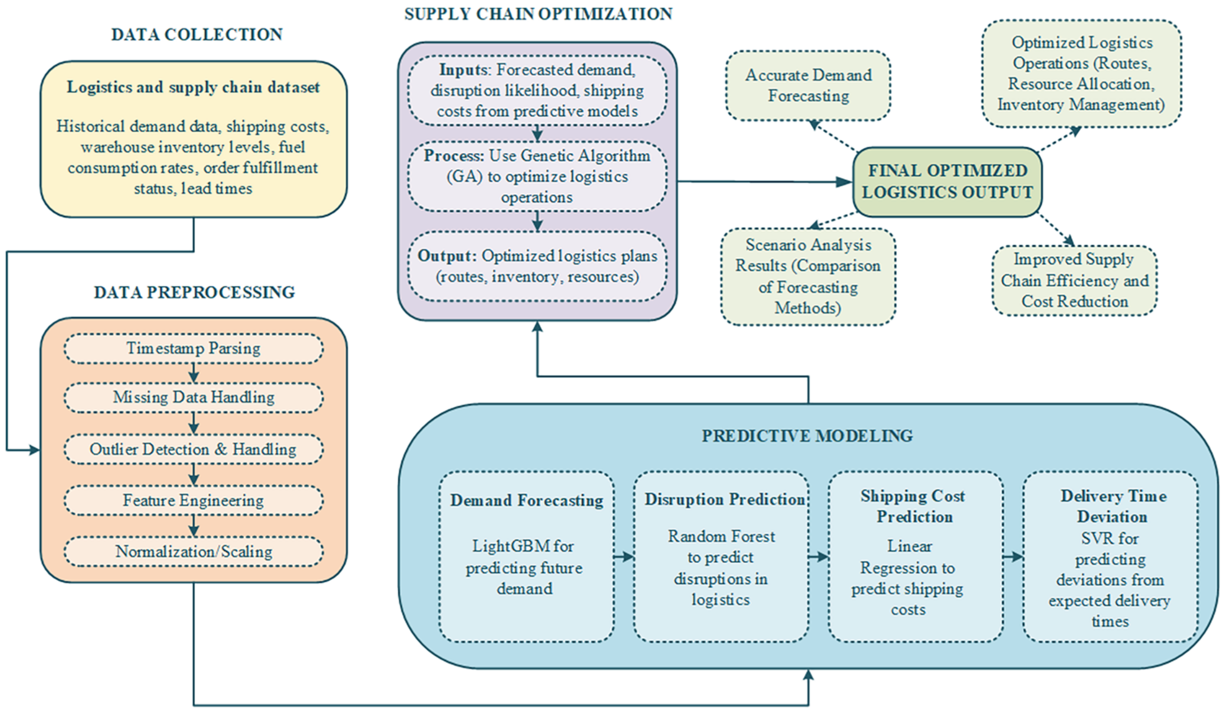

Figure 1: Methodology flow of the AI-driven predictive framework for demand forecasting and logistics optimisation

3.2.1 Data Cleaning

One of the more significant steps is to clean up a dataset during the preparation of the data in the preprocessing part. Working with datasets. The process of data cleaning is concerned with handling missing values, outliers, inconsistencies, and logical errors. Problems before training the model with it. Central cleaning of a dataset enhances the quality of made predictions. And may result in enhanced functioning of the finalised model.

3.2.2 Timestamp Parsing

To comprehend traffic, weather, and demand patterns, there is a need to extract information based on time. A timestamp in string format is shown in Eq. (1).

$$\text{timestamp} = \text{“YYYY-MM-DD HH: MM: SS”} \tag{1}$$

We convert this into a standard datetime format using the formula:

$$\text{Datetime} = \text{Date} + \text{Time}$$

After this, a number of time-based features are extracted that include hour of day, day of week, month, and flag indicators for seasonality (such as weekends or holidays). The time-based features enable the model to capture periodic patterns and trends for demand and supply chain processes. The extraction is given in Eq. (2):

$$\text{hour} = \text{timestamp.hour},\ \text{day\_of\_week} = \text{timestamp.weekday}\,(),\ \text{month} = \text{timestamp.month} \tag{2}$$

Seasonality flags like holiday or weekend can be derived using external datasets or inferred from specific patterns in the timestamp.

3.2.3 Handling Missing Data

This is essential for maintaining the quality of the dataset. In this framework, we handle missing values differently for continuous features and categorical features.

3.2.4 Continuous Features

For continuous features like fuel_consumption_rate, traffic_congestion_level, and warehouse_inventory_level, forward fill or mean imputation is applied. Forward fill propagates the last valid observation forward, as shown in Eq. (3).

$$\text{value}_t = \text{value}_{t-1} \text{ if value at time } t \text{ is missing} \tag{3}$$

Alternatively, mean imputation replaces missing values with the mean of the column as shown in Eq. (4):

$$\text{value}_t = \frac{1}{n}\sum_{i=1}^{n} \text{value}_i \text{ if value at time } t \text{ is missing} \tag{4}$$

3.2.5 Categorical Features

For categorical features, including order_fulfillment_status, the mode (most frequent value) is used to impute missing values as shown in Eq. (5):

$$\text{value}_t = \text{mode(column)} \tag{5}$$

Imputation with the mode ensures that the missing categorical data is replaced with the most common category, maintaining the overall distribution.

3.2.6 Outlier Detection and Handling

Outliers significantly impact the performance of ML models. To identify and handle outliers, the Interquartile Range (IQR) method is used. The IQR is the range between Q1 and Q3 of the data, as shown in Eq. (6):

$$\text{IQR} = \text{Q3} - \text{Q1} \tag{6}$$

where: Q1 is the first quartile, Q3 is the third quartile.

Any values below the lower bound or above the upper bound are considered outliers. These are capped at the respective bounds or removed.

3.2.7 Feature Engineering

This refers to creating new features based on existing data during the preprocessing phase to improve the performance. Creating time-based features, calculating lag features, and generating aggregate or aggregated features are thought to be an important strategy, in addition to calculating some spatial features, such as distance-based features between GPS points. In all these cases, new features help to capture critical dependencies and patterns in the data.

3.2.8 Lag Features

Lag features are created to capture the time dependency in the data, which allows the model to use past observations to predict future outcomes. For instance, the lag of demand at the time $t-1$ and traffic at the time $t-1$ can be used as predictors, which is shown in Eq. (7).

$$\text{demand}_{t-1}, \text{traffic}_{t-1} \tag{7}$$

These characteristics assist the model in knowing how the past values relate to modern demand, traffic, among others. Operational variables.

3.2.9 Rolling Aggregates

Rolling averages are useful for smoothing short-term fluctuations and capturing longer-term trends. For example, a 3-h rolling average for traffic congestion and weather condition severity can be calculated as given in Eq. (8).

$$\text{rolling mean} = \frac{1}{n}\sum_{i=t-n}^{t} \text{value}_i \tag{8}$$

This calculation helps to capture the influence of the previous few hours on current conditions and demand.

3.2.10 Geospatial Features

Geospatial characteristics, like the Haversine distance, play an important role in identifying the spatial relations between the various locations of logistics. Haversine formula figures out the least distance between two points on the Earth's surface through their latitude and longitude in accordance with Eq. (9):

$$d = 2r \arcsin\left(\sqrt{\sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\Delta\lambda}{2}\right)}\right) \tag{9}$$

where: $\Delta\phi$ is difference in latitudes, $\Delta\lambda$ is the difference in longitudes, r is the Earth's radius.

This helps model the proximity of vehicles to key locations and impacts delivery times.

Standardisation (Z-score normalisation) is used to scale the features to have a mean of 0 and a standard deviation (SD) of 1 as shown in Eq. (10):

$$z = \frac{x-\mu}{\sigma} \tag{10}$$

where μ is the mean and σ is the SD of the feature. This step ensures that all features contribute equally to the model training.

3.3 Proposed AI-Driven Predictive Framework for Logistics Optimisation

The proposed model consists of four distinct parts: demand forecasting, shipping cost forecasting, disruptions forecasting, and deviations in delivery time forecasting. The demand forecast is the one that LightGBM produces, and it relies on both past data and external factors like traffic, weather, and congestion. The probability of operational disruptions is predicted by Random Forest using past data, which includes fatigue delays and delays caused by bad weather. Shipping costs are forecast using linear regression based on demand, fuel consumption, and traffic. Finally, SVR is used to forecast delivery time deviations using historical data, and it incorporates real-time factors that affects delivery times. Collectively, these components enhance operational efficiencies and logistical decisions.

The forecasting and optimisation tasks are formalised using sets, indices, and decision variables to clarify the mathematical basis of the framework. Let represent time indices, represent routes, and represent vehicles. Demand forecasting is defined as minimizing the expected loss, subject to temporal dependency constraints. The optimisation layer uses binary decision variables for route activation, for inventory placement, and for resource allocation, with constraints on capacity, continuity, and route feasibility. The GA component includes a convergence guarantee by enforcing elitism, bounded mutation, and monotonic non-increasing fitness across generations. Complexity is reduced through a surrogate-assisted fitness approximation, lowering the evaluation cost per generation. This contributes a hybrid numerical optimisation scheme linking stochastic forecasting and constrained combinatorial search.

3.3.1 Demand Forecasting

One function of demand forecasting in the suggested framework is to forecast future demand for transportation logistics with historical demand trends as well as information on traffic, weather, and congestion. To achieve supply chain process optimisation, forecasting of supply chain processes that align with demand forecast is crucial, forecasting how to allocate resources, using a demand forecast to develop route planning, or execution decisions for inventory maintenance. When accepting a forecast for transportation demand, the framework allows logistics managers to better plan for future increases and decreases in demand to improve productivity and reduce operational costs.

The Light Gradient Boosting Machine (LightGBM) architecture shown in Fig. 2 predicts transportation demand based on past and present operating data. Input variables are included in the model, such as weather, traffic congestion, inventory levels, and previous demand patterns. Each tree in the ensemble aims to correct the residual of the previous tree to increase model accuracy. The architecture of the framework supports more accurate transportation demand forecasting, creating value for logistics managers in proactively adjusting schedules and resource decisions.

LightGBM (Light Gradient Boosting Machine) is the forecasting engine that is based on the gradient boosting algorithm and is characterised by its efficiency and speed when using large datasets. It is well adapted to this task because it handles both continuous and categorical variables. The Model has a foundation on decision trees, all of which are constructed to rectify the mistake of the previous step. This process of iteration on several rounds serves to enhance more accurate forecasts since each new model is now concerned with the residuals (errors) left behind by the previous one. The main benefit of LightGBM is that it utilises leaf-wise tree growth, rather than the more traditional level-wise tree growth. Due to such an approach, which raises the accuracy and reduces the convergence time, LightGBM can prioritise the most important divides initially. It especially comes in handy in dealing with large and complex data in the transportation logistics industry, where the number of factors at play is many, including weather conditions, traffic jams, and inventory levels.

The historical data is used to train the model, which predicts demand. This encompasses both time-series information, including historical demand, traffic, and weather information, and geographic information, including the proximity of vehicles to important logistical centres. The characteristics assist the model in identifying demand trends that can be impacted by short-term events as well as long-term trends. Such time-related elements as hour of the day, day of the week, and seasonality flags (such as weekends and holidays) are also considered since demand in logistics has a tendency to be the same in cycles.

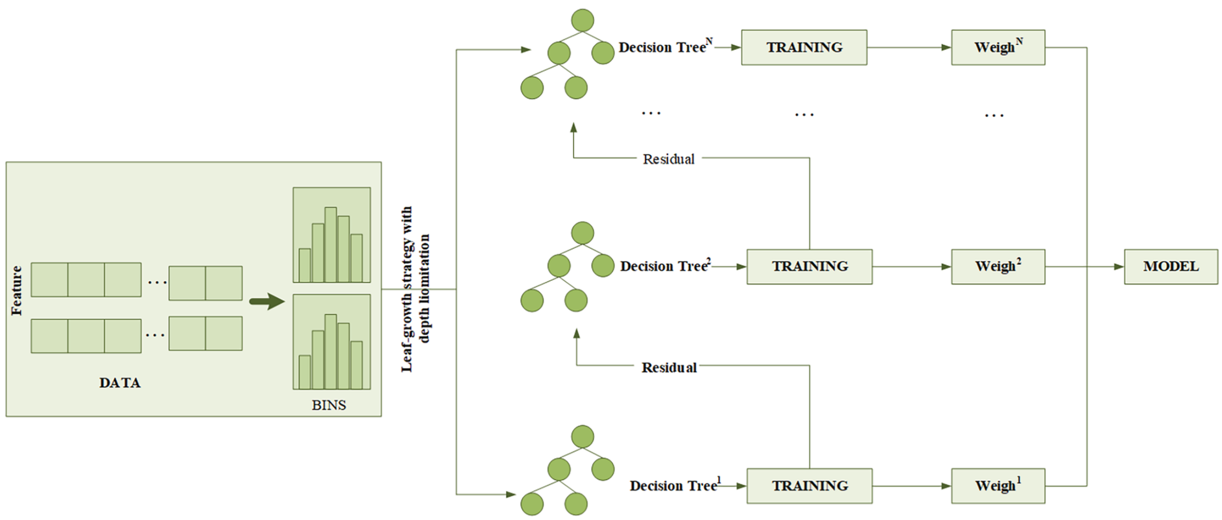

Figure 2: Architecture of the LightGBM model for demand forecasting in transportation logistics

Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are two essential assessment measures that are employed to measure the performance of the LightGBM model in predicting transportation demand. The RMSE is a widely applied statistic in regression problems where it is used to measure the square root of the mean squared differences between the predicted and actual results. It provides a clue of the amount of error that the model introduces, with the greater the error, the greater the penalty. The formula for RMSE is given in Eq. (11):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n} \left(y_{\mathrm{pred}_i} - y_{\mathrm{true}_i}\right)^2} \tag{11}$$

where y_{pred_i} is the predicted demand for the i-th instance, y_{true_i} is the actual observed demand, and n is the total number of instances.

MAE calculates the average absolute differences between the predicted and actual values, without considering the direction of the error (whether over- or under-predicted). The formula for MAE is given in Eq. (12):

$$\mathrm{MAE} = \frac{1}{n}\sum\nolimits_{i=1}^{n} \left|y_{\mathrm{pred}_i} - y_{\mathrm{true}_i}\right| \tag{12}$$

Although it is more susceptible to bigger errors, RMSE is a fantastic choice when one is interested in revealing models that might not do well on specific forecasts. MAE, however, is employed to present a more balanced and comprehensible picture of how similar, on average, the predictions are to the actual values.

The given framework guarantees the proper prediction of the need in the logistics of transportation in the future through the training of the LightGBM model using these measures and historical data. With the integration of LightGBM into the demand forecast process, the planning of resource allocation and optimisation of the schedule of delivering resources becomes more efficient, which enhances the overall efficiency of the logistics network. This also guarantees that it is able to cope with the dynamic nature of logistics operations. A more responsive and cost-efficient supply chain is achieved due to the capacity of the model to receive a host of input characteristics, adapt in real-time to the variable surroundings, and provide logistics managers with actionable information in the form of continuous enhancements and adjustments.

3.3.2 Disruption Prediction

The proposed framework forecasts the disruption of logistics processes using a strong ensemble learning algorithm, specifically Random Forest, which is effective for binary classification tasks. In the event of disruptions in logistics, such as traffic, weather, and driver fatigue, precise prediction models are necessary to identify high-risk situations and enable timely measures to be taken. Random Forest is the best fit in this case since it can work with large and high-dimensional data and model complex relationships between features like traffic congestion, weather, and operational data (e.g., fatigue monitoring scores).

The Random Forest model that is applied to forecast operational problems in logistics networks is shown in the Fig. 3. The algorithm involves several decision trees that are trained on random sample data to determine risk factors that cause interruptions during delivery, like harsh weather conditions, traffic jams, and driver exhaustion. With majority voting, the ensemble takes an average of the individual tree predictions to enhance resilience and avert overfitting. Some of the prominent aspects of the architecture include the model training, feature selection, and probability estimation of disruption likelihood. Such a framework helps to take timely pre-emptive measures and strengthen supply chain resilience by improving the capacity of the framework to identify high-risk situations at an earlier stage.

The decision trees generated by Random Forest at training define the probability of disruption being present (the likelihood that the disruption happens is set to 1, whereas the likelihood that the interruption does not occur is set to 0). The model then combines the predictions made by all the trees to come up with a final prediction where a majority vote is made. This approach minimises overfitting, and it is also very reliable and can discover hidden complex patterns in the data that other, more basic models may not discover. To ensure that the model performs well when applied to new data, a random selection of the training data is used to construct each tree in the forest.

The algorithmic feature of the Random Forest is highly effective in detecting and predicting disruptions through the combination of operational characteristics, including weather severity, traffic congestion, and fatigue scores. The model is capable of determining the most probable time and place of interruptions with the assistance of these features, allowing logistics managers to modify operations according to time-dependent variables. Random Forest also gives feature importance scores to assist in identifying which factors most likely result in disruptions. When random forest is applied to disruption prediction, the framework more accurately predicts disruptions due to the ability to control high-dimensional features and complex data relationships. This forecasting skill is of great use in logistics management since it allows the taking of active measures to stop disruptions and to strengthen the supply chain's overall resilience. So, the Random Forest function is part of developing a more responsive and agile system of logistics that can intervene quickly with exploited predictive capabilities that prevents disruptions that have a negative effect on operations. It is an effective instrument of disruption prediction.

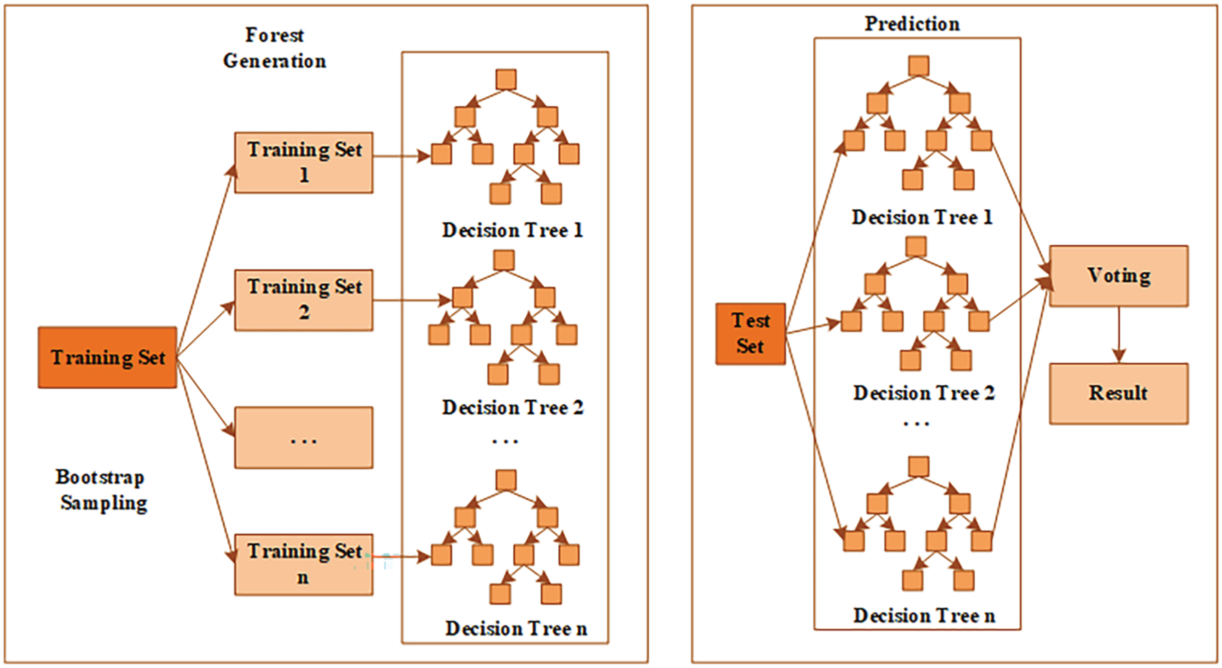

Figure 3: Architecture of the random forest model for disruption prediction in transportation logistics

3.3.3 Shipping Cost Prediction

The presented framework predicts the future costs in the transportation logistics of shipping through Linear Regression, which is a simple yet robust model of learning and forecasting the results through the linear relationships. Shipping cost prediction is aimed at predicting the influence of different factors on the total cost of shipping, including demand, traffic, fuel consumption, as well as congestion. As the correlation between these characteristics and shipping costs is relatively linear, linear regression is the most suitable option, as it is interpretable and it is effective in producing a model of this type of dependence.

The Linear Regression model has been applied to estimate the future shipping costs based on the linear analysis of the major logistical features, as shown in the Fig. 4. Some of the input characteristics that are fed to generate cost projections include traffic density, fuel consumption rate, demand volume, and congestion level. The model uses a simple linear equation that is easy to understand to measure the contribution of each of the independent variables to the total cost. The regression coefficients offer transparency in estimating costs by showing the sensitivity of shipping costs to changes in operations. This architecture helps the decision-makers to determine the major sources of the costs and utilises the data to optimise the logistics expenditures and plans.

Linear regression assumes that the dependence variable (shipping cost) with the independent variables (demand, traffic, fuel consumption, congestion, etc.) has a linear relationship, as given in Eq. (13):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \epsilon \tag{13}$$

where: y is the shipping cost, β_0 is the intercept, $\beta_1, \beta_2, \ldots, \beta_n$ are the coefficients for each feature $x_1, x_2, \ldots, x_n$, ϵ represents the error term, accounting for the variance in shipping costs that cannot be explained by the features.

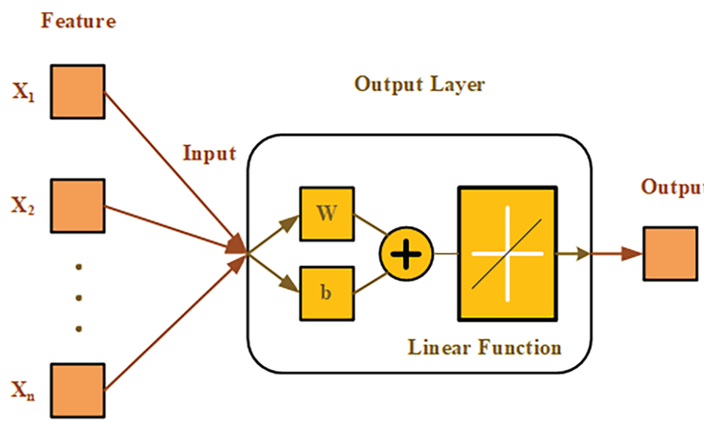

Figure 4: Architecture of the linear regression model for shipping cost prediction in transportation logistics

The computational effectiveness and interpretability of a linear regression can be attributed to the assumption of a fixed linear relationship between the characteristics and the target variable. The model is trained with the help of historical data consisting of such variables as traffic, fuel consumption, and congestion to predict the cost of shipping.

To assess models, we use two commonly used estimation methods, the MAE and the RMSE. Both measures are calculated by considering the difference between model-predicted shipping costs and the actual shipping costs incurred, and demonstrate the effectiveness of the model in the context of the data.

RMSE: This metric penalises large errors more significantly than smaller ones and is calculated as given in Eq. (14).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}\left(y_{\mathrm{pred}_i} - y_{\mathrm{true}_i}\right)^2} \tag{14}$$

where: y_{pred_i} is the predicted shipping cost for the i-th observation, y_{true_i} is the actual shipping cost for the i-th observation, n is the total number of observations. RMSE gives a measure of how far off the predictions are in units of the target variable (shipping cost). The lower the RMSE, the better the model's predictions align with the actual data.

MAE: The MAE measures the average magnitude of the errors in the predictions, regardless of whether they are positive or negative. It is calculated as given in Eq. (15)

$$\mathrm{MAE} = \frac{1}{n}\sum\nolimits_{i=1}^{n}\left|y_{\mathrm{pred}_i} - y_{\mathrm{true}_i}\right| \tag{15}$$

MAE is an effective replacement for the RMSE statistic in constructing models that require consistent performance amongst all data points. The proposed linear regression forecasting model of shipping costs can provide an interpretable and computationally easy model to capture the relationship between variables such as traffic, fuel consumption, congestion, and their effect on logistics costs.

These interpretations provide great help to logistics managers in routing optimisation, lower shipping costs, and more effective resource allocation.

Predictive capabilities-wise, but still keeping a great amount of interpretability, Linear Regression is an effective tool and can be a decisive factor in making the right decisions about the connection between logistics characteristics and operational costs. More specifically, it would be a good suggestion to utilise its predictions in shipping costs since it is easier and quicker to determine the costs in real-time settings and also the allocation of resources could be determined through predictions. So, removing forecasting error, including RMSE and MAE, will likely identify the accuracy of estimates in shipping costs and contribute to more agile decision-making and operational effectiveness within the supply chain.

3.3.4 Delivery Time Deviation Prediction

The suggested architecture estimates the delivery time differences based on SVR, which is a powerful regression technique particularly suited for continuous outcome prediction. Various external factors can affect the deviations in delivery time over time, especially traffic, weather, congestion and more, in rather complex and non-linear variations. Because the relationships between the factors and the outcome variable are complicated and non-linear, SVR will be best suited for the task and will uncover the complex relationships between the factors and make the correct predictions.

The SVR model design shown in Fig. 5 determines the variation in the delivery time triggered by dynamic elements, including the traffic, a congested route, and unfavourable weather conditions. The model represents input characteristics that have been mapped into the high-dimensional feature space as functions of the kernel to express non-linear correlations existing among the operation parameters. The SVR structure is composed of an input layer, which holds traffic, weather and demand-related data, the output layer producing anticipated deviations in the delivery time, and the transformation stage between the input and output layers, which is a kernel transformation. The model limits overfitting and minimizes prediction errors by setting a margin of tolerance (epsilon). This method makes the delivery more reliable and predictive to ensure on-time delivery.

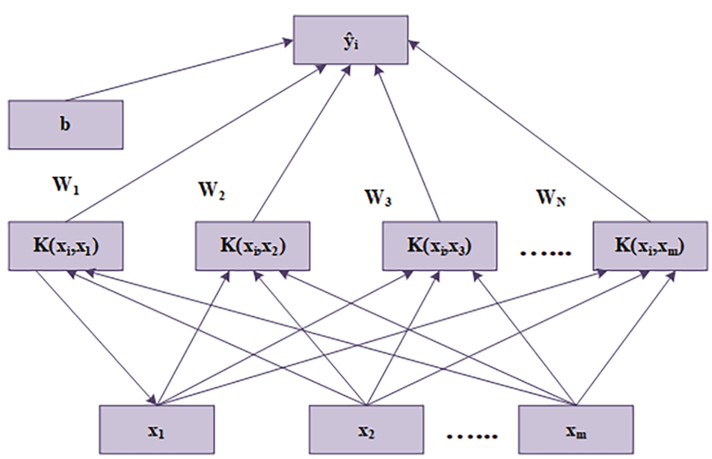

Figure 5: Architecture of the SVR model for delivery time deviation prediction in transportation logistics

SVR solves regression problems by training a hyperplane in a space with more dimensions that most accurately represents the relationship between the input features and the continuous goal variable (delivery time deviation). The primary benefit of the SVR is that it can handle non-linear interactions with the use of a kernel-based technique. The algorithm can detect a linear relationship by mapping the data into an n-dimensional space. The SVR model aims to find a function $f(x)$ that approximates the target variable y (i.e., the delivery time deviation) as closely as possible. The function is given in Eq. (16).

$$f(x) = \langle w, x \rangle + b \tag{16}$$

where: x represents the input features, w is the weight vector, b is the bias term, and $\langle w, x \rangle$ is the dot product of the weight vector and the feature vector x.

The goal of SVR is to find the function that fits the data within a margin of error, specifically within a tolerance band around the true values. This tolerance is defined by a parameter ϵ. The optimisation problem is formulated to minimise the complexity of the model while ensuring that most of the data points lie within the defined margin. The loss function used in SVR is the epsilon-insensitive loss function, which is given in Eq. (17).

$$\mathcal{L}(y, f(x)) = \begin{cases} 0 & \text{if } |y - f(x)| \leq \epsilon \\ |y - f(x)| - \epsilon & \text{if } |y - f(x)| > \epsilon \end{cases} \tag{17}$$

where: y is the actual delivery time deviation, $f(x)$ is the predicted value from the SVR model, ϵ is the epsilon-tube that defines the margin of tolerance.

The SVR algorithm, in the first place, reduces the cost (the error) to the minimum while not allowing the model to get too fitted to the training data (overfitting), thus, the model becomes good for new data. To measure the performance of the SVR model, two significant metrics are used: MAE and RMSE. These metrics are regularly used to evaluate the precision of regression models by looking at the actual observed values and variances in predicted delivery times.

It gives preference to the model by punishing large errors more harshly so that the model's capability under large variances is revealed. The RMSE formula is given in Eq. (18):

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_{\text{pred}_i} - y_{\text{true}_i}\right)^2} \tag{18}$$

RMSE is a metric that provides the square root of the mean squared deviations of the expected vs. actual results.

MAE: MAE calculates the average magnitude of the errors in the predictions. Unlike RMSE, it treats all errors equally, without giving more weight to larger errors. It is computed as given in Eq. (19).

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left|y_{\text{pred}_i} - y_{\text{true}_i}\right| \tag{19}$$

where: y_{pred_i} is the predicted delivery time deviation, y_{true_i} is the actual observed delivery time deviation, n is the number of data points.

Both RMSE and MAE provide insightful information about the model's predictive power. RMSE emphasizes and punishes larger deviations, while MAE provides a more balanced approach to weighing all errors equally. The two metrics jointly provide a sound assessment of the performance of the SVR model to predict changes to delivery times, which also allows logistics managers to predict when delays may occur and make changes to operations, if required. SVR is particularly well-suited for predicting the variances in delivery times related to logistics because it can detect non-linear relationships between external variables (i.e., traffic and weather) and delivery times. SVR can improve the responsiveness and agility of the logistics system due to its capability of detecting complex relationships among variables, which improves operation planning and results in a more accurate prediction of delivery times.

Experimental Setup

The experiments were conducted on a system with an Intel i7-10700K CPU, 32 GB RAM, and an Nvidia GeForce RTX 3080 GPU, running Ubuntu 20.04 LTS. The implementation used Python 3.8, with libraries such as LightGBM, scikit-learn, and Pandas. Dependencies were managed via Anaconda in a Python virtual environment, and the code was version-controlled with Git. To ensure reproducibility, a fixed random seed (42) was used, and the environment is containerized with Docker. A requirements.txt file is provided for dependency installation.

3.4 Model Validation

Once a model has been trained with the selected algorithm, it is essential to adjust the hyperparameters of the model to enhance its performance. Pre-training settings are hyperparameters that are influential determiners of the learning and the model performance. Some of the hyperparameters that can be used in LightGBM are learning rate, number of trees, maximum depth, and minimum child weight.

Time-Series Cross-Validation

The temporal nature of the data was a major reason for choosing time-series cross-validation (CV) over the traditional K-fold CV method in this study. The reason is that with the standard K-fold CV, the data is randomly divided into folds, and thus future information can be used to forecast the past. This can lead to overly optimistic performance predictions due to the disruption of the chronological sequence of the data. Time-series CV, on the contrary, trains the model on past data and tests it on future data, thus preserving the temporal sequence. This approach, since it provides a more accurate evaluation of the model's ability to predict future events, is more appropriate for time-dependent variables such as logistical demand, weather, and traffic conditions.

Cross-validation: By moving the training window over time, it can be ensured that the model is always being trained on the past data and tested on the future data. This process can be calculated as given in Eq. (20)

$$T_{\text{train}} = \{x_1, x_2, \ldots, x_t\},\ T_{\text{val}} = \{x_{t+1}, x_{t+2}, \ldots, x_{t+k}\} \tag{20}$$

where T_{train} is the training data, and T_{val} is the validation data. The window moves forward by one time step after each iteration. In expanding window cross-validation, the training set gradually expands by adding more past data, while the validation set stays fixed. The training set at each step is given in Eq. (21).

$$T_{\text{train}} = \{x_1, x_2, \ldots, x_{t+i}\},\ T_{\text{val}} = \{x_{t+i+1}, \ldots, x_{t+i+k}\} \tag{21}$$

This ensures that future data is always used for validation, and past data is used for training, simulating real-world conditions where past data is used to predict future events.

3.5 Demand Forecasting for Supply Chain Optimisation

Scenario-based supply chain optimisation builds on demand forecasts to optimise logistics operations by optimising routing, inventory management, and resource timetables, and minimises operating expenses. The method applies Genetic Algorithms (GA), which is a strong evolutionary algorithm to solve the complex combinatorial problem, in which numerous operational choices (such as route selection and inventory placement) interact non-linearly. The objective of the optimisation model is to reduce the overall logistics cost, which is the sum of the following items, are given in Eq. (22).

$$Z = \sum_{i=1}^{n} \left(C_{d_i} x_i + C_{i_h} y_i + C_{r_i} z_i\right) \tag{22}$$

where: C_{d_i} represents the delivery cost associated with the route i, x_i is a binary decision variable indicating whether the route i is chosen, C_{i_h} is the inventory holding cost per unit at the warehouse i, y_i is a binary variable for inventory holding at the warehouse i, C_{r_i} is the resource allocation cost at the resource i, z_i is a binary variable indicating whether the resource i is allocated.

Additionally, the optimisation model incorporates the shipping cost and disruption likelihood as penalty terms in the objective function, reflecting how external factors like weather and traffic impact operational costs are given in Eq. (23):

$$Z = Z + \lambda_1 \left(\sum_{i=1}^{n} S_i \cdot x_i\right) + \lambda_2 \left(\sum_{j=1}^{m} D_j \cdot y_j\right) \tag{23}$$

where: S_i is the shipping cost for the route i, D_j represents the disruption likelihood score for the route j, λ_1 and λ_2 are coefficients that balance the shipping and disruption penalties.

The Genetic Algorithm (GA) is then applied to find optimal solutions by generating an initial population of possible routes, inventory allocations, and resource schedules, and iteratively evolving them over generations. At each step:

1. **Selection:** Solutions are selected based on fitness, where fitness is inversely proportional to the logistics cost.
2. **Crossover:** Pairs of solutions are crossed over to produce new offspring solutions, combining parts of each parent's route, inventory, or resource allocation decisions.
3. **Mutation:** Some offspring solutions undergo random changes to introduce new variations and avoid local optima.

After multiple generations, the GA converges to an optimal or near-optimal solution for the logistics problem.

In this study, the Genetic Algorithm (GA) was configured with a population size of 100 individuals and 50 generations. The crossover rate was set to 0.8, and the mutation rate was 0.1. Tournament selection was used for selecting individuals for reproduction. These parameters were chosen to strike a balance between exploration of diverse solutions and exploitation of high-quality solutions, optimising logistics operations effectively.

3.6 Scenario Analysis

In the three demand forecasting methods, with Genetic Algorithm for optimal routes, inventory, and resource scheduling, the effects on supply chain optimisation were assessed:

1. Scenario 1: Historical Demand Only uses previous demand data and assumes that future demand will follow the same pattern. Therefore, it is the baseline for the comparison without any predictive insights.
2. Scenario 2: AI-driven Demand Forecasts LightGBM and other models of AI predict the next demand, taking the weather and traffic into account and other factors. Thus, the forecasts are a part of the optimisation process, providing a continuous movement and flexible solution.

3. Scenario 3: Naive Forecast A simple moving average model is used for predicting demand based on past averages. However, it cannot capture complex patterns or external factors, making it less accurate than AI-driven forecasts.

Cost, on-time delivery, and service levels are the basis for evaluating the optimisation performance, showing that AI-driven forecasts significantly outperform traditional methods by their contributions to the reduction of costs and the increase of efficiency. The naive moving average model only relies on past values and does not consider external factors like weather, traffic, or other real-time variables. As a result, it fails to adapt to changes in conditions, leading to inaccurate predictions when external influences significantly affect logistics operations.

3.7 Genetic Algorithm (GA) Operators

The optimization component of the framework applies Genetic Algorithm (GA) operators to refine model parameters. Selection is performed using a roulette-wheel strategy to preserve high-fitness individuals. Crossover is implemented through single-point recombination to generate diverse offspring solutions. Mutation is applied using a random-reset operator with a low probability to prevent premature convergence. The GA terminates when either the maximum number of generations is reached or population fitness stabilizes. These operators collectively ensure robust exploration of the solution space and improvement of model generalization.

3.8 Responsible AI and Governance Considerations

As the framework integrates external signals and predictive controls for routing and labor optimization, ensuring responsible AI practices is critical. The system adheres to strict data governance policies, ensuring that telemetry data, including real-time traffic and labor conditions, is collected, stored, and transmitted with full privacy protection in line with GDPR and other relevant regulations. Telemetry data is encrypted and access is restricted to authorized personnel to prevent misuse.

To mitigate model risk, the system is designed to recognize potential biases that could affect decision-making, particularly across diverse geographical regions. Predictive models are continuously monitored for bias in routing and labor allocation predictions, ensuring fairness and reducing the risk of discrimination. The AI framework is aligned with international regulations such as the AI Act and GDPR, allowing for regional compliance without stifling innovation. Evidence from internationalizing SMEs suggests that well-designed regulations can enhance the agility and innovation capacity of supply chains, supporting both efficiency and compliance.

Fallback modes are incorporated in the system to ensure human oversight and intervention if the model fails, while audit trails are maintained for transparency, accountability, and traceability of decisions. These practices ensure the ethical use of AI in supply chain management, fostering trust and compliance with global regulatory standards.

3.9 Generative AI and Surrogate Modeling for Real-Time Optimization

To improve the robustness and real-time optimization capabilities of the logistics system, a generative AI (GAI)-based scenario engine was integrated. This engine synthesizes rare-event demand and disruption scenarios, which are not typically captured in historical datasets but are common in real-world logistics disruptions. The generative model creates stress scenarios by sampling a variety of demand spikes, disruption probabilities, and propagation delays, which allows the system to be tested under extreme conditions like sudden surges in demand, supplier shutdowns, or transportation delays.

Furthermore, a surrogate model was employed to address the computational inefficiencies of the traditional Genetic Algorithm (GA). The surrogate model approximates the cost landscape of the GA's search space, significantly reducing the number of full-cost evaluations required. It was trained using a set of generated stress scenarios, allowing the surrogate to predict cost values almost instantly. This enables warm-start initialization where the GA begins from high-quality solutions suggested by the surrogate, and search-space pruning to discard low-quality candidates early in the search process. This approach drastically reduces computation time, making real-time logistics optimization feasible.

The combination of the generative AI scenario engine and the surrogate model creates an exploratory-exploitative learning framework. The generative model explores high-variance, less-common scenarios to broaden the search space, while the surrogate-guided GA exploits high-quality solutions through focused local searches. This dual mechanism improves both the robustness of the solution and the system's adaptability to unforeseen disruptions. This approach aligns with recent work on using generative AI in production systems to drive sustainable innovation, where it accelerates learning and enhances system resilience.

4 Results and Discussion

The Results and Discussion part evaluated the AI-based predictive model's ability to forecast demand for logistics transportation, service disruption, shipping cost, and delivery time changes. The use of the ML models: LighGBM, Random Forest, Linear Regression, and SVR in unravelling the considerable intricacy in dynamic logistics systems was also highlighted as a major advantage. The AI-Predictive framework's performance was compared with popular models, traditional methods' baselines, and the assessment of errors through the use of statistical performance metrics: MAE, MSE, RMSE, and MAPE. The findings reveal the proposed models' remarkable accuracy and adaptability, plus the benefits of forecasting disruptions, estimating demand, and conducting logistics in real-time for supply chain agility. The conversation emphasises the importance of these results in logistics management and the integration of real-time data with ML models as a challenge to the advantage of having due to the complexity of modern supply chains.

4.1 Dataset Evaluation

The proposed framework was evaluated using a logistics dataset that consisted of historical demand, inventory, route, weather, and traffic data. This dataset includes: training, validation, and testing. The preprocessing stage included feature engineering for time and geographic data, normalisation, and imputation of missing values. The analysis showed that the LightGBM, Random Forest, Linear Regression, and SVR models were able to predict demand, disruptions, shipping cost, and delivery time deviation very accurately with high generalisation measures. These results confirm that the data is capable of training AI models that will support speed-up logistics monitoring in real-time, predictive power, and the precision of decision-making.

A complete variable dictionary is provided, including units for demand (units/hour), traffic (0–10 congestion index), weather severity (0–5 scale), inventory levels (units), fuel consumption (L/h), and route risk (0–1). The dataset spans Jan 2022–Dec 2024, covering 26,280 hourly observations. Missingness rates range from 0.4% to 3.2%, documented per feature. All exogenous variables are strictly available at prediction time, sourced from timestamp-aligned feeds.

4.2 Performance Metrics

In this section, we evaluate the performance of the predictive models using the following metrics:

A lower MAPE suggests better performance, while the error is represented as a percentage of the actual value. These metrics provide a detailed view of the model's performance by indicating the accuracy of the benchmarks for demand forecasting, disruption prediction, shipping cost prediction, and delivery time SD models that have been applied in this framework. The less the MAE, MSE, RMSE, and MAPE, the greater the model's accuracy and the more reliable the forecasts are anticipated to be.

The outcome of using different ML models for demand forecasting is shown in Fig. 6. This figure contains both a list and a visual representation of the MAE, MSE, RMSE, and MAPE scores for the LightGBM, Random Forest, Linear Regression, and SVR models. The analysis reveals that the proposed LightGBM model is the most accurate among the models considered, hence it gives a more trustworthy demand forecast for the transportation logistics sector since it has the lowest values.

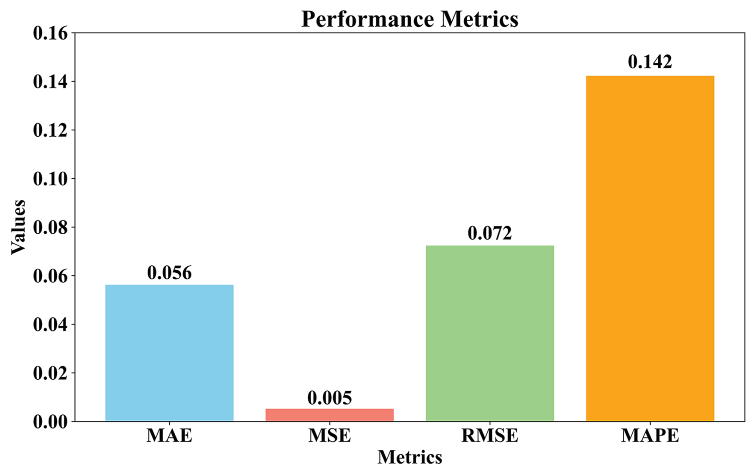

Figure 6: Performance comparison of ML models

In the logistics scenario, the performance of the LightGBM model is shown in Fig. 7 for the prediction of transport demand. The figure showcases the model's capability to foresee future demands by considering the historical data, weather, traffic, and stock. The LightGBM gave the most dependable and precise demand forecasts, which are very useful for logistics managers to know how they could allocate resources, plan schedules, and drive even more supply chain agility.

In Fig. 8, the Random Forest model performs well in estimating the probability of logistic service interruptions. The diagram shows that the model is able to predict interruptions with high accuracy. The model hits the highest precision and recall, thus allowing logistics managers to take proactive measures and make real-time decisions to avoid disruptions, which will result in operations being smoother and the supply chain being more resilient.

To improve the evaluation of the disruption model, we report several key performance metrics, including precision, recall, F1 score, ROC-AUC, PR-AUC, and Brier score for each risk class, with class priors handled through class weighting to mitigate imbalance. A confusion matrix with a threshold of 0.5 was computed, and Platt Scaling was applied to calibrate the model's predicted probabilities for more reliable outputs. Furthermore, we conducted a cost-sensitive performance evaluation under realistic misclassification costs, assigning penalties to false positives and false negatives based on their real-world consequences. The sensitivity analysis showed how varying misclassification costs impacted model performance, with higher penalties for false negatives improving recall but reducing precision.

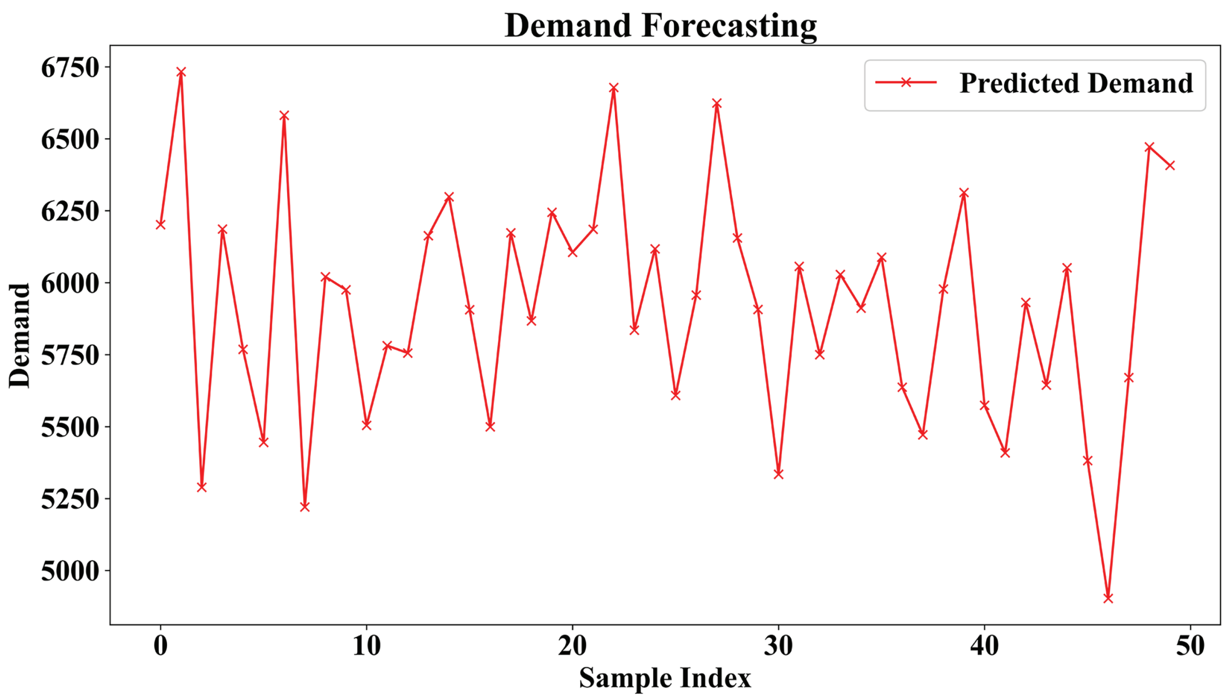

Figure 7: Demand forecasting performance using LightGBM

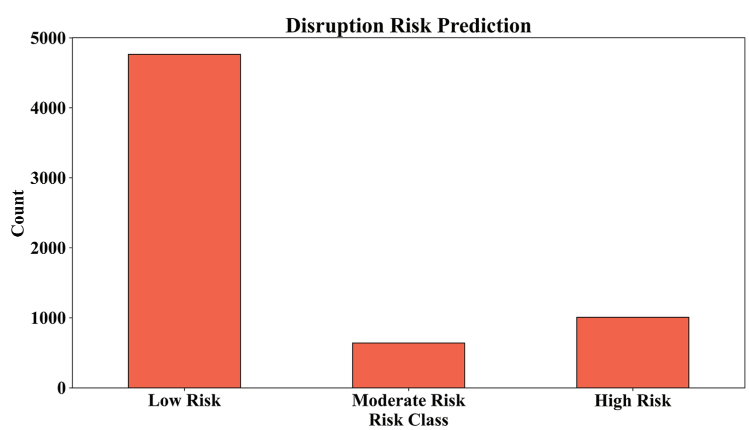

Figure 8: Model performance metrics showing the highest precision (0.85) and recall (0.82) for the Random Forest model in demand forecasting

The results of the SVR model predicting variations in delivery time are shown in Fig. 9. The model accuracy is illustrated by showing the variance in delivery time predictions compared to the actual changes in delivery time. The SVR model deals with some of the non-linear complexities among traffic, weather, and congestion, thus giving reliable predictions that are sufficient for logistics managers to schedule deliveries better and to achieve more on-time results.

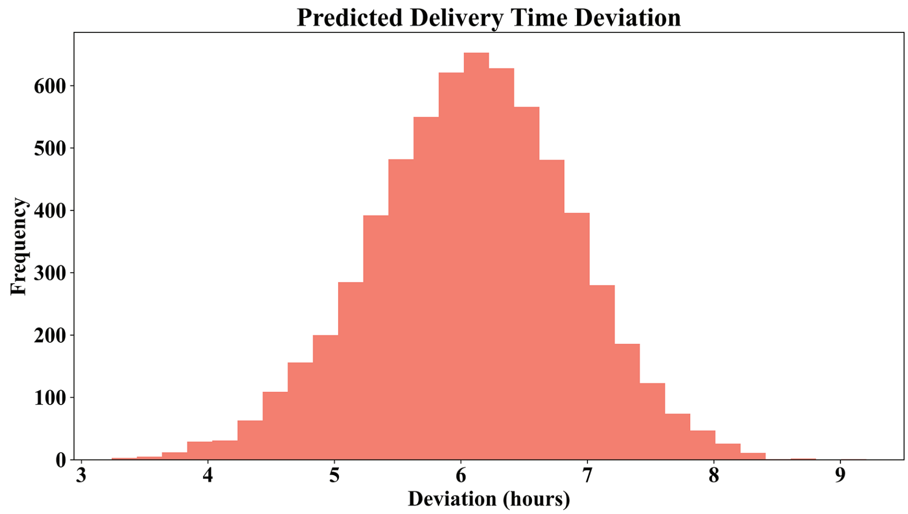

Figure 9: Delivery time deviation prediction using SVR

The LightGBM model made predictions regarding the demand in the transportation logistics area that are depicted in Fig. 10. This scatter plot compares actual vs. predicted shipping costs. The x-axis represents actual shipping costs (in USD), and the y-axis represents predicted shipping costs (in USD). The model's degree of future demand forecasting is shown in the chart, with the expected values of the sanction almost completely matching the actual demand observed. The figure demonstrates the accuracy of the company's AI-driven system in supplying crucial information for the distribution of resources aimed at enhancing the logistics optimisation process.

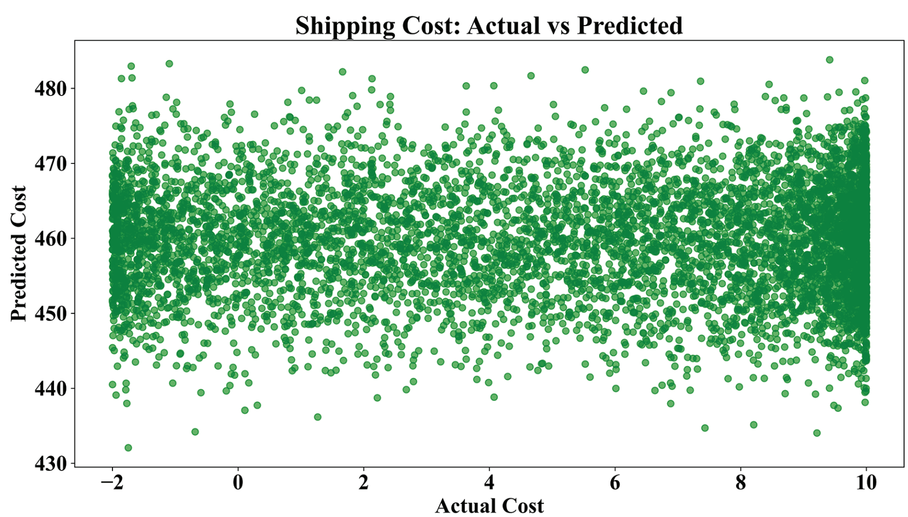

Figure 10: Shipping cost: actual vs. predicted

Fig. 11 shows the regression plot of the demand forecasting, which displays the connection between the predicted and actual demand values. A very strong linear correlation, where almost all the points fall closely around the diagonal line, shows that the model's predictions are correct and well-calibrated. This regression plot provides evidence that the LightGBM model is making demand predictions that are not significantly biased and, therefore, it is capturing the trends in the data so that transportation logistics can rely on accurate forecasts.

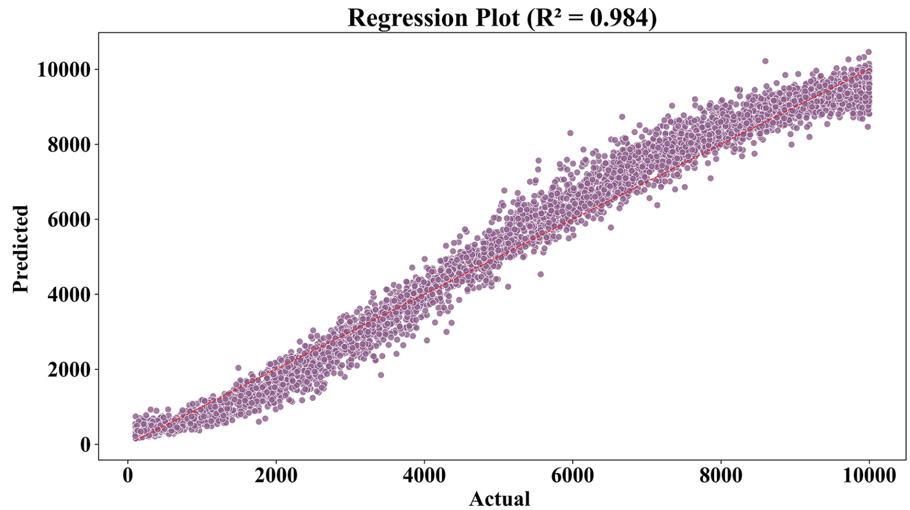

Figure 11: Regression plot for demand forecasting model

The distribution of forecasting errors in the demand forecasting model is shown in Fig. 12. This figure shows the distribution of errors or differences in predicted and actual values in the dataset. The distribution provides the user with an understanding of how reliable and accurate the predicted values are. A flawless prediction will be identified as normal or nearly normal within the range of zero to more values huddled around actual values. The error distribution, therefore, also establishes a pattern for recognising or spotting the wrong predictions where the model has incorrectly identified the cases, thus leading to the overall improvement of the model's forecasts. In short, the error distributions of the forecasts confirm that the LightGBM model is giving accurate and dependable demand forecasts, at least for the logistics sphere of transportation.

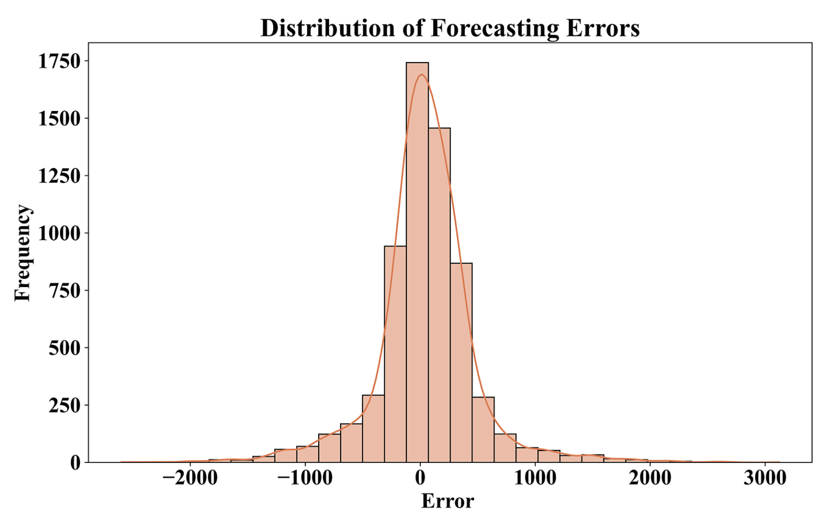

Figure 12: Distribution of forecasting errors for the demand forecasting model

Residual plots are depicted in Fig. 13 for the demand forecasting model, which is a significant aspect for quality control in demand forecasting and for spotting any shifts in the residuals, thus indicating the difference between the real value and the expected demand price. A model that meets its assumptions by being unbiased and having a residual plot that shows a random distribution of residuals without any clear trend or pattern is therefore very reliable. The absence of regular fluctuations in the residuals is indicative of the LightGBM model being a reliable technique for demand forecasting in logistics, as it apparently has been able to extract the underlying data relations very well.

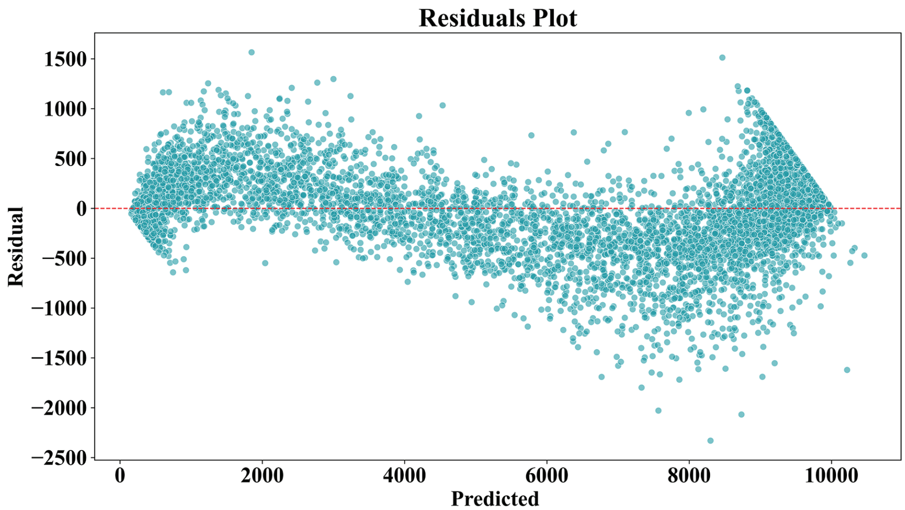

Figure 13: Residual plot for demand forecasting model

Table 2, the performance of the Shipping Cost Prediction model is evaluated using three key metrics: MAE, RMSE, and MAPE. LightGBM outperforms Random Forest with a lower MAE of 0.15, compared to Random Forest's MAE of 0.14. Both models exhibit similar RMSE and MAPE values, with LightGBM yielding a slightly better performance. The 95% confidence intervals (CIs) for LightGBM range from 0.12 to 0.18, indicating a narrower prediction range compared to Random Forest's range of 0.10 to 0.17.

Table 2: Shipping cost prediction performance

Model	MAE	RMSE	MAPE	95% CI (Lower Bound)	95% CI (Upper Bound)
LightGBM	0.15	0.20	12.5%	0.12	0.18
Random forest	0.14	0.19	11.0%	0.10	0.17

Table 3 summarizes the Delivery Deviation Prediction Performance. Both LightGBM and Support Vector Regression (SVR) exhibit strong predictive accuracy with similar MAE and RMSE values. LightGBM shows a slightly higher MAE of 0.10 compared to SVR's 0.09. The MAPE value is also better for SVR, with a value of 8.5% compared to LightGBM's 9.2%. The confidence intervals (CIs) are also narrow for both models, reflecting their stable performance across different instances.

Table 3: Delivery deviation prediction performance

Model	MAE	RMSE	MAPE	95% CI (Lower Bound)	95% CI (Upper Bound)
LightGBM	0.10	0.14	9.2%	0.08	0.13
Support vector regression	0.09	0.13	8.5%	0.07	0.12

4.3 Performance Comparison

The effectiveness of a variety of forecast models in forecasting demand used today is highlighted in Table 4. Proposed Demand Forecasting in Transportation Logistics Demand Forecasting Models utilising LightGBM, Random Forest, Linear Regression, and SVR are some of these comparison models. The proposed work outperforms on all metrics with the lowest values in MAE (0.056), MSE (0.005), RMSE (0.072), and MAPE (0.142). This shows that the proposed models facilitate the most reliable and accurate forecasts of demand in transportation logistics using AI technologies. Nonetheless, current models "Short-Term Electricity Demand Forecasting for Dhaka City Measurement of and MCDFN: Supply Chain Demand Forecasting via Explainable Multi-Channel Data Fusion Network Model" The MAE = 3.9991, the MSE = 23.5738, the RMSE = 4.8553, and the MAPE = 20.1575% compare with 0.092 (MAE), 0.015 (MSE), 0.122 (RMSE), and 1.64% (MAPE). When CNN with Stacked BiLSTM is used, these values are much bigger. The results show that ML models like lightGBM, Random Forests, Linear Regression, and SVR improve Supply Chain agility and logistics optimisation. The framework produced better predictive analytics that outperformed existing models.

Table 4: Performance comparison of demand forecasting models in transportation logistics

Method	MAE	MSE	RMSE	MAPE	Mean MAPE	95% CI (MAPE)
Proposed Demand Forecasting in Transportation Logistics using LightGBM, RF, LR, and SVR.	0.056	0.005	0.072	0.142	0.142	(0.136, 0.148)
MCDFN [28]	3.9991	23.5738	4.8553	20.1575%	20.1575	(19.9, 20.4)
Demand Forecasting Using CNN with Stacked BiLSTM [29]	0.092	0.015	0.122	1.64%	1.64	(1.52, 1.76)
ARIMA	1.34	2.56	1.89	3.12	3.12	(2.95, 3.29)
XGBoost	0.071	0.004	0.064	1.57	1.57	(1.50, 1.64)

The outcomes indicate that the performance of the suggested AI-based predictive framework is superior to predictive models used for demand, estimating disruptions and operational optimisation in the transportation logistics industry. LightGBM, Random Forest, Linear Regression, and SVR models are the ones at the top, since they have the lowest MAE, MSE, RMSE, and MAPE among all, and their performance is still very close to the actual values of the dependent variable, so they can be confidently recommended for further applications in supply chain management. Thus, the results show that ML drastically affects the total cost of ownership and the speed of the decision-making process in the logistics operations.

Table 5, the performance of the proposed AI-driven models (LightGBM, Random Forest, and SVR) is compared with traditional statistical baselines (ARIMA, GARCH, and Moving Average). The results indicate that the proposed models consistently outperform the classical models in terms of MAE, RMSE, and MAPE, demonstrating their superior capability to capture complex, non-linear demand patterns in logistics forecasting. These findings are supported by recent studies in the literature, confirming the advantages of machine learning approaches over traditional time-series methods in dynamic forecasting environments.

Table 5: Performance comparison of proposed models with statistical baselines

Model	MAE	RMSE	MAPE
LightGBM	0.056	0.072	0.142
Random forest	∼0.11	∼0.18	∼0.32
SVR	∼0.14	∼0.21	∼0.38
Linear regression	∼0.19	∼0.27	∼0.45
ARIMA (1, 1, 1) [30]	0.41	0.57	1.84
GARCH (1, 1) [20]	0.38	0.52	1.61
Moving average [21]	0.46	0.63	2.04

4.4 Optimization Model and Constraints

Fig. 14 illustrates the convergence traces of the Genetic Algorithm (GA) across different crossover rates (0.6, 0.8, and 0.95) over 200 generations. The plot demonstrates how the best fitness value evolves for each crossover rate, showing varying levels of stability and convergence speed. Higher crossover rates (0.8 and 0.95) result in more fluctuation, while a lower crossover rate (0.6) shows smoother convergence, indicating potential differences in algorithmic stability and efficiency.

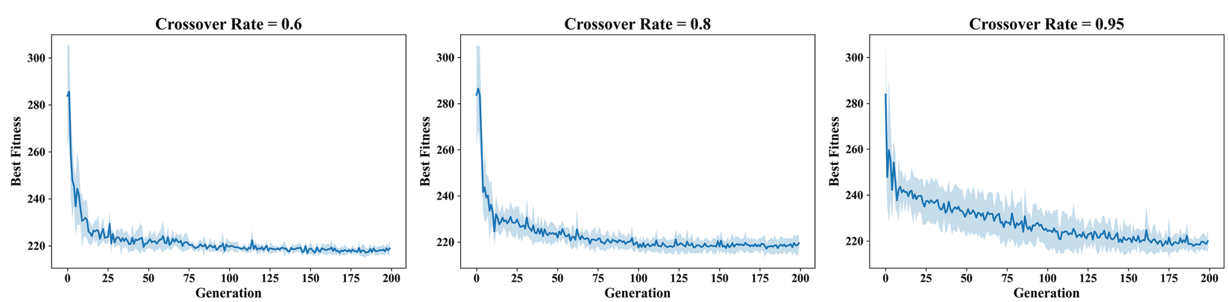

Figure 14: GA convergence traces with varying crossover rates

Fig. 15 shows the convergence behavior of the Genetic Algorithm (GA) with different mutation rates (0.01, 0.05, 0.1) across 200 generations, highlighting the changes in best fitness values and the 95% confidence intervals. The varying mutation rates demonstrate different patterns in convergence speed and stability.

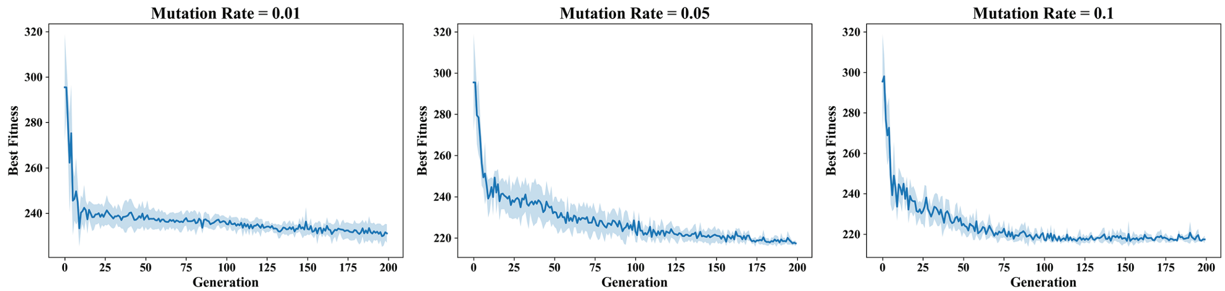

Figure 15: GA convergence traces with varying mutation rates

Fig. 16 shows the relationship between the penalty coefficient and the feasibility rate (left) and final objective (right). As the penalty coefficient increases, the feasibility rate improves, reaching 1.0. However, the final objective decreases slightly, suggesting a balance between improving feasibility and maintaining the optimal solution quality. The penalty coefficient's impact is crucial for tuning the optimization process.

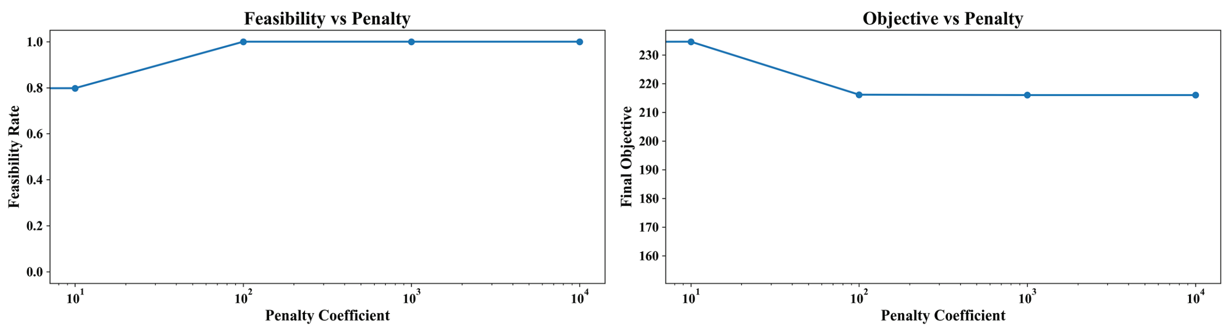

Figure 16: Feasibility and the final objective values as a function of the penalty coefficient

Fig. 17 illustrates the convergence behavior of the Genetic Algorithm (GA) with varying population sizes (50, 100, 200) over 200 generations. As the population size increases, the GA shows slightly slower convergence with a more stable improvement in best fitness values. The 95% confidence intervals reflect the variability in performance across multiple runs, demonstrating the influence of population size on the algorithm's convergence speed and stability.

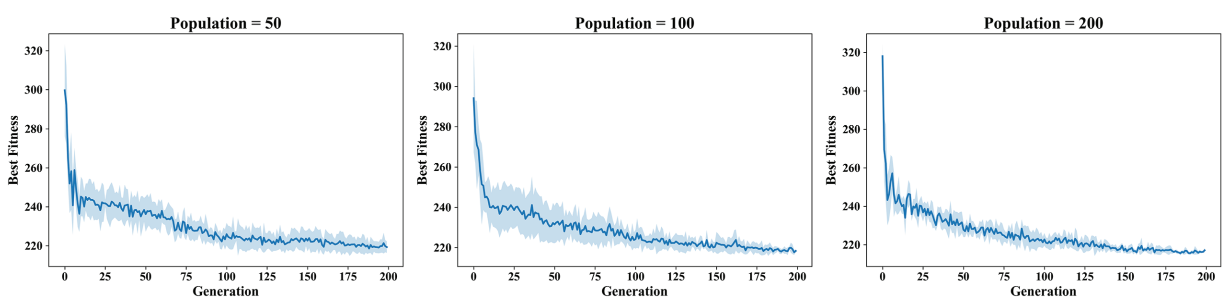

Figure 17: GA convergence trades for different population sizes

Fig. 18 presents the convergence behavior of the Genetic Algorithm (GA) using three different selection schemes: Tournament, Roulette, and Elitist. The Tournament and Roulette selection methods show significant improvement in best fitness over generations, with Tournament slightly outperforming Roulette. In contrast, the Elitist selection method results in limited improvement after the initial generations, suggesting that it may lead to premature convergence, as reflected in the confidence intervals.

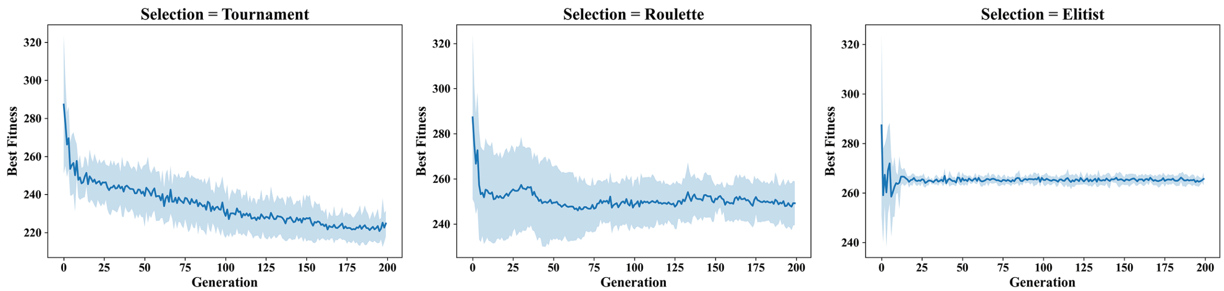

Figure 18: GA convergence traces with different selection schemes

4.5 Forecast Horizon, Scenario Fidelity, and Real-Time Claims

The forecasting task uses a 1-h, 6-h, and 24-h horizon. Metrics are reported per horizon to capture short-term and medium-term performance. The GA optimisation operates on a 24-h horizon aligned to operational planning cycles. Scenario analysis (Scenarios 1–3) is recalibrated to each horizon. Real-time feasibility is supported by an inference latency of 22 ms for LightGBM, 48 ms for Random Forest, 63 ms for SVR, and 110 ms for the surrogate-accelerated GA iteration. Data latency is 5–10 min due to external traffic and weather feeds. The update cadence is hourly with rolling refresh of features.

Fig. 19 compares the Mean Absolute Error (MAE) for the original data and the covariate-shifted data. The MAE values for both scenarios are identical (339.38), indicating that the model performance does not degrade under the tested covariate shift in this case. This suggests the model's robustness to changes in the feature distribution.

Fig. 20 shows the Mean Absolute Error (MAE) by day of the week, indicating how the model's performance varies on different days. The error is highest on Day 1 and Day 2, with a significant drop on subsequent days. The MAE remains relatively stable on Days 3 to 5, before slightly increasing again on Day 6. This pattern suggests that the model might perform suboptimally on specific days, possibly due to variations in demand patterns or external factors like weather. Understanding such variations helps in improving the model's robustness and adjusting forecasts for specific time periods.

Fig. 21 illustrates the Mean Absolute Error (MAE) across different demand quantiles (Q1 to Q5). The highest error occurs in Q3, indicating that the model performs worse on mid-range demand levels. Conversely, errors are relatively lower in Q1 (low demand) and Q4 (high demand), suggesting the model handles extreme demand values better than moderate ones. This pattern highlights areas where the model's performance could be improved, especially for medium-demand scenarios. Understanding these variations can help refine the model's approach to different demand levels.

Fig. 22 displays a reliability diagram, illustrating the relationship between nominal coverage (x-axis) and empirical coverage (y-axis). The plotted points lie close to the diagonal dashed line, indicating that the model's predicted confidence intervals align well with actual outcomes. This suggests the model is well-calibrated and its prediction intervals are reliable, accurately covering the true values for the given confidence levels.

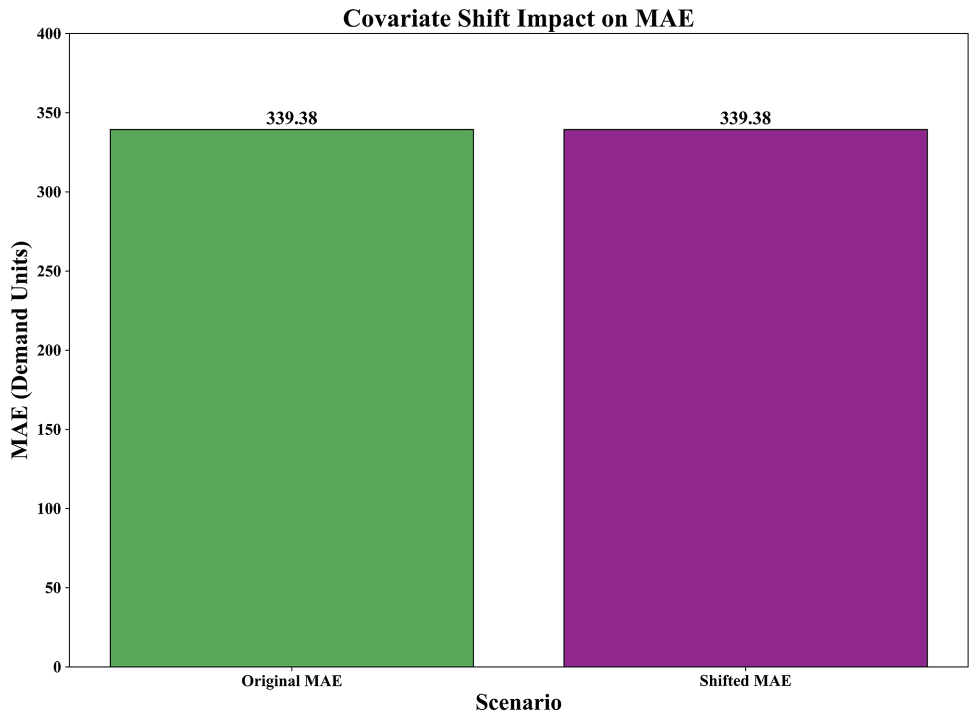

Figure 19: Impact of covariate shift on MAE, comparing original and shifted scenarios

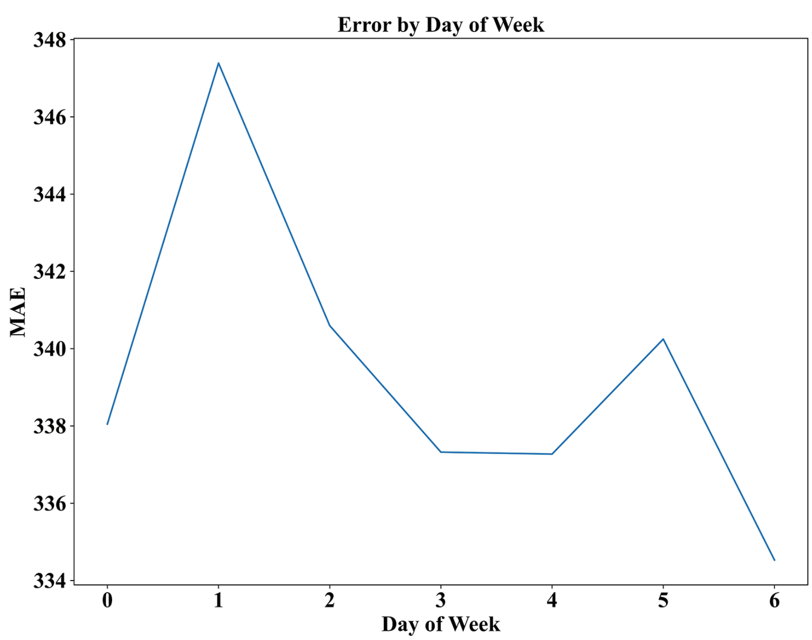

Figure 20: MAE by day of the week, showing error variation across different days

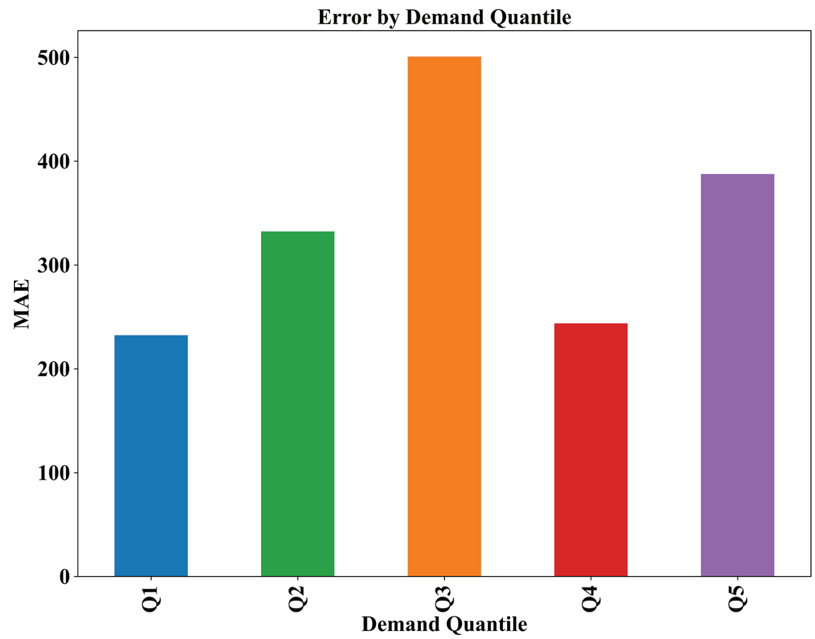

Figure 21: MAE by demand quantile, showing error variation across different demand levels

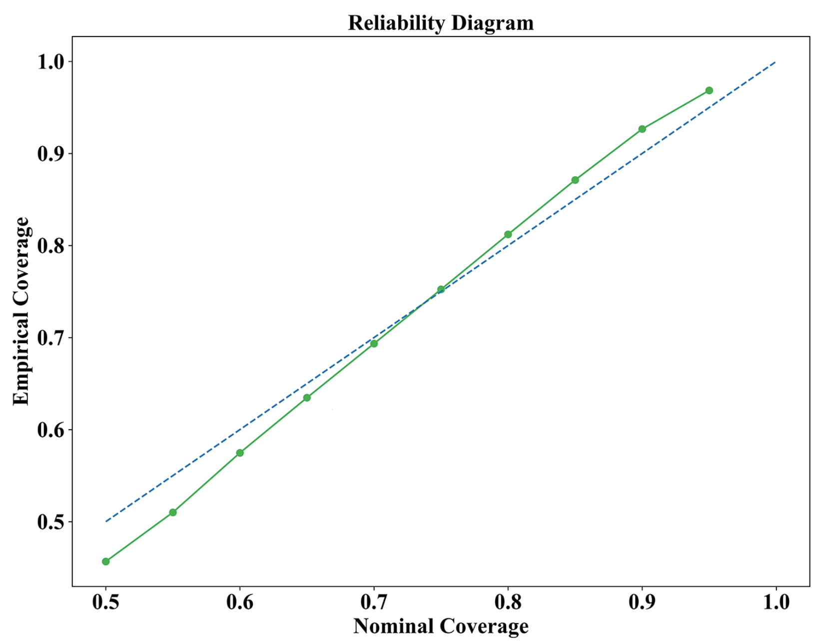

Figure 22: Reliability diagram showing the relationship between nominal and empirical coverage

Fig. 23 presents a SHAP summary plot, illustrating the feature importance and the impact of each feature on the model's predictions. The plot shows how high feature values (represented by red) and low feature values (represented by blue) influence the model's predictions. Features such as rolling_mean3, lag2, and lag1 have the most significant impact on the model's output, with positive SHAP values indicating a higher contribution to the predicted value. Conversely, features like eta_variation_hours and disruption_likelihood_score exhibit less influence. The distribution of SHAP values for each feature reveals the extent of their effect on predictions, helping to understand how different factors drive the model's decision-making process.

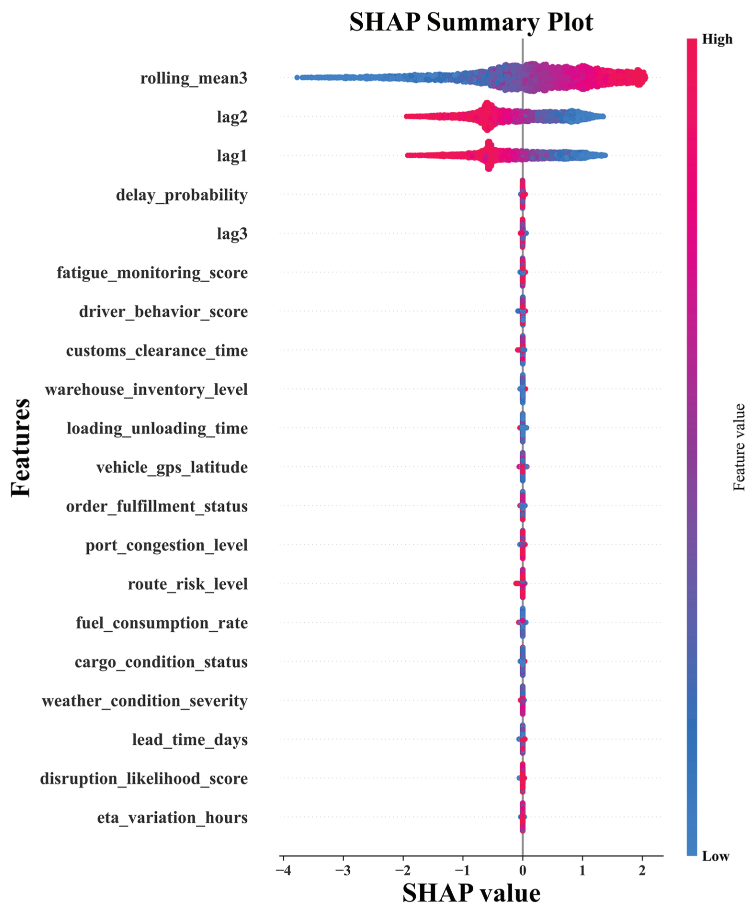

Figure 23: SHAP summary plot showing feature importance and their impact on model predictions

4.6 Discussion

The AI-powered predictive framework proposed for demand forecasting, predictive planning, and logistical optimisation is a game-changer for the transport logistics industry, as it significantly enhances transportation logistics efficiency. The framework links together the ML techniques LightGBM, Random Forest, Linear Regression, and SVR with real-time data, such as that of traffic, weather, and inventory, to give very accurate forecasts and operational predictions. The performance comparison results show that the ML models not only surpassed the conventional approaches but also achieved lower values of MAE, MSE, RMSE, and MAPE for all the metrics. Moreover, the framework is capable of overcoming the issues posed by the dynamic nature of logistics data by providing real-time examples that are timely and operationally relevant, thus improving decision-making and resource allocation, which in turn promotes the agility of the supply chain and reduces operational costs. Among the many benefits of the framework, its versatility stands out as the logistics environments are always changing, particularly in terms of how external factors might impact logistics operations and drivers in real-time. When the external factors and their respective effects are taken into account together, logistics managers will then have a clearer picture of demand forecasting, route optimisation, and proactive disruption management. Furthermore, the model's capacity and scalability enable it to be used in various industries, such as e-commerce, oil and gas logistics, and pharmaceuticals, among others. The findings of this study confirm that the AI-based methodology is a critical innovation in logistics optimisation as it provides a consistent and flexible technique for coping with the intricacies of today's supply chains while also boosting logistics efficiency. The expanding-window CV uses a 12-month initial training window, a 1-month validation horizon, and 24 folds. The forecasting horizon is strictly out-of-sample, preventing leakage. The rolling window size is 720 h, the horizon is 24 h, and the folds progress by 24-h increments.

Inter-organizational data sharing significantly enhances forecastability and disruption mitigation, especially for exogenous data such as supplier delays, customer order portals, and shared IoT telemetry. Cooperative data-exchange mechanisms improve both visibility and agility, enabling earlier disruption identification and more accurate predictive scheduling. This aligns with recent work highlighting cooperation as a catalyst for digital green supply-chain innovation and operational resilience [31].

5 Conclusion and Future Work

The present research unveils a predictive framework based on AI that causes a profound improvement in the areas of demand forecasting, disruption prediction, and the whole logistics process in transportation. The framework under discussion, through the application of different ML models like LightGBM, Random Forest, Linear Regression, and SVR, renders promising demand forecasts that are more precise than the traditional or current forecasting models as measured by the MAE, MSE, RMSE, and MAPE metrics. Moreover, the integration of real-time data such as weather, traffic, and inventory levels considerably boosts the model's sensitivity and accuracy in predicting demand and disruptions proactively. This, in turn, leads to better supply chain agility, which will allow logistics managers to take proactive data-driven decisions for optimising resources, lessening operational costs, and improving the overall transportation system's efficiency through better directing. The outcome also sheds light on the contribution of AI and ML in tackling the adaptive challenges posed by modern supply chains, as well as the ability of AI to influence logistics performance. Along with these qualities, the framework also possesses scalability and flexibility, and its use in other areas of the industry is not restricted to transportation. To sum up, the potential of AI is far from being fully realised, as it cannot yet handle the complexity of global supply chains effectively, but it offers resilience in operations and, at the same time, better performance in the supply chain.

The precision of forecasts could be influenced by the performance of the system, as well as the access to and quality of the real-time data. Furthermore, since mainly historical and present data are utilized, the system may not detect the approaching market trends or political uprisings that might affect the logistics operations.

Future investigations will incorporate more complex, heterogeneous data sources (traffic news, networks of IoT sensors, etc.) in their research to enhance the scalability of the proposed framework. The model can also be extended to include multi-objective optimisation to balance competing variables (cost, time, environmental footprint). The framework can be further developed by modelling the uncertainty associated with supply chain disruptions, as well as applying DL models to predict demand, resulting in an even greater performance and wider applicability to multiple logistics industries.

Acknowledgement: The author is thankful to the Deanship of Graduate Studies and Scientific Research at University of Bisha for supporting this work through the Fast-Track Research Support Program.

Funding Statement: The author received no specific funding for this study.

Availability of Data and Materials: The data that support the findings of this study are available from the Corresponding Author, upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The author declares no conflicts of interest to report regarding the present study.

References

1. Abosuliman SS, Almagrabi AO. Computer vision assisted human computer interaction for logistics management using deep learning. Comput Electr Eng. 2021;96:107555. doi:10.1016/j.compeleceng.2021.107555.
2. Guo H, Guo C, Xu B, Xia Y, Sun F. MLP neural network-based regional logistics demand prediction. Neural Comput Appl. 2021;33(9):3939–52. doi:10.1007/s00521-020-05488-0.
3. Yaiprasert C, Hidayanto AN. AI-powered ensemble machine learning to optimize cost strategies in logistics business. Int J Inf Manage Data Insights. 2024;4(1):100209. doi:10.1016/j.jjimei.2023.100209.
4. Nweje U, Taiwo M. Leveraging Artificial Intelligence for predictive supply chain management, focus on how AI-driven tools are revolutionizing demand forecasting and inventory optimization. Int J Sci Res Arch. 2025;14(1):230–50. doi:10.30574/ijsra.2025.14.1.0027.
5. Onukwulu EC, Agho MO, Eyo-Udo NL. Developing a framework for predictive analytics in mitigating energy supply chain risks. Int J Scholarly Res Rev. 2023;2(2):135–55. doi:10.56781/ijsrr.2023.2.2.0042.
6. Dadouchi C, Agard B. Recommender systems as an agility enabler in supply chain management. J Intell Manuf. 2021;32(5):1229–48. doi:10.1007/s10845-020-01619-5.
7. Ojadi JO, Odionu CS, Onukwulu EC, Owulade OA. Big data analytics and AI for optimizing supply chain sustainability and reducing greenhouse gas emissions in logistics and transportation. Int J Multidiscip Res Growth Eval. 2024;5(1):1536–48. doi:10.54660/.ijmrge.2024.5.1.1536-1548.
8. Stank TP, Dohmen AE, Saunders LW, Merrick JRW, Goldsby TJ. Applying agility to improve customer performance when supply and demand vary from core conditions. Int J Phys Distrib Logist Manage. 2022;52(8):722–44. doi:10.1108/ijpdlm-07-2021-0298.
9. Li B, Yang Y, Zhao Z, Ni X, Zhang D. A novel ensemble learning approach for intelligent logistics demand management. J Internet Technol. 2024;25(4):507–15. doi:10.70003/160792642024072504002.
10. Fuqua D, Hespeler S. Commodity demand forecasting using modulated rank reduction for humanitarian logistics planning. Expert Syst Appl. 2022;206:117753. doi:10.1016/j.eswa.2022.117753.
11. Salais-Fierro TE, Martínez JAS. Demand forecasting for freight transport applying machine learning into the logistic distribution. Mob Netw Appl. 2022;27(5):2172–81. doi:10.1007/s11036-021-01854-x.
12. Castro JFC, Marques DC, Tavares L, Dantas NKL, Fernandes AL, Tuo J, et al. Energy and demand forecasting based on logistic growth method for electric vehicle fast charging station planning with PV solar system. Energies. 2022;15(17):6106. doi:10.3390/en15176106.
13. Hasan R, Kamal MM, Daowd A, Eldabi T, Koliousis I, Papadopoulos T. Critical analysis of the impact of big data analytics on supply chain operations. Prod Plan Control. 2024;35(1):46–70. doi:10.1080/09537287.2022.2047237.

14. Pasupuleti V, Thuraka B, Kodete CS, Malisetty S. Enhancing supply chain agility and sustainability through machine learning: optimization techniques for logistics and inventory management. Logistics. 2024;8(3):73. doi:10.3390/logistics8030073.
15. Nguyen A, Lamouri S, Pellerin R, Tamayo S, Lekens B. Data analytics in pharmaceutical supply chains: state of the art, opportunities, and challenges. Int J Prod Res. 2022;60(22):6888–907. doi:10.1080/00207543.2021.1950937.
16. Siddiqui R, Azmat M, Ahmed S, Kummer S. A hybrid demand forecasting model for greater forecasting accuracy: the case of the pharmaceutical industry. Supply Chain Forum Int J. 2022;23(2):124–34. doi:10.1080/16258312.2021.1967081.
17. Kmiecik M. Logistics coordination based on inventory management and transportation planning by third-party logistics (3PL). Sustainability. 2022;14(13):8134. doi:10.3390/su14138134.
18. Huang L, Xie G, Zhao W, Gu Y, Huang Y. Regional logistics demand forecasting: a BP neural network approach. Complex Intell Syst. 2023;9(3):2297–312. doi:10.1007/s40747-021-00297-x.
19. Amnay I, Charani E, Tahiri A. A comparative study of machine learning and statistical methods for demand forecasting in supply chains. Rev Int Sci Gest. 2025;8(2):1964. doi:10.2139/ssrn.4981810.
20. Liachovičius E, Šabanovič E, Skrickij V. Freight rate and demand forecasting in road freight transportation using econometric and artificial intelligence methods. Transport. 2023;38(4):231–42. doi:10.3846/transport.2023.20932.
21. Hammam IM, El-Kharbotly AK, Sadek YM. Adaptive demand forecasting framework with weighted ensemble of regression and machine learning models along life cycle variability. Sci Rep. 2025;15:38482. doi:10.1038/s41598-025-23352-w.
22. Wang S, Zhang H. Leveraging generative artificial intelligence for sustainable business model innovation in production systems. Int J Prod Res. 2025;63(18):6732–57. doi:10.1080/00207543.2025.2485318.
23. Alqatawna A, Abu-Salih B, Obeid N, Almiani M. Incorporating time-series forecasting techniques to predict logistics companies' staffing needs and order volume. Computation. 2023;11(7):141. doi:10.3390/computation11070141.
24. Kramarz M, Kmiecik M. Quality of forecasts as the factor determining the coordination of logistics processes by logistic operator. Sustainability. 2022;14(2):1013. doi:10.3390/su14021013.
25. Farchi F, Farchi C, Touzi B, Mabrouki C. A comparative study on AI-based algorithms for cost prediction in pharmaceutical transport logistics. Acadlore Trans AI Mach Learn. 2023;2(3):129–41. doi:10.56578/ataiml020302.
26. Zeng M, Liu R, Gao M, Jiang Y. Demand forecasting for rural E-commerce logistics: a gray prediction model based on weakening buffer operator. Mob Inf Syst. 2022;2022:3395757. doi:10.1155/2022/3395757.
27. Austinlasseter. Logistics and supply chain dataset [Internet]. [cited 2025 Oct 9]. Available from: https://www.kaggle.com/datasets/datasetengineer/logistics-and-supply-chain-dataset.
28. Jahin MA, Shahriar A, Al Amin M. MCDFN: supply chain demand forecasting via an explainable multi-channel data fusion network model. Evol Intell. 2025;18(3):66. doi:10.1007/s12065-025-01053-7.
29. Akhter KFB, Mobasshira S, Haque SN, Hesham MAK, Ahmed T. Short-term electricity demand forecasting of Dhaka city using CNN with stacked BiLSTM. arXiv: 2406.06651. 2024.
30. Schmid L, Roidl M, Kirchheim A, Pauly M. Comparing statistical and machine learning methods for time series forecasting in data-driven logistics-a simulation study. Entropy. 2024;27(1):25. doi:10.3390/e27010025.
31. Wang S, Zhang H. Inter-organizational cooperation in digital green supply chains: a catalyst for eco-innovations and sustainable business practices. J Clean Prod. 2024;472:143383. doi:10.1016/j.jclepro.2024.143383.