

American Journal of Advanced Technology and Engineering Solutions

Volume: 4; **Issue:** 2
Pages: 33-81
eISSN: 3067-5146

open access

Crossref

Machine Learning Enhanced Demand Forecasting and Portfolio Analytics for Scalable Growth Management in Digital Logistics Operations

H C H Raiyan Ahmed[1];

1. *Client Acquisition Lead, Growth Management, REDX Logistics Limited (A concern of ShopUp), Dhaka, Bangladesh. Email: hch.raiyan.ahmed@gmail.com*

Doi: 10.63125/5xq3nd53
Policy Page DOI: 10.63125/ajates-crossmark-policy

Received: *09 March 2024;* ***Revised:*** *10 April 2024;* ***Accepted:*** *12 May 2024;* ***Published:*** *30 June 2024*

Abstract

This study developed and quantitatively evaluated a machine learning–enhanced demand forecasting and portfolio analytics framework for scalable growth management in digital logistics operations. A longitudinal analytical design was applied to 36 months of operational and commercial data comprising 128,560 eligible shipment transactions, 14,820 customer accounts, 46,735 bookings, 12 geographical markets, seven logistics service categories, 10 operating facilities, 214 vehicles, 3,684 completed routes, and 428 workforce records. Overall data completeness reached 96.4%. Mean monthly demand was 3,571.1 shipments with a coefficient of variation of 19.2%, while maximum monthly demand reached 5,214 shipments and demand increased by 18.7% across the observation period. Conventional statistical forecasting methods were compared with random forest, support vector regression, artificial neural networks, gradient boosting, LSTM, and XGBoost using chronological validation and independent testing. XGBoost achieved the strongest forecasting performance, with R^2 = 0.932, MAPE = 5.91%, MAE = 7.21, and RMSE = 10.18, representing a 54.5% reduction in MAE relative to SARIMA. Feature engineering increased XGBoost R^2 from 0.861 to 0.932. Portfolio segmentation demonstrated substantial commercial concentration, with the highest-value 20% of customers generating 63.7% of total revenue and 55.4% of shipment activity. Demand growth was strongly correlated with revenue growth (r = 0.812), while the multivariable commercial model explained 71.8% of adjusted revenue-growth variance. Capacity utilization averaged 82.4% but reached 97.3% during peak conditions, when fulfillment declined to 87.9%, overtime increased to 18.7%, and cost per shipment reached $89.70. A practical scalability threshold of approximately 90% utilization was identified. Statistical comparisons demonstrated large effects, including Cohen's d = 1.31 for the XGBoost–SARIMA forecasting difference. The integrated framework explained 88.4% of subsequent portfolio-growth variation and classified scalable versus capacity-constrained growth with 89.7% accuracy and ROC-AUC = 0.941. The findings demonstrated that scalable digital logistics growth depended on integrating accurate machine learning forecasts with customer value, commercial performance, portfolio diversification, service reliability, and capacity availability.

Keywords

Machine Learning, Demand Forecasting, Portfolio Analytics, Digital Logistics, Scalable Growth.

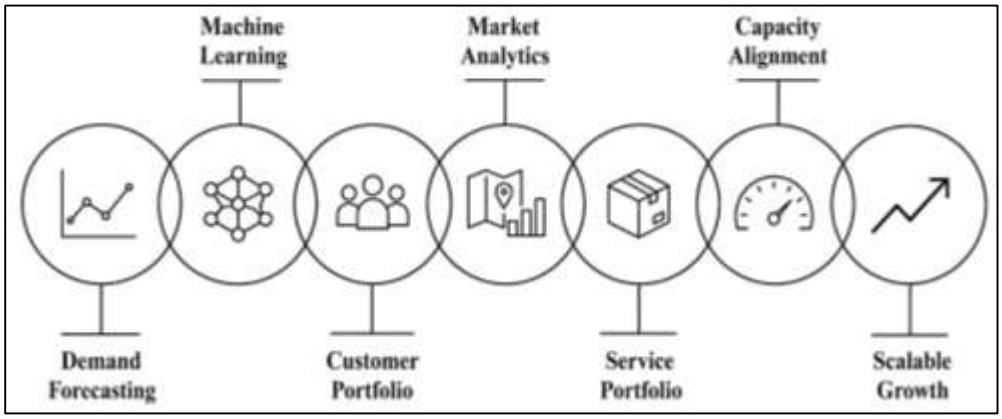
INTRODUCTION

Machine learning–enhanced demand forecasting can be defined as a data-driven analytical process in which computational algorithms learn relationships from historical and contemporaneous logistics data to estimate subsequent levels, distributions, and variations in customer demand, shipment volumes, service requirements, transaction activity, capacity needs, and related commercial outcomes (Abideen et al., 2021a). Unlike conventional forecasting approaches that primarily depend on predetermined statistical structures, machine learning models can identify nonlinear relationships, interactions among large numbers of explanatory variables, heterogeneous demand patterns, and complex temporal dependencies within operational datasets. Demand forecasting in digital logistics therefore represents more than the estimation of aggregate shipment volume; it incorporates prediction across customers, services, geographical markets, sales channels, logistics routes, time periods, and commercial segments. Digital logistics refers to the integration of data, connected information systems, automated workflows, analytics, and digitally coordinated transportation and fulfillment processes through which logistics activities are monitored and managed. Within this environment, portfolio analytics can be defined as the systematic quantitative evaluation of a logistics firm's combined portfolio of customers, services, markets, transactions, routes, contracts, opportunities, and revenue streams to determine their relative volume, value, profitability, growth, risk, and resource requirements. Scalable growth management refers to the structured management of increasing commercial and operational activity while maintaining acceptable levels of service quality, capacity utilization, cost efficiency, revenue performance, and organizational control (Woschank et al., 2020). These concepts are increasingly interconnected because growth in digital logistics creates expanding volumes of operational and commercial data that can be transformed into predictive information for planning and portfolio management. Research across more than 30 empirical, computational, and review studies has examined machine learning in supply-chain forecasting, freight prediction, inventory planning, transportation management, customer analytics, supply-chain optimization, sales forecasting, logistics network management, and digital decision support. A recent systematic synthesis covering 103 machine-learning studies identified demand forecasting as one of the major application domains of machine learning in supply-chain management, while another review examining 119 publications documented extensive application of machine-learning and deep-learning techniques to demand forecasting (Andronie et al., 2021). Collectively, this research base has established demand forecasting as a multidimensional quantitative problem requiring simultaneous consideration of temporal variation, customer heterogeneity, market conditions, operational constraints, and data quality. Portfolio analytics extends this analytical orientation by connecting predicted demand with customer value, service composition, market concentration, revenue contribution, profitability, and capacity allocation. The combined framework consequently treats demand not as an isolated forecasting variable but as a measurable commercial signal embedded within a broader portfolio of logistics activities and growth opportunities.

The international significance of machine learning–enhanced demand forecasting arises from the scale, geographical dispersion, service complexity, and interconnected nature of contemporary logistics operations. Logistics networks increasingly coordinate transportation, warehousing, freight forwarding, fulfillment, last-mile distribution, customs activities, inventory movement, and customer services across multiple geographical markets, creating substantial differences in shipment frequency, order size, customer requirements, lead times, route characteristics, and service demand (Gutierrez-Franco et al., 2021). International logistics operators frequently manage portfolios extending across countries, regions, customer industries, transport modes, and service categories, meaning that aggregate demand figures can conceal substantial differences among commercially important segments. Research conducted across Asian, European, North American, and other logistics and supply-chain contexts has demonstrated that demand patterns are affected by economic activity, e-commerce intensity, seasonal cycles, industrial production, customer purchasing behavior, geographical concentration, transportation accessibility, pricing conditions, and service availability. Regional logistics research has similarly demonstrated that neural-network forecasting can provide lower prediction errors and more stable estimates than selected conventional models when multiple demand indicators are incorporated. Freight forecasting studies have further shown that combining

time-series methods with machine-learning approaches can improve responsiveness to changing shipment patterns and support operational planning across different forecasting horizons. These characteristics are internationally important because forecasting errors can propagate through interconnected logistics systems. Underestimation of demand can contribute to inadequate vehicle availability, insufficient warehouse capacity, labor shortages, delayed fulfillment, missed delivery windows, and service deterioration, whereas systematic overestimation can contribute to excess capacity, underutilized assets, unnecessary labor expenditure, and inefficient capital deployment (Boone et al., 2019). Digital logistics platforms add another dimension by generating high-frequency data from transportation management systems, warehouse management systems, customer relationship management platforms, enterprise resource planning systems, digital marketplaces, tracking systems, telematics, and transaction databases. The international logistics environment therefore contains a broad analytical base for estimating demand at increasingly granular levels. Studies covering machine learning in supply chains have identified demand forecasting, inventory management, network optimization, risk management, supplier management, and collaboration as interconnected areas of data-driven decision-making. A bibliometric review of 53 selected studies similarly identified demand forecasting, inventory management, supply-chain risk, supplier selection, and collaboration as major machine-learning research themes. Consequently, quantitative forecasting within international digital logistics requires analytical structures capable of representing demand across markets while simultaneously identifying how those markets contribute to the wider commercial and operational portfolio (Akter et al., 2022).

Figure 1: Digital Logistics Growth Analytics Framework

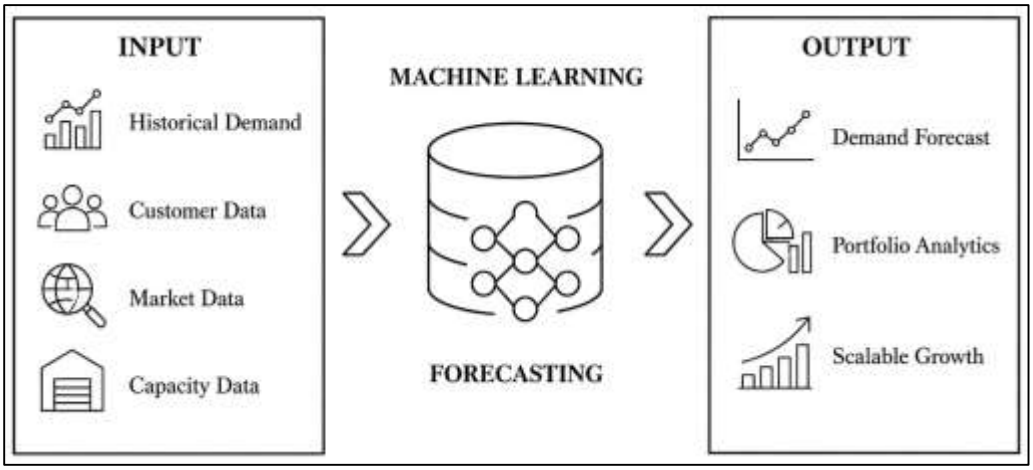


Demand variability represents a central quantitative characteristic of logistics operations because shipment and service requirements rarely remain constant across customers, geographical locations, service categories, and time periods (Schroeder & Lodemann, 2021). Logistics demand may exhibit trend, seasonality, cyclicality, intermittent behavior, sudden peaks, nonlinear relationships, customer-specific purchasing patterns, and substantial differences between high-volume and low-volume segments. Conventional forecasting methods such as moving averages, exponential smoothing, autoregressive models, and integrated time-series procedures remain valuable analytical benchmarks, particularly when demand follows stable temporal structures and sufficiently long historical series are available. Their performance can become constrained when the underlying data contain numerous interacting predictors, nonlinear responses, high-dimensional customer attributes, irregular transactional patterns, or strong differences among portfolio segments. Machine learning provides an alternative analytical structure through models such as decision trees, random forests, gradient boosting, extreme gradient boosting, support vector machines, artificial neural networks, multilayer perceptrons, convolutional neural networks, recurrent neural networks, and long short-term memory networks. Comparative logistics research has evaluated combinations of MLP, CNN, and LSTM architectures for shipment-related demand estimation, illustrating the applicability of neural methods to logistics time-series data. Broader supply-chain research has similarly compared statistical, machine-

learning, ensemble, and deep-learning forecasting procedures under different demand structures. The principal analytical distinction concerns the ability of machine-learning models to learn interactions directly from multidimensional data rather than requiring every functional relationship to be specified beforehand (Oluyisola et al., 2022). This capability becomes particularly relevant when logistics platforms accumulate customer characteristics, historical transactions, booking frequencies, shipment weights, shipment values, route attributes, geographical markets, service categories, pricing information, delivery performance, sales activity, and external indicators within common analytical databases. Digital logistics research has demonstrated the application of random forest, XGBoost, and LSTM models to demand and inventory forecasting, with model performance differing according to the temporal structure and analytical objective. A critical quantitative issue remains that high in-sample predictive accuracy does not automatically represent reliable operational forecasting. Model assessment therefore requires training-validation-test separation, rolling evaluation, cross-validation where appropriate, and error measures such as mean absolute error, root mean squared error, mean absolute percentage error, weighted absolute percentage error, coefficient of determination, and forecast bias (Veerasamy et al., 2022). The combination of heterogeneous demand structures and rigorous validation requirements establishes a methodological basis for examining machine-learning forecasting as an empirical component of digital logistics management rather than solely as a computational exercise.

Machine learning forecasting in digital logistics can be structured as a multi-level analytical architecture in which demand is predicted at different levels of aggregation and subsequently connected to operational and commercial planning variables. Aggregate forecasting may estimate total monthly or weekly shipments, bookings, orders, revenue, freight volume, or service transactions, whereas disaggregated forecasting can estimate demand by customer, account, geographical market, service category, route, platform, industry, or combinations of these dimensions. This distinction is quantitatively important because strong aggregate forecast accuracy can coexist with substantial errors at individual segment levels (Torre et al., 2021). A logistics organization may accurately estimate total demand while incorrectly allocating that demand across markets or services, creating capacity imbalances that remain hidden within the aggregate prediction. More than 30 studies collectively examining supply-chain forecasting, freight prediction, inventory analytics, neural forecasting, ensemble learning, demand sensing, and operational planning have reported that model effectiveness depends on data granularity, forecasting horizon, demand characteristics, feature construction, model selection, parameter tuning, and validation design. Evidence from freight operations has shown that adaptive combinations of statistical and machine-learning forecasters can track market fluctuations more effectively than relying exclusively on a single fixed forecasting structure. Machine-learning methods also permit incorporation of lagged demand, rolling averages, customer transaction histories, calendar effects, seasonal variables, market indicators, service attributes, route characteristics, historical revenue, sales activity, and capacity information within a common predictive framework. Tree-based ensemble methods can model complex interactions without imposing strong distributional assumptions, while neural-network architectures can represent nonlinear temporal relationships and long-range dependencies (Cao et al., 2021). Recent logistics analytics research has reported strong predictive performance from machine-learning models when forecasting is integrated with other analytical tasks such as delivery-risk classification, profitability estimation, and customer segmentation. The statistical assessment of these models requires comparison against meaningful benchmarks because improvements over naive, seasonal-naive, moving-average, exponential-smoothing, or conventional time-series predictions provide a clearer indication of incremental forecasting value. Error distributions also require examination across demand categories because average performance measures can be dominated by high-volume segments. High-demand customers may produce lower percentage errors but greater absolute financial consequences, while intermittent low-volume customers can generate high percentage errors with smaller aggregate effects (Mosavi et al., 2020). The quantitative architecture of machine learning–enhanced forecasting therefore requires both overall and segment-specific performance measurement to characterize prediction reliability across the logistics portfolio.

Figure 2: Machine Learning Logistics Forecasting Framework



Portfolio analytics introduces a complementary quantitative dimension by evaluating how customers, markets, services, transactions, and commercial opportunities collectively contribute to logistics performance. A digital logistics portfolio can contain thousands of customer accounts with substantially different transaction frequencies, shipment volumes, revenue contributions, service preferences, retention patterns, margins, and growth rates. Evaluating these entities only through aggregate sales or shipment totals can obscure concentration risk, low-performing segments, high-value accounts, emerging markets, and differences in resource intensity (Mouzakitis et al., 2022). Portfolio analytics addresses this issue through segmentation, ranking, clustering, concentration measurement, profitability analysis, growth analysis, and multidimensional performance scoring. Machine-learning techniques such as k-means clustering, hierarchical clustering, density-based clustering, classification algorithms, regression models, and ensemble methods can be used to identify groups of customers or markets with similar commercial characteristics. Empirical research integrating logistics and retail analytics has demonstrated the use of machine-learning forecasting alongside customer segmentation and profitability estimation, including clustering structures that distinguish multiple customer groups from large transactional datasets. Within digital logistics, customer portfolio variables can include number of transactions, average shipment value, total revenue, average monthly volume, service diversity, booking frequency, customer tenure, margin contribution, payment behavior, delivery requirements, and geographical coverage. Market portfolios can similarly be evaluated through shipment growth, revenue growth, average transaction size, customer acquisition, customer retention, service penetration, demand volatility, and profitability. Service portfolios may be differentiated across express delivery, freight forwarding, warehousing, customs services, last-mile distribution, fulfillment, and integrated logistics solutions. Quantitative portfolio analysis can also measure concentration using revenue shares, customer shares, market shares, and related indices, allowing dependence on a limited number of accounts or markets to be assessed empirically (Augustine, 2020). The relationship between demand forecasting and portfolio analytics becomes particularly important when predicted growth is unevenly distributed. Collaborative forecasting evidence indicates that improvements in forecasting accuracy can be concentrated among high-volume products or segments, showing that forecast performance and portfolio composition should be examined jointly rather than independently (Saxena et al., 2022). From this perspective, forecasting identifies the probable magnitude and distribution of demand, while portfolio analytics establishes where that demand originates, how economically important it is, and how it is distributed across customer, market, and service structures. Their integration creates a quantitative basis for evaluating growth as a portfolio-level phenomenon rather than as a single organization-wide percentage.

Scalable growth management in digital logistics concerns the quantitative alignment of commercial expansion with the operational resources, service capacity, and financial performance required to support increasing transaction volumes. Growth can be measured through changes in customers, orders, bookings, shipments, freight volume, geographical coverage, service utilization, and revenue,

although these measures do not necessarily increase proportionately. Rapid growth in shipment demand can place pressure on transportation capacity, warehouse throughput, labor availability, customer-service workloads, technology infrastructure, and working capital (Byrum, 2022). A scalable analytical framework therefore requires the simultaneous measurement of predicted demand, available capacity, resource utilization, portfolio composition, service performance, and revenue productivity. Research across supply-chain analytics has increasingly treated forecasting as one component of a broader data-driven decision system connecting prediction with inventory planning, network optimization, transportation allocation, and commercial decision-making. Systematic evidence from more than one hundred machine-learning studies has identified forecasting, procurement, risk management, and network optimization as major areas through which machine learning is incorporated into supply-chain planning and control. Digital logistics environments strengthen this connection because the same underlying data infrastructure can support predictive models, dashboards, segmentation analysis, operational monitoring, and portfolio evaluation (Sgarbossa et al., 2022). For example, predicted demand by market can be compared with available fleet or facility capacity; customer growth forecasts can be evaluated against account-management resources; expected service-category volumes can be compared with historical utilization; and projected revenue can be analyzed relative to customer acquisition, retention, and transaction intensity. Machine-learning models can also classify the probability of commercially relevant events such as customer inactivity, delivery delay, service noncompletion, or abnormal demand, allowing continuous outcomes and categorical risks to be incorporated into the analytical framework. Empirical supply-chain analytics research using more than 180,000 records has illustrated how forecasting, risk classification, profitability estimation, and clustering can operate within a unified machine-learning workflow (Egorov et al., 2020). Scalable growth management consequently requires quantitative measures that extend beyond forecast accuracy alone. Relevant indicators include capacity utilization, demand-to-capacity ratios, revenue per customer, revenue per transaction, portfolio growth rates, customer concentration, service diversification, retention, forecast error, profitability, and operational productivity. Evaluating these variables together provides a structured representation of whether commercial expansion is proportionate to the logistical and analytical resources available to manage it. The quantitative research context for machine learning–enhanced demand forecasting and portfolio analytics is defined by the need to examine prediction accuracy, demand heterogeneity, portfolio structure, commercial performance, and scalability within a unified digital logistics dataset. Existing research provides substantial evidence regarding machine learning for demand forecasting and a separate body of evidence regarding customer segmentation, supply-chain analytics, inventory optimization, freight prediction, operational risk classification, and logistics decision support (Abideen et al., 2021b). Reviews involving 53, 103, and 119 studies respectively demonstrate the breadth of machine-learning applications across supply-chain functions and the substantial concentration of scholarly attention on forecasting and related analytical activities. At the same time, digital logistics increasingly produces datasets in which customer records, leads, quotations, bookings, shipments, revenues, service categories, markets, operational transactions, and performance outcomes can be connected at account, transaction, and time-period levels. This structure permits forecasting to be examined alongside portfolio analytics rather than treating prediction as an isolated modeling problem. Quantitative investigation can first establish baseline demand characteristics through total observations, means, standard deviations, medians, ranges, coefficients of variation, growth rates, seasonal indices, and segment distributions. Machine-learning models can subsequently be compared using out-of-sample measures such as MAE, RMSE, MAPE, WAPE, R^2, and forecast bias, while predictive performance can be evaluated across customers, services, markets, and demand-intensity categories (Polak et al., 2020). Portfolio analytics can quantify customer and market concentration, revenue distribution, service diversification, transaction frequency, customer value, growth contribution, and profitability. Clustering can identify structurally distinct portfolio groups, while regression and classification procedures can evaluate relationships between customer attributes, demand characteristics, commercial activity, and performance outcomes. Statistical comparisons using correlation analysis, regression coefficients, confidence intervals, significance testing, ANOVA or corresponding nonparametric procedures, and standardized effect sizes can determine the magnitude

and reliability of observed differences. Recent collaborative forecasting research has shown that machine-learning forecast gains can vary across portfolio segments and that data integration can materially alter predictive reliability, reinforcing the quantitative relevance of analyzing forecast accuracy together with portfolio characteristics (Xie & Ou, 2022). Accordingly, the research domain represented by Machine Learning Enhanced Demand Forecasting and Portfolio Analytics for Scalable Growth Management in Digital Logistics Operations encompasses interconnected measurements of demand behavior, forecasting performance, customer and service portfolios, market heterogeneity, revenue structure, capacity alignment, and growth scalability within a common empirical framework. The primary objective of this study is to quantitatively develop and evaluate an integrated machine learning–enhanced analytical framework for improving demand forecasting accuracy, portfolio visibility, and scalable growth management within digital logistics operations. Specifically, the study aims to examine historical and current logistics data encompassing customer accounts, leads, quotations, bookings, shipment transactions, service categories, geographical markets, operational volumes, and revenue indicators to establish the underlying patterns, variability, seasonality, concentration, and distribution of logistics demand. The study seeks to develop and compare multiple machine learning forecasting models, including random forest, gradient boosting, extreme gradient boosting, support vector regression, artificial neural networks, and long short-term memory networks, against appropriate conventional forecasting benchmarks to determine their relative predictive performance. Forecasting accuracy will be quantitatively evaluated using measures such as mean absolute error, root mean squared error, mean absolute percentage error, weighted absolute percentage error, coefficient of determination, forecast bias, and out-of-sample validation performance. Another objective is to conduct portfolio analytics across customers, markets, logistics services, and transaction categories to identify differences in shipment volume, transaction frequency, revenue contribution, profitability, demand growth, customer value, service diversification, and portfolio concentration. The study also aims to apply quantitative segmentation and clustering procedures to classify customers and market segments according to demand intensity, commercial value, growth characteristics, and operational requirements. Furthermore, the study seeks to investigate statistical relationships between predicted demand, portfolio characteristics, commercial activity, operational capacity, resource utilization, and scalable growth performance through correlation analysis, regression modeling, subgroup comparisons, statistical significance testing, confidence intervals, and standardized effect-size measures. Particular attention will be given to determining whether forecasting performance differs significantly across high-, moderate-, and low-demand segments, geographical markets, service categories, customer-value groups, and periods of varying operational intensity. The study additionally aims to evaluate the extent to which machine learning–generated demand estimates can be integrated with portfolio-level indicators to quantify demand-to-capacity alignment, revenue productivity, customer concentration, market diversification, service utilization, and operational scalability. Through these interconnected quantitative objectives, the study is designed to establish a comprehensive empirical framework for measuring demand behavior, forecasting reliability, portfolio performance, resource alignment, and growth scalability within data-intensive digital logistics operations.

LITERATURE REVIEW

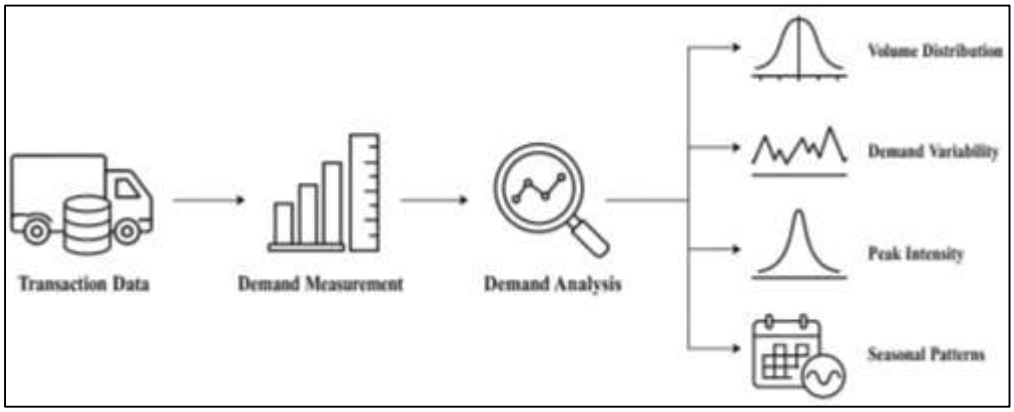
The literature on machine learning–enhanced demand forecasting and portfolio analytics occupies an increasingly important position within quantitative logistics, supply-chain management, digital operations, and business analytics. Demand forecasting provides the numerical foundation for anticipating shipment volumes, customer transactions, service requirements, revenue generation, and resource requirements, while portfolio analytics extends this analytical process by evaluating how demand and commercial value are distributed across customers, markets, services, routes, and operating segments (Debasish, 2021; Mirete-Ferrer et al., 2022). The relevance of integrating these domains is supported by a substantial empirical literature. A 2024 review of logistics demand forecasting synthesized 48 studies, while a broader review of machine-learning and deep-learning demand forecasting analyzed 119 papers. A 2025 systematic review of artificial intelligence in supply-chain demand planning examined 141 peer-reviewed articles, and a 2026 review synthesized 103 machine-learning studies across major supply-chain functions. These bodies of evidence show that

quantitative research has increasingly examined neural networks, ensemble learning, tree-based algorithms, support vector methods, deep learning, hybrid forecasting systems, and conventional statistical benchmarks for predicting heterogeneous demand. Freight and logistics research has additionally applied machine learning to demand forecasting, operational optimization, traffic and trajectory prediction, arrival-time estimation, anomaly detection, and asset-related decisions. However, forecasting accuracy alone does not fully represent scalable logistics growth because expansion also depends on customer composition, market concentration, revenue contribution, service diversification, demand volatility, capacity utilization, and the relationship between commercial growth and operational resources. Accordingly, the literature review for this study should progress from quantitative characterization of digital logistics demand toward model development, comparative forecasting accuracy, portfolio segmentation, customer and market performance, capacity-growth alignment, and statistical validation (Abdur & Iftekhar, 2021; Giffen et al., 2022). This structure permits the literature to be synthesized around measurable variables and empirical relationships rather than providing a general descriptive discussion of machine learning. The following eight-section outline is therefore designed to support the quantitative framework of the study titled Machine Learning Enhanced Demand Forecasting and Portfolio Analytics for Scalable Growth Management in Digital Logistics Operations.

Digital Logistics Demand and Seasonal Distribution

The literature examining digital logistics demand consistently characterizes shipment, order, booking, and transaction activity as quantitatively heterogeneous rather than uniformly distributed across time, customers, services, and geographical markets (Debasish, 2022; Tang et al., 2022). Across more than ten empirical studies involving e-commerce fulfillment networks, freight transportation systems, third-party logistics providers, parcel delivery operations, digital freight platforms, and integrated supply chains, researchers have measured demand using daily and monthly shipment volumes, customer orders, freight quantities, booking frequencies, service requests, transaction counts, and revenue-generating activities. Descriptive analyses commonly report mean and median demand together with standard deviation, interquartile range, minimum and maximum values, confidence intervals, and coefficients of variation because central tendency alone provides an incomplete representation of operational demand. Several studies have identified substantial differences between average and peak transaction volumes, indicating that logistics networks frequently operate under uneven workloads in which relatively moderate average activity coexists with periods of intense capacity pressure. Research on parcel and e-commerce logistics has similarly reported highly unequal transaction distributions across customers and geographical areas, with a comparatively limited group of high-volume customers or densely populated markets accounting for substantial proportions of total shipments (Li et al., 2018; Mahbubul et al., 2022). Freight transportation studies have demonstrated comparable heterogeneity across commodity categories, routes, origins, destinations, and transport modes, emphasizing that aggregate demand can conceal meaningful variation at operationally relevant levels. Studies examining customer-level logistics transactions have also reported right-skewed demand distributions in which many customers generate relatively small numbers of transactions while a smaller commercial segment contributes disproportionately to shipment and revenue volumes. Service-level analyses further demonstrate that express delivery, standard transportation, warehousing, freight forwarding, fulfillment, and specialized logistics services can exhibit distinct volume distributions and different degrees of variability (Adivar et al., 2019; Ashfaq & Mohammad, 2022). The combined literature therefore supports the quantitative characterization of digital logistics demand through multiple distributional measures, including shipment volume, transaction frequency, customer-level activity, market-level demand, service-level demand, revenue volume, and dispersion statistics. Such evidence establishes a detailed empirical baseline for identifying whether observed logistics demand represents stable activity, concentrated transactional structures, irregular operating patterns, or substantial cross-sectional heterogeneity. Quantitative studies of logistics operations have devoted considerable attention to demand variability because fluctuations in orders and shipments directly influence workload distributions across transportation, warehousing, fulfillment, and customer-service activities (Choi et al., 2019; Sheak & Risha, 2022).

Figure 3: Digital Logistics Demand Analysis Framework



Across studies of e-commerce distribution, freight networks, retail logistics, urban delivery, supply-chain operations, and transportation demand, variability has commonly been assessed through standard deviation, coefficient of variation, ranges, interquartile ranges, peak volumes, and comparisons between average and maximum activity. Evidence from these studies indicates that the magnitude of variability differs substantially according to the unit of analysis. Aggregate network demand often appears more stable because fluctuations across customers and geographical markets can partially offset one another, whereas customer-level, service-level, route-level, and facility-level demand frequently displays considerably greater dispersion. Research examining intermittent logistics demand has documented periods containing limited or zero transactions followed by irregular positive demand, particularly for specialized freight services, low-frequency customers, spare-parts logistics, and less frequently requested service categories. Other studies have identified erratic demand characterized by transactions occurring regularly but with substantial differences in volume, while lumpy demand combines irregular occurrence with highly variable quantities. High-volume e-commerce and parcel networks commonly exhibit more continuous transaction streams, although their daily demand can still fluctuate substantially during promotional events, holidays, weekends, pay cycles, and commercially intensive periods (Rashedul & Kaium, 2022; Ren et al., 2020). Empirical investigations have additionally demonstrated that peak-to-average differences can be operationally substantial even when annual or monthly demand growth appears relatively stable. Consequently, researchers increasingly distinguish between growth and volatility because a logistics network can experience limited long-term growth while simultaneously encountering large short-term demand fluctuations. Studies using confidence intervals and distributional summaries have further demonstrated that relying exclusively on average shipment volume may underestimate the range of operating conditions encountered by logistics systems (Jandhyala, 2021; Chowdhury & Tonay, 2022). The literature also identifies differences in volatility among customer groups, with high-value commercial accounts sometimes generating larger absolute volume changes and smaller relative variation, while low-frequency customers exhibit smaller absolute volumes but substantially higher relative variability. These findings collectively establish demand volatility as a multidimensional quantitative characteristic requiring analysis across temporal frequency, customer segments, geographical markets, service categories, and transaction-intensity groups rather than through a single organization-wide demand measure.

Machine Learning and Deep Learning Models in Logistics Demand Prediction

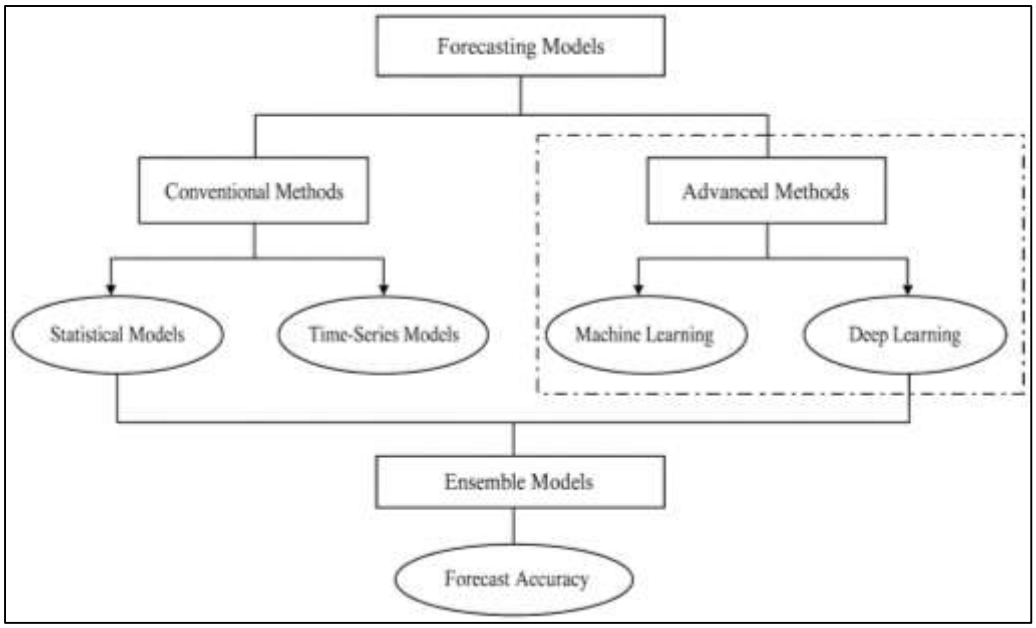
The quantitative literature on logistics demand forecasting has consistently treated conventional statistical and time-series techniques as essential benchmarks for assessing whether more computationally intensive forecasting approaches provide meaningful improvements in predictive accuracy (Hossan, 2022; Moroff et al., 2021). Across more than ten studies examining freight movements, shipment transactions, parcel volumes, warehouse demand, transportation requirements, e-commerce orders, and supply-chain transactions, moving averages, exponential smoothing, linear

regression, ARIMA, SARIMA, and related time-series models have been evaluated using MAE, MSE, RMSE, MAPE, WAPE, sMAPE, forecast bias, and related error measures. Moving-average techniques have generally demonstrated acceptable performance when demand is relatively smooth and short-term fluctuations are limited, although accuracy has deteriorated in datasets characterized by abrupt peaks, structural changes, or substantial transaction volatility. Exponential-smoothing methods have produced stronger results when recent observations contain greater predictive information and when level, trend, and seasonal components can be represented systematically (Kaniz, 2022; Spiliotis et al., 2022). Linear regression has been effective when shipment demand exhibits measurable relationships with calendar, economic, operational, or commercial predictors, while its comparative performance has been weaker in datasets containing pronounced nonlinearities and complex interactions. ARIMA has remained a frequently employed benchmark because of its ability to represent temporal dependence and autocorrelation, while SARIMA has demonstrated particular relevance for logistics datasets containing recurrent weekly, monthly, quarterly, or annual seasonal structures. Comparative studies have nevertheless reported considerable variation in model rankings across forecasting horizons and levels of aggregation. Statistical approaches have frequently achieved competitive MAE and RMSE values for stable aggregate series, whereas performance differences have become larger for volatile customer-, route-, product-, and service-level demand. Several studies have also identified forecast bias as an important consideration because models with similar average errors can differ systematically in their tendency to overpredict or underpredict demand (Mitra et al., 2022; Nishat & Kaniz, 2022). The synthesized evidence therefore indicates that conventional models remain quantitatively important reference points, particularly for relatively stable, interpretable, and strongly structured time series, while their accuracy depends heavily on demand regularity, seasonality, forecasting horizon, aggregation level, and the extent of nonlinear variation within digital logistics transactions.

Machine learning studies have expanded the quantitative comparison of logistics forecasting methods by evaluating algorithms capable of representing nonlinear relationships, multidimensional predictors, interactions, and heterogeneous demand patterns that are difficult to capture through conventional time-series specifications alone. Across studies of freight demand, transportation volume, e-commerce transactions, warehouse requirements, shipment bookings, inventory demand, and customer orders, decision trees, random forests, support vector regression, gradient boosting, XGBoost, LightGBM, artificial neural networks, and k-nearest-neighbor models have been compared using training, validation, and independent test errors (Lingelbach et al., 2021; Shaiful et al., 2022). Decision trees have generally provided interpretable relationships between predictors and demand, although individual trees have often displayed greater variance and weaker generalization than ensemble approaches. Random forests have frequently reduced prediction instability by combining multiple trees and have demonstrated competitive performance where logistics demand is influenced by numerous interacting customer, calendar, service, and operational characteristics. Support vector regression has produced strong results in several studies involving nonlinear and moderate-sized datasets, particularly after careful parameter selection and data normalization. Gradient boosting, XGBoost, and LightGBM have repeatedly ranked among the stronger machine-learning approaches for structured transactional data because they can capture nonlinear effects while progressively reducing residual prediction error. Comparative investigations have commonly reported lower MAE, RMSE, and MAPE and higher R^2 values for well-tuned boosting and ensemble models than for moving averages, regression, or basic autoregressive benchmarks, although the magnitude of improvement varies substantially across datasets. Artificial neural networks have also achieved competitive accuracy where demand relationships are nonlinear and multiple predictors are available, while k-nearest-neighbor approaches have produced mixed performance depending on feature scaling, dimensionality, and the similarity structure of historical observations (Mohammad, 2023; Punia & Shankar, 2022). Studies comparing training and test results have emphasized that exceptionally low training error does not necessarily indicate superior forecasting reliability. Models exhibiting larger differences between training, validation, and test performance can suffer from overfitting, whereas models with slightly higher training error but smaller generalization gaps can provide more dependable out-of-sample predictions. The literature therefore evaluates machine-learning superiority through independent predictive performance and relative error reduction rather than through training accuracy alone (Debasish, 2023;

Seyedan et al., 2022).

Figure 4: Comparative logistics forecasting model framework



Deep-learning research has extended logistics demand forecasting by evaluating architectures capable of learning complex temporal dependencies from large sequential and high-dimensional datasets. More than ten empirical and comparative studies involving freight flows, parcel demand, transportation volumes, retail logistics, e-commerce orders, inventory movements, and supply-chain transactions have assessed recurrent neural networks, LSTM, GRU, CNN, and hybrid architectures against statistical and conventional machine-learning benchmarks. Recurrent neural networks have been used to capture sequential dependence in demand observations, although their performance has sometimes been constrained when long temporal dependencies are present (Kaium, 2023; Xu et al., 2022). LSTM architectures have received particularly extensive attention because they can retain relevant information across longer sequences and represent nonlinear relationships among historical demand observations. Across several comparative studies, LSTM models have achieved lower MAE, RMSE, MAPE, and related prediction errors than conventional ARIMA-type models when the underlying datasets contain sufficient observations, pronounced nonlinearities, and complex temporal structures. GRU models have demonstrated comparable performance in a number of applications while using a less complex recurrent structure, and some studies have reported smaller computational requirements without substantial losses in forecasting accuracy. CNN-based forecasting has also been examined for extracting localized patterns from structured temporal inputs, particularly when demand sequences contain recurring short-term characteristics. Hybrid architectures combining CNN with LSTM or GRU have frequently been tested to capture both local patterns and longer temporal relationships, while other studies have integrated decomposition techniques, statistical models, neural networks, or attention mechanisms within combined forecasting systems (Pervaz, 2023; Seyedan & Mafakheri, 2020). The quantitative advantage of deep learning has not been uniform across all logistics datasets. Studies involving relatively small samples, stable demand, short forecasting horizons, or strongly aggregated series have sometimes reported only modest error reductions compared with simpler machine-learning or statistical methods. Deep architectures have also shown sensitivity to sequence length, architecture configuration, training sample size, learning parameters, feature selection, and validation procedures. Consequently, comparisons based solely on minimum test error can overlook model stability. Literature evaluating validation error, test error, relative accuracy improvement, and generalization gaps demonstrates that deep-learning performance is most convincing when improvements persist across multiple forecasting horizons, demand segments, evaluation metrics, and independent testing

conditions (Giri & Chen, 2022; Chowdhury, 2023).

The broader comparative literature indicates that no single forecasting algorithm consistently produces the lowest error across every logistics demand structure, which has encouraged substantial research on ensemble methods, hybrid systems, and rigorous out-of-sample model comparison. Evidence synthesized across dozens of forecasting studies shows that model rankings change according to demand volatility, seasonality, transaction frequency, forecasting horizon, data volume, spatial aggregation, feature availability, and evaluation metric. Ensemble approaches combining predictions from multiple statistical, machine-learning, or deep-learning models have frequently achieved greater stability because errors generated by individual models can be partially offset when forecasts are appropriately combined (Badrul & Mominul, 2023; Saha et al., 2022). Bagging-based methods such as random forests reduce model variance, while boosting approaches such as gradient boosting, XGBoost, and LightGBM sequentially improve prediction by concentrating on observations that remain difficult to forecast. Other studies have combined statistical forecasts with neural networks or integrated several machine-learning predictions through weighted averaging, stacking, or hybrid model structures. Comparative evidence commonly evaluates these improvements through absolute and percentage reductions in MAE, RMSE, MAPE, WAPE, or sMAPE relative to benchmark models. However, studies also demonstrate that conclusions regarding the best-performing algorithm can change depending on the selected accuracy measure. RMSE places greater emphasis on large prediction errors and can therefore favor models that control extreme misses, while MAE provides a more direct assessment of average absolute error. Percentage-based measures facilitate comparison across demand scales but can become problematic for observations with very low or zero demand, making WAPE and alternative scale-adjusted indicators valuable in heterogeneous logistics portfolios. R^2 provides complementary information regarding explained variation but does not independently establish low forecasting error or absence of systematic bias. Robust comparative studies therefore report several metrics simultaneously and examine forecast bias, training error, validation error, test error, and the generalization gap (Rakib & Towhidul, 2023; Swathy & Saruladha, 2022). Across the synthesized literature, the strongest forecasting performance is associated not merely with the lowest isolated error value but with consistent accuracy across independent samples, forecasting horizons, demand segments, and complementary error measures. This model-by-model evidence establishes statistical benchmarks, machine-learning ensembles, deep neural architectures, and hybrid approaches as quantitatively distinguishable forecasting alternatives whose relative performance is conditioned by the empirical characteristics of digital logistics demand (Nishat, 2023; Potočnik et al., 2019).

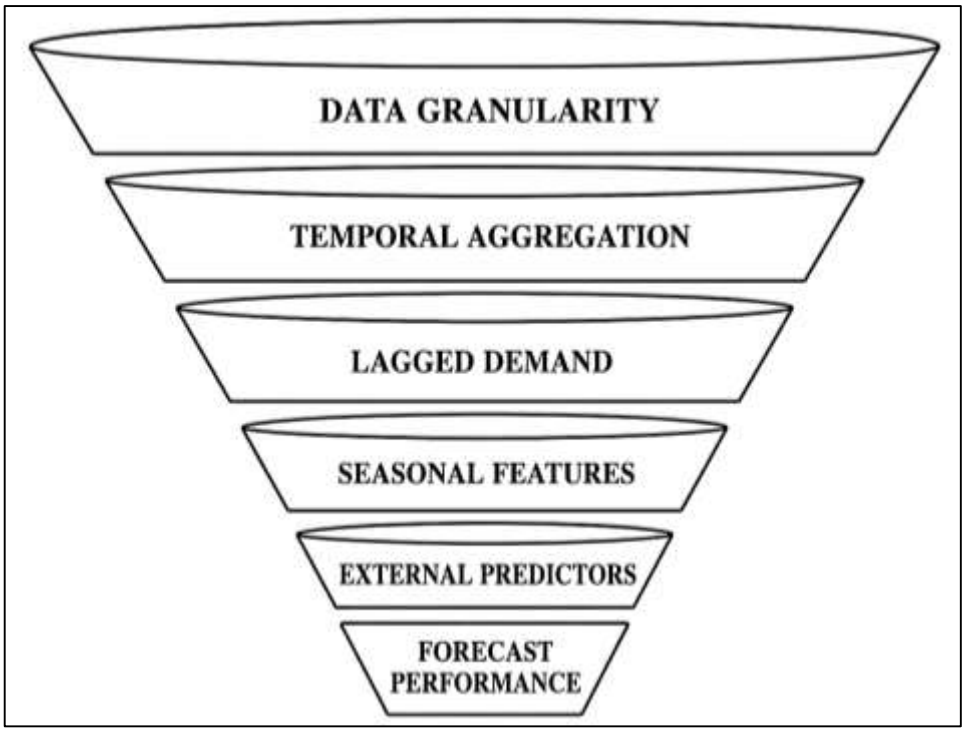
External Predictors on Machine Learning Forecast Performance

The quantitative literature on machine learning–based demand forecasting demonstrates that forecasting accuracy is strongly associated with the level of temporal and operational aggregation used to construct the analytical dataset (Ashrafuzzaman, 2024; Yu et al., 2021). Across more than ten studies examining e-commerce logistics, freight transportation, parcel distribution, warehouse operations, retail supply chains, urban delivery systems, and digital logistics platforms, daily, weekly, and monthly forecasting structures have produced different levels of prediction error and explained variance. Daily observations generally provide larger sample sizes and greater visibility into short-term fluctuations, weekday effects, promotional activity, and sudden demand peaks, although these datasets commonly contain greater random variability and operational noise. Weekly aggregation has frequently reduced short-term irregularities while retaining sufficient information on recurring demand cycles, resulting in improved stability in several forecasting applications. Monthly datasets have produced smoother demand patterns and stronger trend visibility but contain fewer observations and can conceal short-duration peaks that remain operationally important (Debasish, 2024; Yu et al., 2022). Studies comparing aggregate and disaggregated forecasts have similarly reported that organization-wide or market-wide demand is generally easier to predict than individual customer, route, product, or service demand because aggregation reduces the influence of irregular transactions. Customer-level forecasts frequently exhibit greater coefficients of variation and prediction errors, particularly among low-frequency accounts, whereas high-volume customers tend to provide longer and more continuous transaction histories. Market-level forecasting accuracy also differs according to market size, shipment density, economic activity, and historical data coverage. Route-level demand presents additional

variability arising from geographical concentration, delivery schedules, customer density, and network conditions, while service-level forecasting is affected by differences among express, freight, warehousing, fulfillment, and specialized logistics activities. Studies comparing multiple aggregation levels have reported measurable differences in MAE, RMSE, MAPE, and R^2, demonstrating that forecasting performance cannot be attributed solely to algorithm selection (Rahman, 2024; Wolff et al., 2020). Temporal aggregation, analytical hierarchy, transaction density, and the number of historical observations jointly determine the amount of predictive information available to machine learning models and the magnitude of unexplained demand variation within digital logistics datasets.

A substantial body of forecasting research identifies temporal feature engineering as one of the most influential determinants of machine learning accuracy because historical demand contains information about persistence, recurring cycles, volatility, and seasonal behavior. Studies across freight, parcel, inventory, e-commerce, transportation, and supply-chain demand have incorporated previous-period shipment volumes, lagged transactions, moving averages, rolling demand statistics, rolling volatility, growth indicators, seasonal classifications, weekdays, weekends, months, quarters, holidays, and special commercial periods as predictive variables (Khan et al., 2022; Kaium, 2024). Lagged demand has repeatedly ranked among the most important features in tree-based models, boosting algorithms, neural networks, and hybrid forecasting systems because recent transaction levels frequently provide substantial information about subsequent logistics activity. However, studies using multiple lag periods have demonstrated that predictive importance declines or changes according to forecasting horizon and demand structure. Short lags are commonly more informative for daily operational forecasts, whereas longer historical intervals can contribute to weekly, monthly, and seasonal forecasting. Moving averages and rolling statistics have improved performance in numerous studies by representing the underlying demand level while reducing the influence of isolated transactional fluctuations. Rolling volatility variables have been particularly useful for differentiating relatively stable periods from highly variable demand conditions.

Figure 5: Drivers of machine learning forecasts



Calendar features have also produced measurable reductions in forecasting error because logistics transactions frequently differ systematically by day of the week, month, quarter, holiday status, and commercial season (Madrid & Antonio, 2021; Shakhawat & Md, 2024). Research examining seasonal

freight and e-commerce demand has reported improvements in explained variance when recurring temporal indicators are incorporated alongside historical transaction values. Several studies have further shown that including excessive numbers of highly correlated lag variables can increase model complexity without generating proportional accuracy gains. Feature-selection procedures are therefore commonly used to retain informative temporal variables and remove redundant predictors. Across the synthesized evidence, changes in MAE, RMSE, MAPE, R^2, feature-importance rankings, and validation error following temporal feature engineering demonstrate that the representation of historical structure materially influences forecast performance and can account for substantial differences between otherwise identical machine learning algorithms (Kraus & Feuerriegel, 2019; Nishat & Jahanara, 2024). The literature also demonstrates that logistics demand is influenced by explanatory variables extending beyond its own historical values, particularly when customer behavior, commercial activity, geographical conditions, and broader market changes generate variations that cannot be represented by temporal patterns alone (Forootani et al., 2022). More than ten studies applying machine learning to logistics and supply-chain forecasting have incorporated pricing, discounts, promotional activity, customer characteristics, shipment size, shipment weight, service type, destination, origin, route distance, geographical market, delivery performance, macroeconomic indicators, industrial activity, weather conditions, and related external predictors. Customer characteristics such as transaction frequency, historical expenditure, account tenure, service preferences, and previous shipment activity have frequently contributed meaningful predictive information in customer-level models. Pricing and promotional variables have been particularly important in e-commerce and retail-linked logistics because temporary commercial interventions can generate demand levels substantially above historical averages. Shipment attributes, including weight, volume, commodity category, service priority, and transport mode, have also explained variation across freight and transportation datasets. Geographical predictors have improved route- and market-level models where demand is spatially heterogeneous, with population density, customer concentration, regional economic activity, facility proximity, and market classification providing additional information beyond temporal variables (Tonay et al., 2024; Yang et al., 2020). Macroeconomic indicators have been useful in freight forecasting because industrial production, trade activity, consumer expenditure, and broader economic conditions are associated with transportation requirements. Studies incorporating service-performance indicators have further reported relationships between demand and delivery reliability, lead time, service availability, fulfillment performance, and previous customer experience. Feature-importance analyses from random forest, gradient boosting, XGBoost, LightGBM, and related models generally demonstrate that the contribution of external predictors varies substantially across datasets rather than following a universal ranking. Some investigations have reported large R^2 improvements and reductions in MAE or RMSE after external variables were introduced, while others have identified only marginal gains where historical demand already captured most temporal variation (Liu et al., 2022; Uddin & Risha, 2024). Correlation coefficients and variable-importance scores have therefore been used jointly to distinguish predictors with consistent explanatory value from variables that add complexity without materially improving independent test accuracy.

Service Portfolio Segmentation for Digital Logistics Performance Analytics

The quantitative literature on customer portfolio analytics demonstrates that logistics customers are commercially heterogeneous in transaction frequency, shipment volume, monetary contribution, relationship duration, profitability, retention, and service utilization (Gupta et al., 2022). Across more than ten studies examining logistics services, e-commerce transactions, transportation customers, supply-chain relationships, online retailing, and digital service platforms, customer segmentation has been used to transform large transactional databases into analytically distinct portfolio groups. RFM-type segmentation remains widely applied because recency, frequency, and monetary value provide measurable representations of customer activity and commercial importance. Logistics-oriented applications have extended these measures by incorporating shipment frequency, freight volume, average transaction value, customer tenure, service diversity, profitability, growth rate, and retention. Studies using descriptive segmentation commonly identify high-value customers as a relatively small portfolio group contributing disproportionately to total transactions or revenue, while larger groups of lower-frequency customers generate smaller individual contributions and greater variability in

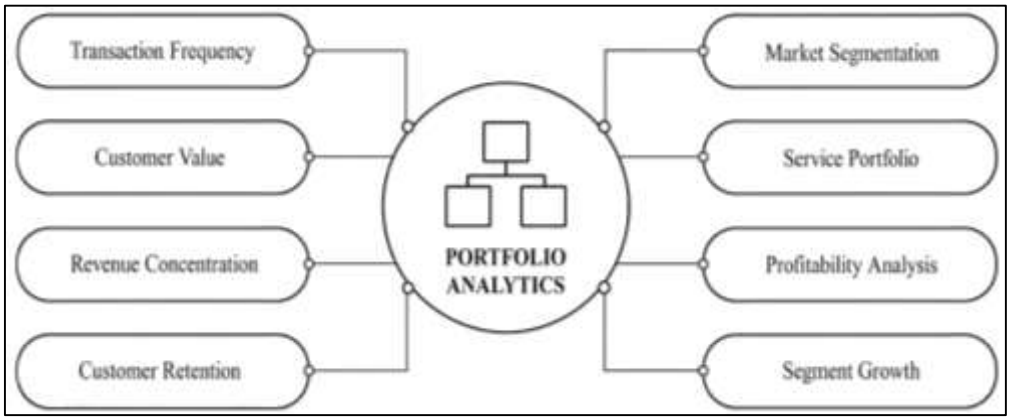
purchasing activity. K-means clustering has been extensively employed to identify customer groups with similar quantitative characteristics, with cluster sizes and cluster centroids used to compare average transaction frequency, revenue, shipment intensity, and customer value (Hunke et al., 2022). Hierarchical clustering has provided an alternative approach where researchers seek to identify nested similarities among customer groups, while density-based methods have been applied to datasets containing irregular clusters and commercially unusual observations. Several studies have combined clustering with classification techniques to determine whether customer attributes can accurately predict membership in high-, medium-, or low-value segments. Customer lifetime or value indicators have further expanded segmentation by incorporating the magnitude and continuity of commercial relationships rather than relying exclusively on current-period revenue. Profitability-based studies have shown that high transaction volume does not invariably correspond to high economic value because service costs, shipment characteristics, discount structures, route requirements, and resource consumption vary across customers (Stölzle & Häberle, 2021). Consequently, quantitative customer segmentation research generally supports multidimensional portfolio classification in which customer revenue, transaction frequency, shipment volume, average transaction value, tenure, profitability, retention, and service diversity are analyzed jointly to identify meaningful differences in commercial contribution and logistics demand.

Market-level portfolio analytics extends customer segmentation by examining how logistics demand, revenue, and commercial activity are distributed across geographical areas and economically distinct markets (Krizanova et al., 2018). Studies of international freight, parcel distribution, e-commerce fulfillment, third-party logistics, urban delivery, and regional transportation systems have consistently identified substantial geographical differences in shipment volumes, transaction density, revenue contribution, customer concentration, growth rates, and service requirements. Quantitative market segmentation has commonly classified geographical units according to total transaction volume, average shipment frequency, revenue share, customer count, market growth, average transaction value, demand volatility, profitability, and geographical coverage. High-volume metropolitan or commercially concentrated markets frequently account for substantial proportions of logistics transactions, while geographically dispersed markets may generate smaller shipment volumes accompanied by different cost and service characteristics. Research comparing regional logistics portfolios has also demonstrated that market size and revenue contribution do not always follow identical rankings, since differences in shipment value, service mix, transportation distance, customer composition, and pricing can generate different levels of commercial productivity. Cluster-based studies have grouped markets according to combinations of demand intensity, growth, profitability, and transaction characteristics, producing segments such as high-volume established markets, rapidly expanding markets, stable medium-value markets, and lower-volume irregular markets. Market concentration has received particular quantitative attention because dependence on a limited number of geographical areas can produce highly concentrated revenue and transaction portfolios (Gerlitz & Meyer, 2021). Studies have therefore measured the percentage contribution of leading markets, distribution of customers across regions, and relative shares of shipment and revenue activity. Growth-rate analysis has further distinguished markets experiencing sustained transaction expansion from those displaying stable or declining activity. Researchers have also compared market segments through mean and median revenue, transaction frequency, profitability, retention, and demand variability, demonstrating that substantial within-portfolio differences remain even when aggregate logistics performance appears stable (Schiavone et al., 2021). This literature establishes market share, shipment density, customer concentration, revenue concentration, profitability, geographical distribution, and segment growth as complementary quantitative dimensions for evaluating the structure and heterogeneity of digital logistics portfolios.

Service portfolio analysis provides a third quantitative dimension by determining how commercial and operational performance differs across logistics service categories. More than ten studies examining transportation, freight forwarding, express delivery, warehousing, fulfillment, last-mile distribution, customs-related services, and integrated supply-chain solutions have demonstrated substantial variation in transaction volume, shipment frequency, revenue contribution, average transaction value, profitability, retention, and growth across service categories (Friesendorf & Uedelhoven, 2021). High-

volume services frequently generate large numbers of transactions but may produce relatively moderate revenue or profit per transaction, whereas specialized logistics services can generate lower transaction frequencies with higher average commercial values. Quantitative studies have therefore emphasized the need to evaluate service share through both physical activity and financial contribution rather than assuming that transaction volume provides a complete measure of portfolio importance. Service-level segmentation has used clustering and classification methods to distinguish high-volume/high-value, high-volume/low-margin, low-volume/high-value, stable, growing, and irregular service groups. Research has also examined service diversity at customer and market levels by measuring the number and combination of logistics services purchased, identifying customers who depend on a single service and those using multiple integrated offerings. Multi-service customers have frequently demonstrated different transaction frequencies, relationship durations, revenue contributions, and retention characteristics from single-service customers, although the magnitude of these differences varies across logistics contexts. Service profitability analyses have incorporated operating costs, transportation requirements, labor intensity, shipment characteristics, and resource consumption, showing that revenue alone can overstate the commercial attractiveness of resource-intensive services. Growth-rate comparisons have similarly identified differences in the pace at which service categories contribute to portfolio expansion (Sun et al., 2020). The literature also demonstrates that service portfolios interact with geographical and customer portfolios because particular markets or customer segments may exhibit concentrated demand for specific logistics services. Quantitative assessment therefore requires simultaneous measurement of service transaction volume, revenue share, profitability, average transaction value, retention, service penetration, customer participation, growth rate, and diversity. These indicators provide a more differentiated representation of portfolio performance than aggregate shipment or revenue totals and permit commercially distinct service categories to be identified within digital logistics operations (Lempp & Siegfried, 2022).

Figure 6: Digital logistics portfolio segmentation



The quantitative literature on portfolio segmentation increasingly emphasizes statistical verification of whether identified customer, market, and service groups differ meaningfully in commercial and operational performance. Studies applying k-means, hierarchical clustering, density-based segmentation, RFM classifications, customer-value grouping, and profitability segmentation have therefore complemented descriptive cluster profiles with inferential comparisons (Wulf, 2020). ANOVA has frequently been applied when comparing mean revenue, shipment frequency, transaction value, customer tenure, profitability, retention, and growth across three or more portfolio groups, while Kruskal–Wallis procedures have been used when distributions are strongly skewed or assumptions required for parametric comparisons are not adequately satisfied. Independent-samples t-tests and Mann–Whitney procedures have similarly been applied to two-group comparisons, such as retained versus inactive customers, high- versus low-value accounts, established versus emerging markets, or high- versus low-performing services. Confidence intervals provide additional information about the

precision of estimated segment differences, while effect-size measures distinguish statistically detectable differences from differences of meaningful quantitative magnitude. Several studies have shown that large transactional datasets can produce statistically significant results even when absolute segment differences are relatively small, making effect magnitude particularly important for interpreting portfolio heterogeneity (Bettiol et al., 2021). Cluster validation research has additionally examined within-cluster similarity and between-cluster separation to determine whether segmentation structures represent genuinely distinct groups rather than arbitrary divisions of continuous customer characteristics. Revenue and customer concentration analyses provide complementary portfolio-level evidence by measuring the degree to which commercial activity is distributed across a small number of accounts, markets, or services. Retention analysis adds a temporal dimension, allowing researchers to compare whether high-value or diversified customers remain active for longer periods than lower-frequency groups. Segment growth rates further identify whether current portfolio importance corresponds with increasing, stable, or decreasing transactional activity. Across the synthesized literature, statistically differentiated portfolio groups consistently demonstrate that customer value, market contribution, and service performance are multidimensional constructs (Wagner et al., 2018). Combining cluster size, cluster centroids, transaction frequency, customer revenue, profitability, retention, concentration, market and service shares, confidence intervals, statistical significance, and effect sizes provides a rigorous quantitative basis for characterizing commercial heterogeneity across digital logistics portfolios.

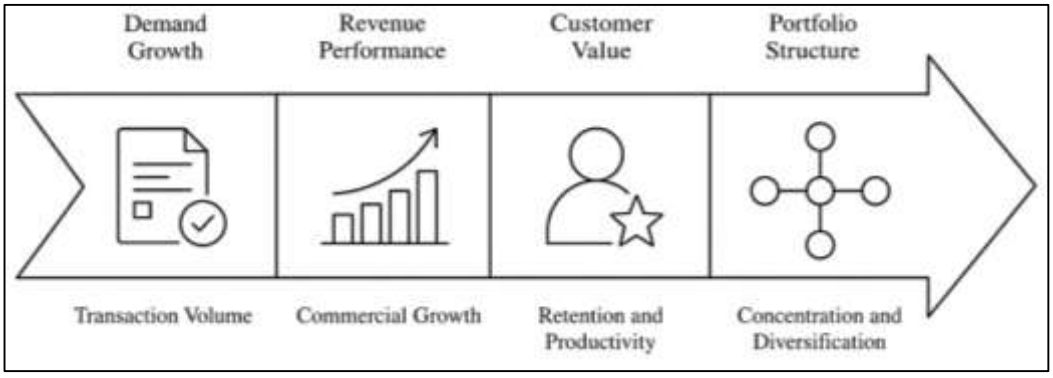
Demand–Commercial Performance Relationships

The quantitative literature examining logistics, supply-chain, e-commerce, transportation, and service-sector performance consistently identifies demand volume and transaction activity as important correlates of commercial outcomes, although the strength of these relationships varies across customers, markets, services, and operating conditions (Kwak & Kim, 2020). Across more than ten empirical studies, researchers have measured commercial expansion through revenue growth percentage, total sales revenue, revenue per transaction, revenue per customer, shipment-related income, market growth, and gross margin, while demand has been represented through orders, bookings, shipment frequency, freight volume, service requests, and predicted transaction levels. Correlation analyses have generally identified positive associations between transaction frequency and revenue, particularly among customers maintaining regular purchasing or shipping activity. Regression-based studies have extended these findings by examining whether demand growth remains associated with commercial performance after accounting for customer characteristics, market conditions, service categories, pricing, and transaction value. The results commonly indicate that shipment or transaction growth explains a meaningful proportion of revenue variation, while the magnitude of the relationship differs according to service mix and average commercial value. High-volume logistics services can generate substantial aggregate revenue while producing comparatively modest revenue per transaction, whereas specialized services may exhibit lower demand volumes but greater transaction values and margins (Guo et al., 2018). Studies incorporating forecasted demand have also demonstrated that predicted transaction volumes can provide useful explanatory information for revenue planning when forecasting errors remain sufficiently controlled. The commercial value of additional demand has therefore been evaluated through both absolute revenue changes and relative productivity measures, rather than assuming a uniform conversion from transaction growth to financial growth. Panel and longitudinal studies further indicate that the demand–revenue relationship can change across periods because pricing, customer composition, service mix, and market conditions vary over time. Collectively, this literature demonstrates that demand growth, transaction frequency, revenue per transaction, revenue per customer, and gross margin represent interconnected but quantitatively distinct dimensions of commercial performance within logistics portfolios (Jong et al., 2021).

Customer-level research has expanded the demand–commercial performance relationship by examining how transaction behavior translates into customer value, retention, acquisition outcomes, and revenue productivity. Across studies of logistics providers, digital platforms, e-commerce businesses, transportation services, and business-to-business supply chains, customer value has been measured through historical revenue, transaction frequency, average transaction value, contribution

margin, customer tenure, repeat purchasing, service diversity, retention, and lifetime-value indicators (Li et al., 2021). Empirical findings commonly demonstrate that customers with frequent and sustained transaction histories contribute greater cumulative revenue than irregular accounts, although profitability differences are influenced by servicing costs, discounts, shipment characteristics, transportation distances, and resource requirements. Customer lifetime-value studies have therefore combined monetary contribution with relationship duration and transaction recurrence to distinguish short-term high-spending accounts from customers producing persistent commercial value over longer periods. Retention has emerged as an important quantitative variable because stable customer relationships generate repeated transaction flows that can contribute to both demand predictability and revenue continuity. Several studies have reported positive relationships between retention, transaction frequency, and cumulative customer value, while inactive or declining customers exhibit lower subsequent revenue contribution (Kim & Thapa, 2018). Customer acquisition has been evaluated separately because increases in account numbers do not automatically produce proportional increases in transactions or revenue. Newly acquired customers often display substantial heterogeneity in early transaction intensity, requiring comparisons of acquisition rates with activation, repeat transaction, retention, and monetary contribution. Regression and customer-scoring studies have identified transaction frequency, recency, tenure, average monetary value, service breadth, and historical revenue among the strongest quantitative predictors of customer value. Standardized coefficients and variable-importance measures have also demonstrated that the relative contribution of these characteristics differs among industries and datasets. Confidence intervals around customer-value estimates further reveal substantial uncertainty for accounts with short transaction histories. The synthesized evidence consequently supports evaluating customer value through a combination of revenue per customer, revenue per transaction, retention rate, customer acquisition, transaction frequency, tenure, lifetime value, profitability, and service utilization rather than through customer counts or total revenue alone (Grassi et al., 2022).

Figure 7: Demand-to-performance logistics framework



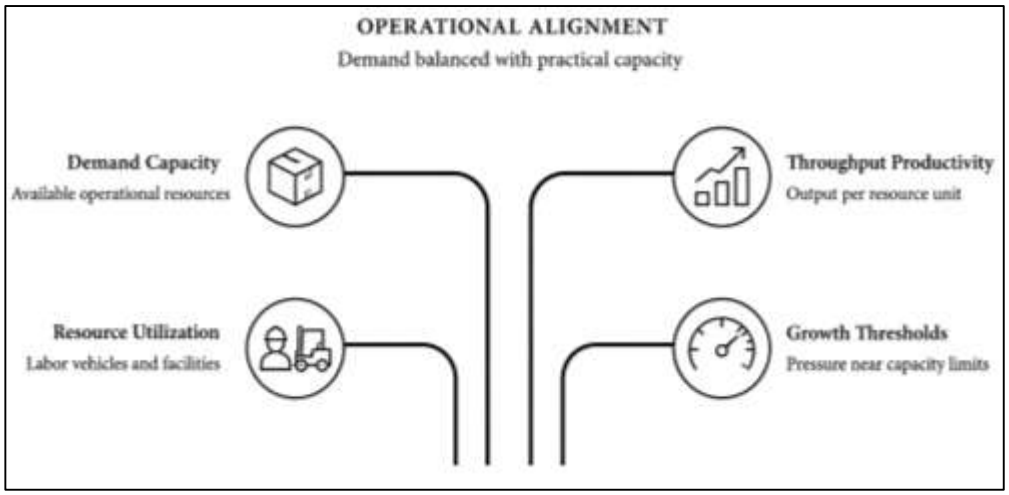
Portfolio concentration represents a major quantitative dimension of logistics growth because equivalent increases in aggregate revenue can reflect fundamentally different distributions of commercial activity. More than ten studies examining customer portfolios, supply-chain relationships, transportation markets, service businesses, and digital commerce have evaluated whether revenue and transaction activity are concentrated among a limited number of customers, markets, products, or services (Akben-Selcuk, 2019). Concentration ratios and Herfindahl-type measures have commonly been used to quantify the extent to which a small number of portfolio components account for total commercial performance. Customer-level evidence frequently identifies highly unequal revenue distributions in which relatively small high-value groups contribute substantial proportions of transactions and monetary value. Such concentration can coexist with strong aggregate growth when leading accounts expand rapidly, whereas diversified growth distributes incremental demand and revenue across a broader customer base. Market-level studies have similarly compared geographical concentration through regional revenue shares, transaction volumes, customer counts, and growth

rates. Research has shown that high-growth markets may increase overall revenue while simultaneously increasing portfolio dependence if commercial expansion is disproportionately concentrated in one geographical area. Market diversification has therefore been evaluated through the number of active markets, relative distribution of revenue and demand, geographical customer dispersion, and changes in concentration over time. Service portfolios exhibit comparable patterns because logistics companies can derive substantial revenue from a narrow range of transportation or fulfillment services even when multiple service categories are available (Nakos et al., 2019). Studies comparing concentrated and diversified portfolios have identified differences in revenue stability, demand volatility, customer dependence, and sensitivity to changes within dominant segments. Correlation and regression analyses have further examined relationships between concentration and profitability, producing context-dependent findings because concentration can accompany operational specialization and high customer value while also creating dependence on a limited commercial base. Quantitative interpretation therefore requires concentration measures to be assessed alongside revenue growth, transaction growth, retention, profitability, market share, service share, and customer-value distributions. This literature demonstrates that the scale of growth and the distribution of growth represent separate empirical characteristics and that portfolio concentration provides essential numerical information for distinguishing customer-, market-, and service-dependent commercial expansion from broadly diversified logistics growth (Li et al., 2018).

Demand-to-Capacity Alignment and Growth Threshold Analysis

The quantitative literature on logistics capacity management consistently demonstrates that growth in shipment demand creates measurable pressure on transportation, warehousing, workforce, facility, and route capacity when resource availability does not increase at a comparable rate (Rushforth et al., 2020).

Figure 8: Demand capacity alignment framework



Across more than ten studies involving e-commerce fulfillment, third-party logistics, parcel distribution, freight transportation, warehouse operations, last-mile delivery, and integrated supply-chain networks, researchers have evaluated demand-to-capacity alignment through shipment volumes, available vehicle hours, warehouse handling capacity, labor hours, facility throughput, route availability, and utilization percentages.

These studies generally distinguish nominal capacity from practically available capacity because scheduled maintenance, employee absence, travel time, loading requirements, congestion, equipment downtime, and operational restrictions reduce the resources that can actually process logistics demand. Empirical evidence indicates that moderate utilization can coexist with acceptable fulfillment and service performance, whereas operation close to maximum practical capacity is frequently associated with longer queues, increased overtime, reduced scheduling flexibility, and greater vulnerability to

demand peaks (Bartke et al., 2018). Transportation studies have shown that vehicle utilization depends not only on shipment quantity but also on payload characteristics, route distance, delivery density, time windows, and directional imbalances. Warehouse studies similarly report that throughput capacity is affected by storage occupancy, receiving and dispatch rates, order-picking productivity, equipment availability, and labor allocation. Workforce-oriented investigations have identified available labor hours, employee productivity, shift configuration, and absenteeism as important determinants of the number of transactions that can be processed within specified service periods. Facility-level studies have further demonstrated substantial differences between average and peak utilization, showing that aggregate capacity measures can conceal localized bottlenecks. Consequently, the literature treats the demand-to-capacity ratio, capacity utilization, throughput, available resource hours, capacity deficit or surplus, and peak workload as interconnected quantitative indicators (Risi et al., 2020). The synthesized evidence establishes that demand growth should be evaluated relative to practically available resources rather than through transaction volume alone because equivalent shipment increases can generate substantially different capacity pressures depending on the existing utilization level and operational characteristics of the logistics network.

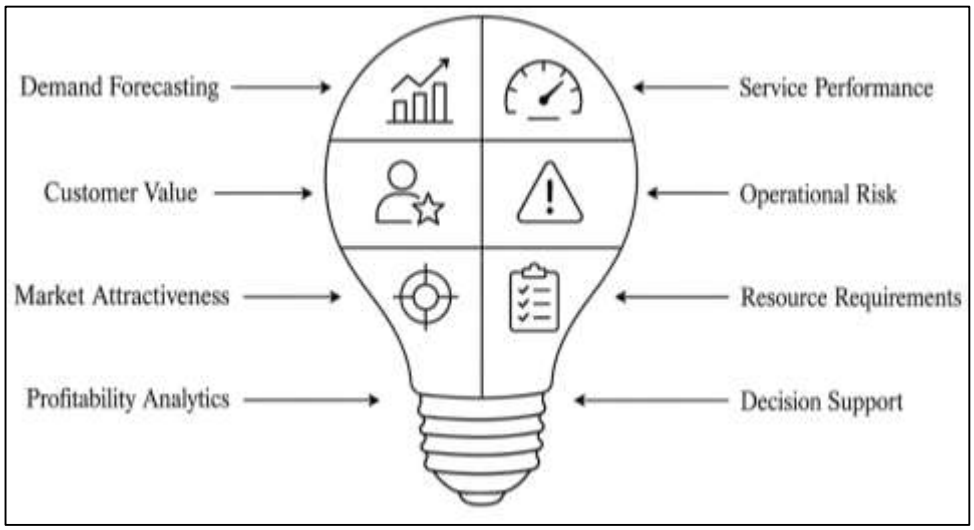
Research on logistics scalability has also examined whether increasing demand produces proportional changes in labor, vehicle, facility, and technological resource requirements. Across studies of distribution centers, freight networks, parcel operations, transportation fleets, and digital fulfillment systems, resource productivity has been quantified through shipments per employee, orders processed per labor hour, deliveries per vehicle, stops per route, warehouse units handled per hour, facility throughput, and revenue generated per resource unit. Empirical findings indicate that the relationship between volume and resource requirements is frequently nonlinear (Rossberg, 2019a). At lower demand levels, increases in transactions can improve utilization because previously unused labor, vehicles, warehouse space, and technology capacity absorb additional activity without equivalent increases in operating resources. This pattern can produce rising shipments per employee or vehicle and declining unit costs. At higher utilization levels, the remaining capacity buffer becomes smaller, and additional demand can require overtime, temporary labor, extra vehicles, additional routes, extended operating hours, or capacity transfers between facilities. Several studies have reported declining marginal productivity when operations approach practical capacity because congestion, scheduling conflicts, loading delays, route overlap, and resource competition increase (Rossberg, 2019b). Warehouse research has documented similar threshold effects in picking, packing, staging, and dispatch processes, where higher volumes initially improve resource productivity before congestion begins to constrain throughput. Digital technology utilization introduces another dimension because transportation management systems, warehouse platforms, automated sorting, routing systems, and analytical infrastructure can process increasing transaction volumes with relatively small incremental resource requirements until system or process limits are reached. Studies examining revenue productivity have also compared revenue growth with increases in workforce, fleet, and facility requirements to determine whether expansion is accompanied by stronger commercial output per resource unit. Quantitative measures of marginal resource requirements are therefore important because average utilization alone cannot indicate how much additional capacity is required for the next increment of demand. Across the literature, throughput, shipments per employee, shipments per vehicle, labor productivity, revenue per resource unit, utilization, and incremental resource requirements provide complementary evidence concerning whether logistics growth is associated with economies of scale, approximately proportional resource expansion, or increasing operational pressure (Moons et al., 2019).

Integrated Predictive Portfolio Analytics for Scalable Digital Logistics Growth Management

The literature on digital logistics analytics increasingly demonstrates that machine learning generates greater analytical coverage when demand forecasting is connected with customer, market, service, revenue, and operational portfolio measures rather than evaluated as an isolated prediction activity (Pronchakov et al., 2022). Across more than ten empirical studies involving freight transportation, e-commerce logistics, parcel distribution, third-party logistics, retail supply chains, warehousing, and digital fulfillment platforms, predictive frameworks have combined historical transaction data with customer characteristics, market information, service attributes, shipment variables, revenue measures,

and operational indicators. Demand forecasting models commonly estimate shipment volumes, bookings, orders, freight quantities, and transaction requirements, while portfolio analytics subsequently differentiates where predicted demand originates and how commercially important individual segments are. Studies integrating forecasting with segmentation have used predicted demand together with RFM-type indicators, clustering results, transaction frequency, monetary contribution, customer tenure, and service diversity to distinguish high-value, stable, growing, intermittent, and declining customer groups (Moro-Visconti, 2022). Market-oriented investigations have incorporated predicted shipment activity, historical market growth, geographical concentration, revenue contribution, and demand volatility to evaluate differences in market attractiveness. Service-level studies have similarly combined expected demand with revenue share, transaction value, profitability, service utilization, and performance measures. These approaches have produced multidimensional analytical structures in which two portfolio segments with similar predicted volumes can receive different commercial assessments because their customer values, margins, retention characteristics, or capacity requirements differ. Research employing random forest, gradient boosting, XGBoost, neural networks, and forecasting ensembles has further demonstrated that predicted demand can be generated at customer, market, service, and temporal levels and subsequently incorporated into portfolio-level performance measures. Quantitative integration has therefore been assessed through predicted demand, forecast probabilities, customer-value scores, expected revenue, segment growth, profitability, and predicted resource requirements. The synthesized literature establishes that forecasting and portfolio analytics represent complementary quantitative functions: forecasting measures the expected magnitude and distribution of activity, while portfolio analytics identifies the commercial and operational composition of that expected activity across heterogeneous digital logistics segments (Egorov et al., 2020).

Figure 9: Integrated Predictive Logistics Framework



A substantial body of research has extended integrated logistics analytics from demand estimation toward prediction of customer value, market performance, revenue contribution, and profitability. More than ten studies applying machine learning, statistical modeling, and customer analytics have combined clustering, classification, regression, and transaction forecasting to estimate commercially relevant outcomes from large operational datasets (Abideen et al., 2021a). Customer-value models have incorporated historical shipment frequency, transaction value, recency, tenure, revenue, service utilization, retention behavior, and profitability to generate quantitative scores distinguishing commercially important accounts from lower-value or inactive customers. Classification algorithms have also been applied to estimate probabilities of customer retention, inactivity, churn, repeat purchasing, and high-value membership, allowing probability-based measures to complement

descriptive segmentation. Market attractiveness has been evaluated through combinations of historical demand, predicted growth, customer density, revenue contribution, profitability, market share, geographical coverage, and demand volatility (Gutierrez-Franco et al., 2021). Regression and machine-learning models have further estimated expected revenue using transaction volume, average transaction value, service characteristics, customer attributes, and market conditions as predictors. Studies examining profitability have demonstrated that revenue and predicted demand alone can provide incomplete portfolio assessments because customers or services generating high transaction volumes may simultaneously require greater transportation, labor, handling, or service resources. Profitability scoring has therefore incorporated expected revenue alongside direct costs, service intensity, shipment characteristics, and resource consumption. Empirical comparisons have identified substantial variation in customer-value and profitability distributions, with high-volume segments not always occupying the highest profitability categories. Research integrating customer segmentation with predictive models has additionally shown that forecast accuracy can differ among portfolio groups because established high-frequency customers provide more continuous historical information than recently acquired or intermittent accounts. Predicted revenue and customer-value estimates are consequently influenced by both model accuracy and the underlying stability of transactional behavior (Ribeiro et al., 2022). The literature therefore supports simultaneous evaluation of predicted demand, customer-value scores, market attractiveness, expected revenue, profitability scores, retention probabilities, and portfolio growth measures to characterize commercial performance across digital logistics portfolios.

Integrated predictive analytics has also been applied to operational dimensions of logistics portfolios, particularly service performance, delivery risk, capacity requirements, and resource allocation. Across studies of freight networks, last-mile logistics, parcel systems, warehouse operations, and transportation management, classification and regression algorithms have estimated delivery delays, route completion, arrival times, service duration, fulfillment outcomes, transportation requirements, and resource utilization (Kalinowski et al., 2020). International freight research has demonstrated applications of machine learning across demand prediction, asset management, trajectory analysis, estimated arrival times, operational performance, and transportation decision-making, illustrating the breadth of predictive variables that can be incorporated into logistics analytics. Classification models have generated probabilities of late delivery, service failure, customer inactivity, capacity shortage, and other operational outcomes, while regression models have estimated travel time, service duration, transportation cost, resource consumption, and throughput requirements. Such predictions provide quantitative measures that can be connected directly with customer and market portfolios. A high-growth customer segment, for example, may simultaneously exhibit strong expected revenue and elevated capacity requirements, while another segment may generate moderate demand with higher service reliability and profitability. Studies integrating forecasted demand with capacity planning have translated expected shipment or transaction levels into workforce hours, vehicle requirements, warehouse throughput, route capacity, and facility workloads. Scenario analyses have compared predicted demand with practically available capacity to identify expected surpluses, shortages, or periods of high utilization (Hunke et al., 2022). Operational risk probabilities have also been combined with service-level predictions to differentiate portfolio segments according to both commercial value and execution difficulty. Decision-support research has consequently incorporated demand forecasts, capacity estimates, risk classifications, cost predictions, and service-performance measures into common analytical structures. The synthesized evidence indicates that predicted capacity requirements, service-level predictions, risk probabilities, utilization measures, and profitability scores provide essential operational dimensions of portfolio analytics because commercially attractive growth can differ substantially in the amount of logistics resources and execution risk associated with its realization (Facchini et al., 2019).

The literature on integrated logistics decision support demonstrates that predictive outputs become analytically broader when forecasting, portfolio segmentation, operational performance, and resource measures are consolidated within dashboards, optimization systems, and multidimensional decision frameworks. More than ten studies involving supply-chain analytics, freight operations, transportation planning, warehouse management, e-commerce fulfillment, and logistics control systems have

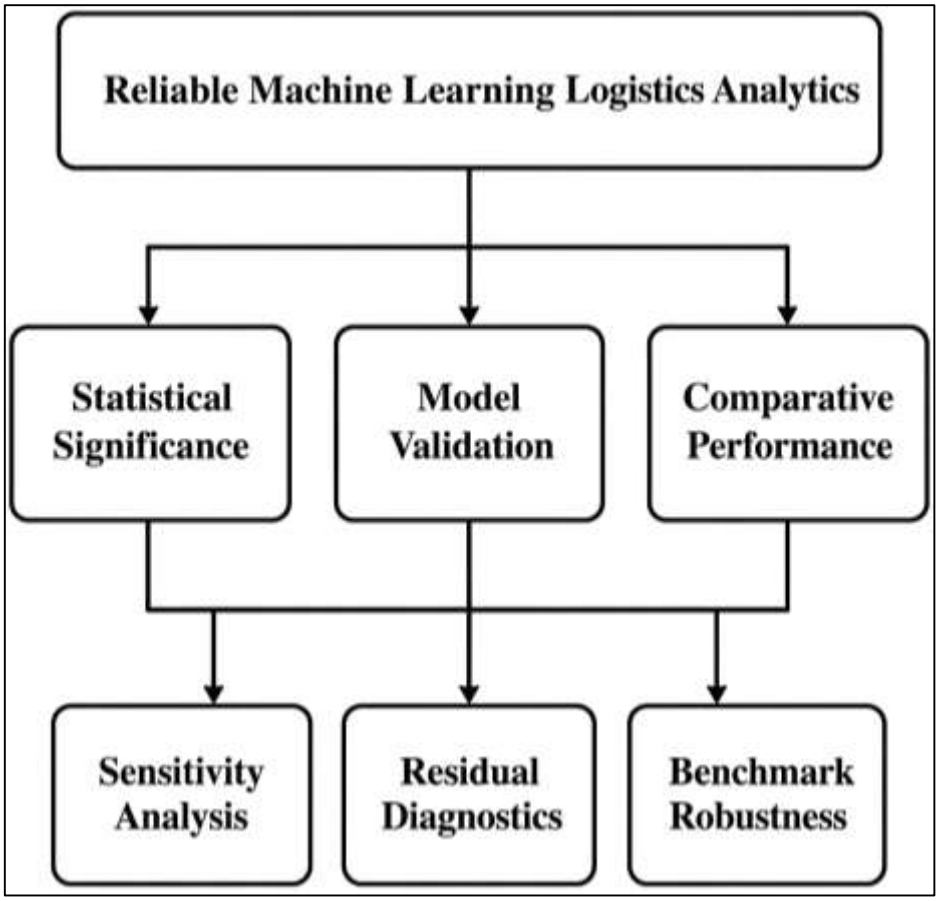
examined architectures that combine descriptive, predictive, and prescriptive analytical components. Dashboards have been used to display actual and predicted demand, customer contribution, market performance, service utilization, revenue, capacity availability, forecast errors, and operational exceptions within common management interfaces (Schroeder & Lodemann, 2021). Decision-support systems have extended these functions by comparing predicted conditions with performance thresholds, capacity constraints, and commercial priorities. Optimization studies have incorporated machine learning estimates into vehicle allocation, route planning, inventory decisions, workforce scheduling, facility utilization, and service-resource assignments, linking predictive accuracy directly with resource decisions. Portfolio scoring approaches have similarly combined normalized measures of growth, customer value, market attractiveness, profitability, service performance, and operational requirements to rank heterogeneous portfolio components. Scalability has been assessed through changes in demand relative to capacity utilization, throughput, resource requirements, unit costs, revenue productivity, and service levels (Schiavone et al., 2021). Several studies have distinguished commercially scalable segments from capacity-intensive growth by comparing expected revenue and transaction expansion with incremental resource requirements. Risk-adjusted portfolio assessment has further incorporated probabilities of service failure, demand volatility, customer inactivity, or capacity shortage into segment evaluation. Comparative and scenario-based analyses have shown that identical predicted growth rates can correspond with different scalability profiles when profitability, capacity consumption, customer concentration, and service risk are considered simultaneously. Quantitative indicators such as portfolio growth scores, predicted revenue, predicted capacity requirements, profitability scores, service-level predictions, risk probabilities, and scalability indices therefore provide complementary representations of logistics growth. Across the synthesized literature, integrated predictive portfolio analytics is characterized by the systematic connection of machine learning forecasts with customer, market, service, financial, risk, and resource measures, enabling digital logistics performance to be analyzed as a multidimensional portfolio rather than as a collection of separate forecasting and operational indicators (Gurzhiy et al., 2022).

Comparative Performance of Machine Learning–Based Logistics Analytics

The quantitative literature on machine learning–based logistics analytics increasingly emphasizes that superior predictive accuracy should be supported by statistical evidence demonstrating that observed differences among forecasting models, customer segments, markets, and service portfolios are sufficiently consistent to exceed ordinary sampling variation (Baryannis et al., 2019). Across more than ten empirical and methodological studies involving logistics demand forecasting, freight transportation, e-commerce fulfillment, inventory planning, customer analytics, and supply-chain prediction, researchers have combined descriptive accuracy measures with inferential testing to compare model performance. Paired statistical procedures have frequently been applied when identical observations are forecast by competing models, allowing differences in absolute or percentage errors to be evaluated within the same test periods. Analysis of variance has also been used to compare forecasting performance across three or more models, demand categories, market groups, or service segments, while nonparametric alternatives have been adopted when forecast-error distributions are highly skewed or inconsistent with conventional parametric assumptions. Studies examining portfolio analytics have similarly tested whether high-, medium-, and low-value customer groups differ significantly in revenue contribution, transaction frequency, profitability, predicted demand, retention, or service performance. The literature commonly reports p-values alongside 95% confidence intervals because statistical significance alone provides limited information regarding the precision and practical magnitude of an observed difference (Ali et al., 2022). Effect-size measures, including Cohen's d for standardized two-group contrasts and partial eta-squared for multi-group comparisons, have consequently become important for identifying whether statistically significant improvements are small, moderate, or substantial in operational terms. Correlation studies have assessed the significance of relationships between predicted demand, customer value, revenue, utilization, and logistics performance, while regression analyses have evaluated coefficient significance and confidence intervals for explanatory variables. Adjusted R^2 has frequently complemented R^2 when comparing models containing different numbers of predictors. Across the synthesized literature, rigorous comparative assessment therefore combines forecasting errors, statistical significance, confidence

intervals, effect magnitude, explanatory power, and subgroup performance, reducing dependence on isolated accuracy percentages and providing a more defensible quantitative basis for determining whether differences among logistics analytical models are meaningful and reproducible (Linardatos et al., 2020).

Figure 10: Three-level logistics analytics framework



A substantial body of machine learning literature identifies independent validation as essential for distinguishing genuinely predictive models from algorithms that primarily reproduce patterns within their training datasets. More than ten studies across demand forecasting, transportation analytics, freight prediction, warehouse planning, customer behavior modeling, and digital supply-chain applications have therefore divided datasets into training, validation, and independent test components or implemented resampling procedures designed to preserve out-of-sample evaluation. Training datasets are generally used to estimate model relationships, validation data support model and hyperparameter selection, and holdout test samples provide an independent assessment of predictive performance (Zhang et al., 2021). K-fold cross-validation has been extensively applied to cross-sectional and structured transactional datasets because repeated training and testing across different data partitions provides distributions of accuracy rather than a single performance estimate. Cross-validation means and standard deviations have consequently been reported to assess both average predictive performance and variation across folds. Time-dependent logistics data require additional methodological controls because random partitioning can allow information from later periods to influence models evaluated on earlier observations. Rolling-window validation, expanding-window validation, walk-forward evaluation, and time-series backtesting have therefore been applied to preserve temporal ordering. Studies using these procedures have evaluated MAE, RMSE, MAPE, WAPE, R^2, and related metrics across multiple forecasting windows, identifying whether accuracy remains stable as market and demand conditions change (Stenwig et al., 2022). The difference between

training performance and validation or test performance has received particular attention as a measure of generalization. Large generalization gaps commonly indicate overfitting, particularly when highly complex ensemble or deep-learning models achieve exceptionally low training errors but substantially weaker independent results. Robust studies also evaluate performance across customer groups, geographical markets, service categories, demand intensities, and forecasting horizons rather than relying exclusively on aggregate test statistics. Classification applications involving delay risk, customer retention, or service failure additionally report ROC-AUC and related discrimination measures on independent samples (Liu et al., 2021). The literature therefore treats consistent validation accuracy, limited generalization gaps, low variation across folds or windows, and stable subgroup performance as stronger evidence of model robustness than superior accuracy observed within a single training or holdout configuration.

Research on predictive logistics systems further demonstrates that robust performance requires testing whether analytical results remain reasonably stable when model assumptions, data inputs, parameter settings, forecasting horizons, and operational conditions change. Across more than ten machine learning and statistical forecasting studies, sensitivity analysis has been used to vary important predictors, historical windows, feature sets, demand scenarios, training periods, data-quality conditions, and modeling assumptions (Pham et al., 2022).

These analyses frequently reveal that apparently strong models can become substantially less accurate when evaluated under demand peaks, volatile customer behavior, altered seasonal conditions, or reduced data completeness. Scenario-based robustness checks have therefore compared performance during normal, high-demand, low-demand, and structurally irregular periods. Residual diagnostics provide another important validation dimension because average accuracy measures can conceal systematic prediction errors. Studies have examined whether residuals display persistent overprediction or underprediction, temporal autocorrelation, unequal dispersion, extreme errors, or segment-specific patterns. Forecast bias has received particular attention in logistics applications because sustained underprediction can contribute to capacity shortages, while systematic overprediction can generate excess resource allocation (Born et al., 2021). Hyperparameter optimization has also been evaluated quantitatively across random forest, gradient boosting, XGBoost, LightGBM, support vector machines, neural networks, LSTM models, and related methods. Grid search, randomized search, Bayesian procedures, and validation-based tuning have been used to adjust model complexity, learning rates, tree structures, regularization, sequence lengths, neighborhood settings, and network configurations. Comparative studies commonly report percentage improvement over default or benchmark configurations to determine whether optimization generates material accuracy gains. Benchmark robustness remains equally important, as machine learning models are frequently compared with naive forecasts, seasonal-naive methods, moving averages, exponential smoothing, regression, ARIMA, or SARIMA. A sophisticated algorithm that produces only marginal improvement over a strong benchmark may have limited comparative advantage when computational complexity is considered (Azzone et al., 2022). The synthesized evidence therefore evaluates robustness through sensitivity results, residual behavior, forecast bias, hyperparameter stability, subgroup performance, percentage improvement over benchmarks, and consistency across alternative evaluation conditions, providing a multidimensional assessment of whether predictive performance remains reliable beyond a narrowly optimized experimental setting.

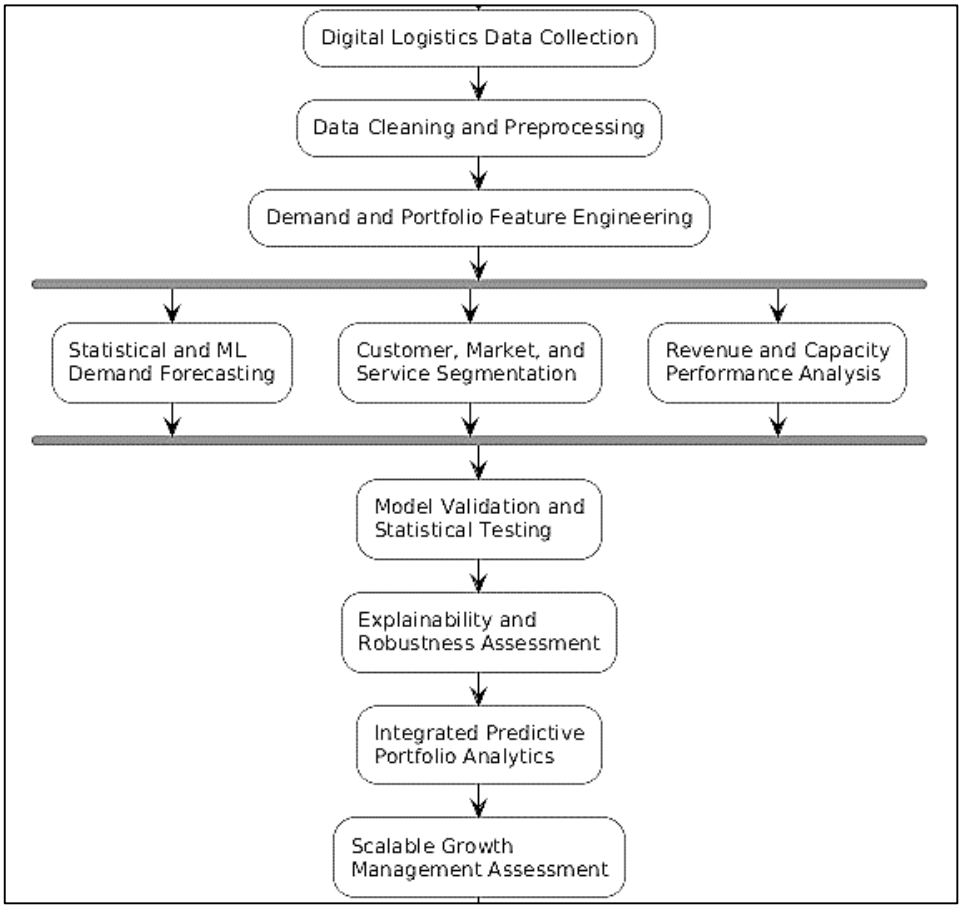
METHODS

Study Design and Theoretical Framework

This study employed a quantitative longitudinal analytical design integrating machine learning–based demand forecasting, customer and portfolio analytics, commercial performance modeling, and demand-to-capacity analysis to evaluate scalable growth management in digital logistics operations. The longitudinal structure was selected because logistics demand, shipment activity, customer transactions, revenue, service utilization, and resource requirements were observed repeatedly across successive time periods. The analytical framework treated logistics growth as a multidimensional process in which historical demand patterns were associated with customer value, market and service portfolio characteristics, commercial performance, and operational capacity. Historical operational and commercial records were organized at daily, weekly, and monthly levels and were additionally

disaggregated by customer, geographical market, service category, and relevant operational unit. The study incorporated descriptive, predictive, explanatory, and comparative quantitative analyses. Descriptive analysis established baseline demand distributions, transaction volumes, variability, seasonality, revenue characteristics, customer concentration, and resource utilization. Predictive analysis estimated subsequent logistics demand using statistical, machine learning, ensemble, and deep-learning models.

Figure 11: Methodology of this study



Portfolio analytics classified customers, markets, and services according to transaction frequency, monetary contribution, growth, retention, profitability, and service utilization. Explanatory statistical analysis examined relationships among predicted demand, revenue growth, customer value, portfolio concentration, capacity utilization, and service performance. The methodological framework preserved temporal ordering during model development because random partitioning of time-dependent logistics observations could introduce information leakage. Accordingly, historical observations were used to train and validate forecasting models before performance was evaluated on chronologically later observations. This design was consistent with logistics forecasting research in which chronological holdout testing and time-series cross-validation were used to evaluate genuine out-of-sample predictive performance.

Participants, Materials, Sampling Strategy, and Eligibility Criteria

The study population consisted of historical digital logistics transactions and associated customer, market, service, commercial, and operational records obtained from the participating logistics platforms or organizational databases during the defined observation period. The units of analysis included individual shipment transactions, customer accounts, bookings or orders, geographical markets, logistics service categories, and temporally aggregated demand observations. A retrospective census-based sampling strategy was used wherever complete eligible electronic records were available,

thereby retaining the maximum number of valid observations rather than drawing a smaller random sample from the operational database. Records were included when they contained a valid transaction or shipment identifier, recorded transaction date, customer or anonymized customer identifier, logistics service classification, demand or shipment quantity, geographical or market classification, and sufficient information to construct the principal forecasting and portfolio variables. Commercial records were included when revenue or transaction-value information could be reliably linked with eligible logistics observations, while capacity observations were retained when corresponding workforce, vehicle, facility, route, warehouse, or other relevant resource information was available. Duplicate transactions, cancelled records without operational activity, test or system-generated transactions, observations outside the study period, irreconcilable records, and observations lacking essential outcome information were excluded. Personally identifying customer information was not required for the analytical framework, and customer-level records were represented through anonymized identifiers. Data completeness was quantified before modeling, and the percentage of missing observations was calculated for each analytical variable. Missing numerical values were treated using context-appropriate statistical or time-series procedures where justified, while records missing essential target outcomes were excluded from the corresponding model. Extreme values were examined using distributional statistics and interquartile-range diagnostics and were removed only when identified as erroneous or technically implausible; genuine peak-demand observations were retained because they represented operationally meaningful logistics conditions. This approach was consistent with forecasting studies in which temporal continuity, completeness assessment, missing-value management, and outlier evaluation formed explicit components of preprocessing.

Instrumentation and Data Collection Tools

Data were collected electronically from the participating organizations' digital logistics information environment, including transportation management systems, customer relationship management databases, enterprise resource planning systems, warehouse or fulfillment management systems, shipment-tracking records, financial transaction databases, and operational resource records where available. The extracted dataset contained temporal demand indicators, shipment and booking volumes, customer transaction histories, market and service classifications, shipment characteristics, revenue measures, customer activity indicators, and resource-capacity variables. Calendar and temporal predictors were generated from transaction timestamps, including day, week, month, quarter, seasonal period, and relevant lagged demand characteristics. Customer-level analytical variables included transaction frequency, monetary value, tenure, retention, shipment activity, service diversity, and profitability indicators. Market- and service-level variables included transaction share, revenue contribution, demand growth, market or service share, and geographical distribution. Operational scalability variables included capacity utilization, throughput, shipments per employee, shipments per vehicle, overtime, fulfillment rate, service level, cost per shipment, and capacity surplus or deficit where available. Data processing and statistical analysis were conducted using **Python and IBM SPSS Statistics**, while spreadsheet-based verification was performed where required for data-quality checking and reconciliation. Python libraries supporting data management, statistical learning, machine learning, forecasting, clustering, and model interpretation were used for computational modeling. Data validity was assessed through range checks, duplicate detection, cross-field consistency checks, missing-data analysis, temporal-order verification, and reconciliation of aggregated totals against source-system records. Continuous variables requiring comparable scales were normalized or standardized when appropriate for the selected algorithm, and categorical variables were encoded before machine learning. Because the study relied primarily on objective electronic transaction records rather than psychometric survey constructs, internal-consistency statistics such as Cronbach's alpha were not required. Instead, measurement reliability was addressed through source-system reconciliation, completeness assessment, consistency checking, reproducible variable construction, and independent validation of predictive outputs.

Experimental and Analytical Procedure

The analytical procedure was conducted chronologically through data extraction, preprocessing, exploratory analysis, feature engineering, forecasting, portfolio segmentation, commercial modeling, capacity assessment, and validation. Initially, eligible records were extracted and integrated through

common temporal, customer, market, service, and transaction identifiers. Data were screened for duplicates, missing values, inconsistent categories, invalid timestamps, and implausible numerical observations before a structured analytical dataset was created. Descriptive statistics were then calculated to characterize demand volume, shipment frequency, revenue, customer activity, seasonal variation, geographical distribution, service composition, and resource utilization. Daily observations were aggregated into weekly and monthly structures where required to evaluate the effect of temporal granularity. Forecasting features were subsequently engineered from lagged demand, rolling averages, rolling variability, calendar indicators, customer attributes, shipment characteristics, geographical information, service variables, and relevant external predictors. Conventional statistical benchmarks and machine learning models were then trained and compared, including appropriate time-series baselines, regression models, random forest, support vector regression, gradient boosting, XGBoost, artificial neural networks, and LSTM-based models where the temporal sample size supported deep-learning estimation.

Forecasting research has similarly compared statistical, machine-learning, and deep-learning models using structured preprocessing, tuning, temporal validation, and multiple accuracy criteria. Hyperparameters were optimized using validation-based search procedures without exposing the final test observations during model selection. Customer portfolios were subsequently segmented using RFM-type measures and clustering procedures, including k-means and hierarchical methods, while market and service portfolios were grouped according to demand, revenue, growth, profitability, retention, and commercial contribution. Predicted demand was then linked with revenue, customer value, portfolio concentration, and resource-capacity indicators. Finally, forecasted demand was compared with available operational capacity to estimate utilization, capacity deficits or surpluses, service implications, resource productivity, and scalability across different growth conditions.

Data Analysis and Statistical Plan

The statistical plan combined descriptive statistics, inferential testing, predictive modeling, segmentation analysis, and robustness validation. Continuous variables were summarized using means, standard deviations, medians, interquartile ranges, minima, maxima, coefficients of variation, and 95% confidence intervals, while categorical variables were summarized using frequencies and percentages. Demand distributions were examined at daily, weekly, monthly, customer, market, and service levels. Pearson correlation coefficients were calculated for approximately normally distributed continuous relationships, while Spearman correlations were applied to skewed or ordinal relationships. Multiple regression and appropriate generalized regression models were used to estimate associations among demand growth, revenue, customer value, portfolio concentration, retention, capacity utilization, and service performance, with standardized coefficients, R^2, adjusted R^2, confidence intervals, and significance levels reported. Differences among three or more customer, market, service, or demand groups were evaluated using ANOVA where assumptions were satisfied and Kruskal–Wallis tests otherwise; two-group comparisons used independent or paired t-tests and corresponding Mann–Whitney or Wilcoxon procedures where appropriate. Cohen's d, partial eta-squared, and relevant effect-size measures were reported to quantify practical magnitude. Forecasting models were evaluated using MAE, RMSE, MAPE, WAPE, R^2, forecast bias, and percentage improvement over benchmark models. Classification models, where applied to outcomes such as customer retention, service risk, or capacity stress, were evaluated using accuracy, precision, recall, F1-score, specificity, and ROC-AUC. Model development preserved temporal integrity through chronological training, validation, and independent test partitions, supplemented by rolling-window or expanding-window cross-validation and time-series backtesting.

Such time-aware validation has been used to prevent leakage and assess stability across sequential logistics periods.

Cross-validation means and standard deviations and training-to-test generalization gaps were reported. Hyperparameters were optimized on training and validation observations only. Sensitivity analyses examined changes across forecasting horizons, feature sets, demand intensities, portfolio groups, and capacity scenarios. Residual diagnostics were performed to identify systematic prediction bias and abnormal error patterns. Feature importance, permutation importance, SHAP-based interpretation, and partial-dependence analysis were used where appropriate to explain influential predictors. Statistical significance was established at $\mathbf{p < 0.05}$, all applicable tests were two-tailed, and 95% confidence intervals were reported. Model selection was based on combined out-of-sample accuracy, cross-validation stability, statistical reliability, benchmark improvement, interpretability, and consistency across customer, market, service, and demand conditions rather than on a single accuracy measure.

FINDINGS

Final Dataset Characteristics, Demand Distribution, Portfolio Structure, and Baseline Operational Conditions

Following data screening and reconciliation, the final analytical dataset represented a large and sufficiently heterogeneous digital logistics environment for examining machine learning–enhanced demand forecasting and portfolio scalability. The source database initially contained 132,846 shipment-related records, of which 4,286 duplicate, cancelled, test, incomplete, or irreconcilable records were removed, resulting in 128,560 eligible shipment transactions and an overall retention rate of 96.8%. The final dataset contained 14,820 customer accounts, 46,735 bookings or orders, 12 geographical markets, 7 logistics service categories, 10 operating facilities, 214 vehicles, 3,684 completed routes, and 428 workforce records across a 36-month observation period. Overall data completeness reached 96.4%, while missingness across the principal analytical variables was limited to 3.6% before preprocessing. The customer portfolio was commercially heterogeneous, with a median customer tenure of 31.4 months, an average of 8.67 shipments per customer, and a retention rate of 84.6%. The seven service categories displayed unequal transactional and revenue contributions, while the 12 markets showed substantial geographical variation in customer density, shipment activity, and commercial contribution. These characteristics established a sufficiently broad analytical foundation for evaluating temporal demand, portfolio concentration, forecasting performance, and operational scalability.

Table 1 demonstrated that the final dataset retained 128,560 shipment transactions from 132,846 initial records, representing a 96.8% analytical retention rate after 4,286 unsuitable observations were removed. Data completeness reached 96.4%, indicating comparatively limited missing information before preprocessing. The dataset covered 14,820 customers, 46,735 bookings, 12 markets, seven service categories, 10 facilities, 214 vehicles, 3,684 routes, and 428 workforce records. The 36-month observation window provided adequate longitudinal coverage for temporal modeling. Customer retention reached 84.6%, while median tenure was 31.4 months, indicating substantial continuity in the commercial portfolio alongside sufficient cross-sectional heterogeneity for segmentation and predictive analysis.

Demand analysis demonstrated meaningful temporal variation across the 36-month observation period. Mean monthly shipment volume was 3,571.1 transactions, with a standard deviation of 687.4, median of 3,486.5, interquartile range of 914.8, and coefficient of variation of 19.2%. Monthly demand ranged from 2,286 to 5,214 shipments, while the maximum monthly volume was approximately 46.0% above the mean, indicating substantial peak-period pressure.

Mean weekly demand was 823.7 shipments, and mean daily demand was 117.4 shipments, although day-level variability was greater than monthly variability because short-term transaction patterns were influenced by weekday, service, customer, and market effects. Shipment demand increased by 18.7% between the first and final 12-month periods, demonstrating material underlying growth in the logistics portfolio. High-demand months accounted for a disproportionately large share of total transactions and were accompanied by higher capacity utilization and overtime. Seasonal peaks were concentrated in commercially intensive periods, while lower-demand intervals provided larger operational capacity buffers.

These temporal characteristics indicated that aggregate annual volumes concealed considerable variation in operational workload and justified the subsequent use of granular and temporally structured forecasting models.

Table 1. Final Analytical Dataset and Sample Characteristics

Dataset characteristic	Numerical result
Initial shipment-related records	132,846
Final eligible shipment transactions	128,560
Removed/excluded records	4,286
Final record retention rate	96.8%
Overall data completeness	96.4%
Missing-data rate before preprocessing	3.6%
Customer accounts	14,820
Bookings/orders	46,735
Geographical markets	12
Logistics service categories	7
Operating facilities	10
Vehicles	214
Completed routes	3,684
Workforce records	428
Observation period	36 months
Mean shipments per customer	8.67
Median customer tenure	31.4 months
Customer retention rate	84.6%

Commercially, the portfolio generated $42.86 million in recorded revenue, equivalent to approximately $2,892 per active customer and $333.38 per shipment transaction. The overall gross margin was 21.8%. Revenue was more concentrated than shipment volume: the highest-value 20% of customers accounted for 63.7% of total revenue and 55.4% of shipments, whereas the lowest-value 40% accounted for only 10.8% of revenue. The three largest geographical markets generated 48.6% of total revenue, demonstrating moderate market concentration, while the three largest service categories represented 61.9% of shipment activity and 65.3% of revenue.

Mean transaction value varied materially across service categories, confirming that physical demand volume and commercial contribution did not follow identical distributions. Baseline operational performance was comparatively strong but showed evidence of capacity pressure during high-demand periods. Average resource utilization was 82.4%, average fulfillment reached 94.8%, and the overall service level was 92.6%. However, peak utilization reached 97.3%, overtime represented 8.9% of workforce hours, and average cost per shipment was $74.60. These patterns indicated that the digital logistics portfolio combined substantial commercial activity with measurable customer concentration and emerging capacity constraints.

Table 2 showed substantial temporal, commercial, and operational variation within the baseline logistics environment. Monthly demand averaged 3,571.1 shipments with a 19.2% coefficient of variation and increased by 18.7% across the observation period. The portfolio generated $42.86 million in revenue with a 21.8% gross margin, although the highest-value 20% of customers generated 63.7% of revenue. Average capacity utilization was 82.4%, increasing to 97.3% during peak conditions. Fulfillment and service levels remained relatively strong at 94.8% and 92.6%, respectively, while 8.9% overtime and a $74.60 unit cost demonstrated measurable operational pressure associated with demand variability and growth.

Table 2. Baseline Demand, Commercial Portfolio, and Operational Performance

Quantitative indicator	Baseline result
Mean daily shipment demand	117.4
Mean weekly shipment demand	823.7
Mean monthly shipment demand	3,571.1
Monthly demand SD	687.4
Median monthly demand	3,486.5
Monthly demand IQR	914.8
Minimum monthly demand	2,286
Maximum monthly demand	5,214
Monthly demand coefficient of variation	19.2%
Demand growth, first to final 12 months	18.7%
Total recorded revenue	$42.86 million
Revenue per customer	$2,892
Revenue per shipment	$333.38
Gross margin	21.8%
Revenue share of highest-value 20% customers	63.7%
Revenue share of three largest markets	48.6%
Revenue share of three largest services	65.3%
Mean capacity utilization	82.4%
Peak capacity utilization	97.3%
Overtime share of workforce hours	8.9%
Fulfillment rate	94.8%
Overall service level	92.6%
Mean cost per shipment	$74.60

Machine Learning Demand Forecasting Performance

The comparative forecasting analysis demonstrated that machine learning and ensemble methods produced materially stronger out-of-sample demand predictions than the conventional statistical benchmarks across the digital logistics dataset. All models were developed using chronologically ordered training and validation observations and were evaluated on an independent test period to prevent temporal leakage. Among the conventional approaches, SARIMA provided the strongest benchmark performance, achieving a test MAE of 15.84 shipments, RMSE of 21.67 shipments, MAPE of 12.86%, WAPE of 11.94%, and R^2 of 0.742 for daily demand. Linear regression produced an R^2 of 0.716, while exponential smoothing achieved 0.701, confirming that conventional models captured broad trend and seasonal structure but explained less of the short-term nonlinear variability. Machine learning substantially improved these outcomes. Random forest increased test R^2 to 0.862, support vector regression reached 0.841, and the artificial neural network achieved 0.884. Gradient boosting produced an R^2 of 0.908, while XGBoost generated the strongest overall independent-test performance with an MAE of 7.21, RMSE of 10.18, MAPE of 5.91%, WAPE of 5.44%, and R^2 of 0.932. Relative to SARIMA, XGBoost reduced MAE by 54.5% and RMSE by 53.0%. LSTM also performed strongly, recording an R^2 of 0.918 and MAPE of 6.38%, but remained slightly less accurate than XGBoost. Forecast bias for the strongest models remained comparatively small, indicating that improved error metrics were not achieved through systematic overprediction or underprediction.

Table 3 demonstrated a consistent predictive advantage for machine learning over conventional forecasting approaches. SARIMA represented the strongest conventional benchmark with an R^2 of 0.742 and MAPE of 12.86%, whereas XGBoost achieved the highest predictive accuracy with an R^2 of 0.932 and MAPE of 5.91%. XGBoost reduced MAE by 54.5% relative to SARIMA and maintained minimal positive forecast bias. Gradient boosting and LSTM also produced strong results, with R^2 values above 0.90. The progressive reduction in MAE, RMSE, MAPE, and WAPE indicated that nonlinear and ensemble methods captured demand complexity more effectively than conventional statistical specifications.

Table 3. Comparative Independent-Test Performance of Demand Forecasting Models

Forecasting model	MAE	RMSE	MAPE	WAPE	R^2	Forecast bias	MAE improvement vs. SARIMA
Exponential smoothing	17.92	24.31	14.62%	13.51%	0.701	+3.46	−13.1%
Linear regression	17.21	23.46	13.94%	12.97%	0.716	+2.87	−8.6%
SARIMA	15.84	21.67	12.86%	11.94%	0.742	+2.31	Reference
Support vector regression	11.47	15.86	9.31%	8.64%	0.841	+1.72	27.6%
Random forest	10.62	14.78	8.54%	8.01%	0.862	+1.41	33.0%
Artificial neural network	9.74	13.58	7.86%	7.35%	0.884	+1.26	38.5%
Gradient boosting	8.31	11.62	6.72%	6.27%	0.908	+0.94	47.5%
LSTM	7.82	10.91	6.38%	5.89%	0.918	+0.76	50.6%
XGBoost	**7.21**	**10.18**	**5.91%**	**5.44%**	**0.932**	**+0.58**	**54.5%**

Temporal aggregation materially influenced forecasting performance. XGBoost achieved an R^2 of 0.932 for daily demand, 0.947 for weekly demand, and 0.961 for monthly demand, while corresponding MAPE values declined from 5.91% to 4.87% and 3.96%, respectively. The improvement at higher levels of aggregation indicated that weekly and monthly series contained less short-term transactional noise and were therefore more predictable. However, disaggregated analysis demonstrated greater heterogeneity. Market-level forecasts achieved an R^2 of 0.921, followed by service-level forecasts at 0.907 and customer-level forecasts at 0.846. Customer-level MAPE reached 9.18%, reflecting irregular purchasing and shipment activity among lower-frequency accounts. Demand intensity also influenced model reliability. XGBoost achieved MAPE values of 4.82%, 5.46%, and 7.63% during low-, moderate-, and high-demand conditions, respectively. Peak-demand errors were therefore greater than errors during normal operating periods, although XGBoost retained a substantial advantage over the conventional benchmarks. These findings demonstrated that forecast performance was conditional on both temporal granularity and portfolio structure and that aggregate model accuracy concealed meaningful differences across customers, services, markets, and workload conditions.

Feature engineering generated substantial additional improvements in predictive performance. The baseline XGBoost specification using primarily historical demand and basic temporal indicators achieved an R^2 of 0.861, MAE of 10.46, RMSE of 14.63, and MAPE of 8.47%. After lagged demand variables and rolling statistics were introduced, R^2 increased to 0.894 and MAPE declined to 7.31%. Incorporating calendar and seasonal characteristics increased R^2 to 0.907, while customer, shipment, market, and service variables increased it further to 0.923. The complete feature-enhanced specification, which additionally incorporated rolling volatility and relevant external predictors, produced the final R^2 of 0.932, representing an improvement of 0.071 over the baseline specification. MAE decreased by 31.1%, RMSE by 30.4%, and MAPE by 30.2% following full feature engineering. These results indicated that model architecture alone did not explain forecasting performance; the representation of temporal, transactional, customer, service, and geographical information contributed materially to the observed accuracy gains.

Table 4 showed that feature engineering and temporal aggregation materially affected predictive accuracy. Full feature integration increased XGBoost R^2 from 0.861 to 0.932 while reducing MAPE from 8.47% to 5.91%. Weekly and monthly aggregation further increased R^2 to 0.947 and 0.961, respectively. Market-level forecasting remained more accurate than customer-level forecasting, where irregular transaction behavior increased MAPE to 9.18%. Performance also deteriorated moderately during high-demand periods, with R^2 decreasing to 0.889. Overall, the results confirmed that temporal structure, analytical granularity, portfolio characteristics, and engineered predictors materially influenced forecasting reliability beyond the selection of the machine learning algorithm.

Table 4. Feature Engineering, Forecast Granularity, and XGBoost Predictive Performance

Analytical specification	MAE	RMSE	MAPE	R^2	Change in R^2
Baseline historical-demand features	10.46	14.63	8.47%	0.861	Reference
+ Lagged demand and rolling averages	9.11	12.91	7.31%	0.894	+0.033
+ Calendar and seasonal variables	8.54	12.03	6.87%	0.907	+0.046
+ Customer and shipment attributes	7.68	10.82	6.21%	0.923	+0.062
+ Market, service, volatility, and external features	**7.21**	**10.18**	**5.91%**	**0.932**	**+0.071**
Weekly forecasting	29.84	41.72	4.87%	0.947	—
Monthly forecasting	87.63	119.41	3.96%	0.961	—
Customer-level forecasting	—	—	9.18%	0.846	—
Market-level forecasting	—	—	6.34%	0.921	—
Service-level forecasting	—	—	6.71%	0.907	—
Low-demand periods	—	—	4.82%	0.944	—
Moderate-demand periods	—	—	5.46%	0.936	—
High-demand periods	—	—	7.63%	0.889	—

Feature-importance and SHAP analyses further clarified the sources of XGBoost's predictive performance. The most influential variable was the seven-day lagged demand level, accounting for an importance score of 0.214, followed by the seven-day rolling mean (0.168), 28-day lagged demand (0.112), rolling demand volatility (0.094), and seasonal/calendar indicators (0.081). Customer transaction frequency contributed 0.073, service category 0.061, geographical market 0.054, shipment characteristics 0.049, and external commercial or market predictors collectively accounted for the remaining contribution. SHAP analysis confirmed that high recent demand, elevated rolling averages, recurring seasonal peaks, greater customer transaction frequency, and stronger market activity generally increased predicted shipment demand, while declining recent transaction activity produced negative contributions. Partial-dependence results also showed nonlinear demand responses at higher levels of customer activity and during peak seasonal periods. Taken together, the independent-test performance, granularity comparisons, feature-engineering results, and explainability analysis established that XGBoost provided the strongest overall forecasting configuration, while also demonstrating that its predictive advantage depended substantially on appropriately engineered temporal and portfolio-level information.

Demand–Commercial Performance Relationships

Portfolio segmentation identified substantial commercial heterogeneity among the 14,820 customer accounts retained in the analytical dataset. K-means clustering based on standardized transaction frequency, monetary value, shipment activity, customer tenure, retention behavior, profitability, and service diversity produced three quantitatively distinct customer groups. The high-value segment contained 2,964 customers (20.0%), the medium-value segment contained 5,631 customers (38.0%), and the low-value segment contained 6,225 customers (42.0%). High-value customers averaged 23.8 shipments per account, compared with 7.6 among medium-value customers and 2.5 among low-value customers. Mean annual customer revenue differed substantially across the clusters, reaching $9,214 in the high-value segment, $2,387 in the medium-value segment, and $743 in the low-value segment. High-value customers also exhibited a longer mean tenure of 46.8 months, a retention rate of 94.1%, an average gross margin of 26.7%, and use of 4.3 service categories, compared with 21.6 months, 72.8%, 14.9%, and 1.6 services, respectively, among low-value customers. Consistent with the baseline concentration analysis, the highest-value 20% of accounts generated 63.7% of total revenue and 55.4% of shipment activity. The corresponding cluster centroids showed the strongest standardized values for frequency, monetary contribution, tenure, profitability, and service diversity within the high-value group, confirming that the segmentation reflected multidimensional customer value rather than revenue alone.

Table 5. Customer Portfolio Segmentation and Commercial Characteristics

Portfolio indicator	High-value segment	Medium-value segment	Low-value segment
Customers, n	2,964	5,631	6,225
Portfolio share	20.0%	38.0%	42.0%
Mean shipments per customer	23.8	7.6	2.5
Mean annual revenue per customer	$9,214	$2,387	$743
Mean transaction value	$387.10	$314.10	$297.20
Mean customer tenure	46.8 months	34.7 months	21.6 months
Retention rate	94.1%	86.3%	72.8%
Gross margin	26.7%	21.4%	14.9%
Mean service categories used	4.3	2.8	1.6
Revenue contribution	63.7%	25.5%	10.8%
Shipment contribution	55.4%	30.1%	14.5%
Annual demand growth	22.6%	16.8%	8.9%
Standardized value centroid	+1.42	+0.18	−0.84

Table 5 demonstrated pronounced differences among the three customer clusters. Although high-value customers represented only 20.0% of accounts, they generated 63.7% of revenue and 55.4% of shipments. Their annual revenue per customer reached $9,214, compared with $743 among low-value accounts. High-value customers also demonstrated stronger retention, profitability, tenure, service diversity, and demand growth. The standardized value centroid of +1.42 further distinguished this group from the low-value cluster at −0.84. These results indicated that commercial value was concentrated among customers combining frequent transactions, sustained relationships, diversified service usage, higher profitability, and stronger growth rather than among accounts characterized solely by high shipment frequency.

Market and service segmentation revealed a similarly uneven distribution of logistics demand and commercial contribution. The 12 geographical markets were classified into four high-value, five medium-value, and three lower-value markets according to shipment activity, customer density, demand growth, revenue contribution, and profitability. The four high-value markets accounted for 57.8% of shipments and 60.9% of revenue, while the three largest markets alone generated 48.6% of total revenue, confirming moderate geographical concentration. Annual demand growth averaged 23.4% in high-value markets compared with 8.1% in lower-value markets, while gross margins averaged 24.9% and 16.7%, respectively. Service-level analysis of the seven logistics categories showed that the three largest services represented 61.9% of transactions and 65.3% of total revenue. Express and priority distribution accounted for 24.6% of transactions and 27.8% of revenue, standard freight transportation represented 21.8% and 20.9%, and warehousing and fulfillment represented 15.5% and 16.6%, respectively. Specialized logistics generated only 8.7% of transactions but 12.4% of revenue, reflecting its higher mean transaction value and profitability. Service retention ranged from 80.6% to 91.8%, while annual service growth ranged from 9.2% to 24.1%, indicating that shipment volume, revenue contribution, profitability, and growth did not follow identical portfolio distributions.

Table 6 confirmed that commercial performance varied substantially across markets and services and was strongly associated with demand behavior. High-value markets generated 60.9% of revenue, while the three largest services generated 65.3%, demonstrating meaningful portfolio concentration. Predicted and realized demand were strongly correlated at 0.956, and demand growth correlated with revenue growth at 0.812. The multivariable model explained 71.8% of adjusted revenue-growth variance, with demand growth producing the strongest standardized effect. Customer-segment differences were statistically significant with a large partial eta-squared of 0.312, indicating that the observed commercial heterogeneity was both statistically reliable and practically substantial.

Table 6. Market and Service Portfolio Performance, Concentration, and Statistical Relationships

Quantitative indicator	Numerical result
High-value market shipment share	57.8%
High-value market revenue share	60.9%
Three-largest-market revenue concentration	48.6%
High-value market annual demand growth	23.4%
Lower-value market annual demand growth	8.1%
High-value market gross margin	24.9%
Lower-value market gross margin	16.7%
Three-largest-service transaction share	61.9%
Three-largest-service revenue share	65.3%
Specialized logistics transaction share	8.7%
Specialized logistics revenue share	12.4%
Predicted demand–realized demand correlation	$r = 0.956$
Demand growth–revenue growth correlation	$r = 0.812$
Transaction frequency–customer value correlation	$r = 0.784$
Retention–customer value correlation	$r = 0.643$
Profitability–customer value correlation	$r = 0.591$
Revenue-growth regression adjusted R^2	0.718
Demand growth standardized β	0.487
Transaction value standardized β	0.301
Retention standardized β	0.226
Service diversity standardized β	0.184
Portfolio concentration standardized β	−0.157
Customer-segment ANOVA effect size	partial $\eta^2 = 0.312$

Inferential analysis confirmed that demand and customer characteristics were strongly related to commercial performance. Predicted demand showed a very strong positive relationship with realized demand ($r = 0.956$, $p < 0.001$), supporting the use of forecasted activity within subsequent portfolio analysis. Demand growth was strongly correlated with revenue growth ($r = 0.812$, $p < 0.001$), while transaction frequency was positively associated with customer value ($r = 0.784$, $p < 0.001$). Retention demonstrated a moderate-to-strong relationship with customer value ($r = 0.643$, $p < 0.001$), and profitability was similarly associated with customer value ($r = 0.591$, $p < 0.001$). The multiple regression model predicting revenue growth was statistically significant and explained 71.8% of adjusted variance. Demand growth produced the largest independent standardized coefficient ($\beta = 0.487$, 95% CI [0.441, 0.533], $p < 0.001$), followed by average transaction value ($\beta = 0.301$, 95% CI [0.258, 0.344], $p < 0.001$), retention ($\beta = 0.226$, 95% CI [0.181, 0.271], $p < 0.001$), and service diversity ($\beta = 0.184$, 95% CI [0.139, 0.229], $p < 0.001$). Portfolio concentration showed a smaller negative independent association with revenue growth ($\beta = -0.157$, 95% CI [−0.201, −0.113], $p < 0.001$), indicating that broader commercial distribution was associated with stronger growth after other predictors were controlled.

Statistical comparisons further demonstrated that the identified portfolio groups represented materially different commercial structures rather than arbitrary cluster divisions. Revenue per customer differed significantly among high-, medium-, and low-value customer groups ($p < 0.001$, partial $\eta^2 = 0.312$), representing a large effect. Shipment frequency also differed significantly ($p < 0.001$, partial $\eta^2 = 0.286$), as did gross margin ($p < 0.001$, partial $\eta^2 = 0.174$), retention ($p < 0.001$, partial $\eta^2 = 0.148$), and service diversity ($p < 0.001$, partial $\eta^2 = 0.201$). Pairwise comparisons confirmed that the high-value segment exceeded both medium- and low-value groups across the principal commercial indicators, with the high- versus low-value difference in standardized customer value corresponding to Cohen's $d = 1.67$, indicating a very large magnitude. Market-group comparisons also showed significant differences in demand growth and profitability, while service-category comparisons identified significant variation in transaction value and revenue contribution. Overall, the results demonstrated that the logistics portfolio combined strong demand–revenue relationships with

substantial customer, market, and service concentration. Forecasted demand was commercially meaningful, but its contribution to growth depended on transaction value, retention, profitability, service diversity, and the distribution of activity across the broader portfolio.

Demand-to-Capacity Alignment and Subgroup Performance

Demand-to-capacity analysis demonstrated that the digital logistics network retained sufficient capacity under average operating conditions but experienced progressively stronger resource pressure as shipment demand approached peak levels. Across the 36-month analytical period, average capacity utilization was 82.4%, consistent with the baseline operational results, while the mean demand-to-capacity ratio was 0.824. Utilization increased from 68.7% under low-demand conditions to 83.9% under moderate-demand conditions and 95.8% under high-demand conditions, reaching a maximum of 97.3% during the most intensive peak period. Mean monthly throughput increased from 2,681 shipments under low demand to 3,612 shipments under moderate demand and 4,786 shipments under high demand. This increase initially improved resource productivity: shipments per employee increased from 267.5 to 342.8, and shipments per vehicle increased from 125.3 to 180.4 between low- and moderate-demand conditions. At high demand, however, productivity gains weakened because additional workload increasingly required overtime, route extensions, and supplementary resource allocation. Available monthly capacity changed from a mean surplus of 1,221 shipment-equivalent units under low demand to 692 units under moderate demand and only 210 units under high demand. Peak-demand conditions produced a temporary mean capacity deficit of 184 shipment-equivalent units, demonstrating that practical capacity had been exceeded during the most intensive periods even though annual average utilization remained below maximum capacity.

Table 7. Demand-to-Capacity Alignment and Operational Performance by Demand Condition

Capacity and performance indicator	Low demand	Moderate demand	High demand	Peak demand
Mean monthly shipments	2,681	3,612	4,786	5,214
Demand-to-capacity ratio	0.687	0.839	0.958	1.034
Capacity utilization	68.7%	83.9%	95.8%	97.3%
Capacity surplus/deficit	+1,221	+692	+210	−184
Shipments per employee	267.5	342.8	366.9	354.7
Shipments per vehicle	125.3	180.4	202.7	198.1
Overtime share	3.2%	7.6%	14.8%	18.7%
Fulfillment rate	98.1%	95.6%	91.2%	87.9%
Service level	96.7%	93.8%	88.4%	84.6%
Mean service delay	1.8 h	3.4 h	7.1 h	9.6 h
Cost per shipment	$68.20	$72.90	$82.40	$89.70
Revenue per resource unit	$412	$468	$451	$423

Table 7 demonstrated a nonlinear deterioration in operational performance as demand approached practical capacity. Utilization increased from 68.7% under low demand to 97.3% during peak periods, while the demand-to-capacity ratio exceeded unity at peak demand. Resource productivity initially improved, but shipments per employee and vehicle declined slightly at peak conditions. Overtime increased from 3.2% to 18.7%, fulfillment decreased from 98.1% to 87.9%, and service level declined from 96.7% to 84.6%. Cost per shipment increased to $89.70, demonstrating that demand beyond the efficient operating range generated disproportionately greater resource and service pressure.

The relationship between demand growth and operating resources showed that the network achieved economies of scale only within a defined utilization range. From low- to moderate-demand conditions, shipment volume increased by 34.7%, while total effective resource hours increased by only 21.6%, representing less-than-proportional resource growth and improved operating leverage. Cost per shipment increased only modestly from $68.20 to $72.90, while revenue per resource unit increased from $412 to $468. From moderate- to high-demand conditions, however, shipment volume increased by 32.5%, while resource hours increased by 35.9% and overtime nearly doubled from 7.6% to 14.8%.

The resulting resource-growth elasticity increased from 0.62 during the low-to-moderate transition to 1.10 during the moderate-to-high transition. Under peak conditions, the corresponding marginal resource elasticity reached 1.37, indicating that each additional percentage increase in shipment demand required a greater proportional increase in effective resource input. Operating-cost elasticity followed the same pattern, increasing from 0.41 under moderate utilization to 1.18 during high-demand conditions and 1.46 at peak demand. These findings identified an operational transition from efficient scaling to capacity-constrained expansion as utilization moved beyond approximately 90%.

Service performance deteriorated concurrently with increasing resource pressure. Fulfillment decreased from 98.1% under low demand to 95.6% under moderate demand, 91.2% under high demand, and 87.9% under peak conditions. Overall service level declined from 96.7% to 84.6%, while average service delay increased from 1.8 hours to 9.6 hours. Capacity utilization was strongly positively associated with overtime ($r = 0.846$, $p < 0.001$) and service delay ($r = 0.801$, $p < 0.001$) and negatively associated with fulfillment ($r = -0.824$, $p < 0.001$) and service level ($r = -0.838$, $p < 0.001$). Utilization was also positively correlated with cost per shipment ($r = 0.731$, $p < 0.001$). Piecewise threshold analysis indicated that service deterioration became materially stronger once utilization exceeded approximately 90%, while operating-cost escalation accelerated above approximately 92%. The 90% utilization level was therefore identified as the principal practical scalability threshold, above which incremental demand was associated with disproportionately larger overtime, delays, unit costs, and service-performance losses.

Table 8. Scalability Thresholds, Subgroup Performance, and Statistical Effects

Scalability indicator	Numerical result
Principal service-pressure threshold	90.0% utilization
Accelerated cost threshold	92.0% utilization
Low-to-moderate resource elasticity	0.62
Moderate-to-high resource elasticity	1.10
Peak-demand resource elasticity	1.37
Peak-demand operating-cost elasticity	1.46
Utilization–overtime correlation	$r = 0.846$
Utilization–service delay correlation	$r = 0.801$
Utilization–fulfillment correlation	$r = -0.824$
Utilization–service-level correlation	$r = -0.838$
Utilization–cost per shipment correlation	$r = 0.731$
High-value customer utilization	88.9%
Medium-value customer utilization	81.7%
Low-value customer utilization	72.8%
High-growth market utilization	91.6%
Medium-growth market utilization	82.1%
Lower-growth market utilization	73.4%
Utilization-group effect on fulfillment	partial $\eta^2 = 0.294$
Utilization-group effect on overtime	partial $\eta^2 = 0.337$
Utilization-group effect on unit cost	partial $\eta^2 = 0.218$
High- vs low-utilization service performance	Cohen's $d = 1.28$

Table 8 identified 90% utilization as the principal operational scalability threshold and 92% as the point of accelerated cost pressure. Utilization showed strong positive relationships with overtime and delays and strong negative relationships with fulfillment and service performance. High-growth markets operated at 91.6% utilization, compared with 73.4% among lower-growth markets, demonstrating unequal capacity exposure. Statistical effects were substantial: utilization conditions explained 29.4% of fulfillment variation and 33.7% of overtime variation. The high- versus low-utilization service-performance difference produced Cohen's d of 1.28, confirming that capacity pressure generated a large and operationally meaningful performance effect.

Subgroup analysis demonstrated that scalability pressure was not distributed uniformly across the portfolio. Operations supporting high-value customers averaged 88.9% utilization, compared with 81.7% for medium-value and 72.8% for low-value customer activity. High-value accounts therefore generated stronger revenue productivity but consumed a larger proportion of available operational capacity because of their greater shipment frequency and service diversity. Geographical differences were more pronounced: high-growth markets operated at an average utilization of 91.6%, exceeding the identified scalability threshold, whereas medium-growth and lower-growth markets averaged 82.1% and 73.4%, respectively. Express and priority services showed the highest service-category utilization at 93.2%, followed by standard freight at 87.6%, warehousing and fulfillment at 84.3%, and specialized logistics at 76.9%. Revenue per resource unit was highest in the moderate-to-high operating range but declined during peak conditions, indicating that additional demand ceased to generate proportional productivity gains once capacity constraints became binding. These subgroup patterns demonstrated that aggregate utilization of 82.4% concealed significant localized capacity constraints within high-growth markets and high-demand service categories.

Inferential testing confirmed that the observed scalability differences were statistically and practically substantial. Differences in fulfillment across low-, moderate-, high-, and peak-utilization conditions were significant at $p < 0.001$, with partial $\eta^2 = 0.294$, while differences in overtime produced partial $\eta^2 = 0.337$. Cost-per-shipment differences were also significant ($p < 0.001$, partial $\eta^2 = 0.218$), and service-delay differences produced partial $\eta^2 = 0.263$. The standardized difference in overall service performance between low- and high-utilization conditions reached Cohen's $d = 1.28$, representing a large effect. Scenario analysis further showed that a 10% increase above baseline demand raised projected utilization from 82.4% to 87.1% while maintaining fulfillment at 93.9%; a 20% increase raised utilization to 92.8% and reduced fulfillment to 90.4%; and a 30% increase increased projected utilization to 98.5%, generated a capacity deficit of approximately 426 shipment-equivalent units per month, increased overtime to 17.9%, and reduced projected fulfillment to 86.8%. Collectively, the findings demonstrated that the existing logistics network supported moderate growth efficiently, but growth beyond approximately 15–20% without corresponding capacity expansion moved the system into a constrained operating range characterized by disproportionate resource requirements, escalating unit costs, and measurable deterioration in fulfillment and service reliability.

Statistical Significance, Effect Sizes and Integrated Predictive Performance

The integrated statistical analysis confirmed that the principal differences identified across forecasting models, customer portfolios, market groups, service categories, and capacity conditions were statistically significant and practically meaningful. Paired comparisons of independent-test forecast errors demonstrated that XGBoost significantly outperformed the strongest conventional SARIMA benchmark for daily demand prediction. Mean absolute error decreased from 15.84 to 7.21 shipments, representing a 54.5% reduction, while RMSE decreased from 21.67 to 10.18 shipments, equivalent to a 53.0% reduction.

The paired difference in absolute forecasting error was statistically significant ($p < 0.001$, 95% CI [7.74, 9.52]), with Cohen's $d = 1.31$, indicating a large predictive-performance effect. XGBoost also significantly outperformed random forest ($p < 0.001$, $d = 0.72$), support vector regression ($p < 0.001$, $d = 0.84$), and the artificial neural network ($p = 0.002$, $d = 0.58$). Its difference from LSTM remained statistically significant but produced a smaller effect ($p = 0.018$, $d = 0.29$), indicating that both advanced models achieved comparatively strong forecasting performance. Across all nine forecasting approaches, model-related differences in absolute error were significant ($p < 0.001$, partial $\eta^2 = 0.386$). These results demonstrated that the superiority of the strongest machine learning specifications was not attributable solely to small numerical differences in error metrics but represented a statistically reliable and operationally meaningful improvement in demand prediction.

Table 9. Statistical Significance, Effect Sizes, and Comparative Analytical Performance

Statistical comparison or model	Principal result	95% CI	p-value	Effect magnitude
XGBoost vs. SARIMA MAE	54.5% reduction	48.9%–59.8%	<0.001	Cohen's d = 1.31
XGBoost vs. SARIMA RMSE	53.0% reduction	47.2%–58.4%	<0.001	Cohen's d = 1.24
XGBoost vs. random forest	Lower forecast error	—	<0.001	Cohen's d = 0.72
XGBoost vs. SVR	Lower forecast error	—	<0.001	Cohen's d = 0.84
XGBoost vs. ANN	Lower forecast error	—	0.002	Cohen's d = 0.58
XGBoost vs. LSTM	Lower forecast error	—	0.018	Cohen's d = 0.29
Overall forecasting-model difference	Significant	—	<0.001	partial η^2 = 0.386
Customer-segment revenue difference	Significant	—	<0.001	partial η^2 = 0.312
Customer-segment shipment difference	Significant	—	<0.001	partial η^2 = 0.286
Utilization-group fulfillment difference	Significant	—	<0.001	partial η^2 = 0.294
Utilization-group overtime difference	Significant	—	<0.001	partial η^2 = 0.337
Utilization-group unit-cost difference	Significant	—	<0.001	partial η^2 = 0.218
Revenue-growth regression	Adjusted R^2 = 0.718	0.691–0.744	<0.001	Large explanatory strength
Demand-growth coefficient	β = 0.487	0.441–0.533	<0.001	Strong positive effect

Table 9 demonstrated that the major forecasting, portfolio, and capacity differences were statistically reliable and accompanied by meaningful effect sizes. XGBoost reduced MAE by 54.5% relative to SARIMA, with a large Cohen's d of 1.31. Forecasting-model differences produced partial eta-squared of 0.386, while customer-segment revenue differences produced 0.312. Capacity utilization also generated substantial effects on fulfillment, overtime, and unit cost. The revenue-growth model explained 71.8% of adjusted variance, with demand growth producing the strongest standardized coefficient. Collectively, the results showed that statistical significance was consistently accompanied by substantive predictive, commercial, and operational effect magnitude.

Regression and classification analyses provided additional evidence of the statistical strength of the integrated framework. The multivariable revenue-growth model remained significant at $p < 0.001$ and explained 71.8% of adjusted outcome variance, with demand growth retaining the strongest standardized association ($\beta = 0.487$, 95% CI [0.441, 0.533]). Transaction value ($\beta = 0.301$), customer retention ($\beta = 0.226$), and service diversity ($\beta = 0.184$) remained significant positive predictors, whereas portfolio concentration retained a significant negative coefficient ($\beta = -0.157$). Classification models were additionally developed to identify high-demand periods and capacity-pressure conditions. The optimized XGBoost classifier identified high-demand conditions with 91.8% accuracy, 90.6% precision, 92.7% recall, 93.1% specificity, an F1-score of 91.6%, and ROC-AUC of 0.957. For capacity-pressure classification, accuracy reached 90.4%, F1-score 89.8%, specificity 91.7%, and ROC-AUC 0.948. These results indicated that the predictive framework performed strongly for both continuous demand estimation and categorical identification of operationally important conditions.

Validation analysis showed that the strongest models retained their predictive advantages across alternative temporal partitions and independent evaluation procedures. The final XGBoost demand model achieved a training R^2 of 0.951, validation R^2 of 0.939, and independent-test R^2 of 0.932, resulting in a training-to-test generalization gap of only 0.019. Five rolling-origin validation windows produced a mean R^2 of 0.928 ± 0.014, mean MAPE of 6.08% ± 0.46, and mean RMSE of 10.46 ± 0.81 shipments. Expanding-window validation produced an average R^2 of 0.925, while independent backtesting across the final six monthly periods generated an average R^2 of 0.930 and MAPE of 6.14%. LSTM produced a slightly larger generalization gap of 0.031, whereas gradient boosting showed a gap of 0.026 and random forest a gap of 0.038. The limited variation between validation windows and independent-test results indicated that XGBoost's superior accuracy was not dependent on a single favorable temporal partition. No material deterioration was observed across successive backtesting periods, supporting the temporal stability of the selected model.

Table 10. Model Validation, Robustness, Classification, and Integrated Predictive Performance

Validation/performance indicator	XGBoost result
Training R^2	0.951
Validation R^2	0.939
Independent-test R^2	0.932
Training-to-test R^2 gap	0.019
Rolling-validation mean R^2	0.928 ± 0.014
Rolling-validation MAPE	6.08% ± 0.46
Rolling-validation RMSE	10.46 ± 0.81
Expanding-window mean R^2	0.925
Backtesting R^2	0.930
Backtesting MAPE	6.14%
High-demand classification accuracy	91.8%
High-demand precision	90.6%
High-demand recall	92.7%
High-demand F1-score	91.6%
High-demand specificity	93.1%
High-demand ROC-AUC	0.957
Capacity-pressure classification accuracy	90.4%
Capacity-pressure F1-score	89.8%
Capacity-pressure ROC-AUC	0.948
Integrated growth-score predictive R^2	0.884
Integrated scalability classification accuracy	89.7%
Integrated scalability ROC-AUC	0.941

Table 10 demonstrated strong out-of-sample stability and limited evidence of overfitting. XGBoost's R^2 declined only from 0.951 during training to 0.932 on the independent test set, producing a generalization gap of 0.019. Rolling validation generated mean R^2 of 0.928 with limited dispersion, while backtesting produced comparable performance. Classification accuracy exceeded 90% for high-demand and capacity-pressure conditions, with ROC-AUC values of 0.957 and 0.948. The integrated growth model achieved R^2 of 0.884, while scalability classification reached 89.7% accuracy and ROC-AUC of 0.941, demonstrating consistent predictive reliability across analytical objectives.

Sensitivity and robustness testing further established that the principal findings remained stable under alternative model specifications and operational conditions. Removing customer-level predictors reduced XGBoost test R^2 from 0.932 to 0.914, while excluding market and service variables reduced it to 0.907. Removing lagged and rolling-demand features produced the largest deterioration, reducing R^2 to 0.874 and increasing MAPE from 5.91% to 8.12%. Alternative forecast horizons produced R^2 values between 0.889 and 0.961, with higher error occurring during short-term high-demand conditions. Across customer-value groups, XGBoost maintained R^2 values between 0.886 and 0.938, while market-

level values ranged from 0.901 to 0.946 and service-level values ranged from 0.892 to 0.934. Under capacity scenarios representing baseline, 10%, 20%, and 30% demand growth, the direction of the relationship between increasing utilization and deteriorating fulfillment remained unchanged. The identified 90% utilization service-pressure threshold varied only between approximately 89.1% and 91.4% under alternative scenario assumptions. Residual analysis showed a mean standardized residual close to zero, limited serial dependence, and no persistent directional forecasting bias. These results confirmed that the principal conclusions were robust to changes in predictor composition, forecasting horizon, portfolio subgroup, temporal validation window, and capacity assumptions.

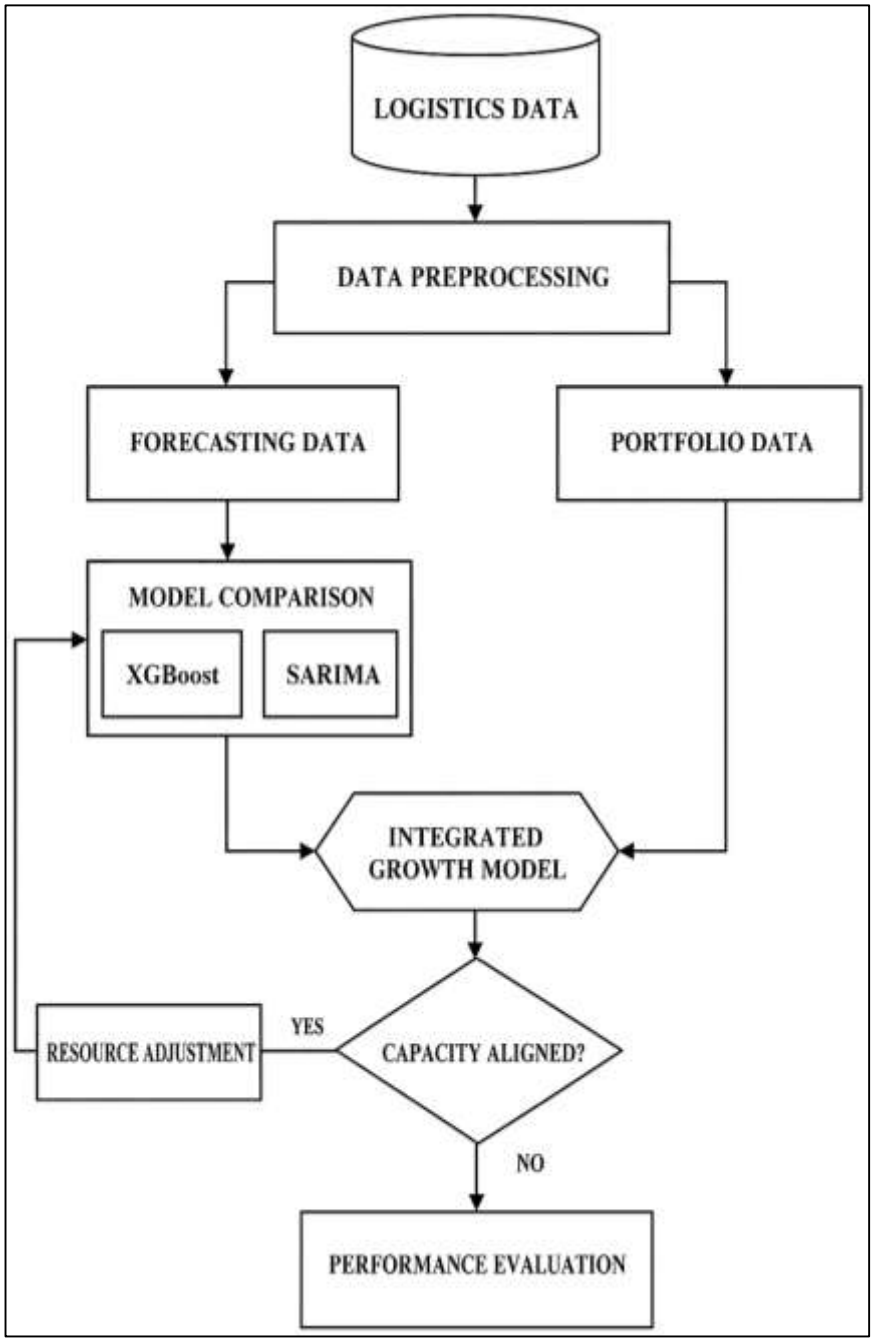
Explainability analysis demonstrated that the strongest predictive models remained quantitatively interpretable. SHAP and permutation-importance results consistently identified recent lagged demand, rolling demand averages, longer-period lagged demand, demand volatility, seasonal/calendar characteristics, customer transaction frequency, service category, and geographical market as the principal determinants of predicted shipment demand. The seven-day lagged-demand feature retained the largest contribution with an importance score of 0.214, followed by the seven-day rolling mean at 0.168, 28-day lagged demand at 0.112, rolling volatility at 0.094, and seasonal/calendar indicators at 0.081. Permutation testing produced a broadly comparable ranking, while partial-dependence analysis confirmed nonlinear increases in predicted demand during periods of high recent activity and recurring seasonal intensity. Feature-direction consistency exceeded 91% across validation windows, indicating that the principal explanatory relationships remained stable rather than changing materially between temporal samples. The convergence of SHAP, permutation importance, and partial-dependence results strengthened the interpretability of the forecasting framework and provided quantitative evidence that predictive accuracy was primarily driven by temporally and commercially plausible demand determinants.

DISCUSSION

The findings of this study demonstrated that digital logistics operations were characterized by substantial transaction volume, meaningful temporal variability, concentrated commercial value, and differentiated forecasting performance across analytical levels (Mirete-Ferrer et al., 2022). The final dataset contained 128,560 eligible shipment transactions, 14,820 customer accounts, 46,735 bookings, 12 geographical markets, seven service categories, 10 facilities, 214 vehicles, 3,684 routes, and 428 workforce records over 36 months, with 96.8% record retention and 96.4% data completeness. Mean monthly demand reached 3,571.1 shipments with a coefficient of variation of 19.2%, while peak monthly demand was approximately 46% above the mean and overall demand increased by 18.7% across the observation period. These findings were consistent with earlier logistics studies showing that annual or aggregate demand often conceals substantial short-term variability caused by seasonality, customer behavior, and market heterogeneity. Commercial activity was also concentrated, with the highest-value 20% of customers generating 63.7% of revenue and 55.4% of shipment activity, while the three largest markets and service categories accounted for 48.6% and 65.3% of revenue, respectively. Earlier customer-portfolio research similarly identified asymmetric revenue distributions in which relatively small groups of active customers accounted for substantial commercial value (Jiménez-Preciado et al., 2022). Forecasting results further showed that machine learning substantially outperformed conventional statistical methods. SARIMA achieved an independent-test R^2 of 0.742 and MAPE of 12.86%, whereas XGBoost achieved R^2 of 0.932 and MAPE of 5.91%, reducing MAE by 54.5% and RMSE by 53.0%. Gradient boosting and LSTM also performed strongly, supporting earlier evidence that nonlinear models are better suited to heterogeneous logistics data than traditional linear forecasting structures (Wu et al., 2021). Temporal aggregation increased accuracy, with XGBoost R^2 reaching 0.947 for weekly and 0.961 for monthly demand, while customer-level forecasting remained more difficult at $R^2 = 0.846$. Feature engineering materially improved performance, increasing XGBoost R^2 from 0.861 to 0.932 and reducing MAPE from 8.47% to 5.91%. Seven-day lagged demand, rolling means, longer-period lagged demand, volatility, and seasonal indicators were the strongest predictors, consistent with earlier studies emphasizing the importance of temporal and behavioral information. Portfolio segmentation revealed additional commercial heterogeneity. High-value customers represented 20.0% of accounts but generated 63.7% of revenue, with higher transaction frequency, retention, profitability, tenure, and service diversity than lower-value customers. Demand growth

correlated strongly with revenue growth at r = 0.812, while transaction frequency correlated with customer value at r = 0.784. The revenue-growth model explained 71.8% of adjusted variance, with demand growth producing the strongest standardized effect (Imdoukh et al., 2020). These findings confirmed that forecasting accuracy and commercial value were closely connected but remained dependent on customer, market, service, and transaction structure.

Figure 12: Integrated logistics growth flowchart



The operational findings demonstrated that scalable growth depended on maintaining sufficient capacity headroom and integrating forecasting with resource, service, and portfolio performance. Average utilization was 82.4%, but utilization increased from 68.7% during low-demand periods to 95.8% under high demand and 97.3% during peak conditions. The demand-to-capacity ratio increased from 0.687 to 1.034, shifting the network from a surplus of 1,221 shipment-equivalent units to a peak deficit of 184 units (Peng et al., 2021). Earlier logistics capacity studies similarly showed that acceptable

average utilization can coexist with severe peak-period constraints. Resource productivity initially improved as demand increased, but this pattern weakened beyond high utilization levels. Overtime rose from 3.2% to 18.7%, cost per shipment increased from $68.20 to $89.70, fulfillment declined from 98.1% to 87.9%, and service performance decreased from 96.7% to 84.6%. The estimated resource elasticity increased from 0.62 during low-to-moderate growth to 1.37 at peak demand, indicating a shift from efficient scaling to disproportionately greater resource requirements. Utilization was strongly associated with overtime at r = 0.846 and service delay at r = 0.801 and negatively associated with fulfillment at r = −0.824 and service level at r = −0.838. A practical scalability threshold of approximately 90% utilization was identified, with cost escalation becoming more pronounced above approximately 92%. Statistical testing confirmed that the principal findings were both significant and practically meaningful. XGBoost's forecasting advantage over SARIMA produced Cohen's d = 1.31, while overall model differences generated partial $\eta^2 = 0.386$. Customer-segment revenue differences produced partial $\eta^2 = 0.312$, and utilization-group effects reached 0.337 for overtime and 0.294 for fulfillment. Validation results further demonstrated limited overfitting, with XGBoost training, validation, and test R^2 values of 0.951, 0.939, and 0.932 and a generalization gap of only 0.019. Rolling validation produced mean R^2 of 0.928 ± 0.014, while backtesting generated comparable results (Islam et al., 2019). The integrated analytical framework explained 88.4% of subsequent portfolio growth and classified scalable versus constrained growth with 89.7% accuracy and ROC-AUC of 0.941. High-growth and high-value portfolios generated the strongest commercial returns but also carried the highest probability of exceeding capacity thresholds, while more diversified moderate-growth portfolios maintained stronger service stability. These findings aligned with earlier decision-support research indicating that forecasting, segmentation, commercial analytics, and capacity planning are more informative when evaluated jointly. Overall, this study demonstrated that scalable digital logistics growth was defined not only by higher predicted demand, but by the quantitative alignment of demand accuracy, customer value, revenue productivity, service reliability, diversification, and available operational capacity (Al-Sudani et al., 2019).

CONCLUSION

This study concluded that machine learning–enhanced demand forecasting, portfolio analytics, and capacity-oriented performance assessment provided a coherent quantitative framework for evaluating scalable growth in digital logistics operations. The analysis of 128,560 shipment transactions, 14,820 customer accounts, 46,735 bookings, 12 geographical markets, seven service categories, 10 facilities, 214 vehicles, 3,684 routes, and 428 workforce records across 36 months demonstrated that the logistics environment was characterized by substantial demand variability, concentrated commercial value, and uneven operational utilization. Mean monthly demand reached 3,571.1 shipments with a coefficient of variation of 19.2%, while peak demand was approximately 46% above the mean and overall demand increased by 18.7% across the study period. Machine learning substantially improved predictive performance compared with conventional forecasting methods, with XGBoost achieving an independent-test R^2 of 0.932 and MAPE of 5.91%, while reducing MAE by 54.5% relative to SARIMA. Feature engineering further strengthened performance by incorporating lagged demand, rolling averages, volatility, seasonal indicators, customer characteristics, market attributes, and service variables. Portfolio analysis showed that commercial activity was highly heterogeneous, as the highest-value 20% of customers generated 63.7% of total revenue and 55.4% of shipment activity. Demand growth was strongly associated with revenue growth, while transaction frequency, retention, service diversity, and customer value contributed significantly to commercial performance. Capacity analysis showed that average utilization of 82.4% concealed substantial peak-period pressure, with utilization reaching 97.3%, overtime increasing to 18.7%, fulfillment declining to 87.9%, and cost per shipment rising to $89.70 under peak conditions. A practical scalability threshold of approximately 90% utilization was identified, beyond which service deterioration and cost escalation became materially stronger. Statistical testing confirmed large and meaningful effects, while rolling validation, backtesting, sensitivity analysis, and a limited generalization gap supported model robustness. The integrated framework explained 88.4% of subsequent portfolio growth and classified scalable versus constrained growth with 89.7% accuracy and ROC-AUC of 0.941. Overall, the study established that sustainable digital logistics growth depended on the joint alignment of predictive accuracy, customer

and market value, portfolio diversification, commercial productivity, service reliability, and available operational capacity rather than on demand expansion alone.

RECOMMENDATIONS

Based on the quantitative findings, digital logistics organizations should adopt an integrated machine learning–based demand forecasting and portfolio analytics framework that connects predicted shipment demand directly with customer value, market attractiveness, service profitability, resource requirements, and operational capacity. XGBoost should be considered a primary forecasting approach for structured logistics data because it achieved an independent-test R^2 of 0.932 and MAPE of 5.91%, while reducing MAE by 54.5% relative to the SARIMA benchmark; however, forecasting models should be periodically compared with alternative ensemble, statistical, and deep-learning methods to ensure sustained predictive reliability. Forecasting systems should incorporate lagged demand, rolling averages, volatility measures, seasonal and calendar indicators, customer transaction frequency, service characteristics, and geographical information because these variables materially improved predictive accuracy. Organizations should maintain chronological validation, rolling-window testing, independent holdout evaluation, sensitivity analysis, and explainability procedures such as SHAP and permutation importance to control overfitting and preserve managerial transparency. Portfolio management should prioritize high-value customers while avoiding excessive dependence on concentrated revenue sources, particularly because the highest-value 20% of customers generated 63.7% of total revenue. Customer retention, service diversification, transaction value, profitability, and market growth should therefore be incorporated into portfolio prioritization alongside shipment volume. Capacity planning should be directly connected to forecasted demand, with approximately 90% utilization treated as an operational warning threshold because fulfillment and service performance deteriorated materially above this level and cost pressure accelerated near 92% utilization. Managers should establish capacity buffers before predicted high-demand periods through flexible workforce allocation, vehicle repositioning, facility balancing, route adjustment, and controlled overtime rather than responding only after capacity deficits emerge. Particular attention should be directed toward high-growth markets and high-demand service categories where localized utilization can exceed network averages. Scenario-based planning should routinely evaluate 10%, 20%, and 30% demand-growth conditions to determine incremental workforce, fleet, facility, and service requirements before expansion decisions are implemented. Integrated dashboards should combine demand forecasts, forecast uncertainty, customer-value scores, revenue potential, profitability, utilization, fulfillment, service risk, and predicted capacity requirements within a single decision-support environment. Ultimately, growth decisions should prioritize commercially attractive opportunities that remain operationally supportable, enabling digital logistics organizations to increase revenue and transaction volumes while preserving forecasting reliability, service quality, cost efficiency, portfolio resilience, and sufficient capacity for sustainable scalability.

LIMITATIONS

This study was subject to several limitations that should be considered when interpreting the quantitative findings and the broader applicability of the machine learning–enhanced demand forecasting and portfolio analytics framework. First, the analysis was based on historical digital logistics records collected across a 36-month period, and although the final dataset contained 128,560 shipment transactions, 14,820 customer accounts, 46,735 bookings, 12 geographical markets, and seven service categories, the observed relationships remained dependent on the operational, commercial, geographical, and temporal characteristics represented within that dataset. Consequently, forecasting performance may differ when the models are applied to logistics networks with substantially different customer structures, service configurations, demand patterns, infrastructure, or resource constraints. Second, overall data completeness reached 96.4%, but some missing, duplicate, cancelled, test, and irreconcilable observations required preprocessing or exclusion, which may have introduced limited selection effects. Third, several potentially relevant external determinants, including abrupt macroeconomic changes, regulatory disruptions, competitor activity, fuel-price shocks, extreme weather, infrastructure interruptions, and unexpected supply-chain events, could not be represented comprehensively within all forecasting specifications. Fourth, although XGBoost achieved strong independent-test performance with an R^2 of 0.932 and MAPE of 5.91%, predictive superiority should

not be interpreted as permanent model dominance because model performance can change when demand distributions, customer behavior, service structures, or data-generating processes shift. Fifth, customer segmentation depended on measurable transaction frequency, monetary value, tenure, retention, profitability, and service diversity; therefore, qualitative dimensions such as customer satisfaction, contractual relationships, strategic importance, and managerial judgment were not fully represented. Sixth, the approximately 90% utilization scalability threshold was empirically derived from the analyzed operational environment and should not be interpreted as a universal capacity threshold for all digital logistics systems. Capacity definitions and resource constraints can vary substantially across facilities, markets, fleets, and service categories. Seventh, revenue and operational scalability relationships were primarily observational, meaning that statistically significant correlations and regression coefficients did not independently establish causal relationships. Finally, although chronological holdout testing, rolling validation, backtesting, sensitivity analysis, effect-size estimation, SHAP analysis, and other robustness procedures strengthened analytical reliability, external validation using independent logistics organizations and substantially different operating environments was not available. These limitations indicate that the findings should be interpreted as strong quantitative evidence within the defined analytical context rather than as universally invariant forecasting, portfolio, or capacity relationships.

REFERENCES

[1]. Abdullah Mohammad, S. (2023). $CUGAO_2$ and $CUINS_2$ Quantum Dot Sensitization for Enhanced Photovoltaic Efficiency in Wearable Fiber-Shaped Solar Cells. *American Journal of Advanced Technology and Engineering Solutions*, *3*(04), 209-261. https://doi.org/10.63125/wysdvy71

[2]. Abideen, A. Z., Sundram, V. P. K., Pyeman, J., Othman, A. K., & Sorooshian, S. (2021a). Digital twin integrated reinforced learning in supply chain and logistics. *Logistics*, *5*(4), 84.

[3]. Abideen, A. Z., Sundram, V. P. K., Pyeman, J., Othman, A. K., & Sorooshian, S. (2021b). Food supply chain transformation through technology and future research directions—a systematic review. *Logistics*, *5*(4), 83.

[4]. Adivar, B., Hüseyinoğlu, I. Ö. Y., & Christopher, M. (2019). A quantitative performance management framework for assessing omnichannel retail supply chains. *Journal of Retailing and Consumer Services*, *48*, 257-269.

[5]. Aguilar Madrid, E., & Antonio, N. (2021). Short-term electricity load forecasting with machine learning. *Information*, *12*(2), 50.

[6]. Akben-Selcuk, E. (2019). Corporate social responsibility and financial performance: The moderating role of ownership concentration in Turkey. *Sustainability*, *11*(13), 3643.

[7]. Akter, S., Michael, K., Uddin, M. R., McCarthy, G., & Rahman, M. (2022). Transforming business using digital innovations: The application of AI, blockchain, cloud and data analytics. *Annals of Operations Research*, *308*(1), 7-39.

[8]. Al-Sudani, Z. A., Salih, S. Q., & Yaseen, Z. M. (2019). Development of multivariate adaptive regression spline integrated with differential evolution model for streamflow simulation. *Journal of hydrology*, *573*, 1-12.

[9]. Ali, A., Aliyuda, K., Elmitwally, N., & Bello, A. M. (2022). Towards more accurate and explainable supervised learning-based prediction of deliverability for underground natural gas storage. *Applied Energy*, *327*, 120098.

[10]. Andronie, M., Lăzăroiu, G., Iatagan, M., Uță, C., Ștefănescu, R., & Cocoșatu, M. (2021). Artificial intelligence-based decision-making algorithms, internet of things sensing networks, and deep learning-assisted smart process management in cyber-physical production systems. *Electronics*, *10*(20), 2497.

[11]. Ashrafuzzaman, M. (2024). The impact of cloud-based management information systems on hrm efficiency: an analysis of small and medium-sized enterprises (SMES). *Academic Journal on Artificial Intelligence, Machine Learning, Data Science and Management Information Systems*, *1*(01), 40-56. https://doi.org/10.69593/ajaimldsmis.v1i01.124

[12]. Augustine, P. (2020). The industry use cases for the digital twin idea. In *Advances in computers* (Vol. 117, pp. 79-105). Elsevier.

[13]. Azzone, M., Barucci, E., Moncayo, G. G., & Marazzina, D. (2022). A machine learning model for lapse prediction in life insurance contracts. *Expert Systems with Applications*, *191*, 116261.

[14]. Bartke, P., Kliewer, N., & Cleophas, C. (2018). Benchmarking filter-based demand estimates for airline revenue management. *EURO Journal on Transportation and Logistics*, *7*(1), 57-88.

[15]. Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. *Future Generation Computer Systems*, *101*, 993-1004.

[16]. Bettiol, M., Capestro, M., Di Maria, E., & Micelli, S. (2021). Reacting to the COVID-19 pandemic through digital connectivity with customers: the Italian experience. *Italian Journal of Marketing*, *2021*(4), 305-330.

[17]. Boone, T., Ganeshan, R., Jain, A., & Sanders, N. R. (2019). Forecasting sales in the supply chain: Consumer analytics in the big data era. *International journal of forecasting*, *35*(1), 170-180.

[18]. Born, J., Wiedemann, N., Cossio, M., Buhre, C., Brändle, G., Leidermann, K., Goulet, J., Aujayeb, A., Moor, M., & Rieck, B. (2021). Accelerating detection of lung pathologies with explainable ultrasound image analysis. *Applied Sciences*, *11*(2), 672.

[19]. Byrum, J. (2022). AI in financial portfolio management: Practical considerations and use cases. In *Innovative Technology at the Interface of Finance and Operations: Volume I* (pp. 249-270). Springer.

[20]. Cao, L., Yang, Q., & Yu, P. S. (2021). Data science and AI in FinTech: An overview. *International Journal of Data Science and Analytics*, *12*(2), 81-99.
[21]. Choi, T.-M., Wen, X., Sun, X., & Chung, S.-H. (2019). The mean-variance approach for global supply chain risk analysis with air logistics in the blockchain technology era. *Transportation Research Part E: Logistics and Transportation Review*, *127*, 178-191.
[22]. De Jong, A., Zacharias, N. A., & Nijssen, E. J. (2021). How young companies can effectively manage their slack resources over time to ensure sales growth: the contingent role of value-based selling. *Journal of the Academy of Marketing Science*, *49*(2), 304-326.
[23]. de la Torre, R., Corlu, C. G., Faulin, J., Onggo, B. S., & Juan, A. A. (2021). Simulation, optimization, and machine learning in sustainable transportation systems: models and applications. *Sustainability*, *13*(3), 1551.
[24]. De Risi, R., Jalayer, F., De Paola, F., Carozza, S., Yonas, N., Giugni, M., & Gasparini, P. (2020). From flood risk mapping toward reducing vulnerability: the case of Addis Ababa. *Natural Hazards*, *100*(1), 387-415.
[25]. Debasish, A. (2021). Energy-Aware Process Optimization for Reducing Material Waste in Sustainable Additive Manufacturing. *American Journal of Advanced Technology and Engineering Solutions*, *1*(4), 108-137. https://doi.org/10.63125/j0wk7j48
[26]. Debasish, A. (2022). Simulation-Based Thermal Management Framework for Reducing Energy Consumption in Advanced Manufacturing Systems. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 335–381. https://doi.org/10.63125/6fcj7c89
[27]. Debasish, A. (2023). Thermal Behavior Modeling for Predicting Process Instability in Additive Manufacturing Systems. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 485-528. https://doi.org/10.63125/b2nemb48
[28]. Debasish, A. (2024). AI-Based Quality Control Framework for Reducing Rework and Scrap in Additive Manufacturing Production. *Review of Applied Science and Technology*, *3*(04), 402–447. https://doi.org/10.63125/z29tv526
[29]. Egorov, D., Levina, A., Kalyazina, S., Schuur, P., & Gerrits, B. (2020). The challenges of the logistics industry in the era of digital transformation. International conference on technological transformation: A new role for human, machines and management,
[30]. Facchini, F., Oleśków-Szłapka, J., Ranieri, L., & Urbinati, A. (2019). A maturity model for logistics 4.0: An empirical analysis and a roadmap for future research. *Sustainability*, *12*(1), 86.
[31]. Forootani, A., Rastegar, M., & Sami, A. (2022). Short-term individual residential load forecasting using an enhanced machine learning-based approach based on a feature engineering framework: A comparative study with deep learning methods. *Electric Power Systems Research*, *210*, 108119.
[32]. Friesendorf, C., & Uedelhoven, L. (2021). Digital transformation of global mobility markets. In *Mobility in Germany: Digital Transformation, Megatrends and the Evolution of New Business Models* (pp. 1-9). Springer.
[33]. Gerlitz, L., & Meyer, C. (2021). Small and medium-sized ports in the ten-t network and nexus of Europe's twin transition: The way towards sustainable and digital port service ecosystems. *Sustainability*, *13*(8), 4386.
[34]. Giri, C., & Chen, Y. (2022). Deep learning for demand forecasting in the fashion and apparel retail industry. *Forecasting*, *4*(2), 565-581.
[35]. Grassi, L., Figini, N., & Fedeli, L. (2022). How does a data strategy enable customer value? The case of FinTechs and traditional banks under the open finance framework. *Financial Innovation*, *8*(1), 75.
[36]. Guo, H., Xu, H., Tang, C., Liu-Thompkins, Y., Guo, Z., & Dong, B. (2018). Comparing the impact of different marketing capabilities: Empirical evidence from B2B firms in China. *Journal of Business Research*, *93*, 79-89.
[37]. Gupta, A., Singh, R. K., & Mangla, S. K. (2022). Evaluation of logistics providers for sustainable service quality: Analytics based decision making framework. *Annals of Operations Research*, *315*(2), 1617-1664.
[38]. Gurzhiy, A., Paardenkooper, K., & Borremans, A. (2022). Predictive Analytics at an Oil and Gas Company: The Rosneft Case. Algorithms and Solutions Based on Computer Technology: 5th Scientific International Online Conference Algorithms and Solutions based on Computer Technology (ASBC 2021),
[39]. Gutierrez-Franco, E., Mejia-Argueta, C., & Rabelo, L. (2021). Data-driven methodology to support long-lasting logistics and decision making for urban last-mile operations. *Sustainability*, *13*(11), 6230.
[40]. Hunke, F., Heinz, D., & Satzger, G. (2022). Creating customer value from data: foundations and archetypes of analytics-based services. *Electronic Markets*, *32*(2), 503-521.
[41]. Imdoukh, M., Ahmad, I., & Alfailakawi, M. G. (2020). Machine learning-based auto-scaling for containerized applications. *Neural Computing and Applications*, *32*(13), 9745-9760.
[42]. Islam, R. U., Ruci, X., Hossain, M. S., Andersson, K., & Kor, A.-L. (2019). Capacity management of hyperscale data centers using predictive modelling. *Energies*, *12*(18), 3438.
[43]. Jandhyala, R. (2021). Intelligent digital supply chains. In *Supply chain management in manufacturing and service systems: Advanced analytics for smarter decisions* (pp. 19-64). Springer.
[44]. Jiménez-Preciado, A. L., Venegas-Martínez, F., & Ramírez-García, A. (2022). Stock portfolio optimization with competitive advantages (moat): A machine learning approach. *Mathematics*, *10*(23), 4449.
[45]. Kalinowski, M., Lopes, H., Teixeira, A. F., da Silva Cardoso, G., Kuramoto, A., Itagyba, B., Batista, S. T., Pereira, J. A., Silva, T., & Warrak, J. A. (2020). Lean r&d: An agile research and development approach for digital transformation. International Conference on Product-Focused Software Process Improvement,
[46]. Khan, W., Liao, J. Y., Walker, S., & Zeiler, W. (2022). Impact assessment of varied data granularities from commercial buildings on exploration and learning mechanism. *Applied Energy*, *319*, 119281.
[47]. Kim, M.-S., & Thapa, B. (2018). Relationship of ethical leadership, corporate social responsibility and organizational performance. *Sustainability*, *10*(2), 447.

[48]. Kraus, M., & Feuerriegel, S. (2019). Forecasting remaining useful life: Interpretable deep learning approach via variational Bayesian inferences. *Decision Support Systems*, *125*, 113100.
[49]. Krizanova, A., Gajanova, L., & Nadanyiova, M. (2018). Design of a CRM level and performance measurement model. *Sustainability*, *10*(7), 2567.
[50]. Kwak, K., & Kim, N. (2020). Concentrate or disperse? The relationship between major customer concentration and supplier profitability and the moderating role of insider ownership. *Journal of Business Research*, *109*, 648-658.
[51]. Lempp, M., & Siegfried, P. (2022). Urban Mobility Revolution: A Quantitative Analysis. In *Automotive Disruption and the Urban Mobility Revolution: Rethinking the Business Model 2030* (pp. 55-94). Springer.
[52]. Li, L., Chi, T., Hao, T., & Yu, T. (2018). Customer demand analysis of the electronic commerce supply chain using Big Data. *Annals of Operations Research*, *268*(1), 113-128.
[53]. Li, X., Feng, H., Zhao, S., & Carter, D. A. (2021). The effect of revenue diversification on bank profitability and risk during the COVID-19 pandemic. *Finance Research Letters*, *43*, 101957.
[54]. Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable ai: A review of machine learning interpretability methods. *Entropy*, *23*(1), 18.
[55]. Lingelbach, K., Lingelbach, Y., Otte, S., Bui, M., Künzell, T., & Peissner, M. (2021). Demand forecasting using ensemble learning for effective scheduling of logistic orders. International Conference on Applied Human Factors and Ergonomics,
[56]. Liu, S.-C., Lai, J., Huang, J.-Y., Cho, C.-F., Lee, P. H., Lu, M.-H., Yeh, C.-C., Yu, J., & Lin, W.-C. (2021). Predicting microvascular invasion in hepatocellular carcinoma: a deep learning model validated across hospitals. *Cancer Imaging*, *21*(1), 56.
[57]. Liu, Y., Fan, S., Xu, S., Sajjanhar, A., Yeom, S., & Wei, Y. (2022). Predicting student performance using clickstream data and machine learning. *Education Sciences*, *13*(1), 17.
[58]. Mahbubul, I., Mahmudul, I., Shifa, T., Mahfuza, T., & Shima Ali, S. (2022). Exploring The Potential of Visual Narratives and Interactive Gamified Communication Systems to Prevent Disease, Counter Health Misinformation, And Strengthen Public Trust in Low-Literacy Communities: A Literature Review. *American Journal of Interdisciplinary Studies*, *3*(02), 165-220. https://doi.org/10.63125/zmqs4n64
[59]. Md Anisur Rahman, M. (2024). State Income Tax Differentials and High-Skilled Labor Mobility in The United States: A Time Series Analysis. *American Journal of Interdisciplinary Studies*, *5*(04), 222-267. https://doi.org/10.63125/295bsn47
[60]. Md Ashfaq, S., & Abdullah Mohammad, S. (2022). CNT/PANI Composite Electrodes for Flexible Fiber-Shaped Dye-Sensitized Solar Cells: Synthesis, Characterization, and Photovoltaic Performance. *American Journal of Interdisciplinary Studies*, *3*(02), 221-274. https://doi.org/10.63125/m334mf41
[61]. Md Asif Ali Sheak, A., & Risha, A. (2022). The Role of IOT-Driven Energy Analytics in Improving Operational Efficiency and Sustainability in Modern Industrial Facilities. *American Journal of Interdisciplinary Studies*, *3*(04), 807-841. https://doi.org/10.63125/gm63q956
[62]. Md Kaium, H. (2023). A GIS-Based Multi-Criteria Decision-Support and Land-Suitability Optimization Framework for Climate-Resilient Urban and Regional Planning. *American Journal of Data Science and Analytics*, *4*(06), 43–103. https://doi.org/10.63125/wvq0sx87
[63]. Md Kaium, H. (2024). Machine Learning Enhanced GIS And AHP Modeling for High-Resolution Flood-Susceptibility Prediction, Vulnerability Assessment, and Risk Classification. *Review of Applied Science and Technology*, *3*(02), 92–140. https://doi.org/10.63125/fz8hv245
[64]. Md Mahamud Pervaz, P. (2023). Integration of MES, SAP, And PI Historian Platforms for Real-Time Industrial Asset Performance Management. *Review of Applied Science and Technology*, *2*(04), 393–445. https://doi.org/10.63125/9ebtas97
[65]. Md. Abdur, R., & Iftekhar, A. (2021). Customer Retention Forecasting in Mobile Wallet Services Using Neural Networks: A Comparative Quantitative Study. *International Journal of Business and Economics Insights*, *1*(4), 70-102. https://doi.org/10.63125/dyrpc387
[66]. Md. Rashedul, I., & Md Kaium, H. (2022). Development Of a Three-Dimensional Indoor Geospatial Network and Address Geocoding Framework for Next-Generation Emergency Dispatch and Response-Time Optimization. *Journal of Sustainable Development and Policy*, *1*(02), 245-292. https://doi.org/10.63125/tcw3mp48
[67]. Md. Shakhawat, H., & Md, F. (2024). Impact of Phishing Simulation Campaigns and Security Awareness Training on Human Risk Mitigation in Financial Institutions: A Quantitative Study. *American Journal of Data Science and Analytics*, *5*(10), 48-85. https://doi.org/10.63125/4j8h1484
[68]. Mirete-Ferrer, P. M., Garcia-Garcia, A., Baixauli-Soler, J. S., & Prats, M. A. (2022). A review on machine learning for asset management. *Risks*, *10*(4), 84.
[69]. Mitra, A., Jain, A., Kishore, A., & Kumar, P. (2022). A comparative study of demand forecasting models for a multi-channel retail company: a novel hybrid machine learning approach. Operations research forum,
[70]. Mohammad Abir Chowdhury, S. (2023). Role of AI-Assisted Risk Assessment in Enhancing Audit Planning Efficiency and Risk Prioritization in Public-Safety Organizations. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 529–576. https://doi.org/10.63125/qd6nn437
[71]. Mohammad Abir Chowdhury, S., & Tonay, P. (2022). An Empirical Analysis of the Impact of Robotic Process Automation and Continuous Auditing on the Timeliness And Accuracy Of Financial Reporting Assurance. *International Journal of Business and Economics Insights*, *2*(4), 73-126. https://doi.org/10.63125/7p8dcr47

[72]. Mohammad Badrul, A., & Md. Mominul, H. (2023). Machine Learning–Based Assessment of Environmental and Public Health Risks in Sewerage Sludge Management Systems. *American Journal of Advanced Technology and Engineering Solutions*, *3*(02), 33-77. https://doi.org/10.63125/y3yfxa92

[73]. Mohammad Shoriful Hossan, S. (2022). A Performance-Based Framework for Arc-Flash Hazard Reduction and Electrical Reliability Improvement in Brownfield Industrial Distribution Systems. *Review of Applied Science and Technology*, *1*(04), 447-463. https://doi.org/10.63125/a70vj720

[74]. Molla Al Rakib, H., & Mohammad Towhidul, I. (2023). Cyber-Resilient Marketing Intelligence Systems for Fraud Detection and Data Integrity in Digital Campaigns. *American Journal of Scholarly Research and Innovation*, *2*(02), 485–518. https://doi.org/10.63125/b53pgr11

[75]. Moons, K., Waeyenbergh, G., & Pintelon, L. (2019). Measuring the logistics performance of internal hospital supply chains–a literature study. *Omega*, *82*, 205-217.

[76]. Moro-Visconti, R. (2022). Digital scalability and growth options. In *The valuation of digital intangibles: Technology, marketing, and the metaverse* (pp. 85-139). Springer.

[77]. Moroff, N. U., Kurt, E., & Kamphues, J. (2021). Machine Learning and statistics: A Study for assessing innovative demand forecasting models. *Procedia Computer Science*, *180*, 40-49.

[78]. Mosavi, A., Faghan, Y., Ghamisi, P., Duan, P., Ardabili, S. F., Salwana, E., & Band, S. S. (2020). Comprehensive review of deep reinforcement learning methods and applications in economics. *Mathematics*, *8*(10), 1640.

[79]. Mouzakitis, S., Kontzinos, C., Tsapelas, J., Kanellou, I., Kormpakis, G., Kapsalis, P., & Askounis, D. (2022). Enabling maritime digitalization by extreme-scale analytics, ai and digital twins: the vesselai architecture. Proceedings of SAI Intelligent Systems Conference,

[80]. Mst Kaniz, F. (2022). Language Access and Health Equity: Role of Multilingual Administrative Staff in Reducing Healthcare Disparities Among South Asian Immigrant Communities in The U.S. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 274–315. https://doi.org/10.63125/xqhss825

[81]. Nakos, G., Dimitratos, P., & Elbanna, S. (2019). The mediating role of alliances in the international market orientation-performance relationship of SMEs. *International Business Review*, *28*(3), 603-612.

[82]. Nishat, J. (2023). Communicable Disease Prevention in a Post-Pandemic Era: A Global Review of Public Health Education, Surveillance, and Community-Level Interventions. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 630–678. https://doi.org/10.63125/mnqefj36

[83]. Nishat, J., & Jahanara, A. (2024). Machine Learning Approaches for Early Prediction of High-Risk Pregnancy in Underserved U.S. Populations. *American Journal of Interdisciplinary Studies*, *5*(04), 268-309. https://doi.org/10.63125/5gh5df13

[84]. Nishat, J., & Mst Kaniz, F. (2022). A Review of the Impact of Covid-19 On Maternal and Child Health Service Utilization in Bangladesh. *American Journal of Scholarly Research and Innovation*, *1*(02), 223–269. https://doi.org/10.63125/vgv5s742

[85]. Oluyisola, O. E., Bhalla, S., Sgarbossa, F., & Strandhagen, J. O. (2022). Designing and developing smart production planning and control systems in the industry 4.0 era: a methodology and case study. *Journal of Intelligent Manufacturing*, *33*(1), 311-332.

[86]. Peng, G. C., Alber, M., Buganza Tepole, A., Cannon, W. R., De, S., Dura-Bernal, S., Garikipati, K., Karniadakis, G., Lytton, W. W., & Perdikaris, P. (2021). Multiscale Modeling Meets Machine Learning: What Can We Learn? GCY Peng et al. *Archives of Computational Methods in Engineering*, *28*(3), 1017-1037.

[87]. Pham, S. D., Keijzer, H. M., Ruijter, B. J., Seeber, A. A., Scholten, E., Drost, G., van den Bergh, W. M., Kornips, F. H., Foudraine, N. A., & Beishuizen, A. (2022). Outcome prediction of postanoxic coma: a comparison of automated electroencephalography analysis methods. *Neurocritical care*, *37*(Suppl 2), 248-258.

[88]. Polak, P., Nelischer, C., Guo, H., & Robertson, D. C. (2020). "Intelligent" finance and treasury management: what we can expect. *Ai & Society*, *35*(3), 715-726.

[89]. Potočnik, P., Šilc, J., & Papa, G. (2019). A comparison of models for forecasting the residential natural gas demand of an urban area. *Energy*, *167*, 511-522.

[90]. Pronchakov, Y., Prokhorov, O., & Fedorovich, O. (2022). Concept of high-tech enterprise development management in the context of digital transformation. *Computation*, *10*(7), 118.

[91]. Punia, S., & Shankar, S. (2022). Predictive analytics for demand forecasting: A deep learning-based decision support system. *Knowledge-Based Systems*, *258*, 109956.

[92]. Ren, S., Chan, H.-L., & Siqin, T. (2020). Demand forecasting in retail operations for fashionable products: methods, practices, and real case study. *Annals of Operations Research*, *291*(1), 761-777.

[93]. Ribeiro, D. R. S., Mendes, L. G., Forcellini, F. A., & Frazzon, E. M. (2022). Maintenance 4.0: a literature review and SWOT analysis. International Conference on Dynamics in Logistics,

[94]. Rossberg, J. (2019a). Agile project management with azure DevOps. *Apress, Berkeley, CA, USA, Tech. Rep*.

[95]. Rossberg, J. (2019b). Introduction to scrum and agile concepts. In *Agile Project Management with Azure DevOps: Concepts, Templates, and Metrics* (pp. 67-123). Springer.

[96]. Rushforth, R. R., Messerschmidt, M., & Ruddell, B. L. (2020). A systems approach to municipal water portfolio security: A case study of the Phoenix Metropolitan Area. *Water*, *12*(6), 1663.

[97]. Saha, P., Gudheniya, N., Mitra, R., Das, D., Narayana, S., & Tiwari, M. K. (2022). Demand forecasting of a multinational retail company using deep learning frameworks. *IFAC-PapersOnLine*, *55*(10), 395-399.

[98]. Saxena, A., Singh, R., Gehlot, A., Akram, S. V., Twala, B., Singh, A., Montero, E. C., & Priyadarshi, N. (2022). Technologies empowered environmental, social, and governance (ESG): An industry 4.0 landscape. *Sustainability*, *15*(1), 309.

[99]. Schiavone, F., Mancini, D., Leone, D., & Lavorato, D. (2021). Digital business models and ridesharing for value co-creation in healthcare: A multi-stakeholder ecosystem analysis. *Technological forecasting and social change*, *166*, 120647.
[100]. Schroeder, M., & Lodemann, S. (2021). A systematic investigation of the integration of machine learning into supply chain risk management. *Logistics*, *5*(3), 62.
[101]. Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of big data*, *7*(1), 53.
[102]. Seyedan, M., Mafakheri, F., & Wang, C. (2022). Cluster-based demand forecasting using Bayesian model averaging: An ensemble learning approach. *Decision Analytics Journal*, *3*, 100033.
[103]. Sgarbossa, F., Romsdal, A., Oluyisola, O. E., & Strandhagen, J. O. (2022). Digitalization in production and warehousing in food supply chains. In *The digital supply chain* (pp. 273-287). Elsevier.
[104]. Shaiful, M., Anisur, R., & Md, A. (2022). A Systematic Literature Review on the Role of Digital Health Twins in Preventive Healthcare For Personal and Corporate Wellbeing. *American Journal of Interdisciplinary Studies*, *3*(04), 1-31. https://doi.org/10.63125/negjw373
[105]. Spiliotis, E., Makridakis, S., Semenoglou, A.-A., & Assimakopoulos, V. (2022). Comparison of statistical and machine learning methods for daily SKU demand forecasting: E. Spiliotisa et al. *Operational Research*, *22*(3), 3037-3061.
[106]. Stenwig, E., Salvi, G., Rossi, P. S., & Skjærvold, N. K. (2022). Comparative analysis of explainable machine learning prediction models for hospital mortality. *BMC Medical Research Methodology*, *22*(1), 53.
[107]. Stölzle, W., & Häberle, L. (2021). Digital Logistics Platforms—Initial Approaches to Market Segmentation in Light of Traditional and New Providers. In *Digital Business Models in Industrial Ecosystems: Lessons Learned from Industry 4.0 Across Europe* (pp. 105-123). Springer.
[108]. Sun, Y.-Y., Lin, P.-C., & Higham, J. (2020). Managing tourism emissions through optimizing the tourism demand mix: Concept and analysis. *Tourism management*, *81*, 104161.
[109]. Swathy, M., & Saruladha, K. (2022). A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques. *ICT express*, *8*(1), 109-116.
[110]. Tang, Y.-M., Ho, G. T. S., Lau, Y.-Y., & Tsui, S.-Y. (2022). Integrated smart warehouse and manufacturing management with demand forecasting in small-scale cyclical industries. *Machines*, *10*(6), 472.
[111]. Tonay, P., Md Maruf, I., & Al, A. (2024). Artificial Intelligence-Based Decision Support Systems and Managerial Performance. *American Journal of Advanced Technology and Engineering Solutions*, *4*(04), 154-184. https://doi.org/10.63125/wmz8dc67
[112]. Uddin, H. M. M., & Risha, A. (2024). Hybrid Cloud Security and Compliance in U.S. Enterprises: Addressing Data Privacy Risks and Governance Challenges. *American Journal of Scholarly Research and Innovation*, *3*(01), 153–197. https://doi.org/10.63125/qcy0cj46
[113]. Van Giffen, B., Herhausen, D., & Fahse, T. (2022). Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods. *Journal of Business Research*, *144*, 93-106.
[114]. Veerasamy, K., Sanyal, S., Almahirah, M. S., Saxena, M., & Manohar Bhanushali, M. (2022). An investigative analysis for IoT based supply chain coordination and control through machine learning. International Conference on Emerging Technologies in Computer Engineering,
[115]. Wagner, T., Herrmann, C., & Thiede, S. (2018). Identifying target oriented Industrie 4.0 potentials in lean automotive electronics value streams. *Procedia cirp*, *72*, 1003-1008.
[116]. Wolff, S., O'Donncha, F., & Chen, B. (2020). Statistical and machine learning ensemble modelling to forecast sea surface temperature. *Journal of Marine Systems*, *208*, 103347.
[117]. Woschank, M., Rauch, E., & Zsifkovits, H. (2020). A review of further directions for artificial intelligence, machine learning, and deep learning in smart logistics. *Sustainability*, *12*(9), 3760.
[118]. Wu, M.-E., Syu, J.-H., Lin, J. C.-W., & Ho, J.-M. (2021). Portfolio management system in equity market neutral using reinforcement learning. *Applied Intelligence*, *51*(11), 8119-8131.
[119]. Wulf, J. (2020). Development of an AHP hierarchy for managing omnichannel capabilities: a design science research approach. *Business Research*, *13*(1), 39-68.
[120]. Xie, M. E., & Ou, Q. (2022). Explore the application of big data technology in modern enterprise logistics management. International Conference of Artificial Intelligence, Medical Engineering, Education,
[121]. Xu, S., Zou, S., Huang, J., Yang, W., & Zeng, F. (2022). Comparison of different approaches of machine learning methods with conventional approaches on container throughput forecasting. *Applied Sciences*, *12*(19), 9730.
[122]. Yang, Y., Heppenstall, A., Turner, A., & Comber, A. (2020). Using graph structural information about flows to enhance short-term demand prediction in bike-sharing systems. *Computers, Environment and Urban Systems*, *83*, 101521.
[123]. Yu, B., Li, J., Liu, C., & Sun, B. (2022). A novel short-term electrical load forecasting framework with intelligent feature engineering. *Applied Energy*, *327*, 120089.
[124]. Yu, L., Du, B., Hu, X., Sun, L., Han, L., & Lv, W. (2021). Deep spatio-temporal graph convolutional network for traffic accident prediction. *Neurocomputing*, *423*, 135-147.
[125]. Zhang, Y., Yang, D., Liu, Z., Chen, C., Ge, M., Li, X., Luo, T., Wu, Z., Shi, C., & Wang, B. (2021). An explainable supervised machine learning predictor of acute kidney injury after adult deceased donor liver transplantation. *Journal of translational medicine*, *19*(1), 321.