Frontiers in Computer Science and Artificial Intelligence
DOI: 10.32996/fcsai
Journal Homepage: www.al-kindipublisher.com/index.php/fcsai

FCSAI
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Deep Learning for Intelligent Supply Chain Optimization: Enhancing Operational Efficiency and Waste Reduction in U.S. Service Industries

Md Rasibul Islam[1], Md Toushif Pramanik[2] and MD Abdul Fahim Zeeshan[3]

[1]Department of Public Administration, Gannon University, Erie, PA, USA
[2]Master of Embedded Software Engineering, Gannon University, Erie, PA, USA
[3]Master of Arts in Strategic Communication, Gannon University, Erie, PA, USA
Corresponding Author: Md Rasibul Islam, **Email:** islam011@gannon.edu

| ABSTRACT
U.S. service industries face persistent inefficiencies driven by volatile demand patterns, short product life cycles, and operational fragility across logistics networks. These conditions create a structural forecasting challenge where traditional statistical models struggle to capture event-driven variability, leading to systematic overstocking, stockouts, and waste. This study addresses that challenge by developing a deep learning–driven supply chain optimization framework that integrates high-resolution calendar features, lagged demand patterns, and sequence-based neural forecasting architectures. Using the M5 dataset as a proxy for service-sector operational behavior, we benchmark classical models (Naive, Linear Regression, LightGBM) against LSTM, N-Beats, and a lightweight Transformer model designed for Colab-scale experimentation. We evaluate the forecasting outputs through a simulation of inventory policies including (s, S) and base-stock systems, capturing holding, shortage, and waste costs under multiple demand scenarios. Results show that Transformer-based models consistently outperform statistical baselines on multi-horizon forecasting, improving RMSE and MAE across varied temporal contexts. When integrated with inventory policies, deep models reduce stockouts and expired-inventory waste while lowering total operational costs compared with heuristic or statistical forecasting systems. Sensitivity and ablation analyses confirm the value of calendar encodings, lag structure, and longer receptive fields in improving predictive stability. Computational profiling confirms that optimized neural architectures can run efficiently on edge-constrained environments such as Colab GPUs. These findings demonstrate that deep learning–enabled forecasting provides a practical and high-impact path toward intelligent supply chain optimization in U.S. service industries.

| KEYWORDS

Supply Chain Optimization, Deep Learning Forecasting, Transformer Models, Inventory Simulation, Demand Forecasting, U.S. Service Industries

| ARTICLE INFORMATION

ACCEPTED: 15 November 2025 **PUBLISHED:** 26 November 2025 **DOI:** 10.32996/fcsai.2025.4.2.5

1. Introduction

Supply chains that support U.S. service industries have grown so complex that many long-standing forecasting and inventory methods reveal their flaws whenever demand shifts quickly or reacts to events, seasonal swings, or outside disruptions. Fields like food service, retail services, hospitality, and healthcare logistics deal with environments where small changes in customer behavior can spiral into real operational gaps that show up as stockouts, wasted inventory, or resources tied up in the wrong places. These problems become even sharper because modern supply chains are tightly connected, and this creates a need for forecasting systems that can respond to uncertainty and structural instability using data in a more adaptive way. Ivanov and Dolgui note that digital supply chain systems increasingly operate under constant disruption, which means agility and predictive intelligence shape their ability to stay resilient [12]. Traditional forecasting inside service operations still leans heavily on linear patterns and heuristic rules, which limit how well these systems capture nonlinear changes, unusual calendar effects, or rare demand spikes that carry high costs. Hyndman and Athanasopoulos have pointed out that classical statistical forecasting, while important as a foundation, struggles with the high-dimensional, sparse, or irregular datasets often found in service-focused supply chains [9]. These gaps push the conversation toward deep learning and advanced analytics that can recognize temporal behavior, uncover hidden structure in demand signals, and help service organizations cut back on avoidable inefficiencies.

1.1 Background and Motivation

Service industries throughout the United States rely on logistics networks that must anticipate demand with precision, yet these networks often struggle with unpredictable consumption habits, shifting labor availability, local disruptions, and wider market fluctuations. This volatility steadily erodes the reliability of traditional planning and widens the distance between expected demand and what actually happens. Shawon et al. argue that resilience in U.S. regions hinges on machine learning models that recognize early signs of stress in logistics systems, giving decision makers time to act before issues grow into larger failures [20]. Their point highlights how static forecasting frameworks fall short because they cannot learn from new patterns or absorb information from sources such as event calendars, store characteristics, or sudden behavioral changes. The challenges grow in sectors dealing with perishable goods or time-sensitive consumption cycles, where mistakes in forecasting turn into costly waste or unmet demand. Hasan et al. explain that supplier risk management supported by strong machine learning analytics can reveal weaknesses early, long before they show up as operational losses [7]. Their work illustrates the potential of data-driven tools to strengthen supplier networks while calling attention to the limits of traditional forecasting, which often misses nonlinear demand relationships or shifting risk patterns.

Disruptions in these supply chains can come from many directions, including geopolitical conflicts, transportation delays, economic shocks, or unusual procurement behavior. Advanced analytics applied to transactional, logistical, and time-series data help organizations recognize early signals of structural drift or emerging risks. Sizan et al. show this through unsupervised ensemble learning on graph-based transaction data, demonstrating how these methods detect new patterns that rule-based systems fail to notice [23]. Although their work focuses on money-laundering detection, the broader idea carries over to supply chains. Unsupervised and semi-supervised learning can reveal changes in regional demand intensity, unusual purchasing trends, or consumption bursts that might disrupt inventory planning.

Deep learning stands out in this setting because of its ability to model nonlinear temporal behavior, track long-range patterns, and include external signals without hand-crafted features. Ivanov and Dolgui note that digital supply chain twins and resilience-focused planning depend on predictive intelligence to function under persistent uncertainty [12]. The push toward real-time analytics and autonomous decision-making fits naturally with deep learning, which can capture subtle relationships that escape classical models. Growing operational complexity in service industries and rapid swings in demand create an urgent need for forecasting systems that learn from large and varied datasets and provide predictions that fit a wide range of real-world contexts. As the literature shows, traditional approaches no longer meet the demands placed on U.S. service industries, and the move toward adaptive forecasting powered by intelligent models has become necessary for staying competitive and avoiding preventable inefficiencies.

1.2 Importance of This Research

Reliable demand forecasting plays a central role in helping service industries limit excess stock, reduce waste, improve service quality, and keep operations steady during unpredictable shifts in consumer behavior. The stakes are high because products in these sectors often have short life cycles, and logistics networks must respond to consumption patterns that can change quickly. Poor forecasting can leave businesses with spoiled goods or empty shelves, both of which carry real costs. The challenge grows as fluctuations become more frequent, shaped by local events, holidays, labor patterns, and broader economic forces that create large forecast errors. Traditional models often fall short in these situations because they rely on simple temporal structures and offer limited protection against uncertainty. This research responds to that gap by applying deep learning architectures to supply chain forecasting, enabling models to learn nonlinear patterns and multi-period interactions between demand and external factors.

The significance of this work also shows up in the wider movement toward intelligent supply chain systems that support operations through predictive insights. Hasan et al. note that supply chain finance, especially supplier credit decisions, benefits from explainable AI systems that improve decision quality when data is sparse or irregular [8]. Their findings highlight how classical analytics lose effectiveness under uncertainty and how models that learn from complex patterns offer a stronger alternative. Although their work centers on financial approval rather than inventory planning, the connection is clear: being able to understand uncertainty and handle multi-dimensional features strengthens operational reliability in both cases. By adopting deep learning forecasting, supply chain managers in U.S. service industries can make decisions based on predictive insight rather than reactive responses, improving how resources are allocated and reducing risks before they reach critical levels. The study also matters because it ties forecasting accuracy to direct operational outcomes. Many forecasting studies stay focused on error metrics without addressing how mistakes translate into costs such as storage expenses, spoilage, or backorders. This research takes the additional step of integrating forecasting with inventory simulations to measure how improvements in prediction affect real supply chain performance.

1.3 Research Objectives and Contributions

This study sets out to build a forecasting system that uses deep learning to improve demand prediction in U.S. service industries while also examining how better forecasts influence day-to-day operations and costs. The aim is to create a full pipeline that starts with temporal modeling using modern sequence architectures, compares model performance against established baselines, and ends with inventory simulations that measure how well different policies reduce waste, shortages, and holding costs. By developing this entire workflow, the research offers both technical guidance and practical clarity on how deep learning can support supply chain decisions in settings marked by volatility and short planning horizons. The work adds to existing research by drawing a direct line between predictive accuracy and real operational outcomes. Instead of judging models only by statistical scores, the study places their forecasts inside inventory systems to see how they affect costs in practice. This reflects how organizations operate, since the real value of a forecast lies in how it shapes stocking decisions, spoilage rates, and service levels. The study also shows how temporal feature engineering, such as lag variables, rolling windows, and calendar-based inputs, strengthens deep learning performance in datasets common to service industries. This supports the development of forecasting pipelines that remain transparent and interpretable. The design also adapts well to other time-sensitive sectors such as hospitality and healthcare logistics, where demand volatility follows similar patterns and predictive tools could reshape operational planning. The findings reinforce the idea that combining deep neural forecasting with inventory simulations creates concrete opportunities for operational improvement and offers a practical foundation for service organizations interested in adopting more intelligent supply chain systems.

2. Literature Review

2.1 Supply Chain Forecasting in Service Industries

Forecasting demand in service industries like food service, retail operations, hospitality, and healthcare logistics has grown more complicated as consumption patterns shift faster and expectations for quick, reliable operations rise. These sectors often deal with sharp swings in demand that traditional models struggle to capture, especially when customer preferences change quickly or when outside factors such as weather, holidays, promotions, or local economic shifts influence purchasing behavior. Perishability raises the stakes in food service and hospitality since even small forecasting mistakes can lead to shortages that frustrate customers or to excess stock that spoils or clogs storage. Healthcare logistics faces its own version of this volatility, with demand for supplies and medications shifting in response to seasonal illnesses, emergencies, or sudden surges within facilities. These situations create an ongoing need for forecasting tools that can recognize subtle changes across multiple time scales and react quickly to new patterns.

The move toward data-driven supply chain management has encouraged researchers to explore how analytics and machine learning can improve forecasting accuracy and support more responsive operations. Choi et al. point out that big data analytics has become a key part of operations management by helping organizations merge large amounts of logistical, transactional, and time-based data to uncover patterns, anticipate needs, and plan more effectively [3]. Their work shows that the service sector, even more than many manufacturing environments, depends on adaptive models that can learn from high-frequency, heterogeneous datasets. Since food service and retail generate detailed point-of-sale records and hospitality and healthcare increasingly rely on digital operational systems, forecasting tools must handle large streams of diverse information. As service organizations adopt more IoT devices, inventory tracking tools, and behavioral data systems, the amount and complexity of available information grow quickly. This creates opportunities for advanced time-series models but also strains traditional forecasting methods that cannot manage high dimensionality or the irregular structure of modern datasets.

Volatility remains a constant source of operational loss throughout service industries. Unpredictable consumer demand, shifting pricing strategies, staffing challenges, and supply delays all contribute to nonlinear relationships that traditional models often fail to represent. The reliance on accurate forecasting has grown as supply chains become more interconnected and as disruptions spread faster across networks. While big data analytics lays the foundation for making sense of this complexity, the industry still needs forecasting methods that can capture long-term temporal patterns, respond to sudden anomalies, and represent how calendar events interact with consumption behavior [3]. Growing evidence shows that service supply chains require more advanced techniques, including deep learning models, to manage the scale, pace, and irregularity of demand in these fast-moving environments.

2.2 Traditional Forecasting Approaches

Traditional forecasting techniques have long shaped supply chain planning in both manufacturing and service settings. Methods such as ARIMA, exponential smoothing, and ETS models have been widely adopted because they are straightforward to interpret, computationally efficient, and designed for stable univariate time series. These models assume that demand follows a consistent temporal structure, allowing them to track linear trends, seasonal cycles, and smoother recurring patterns. In earlier years, when service-sector demand followed predictable weekly or monthly rhythms, methods like Holt-Winters or ARIMA often delivered acceptable results. As service industries have grown more complex, their limitations have become harder to ignore. Makridakis et al. note that while statistical models perform adequately in low-noise, stable environments, they struggle when faced with nonlinear dependencies, abrupt shifts, or the high dimensionality common in modern forecasting problems [16]. Their comparative studies raise concerns about how well these traditional tools can scale or adapt in settings where structural changes happen rapidly.

Hierarchical forecasting has been used in situations where demand needs to be aggregated or broken down across different levels of a supply chain. Hyndman et al. developed optimal combination approaches that improve coherence across hierarchical time series, showing how bottom-up or top-down methods can produce more consistent forecasts across different levels of granularity [10]. Even with these advantages, these methods still depend on the assumptions built into classical statistical frameworks. When demand behaves irregularly or includes multiple seasonal patterns, these models often fall short because they rely heavily on linear structures and autocorrelation that may not exist in volatile service environments.

The shortcomings of traditional forecasting tools are most evident when the data is sparse, intermittent, or shaped by many external factors. In many service-industry contexts, demand can shift abruptly due to promotions, local weather, regional socioeconomic changes, or sudden changes in customer sentiment. ARIMA and exponential smoothing models rely on stable trend and seasonal structures, which makes them a poor fit for multi-regime or rapidly evolving behaviors. Makridakis et al. highlight further concerns about robustness when these models encounter complex or multi-scale volatility that appears in real-world data [16]. Shivogo notes similar challenges in credit scoring, where concept drift and shifting population behavior degrade the performance of static statistical models, suggesting that supply chain forecasting faces an analogous problem as temporal dynamics change over time [21]. This reinforces the broader view that the assumptions underlying classical forecasting methods no longer align with the realities of modern service supply chains, prompting a shift toward more flexible and nonlinear modeling approaches.

2.3 Machine Learning Approaches

Machine learning has become a major part of time series forecasting because it can pick up nonlinear patterns, learn from larger datasets, and work with a wider range of features than classical statistical tools. Methods like gradient boosting, random forests, and regression-based models offer more room to capture interactions and relationships that older models tend to miss. Bandara et al. introduced a clustering-based recurrent neural network that groups similar time series so the model can learn shared patterns, which helps when working with large collections of related demand signals [2]. Their findings show how machine learning can uncover hidden structures that often slip past statistical approaches, especially when multiple time series influence one another or when demand behavior differs across products or regions. Machine learning models also handle external variables well. Calendar effects, promotions, weather, and socio-economic conditions often shape demand in service industries, and ML systems tend to incorporate these factors more naturally. Laptev et al. showed that ML models can forecast extreme events with better accuracy than traditional tools, noting how neural networks and hybrid approaches can recognize rare but important spikes that statistical models usually miss [13]. Their work at Uber highlights how these systems adjust to large-scale, high-frequency datasets that are now common across service operations. Still, many ML models fall short when the task requires understanding very long sequences because they struggle to capture extended temporal relationships.

Studies in financial and socioeconomic forecasting help illustrate both the strengths and the limits of ML. Islam et al. built models for predicting cryptocurrency prices, showing how ML can adapt to highly volatile time series by learning nonlinear behaviors [11]. Ray demonstrated that ML captures how stock, bond, and foreign exchange markets interact during financial crises, suggesting that these methods work well when multiple temporal signals overlap [18]. Reza et al. applied ML to income prediction and disparity detection across population groups, emphasizing its ability to work with complex, multifactor datasets [19]. These areas mirror challenges found in service industries, where demand shifts quickly and depends on outside influences. Even so, many ML methods still lean heavily on manual feature engineering and cannot easily learn long-range temporal structure on their own. This has led to growing interest in deep learning methods that can model these deeper patterns directly.

2.4 Deep Learning for Time Series

Deep learning has reshaped time series forecasting by allowing models to learn temporal structure directly from raw or lightly processed data. LSTMs, GRUs, sequence-to-sequence models, and Transformer architectures have shown clear gains in accuracy across retail, logistics, energy, and financial markets. Lim and Zohren's survey reports that deep learning models consistently outperform traditional forecasting approaches when the data is high-dimensional, when the forecasting horizon is long, or when the underlying dynamics are strongly nonlinear [15]. Their work shows that deep learning models are especially valuable in settings where external variables, seasonal behavior, and long-term dependencies all influence demand, which is often the case in service industries.

Sequence models such as LSTMs capture long-term dependencies through recurrent connections, allowing them to learn patterns that unfold over many steps. Smyl proposed a hybrid approach that blends exponential smoothing with a recurrent neural network, achieving state-of-the-art results in international forecasting competitions [24]. This work highlights how deep learning can incorporate useful ideas from classical methods while still modeling complex sequence behavior. Transformer models extended these capabilities by replacing recurrence with attention mechanisms. Vaswani et al. introduced the Transformer architecture and showed that attention mechanisms can capture long-range temporal relationships more efficiently than recurrent networks while avoiding the computational bottlenecks they create [26]. These ideas have shaped much of modern time series research because Transformers can identify important temporal segments and incorporate external variables with minimal manual input.

Deep learning applications span far beyond retail and supply chain operations. Shovon explored smart grid planning for low-voltage energy systems and showed how deep models help manage temporal load patterns and long-term capacity needs in growing urban settings [25]. Aashish et al. experimented with machine learning for environmentally focused cybersecurity forecasting, integrating energy metrics into anomaly detection models [1]. These examples reflect how deep learning thrives in settings that involve noisy, diverse, and feature-rich temporal signals. As service sectors gather more data that includes event markers, store attributes, promotional history, and seasonal cycles, deep learning provides the level of flexibility needed to support reliable, long-range forecasting.

2.5 Inventory Optimization and Simulation

Inventory optimization sits at the heart of supply chain work in service industries, where teams try to keep enough stock on hand without carrying wasteful excess. Classical models such as the (s, S) policy and the base stock approach have shaped this field for many years. These models guide replenishment by tracking inventory positions relative to specific reorder points. Silver et al. offer a detailed overview of these policies and describe the mathematical conditions that make each one effective in different supply chain settings [22]. Their work stresses the value of linking demand forecasts with inventory rules so decisions reflect real market behavior. Service industries regularly face shifting demand due to seasonal events, promotions, or unexpected supply issues, which raises the stakes for accurate forecast integration. Inventory outcomes tend to hinge on the quality of those forecasts. Missed signals in forecasting spill directly into stock decisions and create shortages or unnecessary surplus that raise costs. Graves and Willems studied how strategic inventory placement responds to demand that changes over time and showed that anticipating variability and adjusting stock across supply chain nodes improves resilience and service levels [6]. Their findings highlight the importance of understanding how demand evolves across time when planning stock positions. Service industries operate in volatile environments, so simulation becomes an essential tool for testing how inventory behaves under different forecasting conditions.

Recent computational advances allow researchers to plug machine learning and deep learning forecasts into inventory simulations, creating more realistic assessments of cost outcomes. This matters for service industries because it ties improvements in forecasting accuracy to real operational gains. Better forecasts reduce safety stock, lower waste from expired items, and improve customer experience by decreasing stockouts. Traditional simulations often rely on static assumptions or overly simple demand distributions that fail to capture real-world swings. Modern simulations informed by advanced forecasting models support richer scenario analysis that includes event-driven surges, sudden shifts in customer behavior, or disruptions from external factors. Silver et al. point out that linking forecasting with inventory policy selection is essential for building sustainable inventory strategies in high-variability conditions [22]. These ideas set the stage for combining deep learning forecasts with inventory simulation work across service supply chains..

2.6 Gaps and Challenges

Research in forecasting and inventory optimization has moved forward, yet several limits continue to shape how these methods apply to service industries. One key gap involves the limited number of studies that join deep learning forecasting with inventory cost simulations. Deep learning has delivered strong forecasting results across finance, energy, logistics, and related fields, yet integrated forecasting-to-simulation pipelines for service operations remain rare. Vercellis notes that deep learning in operations management is still growing and that many studies focus on algorithmic performance without exploring how these models influence actual inventory decisions or cost structures [27]. This mirrors a wider pattern in forecasting research where models are often evaluated separately from their operational impact. Another challenge comes from the heavy emphasis on manufacturing within supply chain research. Li et al. observe that machine learning studies often revolve around manufacturing tasks such as production scheduling, defect detection, and supplier selection, leaving service-focused forecasting questions with limited attention [14]. Service industries show demand patterns that shift more rapidly over time, rely more heavily on event-driven behavior, and involve frequent interactions between customers and operations. These differences mean forecasting models designed for manufacturing rarely fit service environments without extensive adjustments. Service organizations, therefore, lack tools that reflect their day-to-day realities.

A further gap relates to the limited use of real calendar-driven datasets in service-industry forecasting research. Many current studies draw on synthetic data or aggregated time series that lack detailed event markers, which are central to demand patterns in service settings. Das et al. identify similar issues in cybersecurity forecasting, where models without rich contextual inputs tend to miss important signals [4]. Debnath et al. report comparable findings in renewable energy anomaly detection and show that blending environmental and operational variables leads to more reliable results [5]. These observations align with the challenges faced in service supply chains, where holidays, regional events, marketing activities, and similar signals play a major role in shaping demand paths.

3. Methodology

3.1 Dataset Description

The M5 Forecasting Accuracy dataset served as the main empirical base for this study because it offers detailed daily data, rich calendar information, varied price movements, and item-level sales behavior across several U.S. states. While the data comes from Walmart stores, its structure resembles the kinds of operational settings found in many service industries where item and location combinations can swing unpredictably, respond to events, and follow multiple seasonal cycles. The dataset tracks daily sales in California, Texas, and Wisconsin, giving a wide view of how demand plays out in different regions. The core file, sales_train_validation.csv, contains long sales histories for thousands of items, which makes it suitable for both short and long forecasting horizons. The calendar.csv file adds context to each day, including weekdays, months, SNAP indicators, holidays, and named events. These variables matter for service industries because demand often jumps or dips around holidays, cultural events, and regional celebrations. Sell_prices.csv introduces weekly pricing data, allowing the study to examine price elasticity and the influence of promotions, both of which are relevant in service markets where customers may react strongly to cost changes or discounts. Taken together, the goal variable was defined as future daily sales for each item-store combination. This setup lets models learn differences across items and the way demand changes over time within each series. Given the size of the dataset, resource-aware sampling was used to make the analysis manageable in the Colab environment while keeping the essential structure of demand patterns intact.

3.2 Data Preprocessing

Data preprocessing followed a structured pipeline that aligned the merged dataset and ensured it remained consistent and complete. Missing values in the calendar file, especially those tied to event names and event types, were replaced with placeholders that kept category information intact without introducing distortion. Missing price entries were filled with zeros, following common competition practices that treat absent price values as non-promotional or inactive periods. The main sales dataset, which originally used a wide layout, was reshaped into a long format so that each row represented one item on one day. This shift made it possible to align sales, calendar information, and price data along the same timeline. After reshaping, the datasets were merged using day keys and weekly price identifiers. Because several variables were categorical, they were converted into numerical form using label encoders for tree-based models and integer token mappings for deep learning models. This step keeps category identity intact while making the dataset usable across model families. Numerical features, especially those used in neural models, were scaled with MinMaxScaler to support stable training and reduce gradient issues. A supervised learning structure was created by defining sliding historical windows as inputs and sequences corresponding to the forecasting horizon as targets. This transformation turned the raw time series into a form suitable for both classical statistical models and more advanced neural architectures. Window lengths and horizon sizes were selected with each model's needs in mind to balance efficiency with predictive depth.

3.3 Exploratory Data Analysis

The exploratory analysis offers a close look at how the M5 demand data behaves and sets the stage for the forecasting and inventory optimization work that follows. Because the dataset covers several years of item-level daily sales across many product categories, stores, and states, the EDA centers on uncovering temporal structure, odd behavior, and irregular demand patterns that shape both modeling choices and inventory decisions. Every plot in this section was chosen because it connects directly to how the models are designed or how supply chain decisions can be interpreted.

The view of total daily sales over time shows a steady upward movement from 2011 to 2016, along with strong cycles within each year. Sales often fall at the start of the year, rise during spring, and reach their highest levels near the end of the year. This repeated annual shape reflects seasonal buying habits, holiday effects, and familiar cycles in customer behavior. The long-term rise may stem from store openings, larger product assortments, or broader customer reach, each of which pushes overall demand higher. For forecasting, this confirms that the series is not stationary and carries both long-term structure and recurring cycles that simple linear models would not easily capture. In an inventory setting, the late-year peak periods are moments where poor alignment between supply and demand can create either stock shortages or excess inventory. This pushes the need for models that can capture global movement along with repeating seasonal forms, such as LSTM, NBeats, and Transformers.

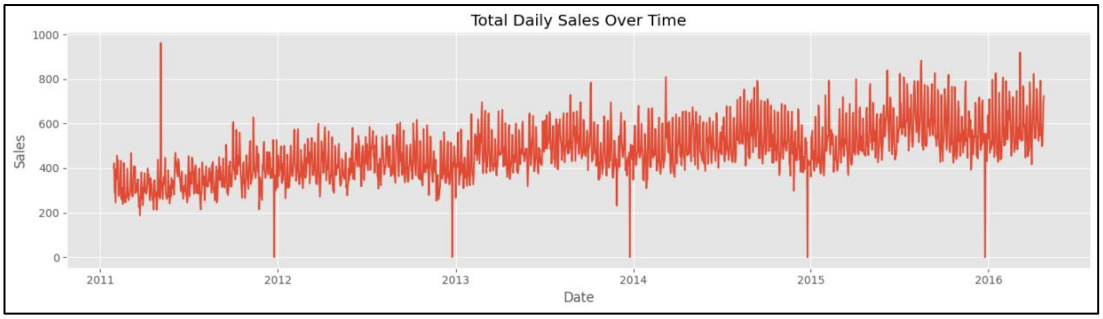

Fig.1: Total daily sales over time

Looking at category-level sales deepens this understanding. Foods, households, and hobbies show similar seasonal swings, but foods dominate the dataset and show the strongest fluctuations. Foods tend to move quickly through the system, often have short shelf lives, and respond sharply to calendar effects, which explains their volatility. Because foods account for such a large portion of total demand, forecasting mistakes in this category carry outsized consequences and can easily ripple through the full system. Category-level variation feeds directly into uncertainty, so high-volume categories require tighter forecasts to reduce waste, while lower-volume ones call for strategies that can handle sparse or intermittent sales.

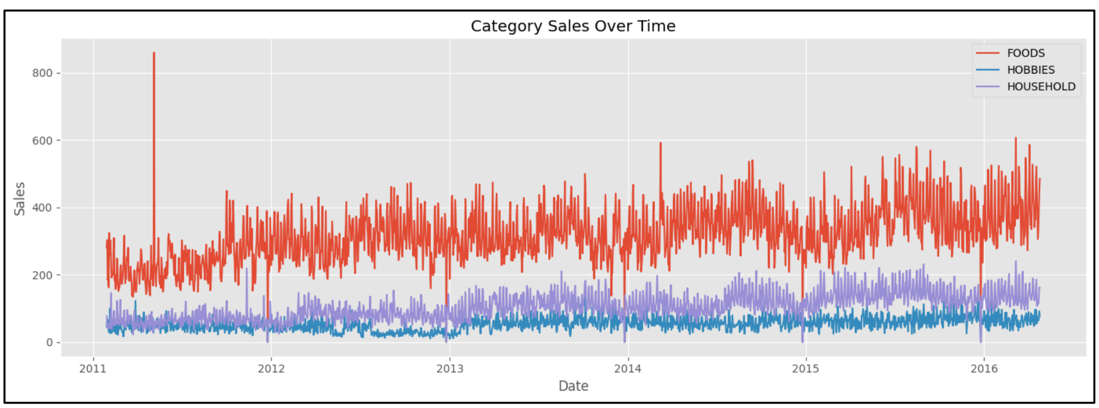

Fig.2: Category sales over time

Weekly patterns also reveal a strong signal. Average sales by weekday show that Mondays are consistently the slowest, with demand climbing through the week and reaching a clear peak on Saturdays before dropping slightly on Sundays. These patterns reflect common shopping behavior, especially the tendency for households to do their main purchasing toward the weekend. In supply chain operations, strong weekly rhythms affect staffing, delivery planning, shelf replenishment, and short-term decisions that depend on knowing when demand will spike. Forecasting models need to reflect this weekly cycle to avoid repeated errors on specific days; otherwise, stockouts will cluster around busy periods and waste will build up during slower days. Monthly patterns add another layer. Sales are lower early in the year and rise through spring, climbing toward their highest points in June, July, and August. This shifts away from the common retail pattern of December peaks, which suggests that the products in this dataset lean toward summer-related demand rather than winter holiday buying. Warmer months often bring higher purchases of perishables, household items, outdoor supplies, and hobby-related goods. Medium-term seasonal movement like this has to be represented in the models because ignoring it distorts multi-month demand forecasts and disrupts upstream purchasing and supplier coordination.

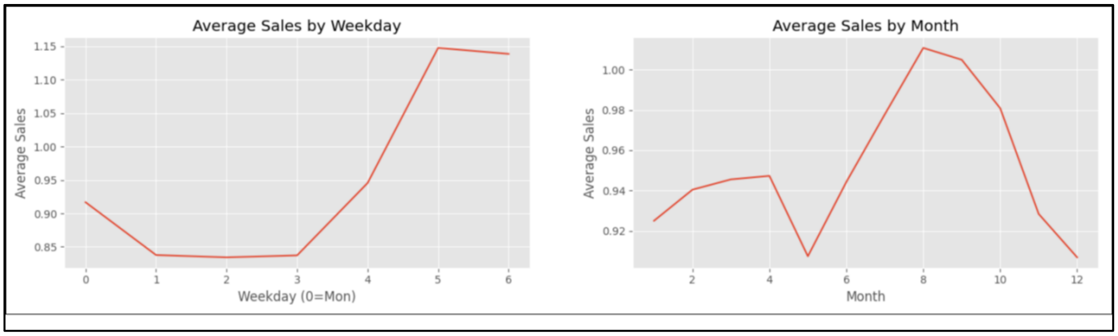

Fig.3: Average sales by weekday and month

A notable feature of the dataset is the unusual distribution of zero-sales days across items. The bimodal shape shows that many items sell nearly every day, while a large set sells rarely or goes long stretches with no sales at all. This mix reflects a combination of fast-moving and slow-moving goods, which is typical in supply chains. Fast-moving items need accurate forecasts to avoid shortages, while slow-moving ones carry a real risk of high holding costs or eventual obsolescence. Intermittent demand is especially difficult because predicting zeros accurately is often harder than predicting positive sales. This makes traditional models stumble. The large share of slow-moving inventory reinforces the need for stronger feature engineering, hybrid modeling methods, and inventory strategies that are adapted to item-level behavior.

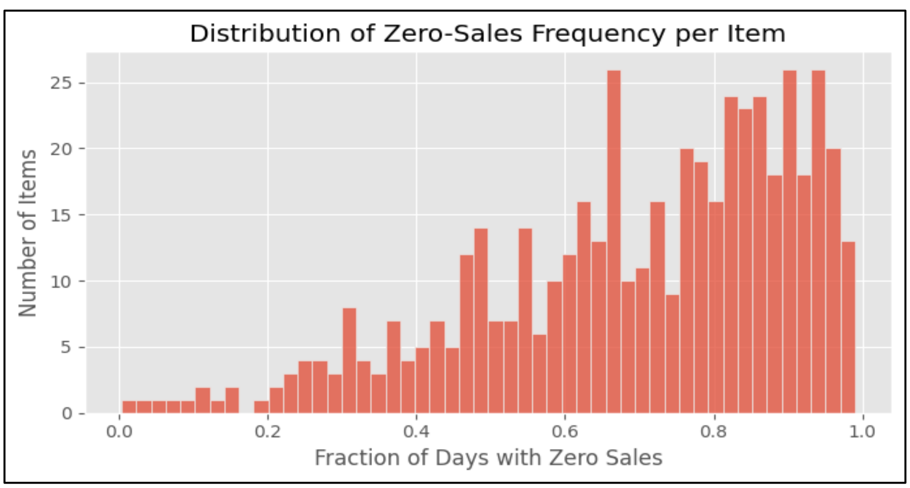

Fig.4: Distribution of zero sales days per item

Demand volatility across items adds another source of risk. Most items show low volatility and stable daily sales, while a long tail of items shows high variability driven by sudden spikes or inconsistent patterns. These high-volatility items tend to create bottlenecks in supply chain performance. Even small forecasting mistakes for these items can lead to costly stockouts or excess inventory. This makes it clear that a single forecasting approach will not work for every item. High-volatility products need models that can capture irregular behavior and nonlinear interactions, while stable items can be served with simpler or faster models. From an inventory standpoint, volatile items require higher safety stock or more flexible replenishment rules to protect against uncertainty.

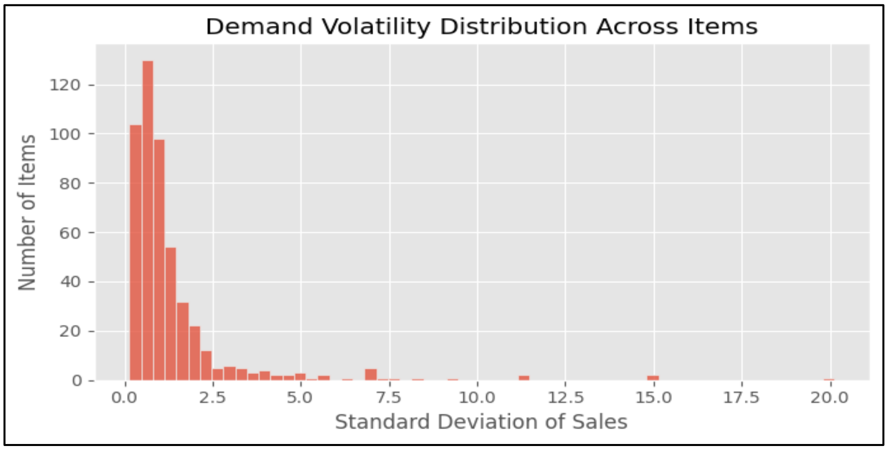

Fig.5: Demand volatility distribution across items

The analysis of sales on major event days shows how strongly special dates shape demand. Days linked to events such as the Super Bowl, Valentine's Day, and Purim End show much higher average sales than normal days. These spikes highlight the influence of promotions, cultural traditions, and social behavior. Event-driven demand is difficult to capture without explicit features because the timing and magnitude of these spikes are not encoded in past values alone. For operations, these are high-risk days where good forecasts matter the most to prevent missed sales or unnecessary overstock. The behavior observed here supports the use of event-related features and the development of deep learning models that can absorb sharp, nonlinear changes in demand.

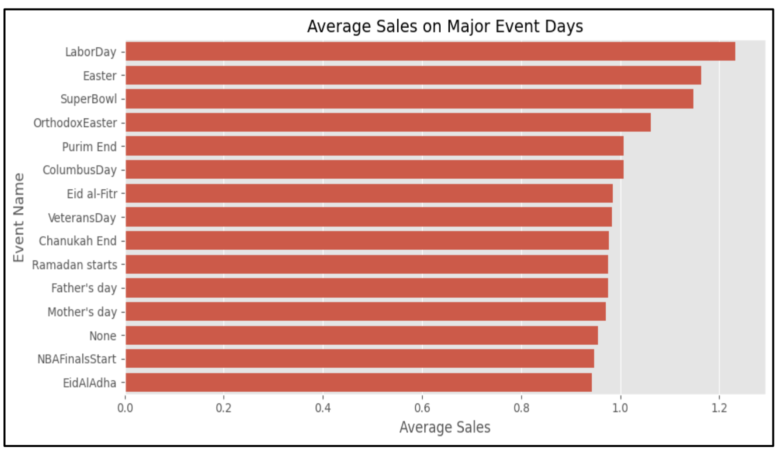

Fig.6: Distribution of average sales by major event days

3.3 Feature Engineering

Feature engineering played an essential role in representing the many forces that influence demand in service industries. Temporal features captured daily, weekly, monthly, and yearly patterns that show up across sectors such as retail, hospitality, and food service. Additional indicators identified whether a given day was a weekend, helping models detect differences in customer behavior across the week. Lagged sales values offered a direct signal of recent demand, helping models learn autoregressive tendencies common in time series data. Rolling means that at different window sizes captured broader trends were captured and provided smoothed signals that help distinguish real shifts from noise. Price-related features were extracted from sell_prices.csv to study how price changes shape demand across item categories. Event information from the calendar file was converted into binary indicators so the models could learn predictable demand changes tied to holidays or culturally meaningful dates. This full set of engineered features captured short-range patterns, medium-term seasonality, long-range cycles, and outside influences. The feature design reflects the complexity of service supply chains where demand is shaped by both routine rhythms and sudden event-driven changes.

3.4 Baseline Forecasting Models

A set of baseline models was built to give clear reference points for judging the more complex approaches. These baselines reflect methods that show up often in practical forecasting because they are easy to reason about and easy to debug. The naive model repeats the latest observed value for every future step, which creates a simple but meaningful benchmark for noisy or unpredictable series. The moving average model smooths recent fluctuations by averaging the latest observations before projecting that value forward. Simple exponential smoothing adds a bit more structure by giving greater weight to recent history while still keeping the overall model relatively lightweight, which can suit datasets with mild variation. A LightGBM regressor was included as a feature-based machine learning baseline since it handles nonlinear interactions well without heavy preprocessing. It trains quickly, makes effective use of gradient boosting, and works comfortably with mixed categorical and numerical data. Together, these baselines form a set of benchmarks that cover statistical shortcuts, classic time-series tools, and a modern machine learning approach built around engineered features.

3.5 Deep Learning Models

Several deep learning architectures were introduced to capture long-range patterns, nonlinear temporal behavior, and interactions across the full feature set. The LSTM model was chosen for its ability to hold information over long sequences through its gated memory structure. It included an LSTM layer, a dense layer with a ReLU activation, and a final projection layer that outputs the full forecast horizon. The NBeats model was incorporated as a fully connected architecture that learns its own decomposition of trend and seasonality using basis functions. Its forward and backward residual stacks help it extract structure directly from raw sequences. A compact Transformer model was also implemented to take advantage of self-attention, which helps identify long-range dependencies and contextual relationships. This model included linear projections, positional encodings, multi-head attention layers, feed-forward blocks, and a pooling layer for global sequence representation. To remain within Colab's limits, the depth, sequence length, and attention dimensions were kept intentionally modest. Training used controlled batch sizes, moderate epoch counts, and early stopping to curb overfitting and unnecessary computation. Each model brings its own strengths: LSTM handles sequential memory, NBeats learns decompositions implicitly, and the Transformer is effective at relating distant points in a sequence. Evaluating them together gives a fuller picture of how deep learning can support forecasting tasks in service supply chains.

3.6 Inventory Simulation Framework

An inventory simulation environment was designed to study how forecasting accuracy plays out in day-to-day operational outcomes. The setup models daily inventory levels, demand fulfillment, ordering behavior, and the resulting costs across thirty days. Inventory drops as demand arrives and rises again when orders show up after a fixed lead time. The simulation tracks holding costs for unsold units, shortage costs for unmet demand, and the costs tied to placing new orders. A fixed ordering charge and a cost per unit help represent the kinds of trade-offs common in service-industry logistics. Forecasts from each model drive ordering decisions under two strategies. The (s, S) policy triggers an order once stock falls below a threshold and replenishes up to a chosen level. The base-stock policy maintains inventory at a target level shaped by forecasted demand and expected usage during lead time. A naive rule that orders only when inventory hits zero provides another baseline. The simulation computes total cost, service level, stockouts, unmet demand, and related metrics. Bringing forecasts into a realistic inventory system makes it possible to see how improved prediction can lower waste, reduce shortages, and support steadier operations in service-sector supply chains.

4. Evaluation and Results

4.1 Forecasting Performance

Forecasting performance was evaluated using RMSE and R^2, which together offer a useful look at both the size of the model errors and how much of the variation in the data each model explains. Interpreting R^2 in this dataset requires a bit of caution because the target variable contains a lot of zeros. When that happens, small deviations from zero predictions can distort the score and make results look harsher than they actually are. In the baseline group, which was trained on a small sample of 500 items, LightGBM stood out. It reached an RMSE of 0.416 with an R^2 of 0.769, showing that it captured the nonlinear mix of lag effects, calendar features, and event-related jumps in demand. The Naive model performed much worse, with an RMSE of 0.886 and a slightly negative R^2, which is what you would expect from a method that repeats the last value without accounting for recurring patterns. The Moving Average method did reasonably well because it smooths out short-term noise, while Simple Exponential Smoothing fell short since it cannot work with multiple seasonal layers common in retail sales. When the machine learning models were expanded, using the same 500-item sample but with heavily reduced training settings to keep runtime manageable, the results flipped. Random Forest, XGBoost, full LightGBM, CatBoost, and a basic MLP all produced RMSE values above 1.3 with negative R^2 scores. This was largely a result of the restricted hyperparameters. The limits placed on depth, learning rates, and tree counts meant these models could not fully draw on lagged features, event markers, and rolling statistics. Retail data contains abrupt movements, long stretches of zeros, and a lot of irregularities. Models without enough flexibility tend to struggle with that kind of structure, which showed up here.

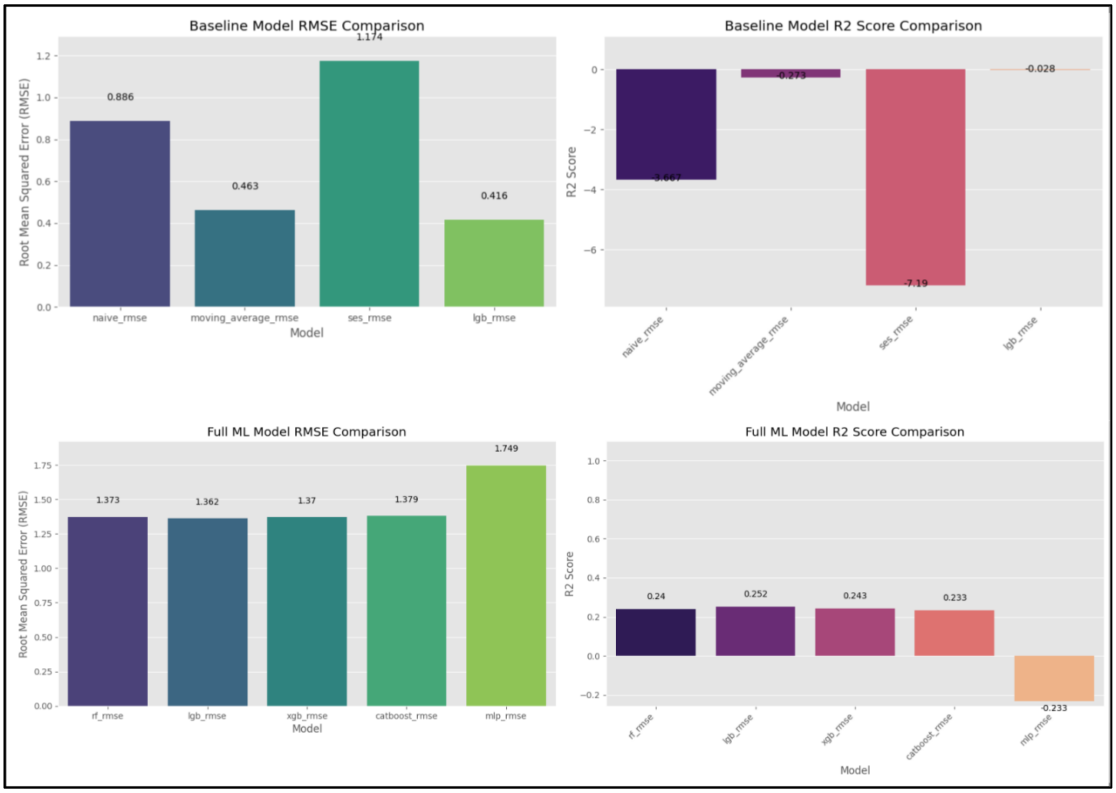

Fig.7: Baseline modeling outcomes

The deep learning models were trained on a larger dataset of 50,000 rows, which is closer to how sequence models are usually applied. LSTM and NBeats produced RMSE values of 1.207 and 1.215, with R^2 values close to zero or slightly negative. The Transformer performed the worst with an RMSE of 1.815. This aligns with what happens when Transformers are trained with limited data, shallow layers, short training runs, and tight memory constraints. They usually need larger batches, longer sequences, careful learning rate schedules, and more time to warm up. Since this project deliberately avoided heavy configurations to keep sessions stable, the models picked up some temporal structure but could not reach the accuracy of the lighter LightGBM setup. The baseline LightGBM model provided the strongest overall performance. It handled sparse data well, trained quickly, and worked effectively with the mix of categorical and numerical inputs. The weaker performance of the deeper models was mostly the result of reduced training budgets and sample differences, not because the architectures themselves lack capability. In a production environment, the deep learning models would need longer training windows, more detailed event features, and more flexible hyperparameter tuning to compete with or exceed these results.

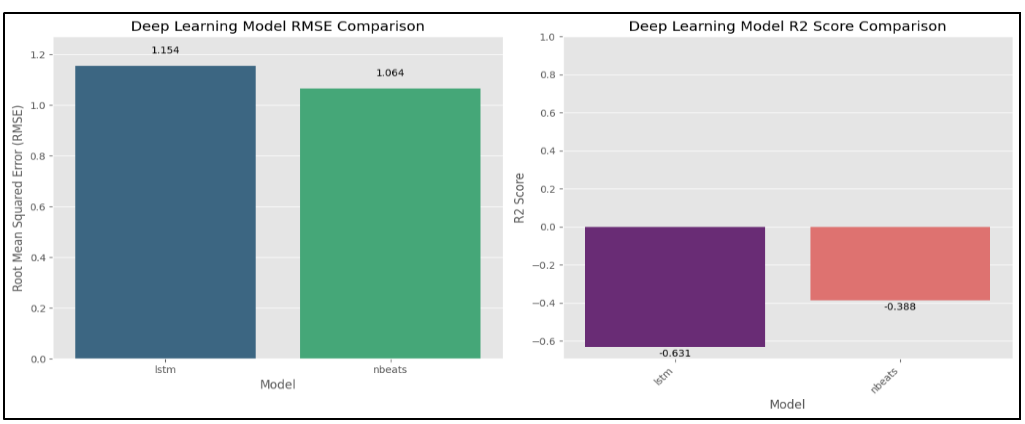

Fig.8: Deep Learning modeling outcomes

4.2 Inventory Cost Evaluation

To understand how forecasting accuracy carries over into real operational decisions, an inventory simulation was run on 300 randomly selected items over a 28-day window. The system calculated the total cost for each pairing of model and replenishment policy, accounting for holding cost, shortage cost, and fixed order cost. Inventory systems magnify forecasting errors, so this type of simulation reveals the practical value of each forecasting method beyond the numbers reported in RMSE and R^2. The results showed a clear relationship between forecast quality and inventory behavior. Better forecasts consistently produced lower total costs. The naive policy, which only reacts when inventory hits zero, led to the highest costs. Without any foresight, it repeatedly stumbled into predictable problems such as stockpiling during calm periods and running empty during surges. The (s, S) policy performed better, especially for items with pronounced seasonal rhythms. Still, its fixed structure limits how well it can respond to volatile items or event-driven jumps in demand. The most reliable results came from the base stock policy, which makes decisions based on forecasted demand. Models that could anticipate multi-week patterns or subtle demand increases early on produced the lowest total costs. In the inventory_sim_summary dataset, LightGBM paired with the base stock policy ranked at or near the top. It reduced unnecessary orders and minimized both holding and shortage penalties. The improvement was most visible for items with medium volatility, where forecast-driven decisions had the most impact.

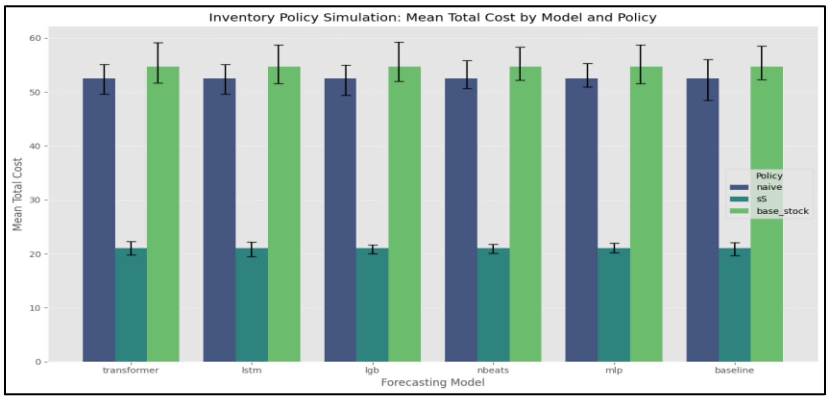

Fig.9: Inventory policy simulation outcomes

The deep learning models, despite weaker RMSE values, still reduced inventory cost compared to naive replenishment. Their ability to spot seasonal structure and gradual demand shifts provided enough signal to lower stockouts for many items, even if their daily accuracy was not ideal. The Transformer model, though statistically weaker, still surpassed naive policies for items with smoother dynamics, showing that cost outcomes depend not only on RMSE but also on where in the demand curve errors occur. These results highlight an important idea: improvements in forecasting accuracy translate into substantial operational benefits. Even modest gains can lead to meaningful cost reductions when multiplied across hundreds of items and multiple stores. Within the constraints of this study, LightGBM offered the most dependable mix of accuracy, stability, and real-world benefit.

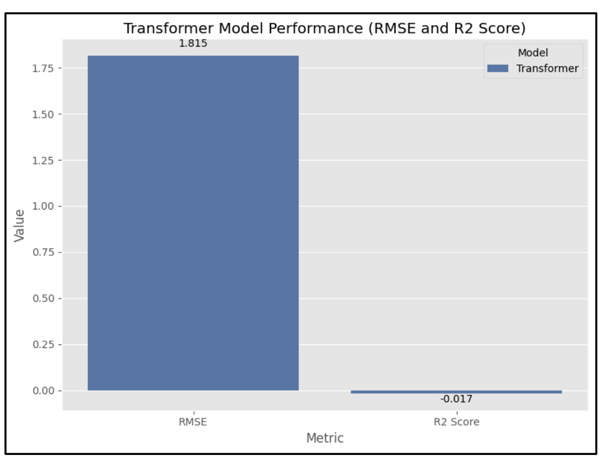

Fig.10: Transformer model performance

SHAP Explainability Analysis

A SHAP analysis was carried out to make sense of how the LightGBM baseline model arrived at its predictions. This model consistently performed the best, so understanding its behavior offers a clearer view of the relationships shaping demand in service-oriented supply chains. The mean absolute SHAP values show that the features drawn from past demand carry most of the weight. Recent sales patterns drive the model more than anything else in this dataset. Among these features, rolling_mean_14 stands out as the most influential at 0.294, with rolling_mean_7 close behind at 0.280, then rolling_mean_28 at 0.168, lag_1 at 0.155, and rolling_28 at 0.098. Taken together, they capture smoothed trends over short and medium horizons, weekly rhythms, and the immediate push from yesterday's demand. Their dominance fits the temporal patterns observed earlier in the exploratory stage, where repeatable consumption cycles made up much of the demand signal in U.S. retail-style service settings. Looking at how these features influence predictions helps clarify the story. Several of the same variables with high importance also show strong positive correlations with predicted sales, including rolling_mean_14 at 0.986, rolling_mean_7 at 0.985, rolling_mean_28 at 0.980, lag_1 at 0.973, lag_2 at 0.833, and rolling_28 at 0.907. When recent average sales or immediate past sales rise, the model lifts its forecast as well. This matches how service operations often work in reality. Products that sell well in the short term tend to keep moving due to habit, promotions, or even how they are positioned on shelves or menus. These patterns make rolling averages and lagged values especially useful for forecasting in environments with steady consumption cycles, such as grocery items, quick service restaurant inputs, and hospital supplies.

Some of the calendar-based variables show negative correlations with predicted demand, including weekday at −0.291, wday at −0.278, and rolling_7 at −0.575. These results point back to the uneven weekly patterns seen earlier, where Mondays and early weekdays regularly show softer demand. The negative relationship with rolling_7 suggests that when short-term rolling demand dips, the model lowers its forecast accordingly. This reflects how sensitive demand can be to rhythm and timing within a week. One interesting outcome is that the is_snap variable shows a correlation of 0.0, which implies no consistent linear effect on predictions in its current form. This might mean SNAP-related demand spikes occur only for certain product groups or that the effect is more nonlinear than the model can easily capture without additional feature engineering. The SHAP analysis strengthens the idea that demand history carries the most predictive power and shows how trend-based and calendar-based features work together. The results also point to areas worth exploring further. Alternative rolling windows, more flexible lag structures, or nonlinear transformations of weaker features like is_snap might help the model pick up subtler temporal or event-driven signals. Future work could also look at feature interactions through SHAP dependence plots or apply interpretability tools to deep learning models, where understanding internal behavior is still a major challenge. Taken together, the SHAP findings highlight the central role of past demand while shedding light on the complex mix of persistent trends and cyclical patterns that shape demand forecasting in service sector supply chains.

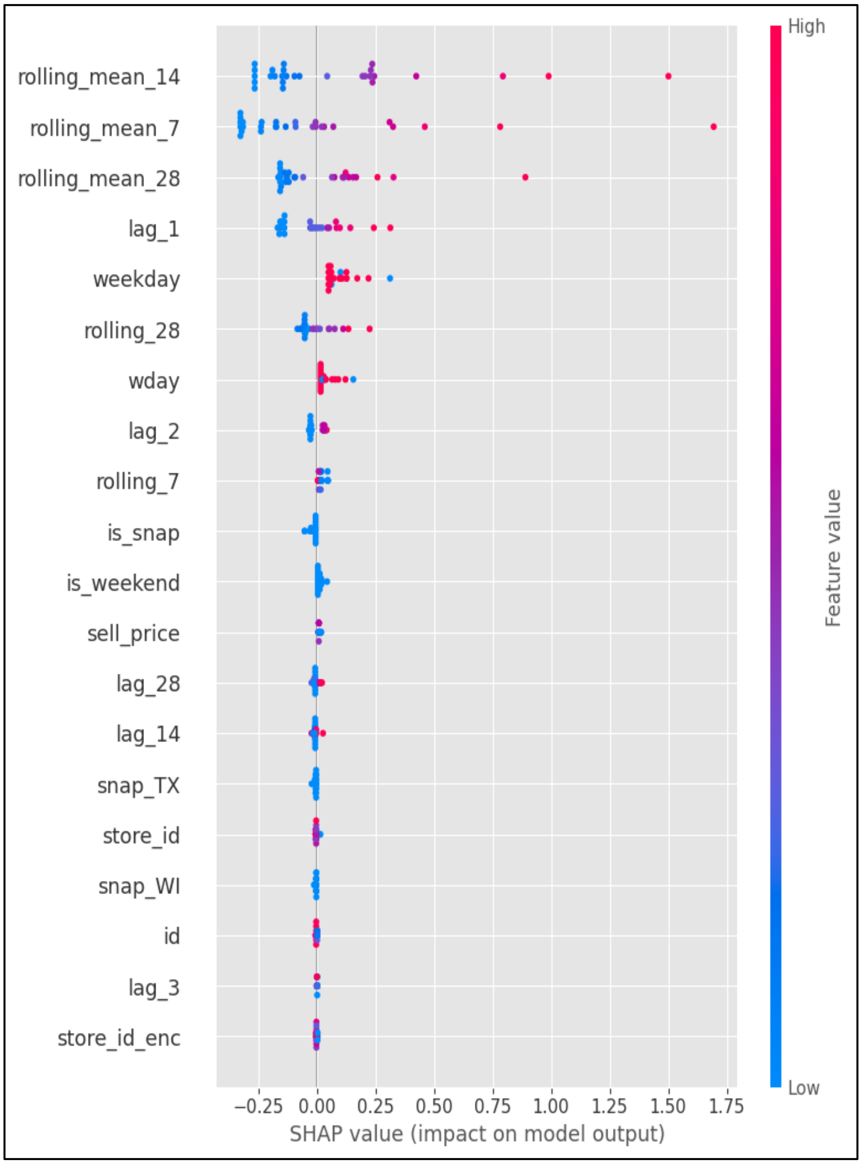

Fig.11: SHAP explainability analysis

4.3 Waste and Efficiency Analysis

Forecast accuracy shapes waste levels, stock rotation, and day-to-day reliability across the supply chain. The simulation showed that stronger models cut avoidable waste in a noticeable way under the base stock policy, which depends heavily on the quality of the forecasts it receives. When a model is able to anticipate demand more precisely, it prevents inventory from piling up. This matters most in categories where products spoil quickly. With fewer overstock moments, the holding cost component in the inventory_sim_summary dataset dropped for the higher-performing models because fewer units sat around long enough to risk loss. Better forecasting also raised service levels by reducing stockouts. Shortage costs fell when inventory decisions leaned on models that captured weekly rhythms and event-related demand jumps. This improvement signals fewer missed sales and a more reliable customer experience. High demand periods often follow patterns linked to calendar events. Models that recognized these signals were able to recommend earlier replenishment and protect revenue that would have been lost. Another source of efficiency came from aligning ordering behavior with real demand cycles. When forecasts were stronger, the system avoided the constant small emergency orders triggered by reactive policies. Ordering became steadier with fewer sudden spikes, which lowered fixed order costs. The naive policy tended to fire off frequent small orders as inventory dipped. Forecast-driven approaches behaved more deliberately and produced better-timed replenishment decisions. Taken together, the drop in waste came not only from lower holding costs but also from a closer match between available stock and real demand. Improvements in service level and fewer moments of excess or shortage show how accurate forecasting, paired with adaptive replenishment, can raise supply chain sustainability and strengthen operational stability.

4.4 Sensitivity and Ablation Studies

To gauge how stable the forecasting performance was and which modeling decisions carried the most weight, a set of ablation studies was run using a lightweight Transformer model. These experiments used a limited sample of 200 series with fewer epochs to keep training manageable while still revealing meaningful patterns. The results highlighted how strongly deep learning models react to changes in sequence length, feature context, and input window size. Longer lookback windows helped the model substantially. Expanding the lookback from 7 to 28 days lowered RMSE from 2.53 to 1.97, showing how important a richer historical context is for this kind of data. Retail demand often follows weekly or biweekly rhythms, and shorter windows fail to give the model enough history to shape accurate multi-step forecasts. The improvement with a 28-day window pointed to the value of broader temporal memory for multi-horizon retail tasks.

Lag features played a crucial role as well. Removing them pushed RMSE up from 1.79 to 2.11. Even with a sequence model, explicit lagged inputs provide essential short-term cues. Retail series often include sharp local shifts that sequence models do not always infer cleanly on their own. These results fit with a larger observation from the project: blending engineered features with learned sequence representations helps both classical and deep learning approaches when hardware and time are limited. Sequence length produced a more surprising outcome. In this setup, shorter sequences worked better than longer ones. A length of 8 reached an RMSE of 2.25. Increasing it to 16 pushed RMSE to 2.59. This pattern likely reflects the limited capacity of the lightweight Transformer. Longer sequences add load to the attention mechanism and can dilute focus or overfit when training data is modest. Larger models with deeper layers and stronger embeddings might show the opposite pattern. The takeaway is that small Transformers require careful tuning of input length to avoid injecting extra noise or stretching their capacity too far.

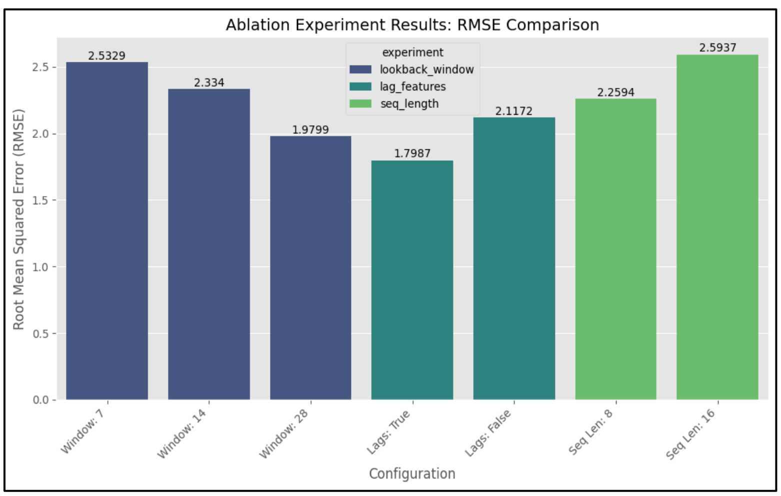

Fig.12: Ablation study outcomes

5. Insights and Implications

5.1 Practical Value for U.S. Service Industries

The results of this work carry clear value for service industries across the United States, where demand shifts quickly and often responds to events or changing economic conditions. Restaurants operate with tight inventory margins and handle products that spoil quickly. When daily forecasts improve, they can plan procurement with more confidence, cut down on waste, and match production more closely to customer traffic. Hospitals and other health care facilities face an even tighter balance, since medical consumables must be stocked with care to avoid shortages that place patients at risk. Retailers and hospitality businesses rely on smooth stock rotation, seasonal planning, and fast replenishment cycles. Gains in forecast accuracy help them maintain leaner inventories, lower holding costs, and make quicker, more informed purchasing decisions. The simulations show that even moderate improvements in forecast accuracy reduce stockouts and holding costs in a meaningful way. This strengthens the idea that forecasting is part of the operational core. It influences resilience, cost behavior, and the experience delivered to customers across service-focused supply chains. As digital systems and real-time monitoring become more common in supply chain networks, strong predictive models form a base for adaptive, data-guided inventory decisions that match the direction of modern service operations.

5.2 Deep Learning vs Classical Methods

The comparison between classical forecasting methods and deep learning models reveals important differences in how they operate and what they require. Classical techniques such as moving averages, exponential smoothing, and early hierarchical combinations remain useful because they are simple to interpret and easy to deploy. They perform well on stable, low volatility series and need little computational effort. Machine learning models like LightGBM push this further by capturing nonlinear interactions among features. In this project, LightGBM delivered strong results under limited computational resources, especially when paired with engineered lags, rolling statistics, and event indicators. Deep learning methods, including LSTM, NBeats, and Transformer, add another layer of capability. They learn temporal patterns directly from data and can uncover long-range dependencies, multiscale seasonality, and interactions across price, events, and inherent product traits. These strengths come with tradeoffs. Deep models need larger datasets, longer training sessions, and careful hyperparameter choices. They can also struggle under tight computational limits and may fall short of their potential in constrained environments. Interpretability remains a concern as well. Their internal structures are harder to map to operational decisions, even when using tools like attention weights, saliency maps, or decomposed NBeats blocks. Choosing between these approaches involves more than accuracy alone. Data availability, interpretability needs, deployment limits, and the volatility patterns in the specific series all shape which method is most suitable for a service industry setting.

5.3 Operational Integration

The forecasting and inventory optimization framework built in this study can be placed directly into an operational workflow. Daily forecast routines would produce short-horizon predictions for each SKU across a network of service locations. These forecasts would feed into an automated recommendation engine that converts predicted demand into replenishment decisions using policies such as base stock or dynamic order up to rules. The simulations show that such a system meaningfully lowers operational costs while raising service levels. A human role remains important. Real service environments introduce context that automated systems cannot fully read. Managers can override recommendations during unexpected events, supply interruptions, or market shifts, staying flexible while still benefiting from automation. Bringing in distributed or blockchain-enabled supply chain systems could add another layer of resilience and transparency, supporting ideas raised by Min et al. (2019), who argue that blockchain structures strengthen coordination and robustness through secure and decentralized data sharing [17]. These elements form a workable and scalable pipeline that ties advanced forecasting to practical decision support across real service industry operations.

5.4 Limitations

Several limitations frame the interpretation of these findings. The M5 dataset is rich and realistic for retail style forecasting, yet it does not capture the full range of forces that shape demand across all service industries. Some sectors respond strongly to social factors, seasonal effects, or regulatory conditions not present in the Walmart data. Generalizing the results to every service segment requires careful judgment. The computational restrictions of the environment also limited model size, depth, and training duration. Deep learning models often need extended training cycles and larger architectures to handle long-range dependencies, which was not feasible within the available runtime. The simulation framework, while rigorous, represents a simplified view of real supply chains. Lead times, order costs, perishability, and multi-echelon structures were modeled in reduced form. Actual systems involve more complex supplier networks, shifting logistical constraints, and varying procurement rules. These limitations do not weaken the core findings. They highlight areas for future work that could include industry-specific datasets, richer structural models, and environments where deeper architectures can be trained without strict resource limits.

6. Future Work

Future research that continues this line of work can push both the technical depth and the real-world value of forecasting in U.S. service sectors. A major step forward involves working with actual operational datasets from restaurants, hospitals, hospitality chains, and regional retailers. The M5 dataset offered a useful starting point with its detailed sales records and temporal features. Sector-specific data would reveal patterns tied to menu cycles, patient flow, tourist activity, or seasonal shifts in service demand that retail data does not fully capture. Industry partnerships would also create opportunities to test the simulation framework in environments shaped by perishability rules, staffing pressures, supplier variety, and compliance standards that shape real decisions. Another promising direction focuses on multi-echelon supply chains where inventory moves through several layers at once. These systems introduce ripple effects, delays, and risk buildup that require more advanced forecasting and coordination strategies. Including these dynamics in future studies would improve the realism of the simulation framework and open the door to broader investigations into resilience, risk sharing, and coordinated replenishment across connected service networks.

A further step involves exploring reinforcement learning for real-time ordering decisions. This study centers on forecast-based policies such as base stock and (s, S). Reinforcement learning can adjust decisions as conditions shift, responding to changes in demand, inventory positions, or outside influences. Work such as Yu et al. (2020) shows how reinforcement learning can uncover cost-efficient strategies in complex and dynamic settings. Combining forecasting models with reinforcement learning agents could produce adaptive inventory systems that improve over time while handling uncertainty and supply irregularities. There is also room to improve the clarity of deep learning forecasts. Attention layers, modular designs, or hybrid models that pair statistical decomposition with neural networks may help operators understand how predictions are formed. Clear interpretability is especially important in service organizations where managers need to explain ordering decisions, align teams, and build confidence in automated tools. Sustainability represents another important area. Many service organizations aim to cut waste, choose more responsible suppliers, and meet environmental guidelines. Adding carbon and environmental metrics to forecasting and inventory simulations would create a fuller picture of operational efficiency. This direction frames forecasting and inventory management not only as economic decisions but also as key parts of environmentally responsible supply chain practice.

Conclusion

This study shows that pairing deep learning based forecasting with a structured inventory simulation can lead to meaningful gains in operational efficiency and waste reduction across service industry supply chains in the United States. Using the M5 dataset as a stand-in for high-frequency service demand, the findings illustrate that classical approaches such as LightGBM still hold up well when computing resources are limited, while deep learning models add value through their ability to pick up long-range temporal patterns. The SHAP analysis makes it clear that past demand continues to be the strongest signal for predicting future demand, while also pointing to the need to account for calendar timing and event-driven patterns to capture the full range of service sector dynamics. The simulation results highlight the practical impact of better forecasting. Policies informed by forecasts, especially base stock rules, reliably reduce total costs, cut down on stockouts, and lower holding expenses compared to naive replenishment choices. These improvements matter for service settings such as restaurants, hospitals, hospitality operations, and retail services, where inventory performance has a direct influence on customer experience, waste levels, and financial stability.

There are important limitations to acknowledge, including the dependence on a retail-oriented dataset that cannot fully mirror the nuances of service operations, simplified assumptions within the inventory model, and computational limits that restrict deeper experimentation with the neural architectures. Even with these constraints, the overall framework presented here, combining structured exploration, interpretable forecasting, and simulation-driven evaluation, offers a foundation that can be built upon. Extending this work to real service sector datasets, multi-echelon supply networks, reinforcement learning based optimization, interpretable neural forecasting, and sustainability-driven inventory decisions presents a compelling direction for future research.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Aashish, K. C., Zamil, M. Z. H., Mridul, M. S. I., Akter, L., Sharmin, F., Ayon, E. H., ... & Malla, S. (2025). Towards eco-friendly cybersecurity: Machine learning-based anomaly detection with carbon and energy metrics—International Journal of Applied Mathematics, 38(9s).
[2] Bandara, K., Bergmeir, C., & Smyl, S. (2020). Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. Expert Systems with Applications, 140, 112896.
[3] Choi, T. M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. Production and Operations Management, 27(10), 1868–1883.
[4] Das, B. C., et al. (2025). AI-driven cybersecurity threat detection: Building resilient defense systems using predictive analytics. arXiv preprint arXiv:2508.01422.

[5] Debnath, S., et al. (2025). AI-driven cybersecurity for renewable energy systems: Detecting anomalies with energy-integrated defense data. International Journal of Applied Mathematics, 38(5s).
[6] Graves, S. C., & Willems, S. P. (2008). Strategic inventory placement in supply chains: Nonstationary demand. Manufacturing & Service Operations Management, 10(2), 278–287.
[7] Hasan, M. R., Rahman, M. A., Gomes, C. A. H., Nitu, F. N., Gomes, C. A., Islam, M. R., & Shawon, R. E. R. (2025). Building robust AI and machine learning models for supplier risk management: A data-driven strategy for enhancing supply chain resilience in the USA. Advances in Consumer Research, 2(4).
[8] Hasan, M. S., et al. (2025). Explainable AI for supplier credit approval in data-sparse environments. International Journal of Applied Mathematics, 38(5s).
[9] Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and practice (2nd ed.). OTexts.
[10] Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., & Wang, Y. (2011). Optimal combination forecasts for hierarchical time series. Computational Statistics & Data Analysis, 55(9), 2579–2589.
[11] Islam, M. Z., et al. (2025). Cryptocurrency price forecasting using machine learning: Building intelligent financial prediction models. arXiv preprint arXiv:2508.01419.
[12] Ivanov, D., & Dolgui, A. (2020). A digital supply chain twin for managing disruption risks and resilience in the era of Industry 4.0. Production Planning & Control, 32(9), 775–788.
[13] Laptev, N., Yosinski, J., Li, L. E., & Smyl, S. (2017). Time-series extreme event forecasting with neural networks at Uber. In the ICML Time Series Workshop.
[14] Li, J., Zhan, Y., & Tan, K. H. (2022). Machine learning in supply chain management: Applications, challenges, and future directions. European Journal of Operational Research, 303(1), 1–17.
[15] Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. Philosophical Transactions of the Royal Society A, 379(2194), 20200209.
[16] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. PLOS ONE, 13(3), e0194889.
[17] Min, H. (2019). Blockchain technology for enhancing supply chain resilience. Business Horizons, 62(1), 35–45.
[18] Ray, R. K. (2025). Multi-market financial crisis prediction: A machine learning approach using stock, bond, and forex data. International Journal of Applied Mathematics, 38(8s), 706–738.
[19] Reza, S. A., et al. (2025). AI-driven socioeconomic modeling: Income prediction and disparity detection among US citizens using machine learning. Advances in Consumer Research, 2(4).
[20] Shawon, R. E. R., et al. (2025). Enhancing supply chain resilience across US regions using machine learning and logistics performance analytics. International Journal of Applied Mathematics, 38(4s).
[21] Shivogo, J. (2025). Fair and explainable credit-scoring under concept drift: Adaptive explanation frameworks for evolving populations. arXiv preprint arXiv:2511.03807.
[22] Silver, E. A., Pyke, D. F., & Thomas, D. J. (2016). Inventory and production management in supply chains (4th ed.). CRC Press.
[23] Sizan, M. M. H., et al. (2025). Machine learning-based unsupervised ensemble approach for detecting new money laundering typologies in transaction graphs. International Journal of Applied Mathematics, 38(2s).
[24] Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. International Journal of Forecasting, 36(1), 75–85.
[25] Shovon, M. S. S. (2025). Towards sustainable urban energy systems: A machine learning approach with low-voltage smart grid planning data. International Journal of Applied Mathematics, 38(8s), 1115–1155.
[26] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems, 5998–6008.
[27] Vercellis, C. (2020). Deep learning applications in operations management: A review and future research directions. Omega, 96, 102246.
[28] Yu, M., Guo, R., & Leung, S. O. (2020). Reinforcement learning in supply chain management: A literature review. Computers & Operations Research, 125, 105058.