Scope International Journal of Science, Humanities, Management and Technology (SIJSHMT) - ISSN : 2455-068X
Vol.11 Issue 4 (2025) 58 - 67. Submitted 25/10/2025. Published 24/11/2025

Predictive Demand Forecasting Analysis

Tamidalapati Muni Poojitha
2nd Year, M.S. Data Science,
EdTech Division, Exafluence Inc.,
SVU College of CM & CS,
Sri Venkateswara University,
Tirupati, India
munipoojitha4@gmail.com

Vijaya Lakshmi Kumba
Professor, Dept. of Computer Science,
SVU College of CM & CS,
Sri Venkateswara University,
Tirupati, India
vijayalakshmik4@gmail.com

Abstract
Managing inventory, particularly during periods with unpredictable demand, continues to remain a significant challenge in the retail sector. This work focuses on building a predictive analytics framework for demand forecasting and order optimization within retail processes. Using historical sales data, advanced machine learning and time-series models—including XGBoost, CatBoost, ARIMA, and Prophet—are applied to predict demand trends for specific time periods. These forecasts then feed into an optimization module designed to recommend actionable reorder points and replenishment quantities. The combined approach reduces stockout and overstock scenarios, improving both operational cost efficiency and customer satisfaction. Power BI and Tableau dashboards are used for real-time visualization of sales trends, inventory status, and forecasting accuracy. Unlike traditional statistical models, the proposed method significantly improves forecast precision, thereby enhancing data-driven decision-making in inventory management. This work highlights the predictive intelligence necessary to support operational sustainability in retail.
Keywords— Predictive Analytics, Machine Learning, Demand Forecasting, Retail Operations, Supply Chain Optimization, Data Visualization

I. INTRODUCTION

Technological advancements, evolving consumer preferences, and dynamic market environments have significantly transformed modern retail operations. Retailers face substantial fluctuations in customer demand across seasons, promotions, and festive periods. Traditional forecasting methods, though foundational, often fail to capture the nonlinear, irregular patterns inherent in retail sales data. These limitations amplify inefficiencies such as stockouts and excessive inventory holding.

Predictive analytics, powered by AI and ML, offers the ability to model complex demand patterns with higher accuracy. By integrating time-series forecasting with machine learning models, retailers can anticipate demand shifts, optimize inventory decisions, reduce operational risks, and improve customer satisfaction levels. The objective of this study is to establish a predictive framework combining machine learning models and time-series forecasting to support optimized inventory and supply chain decisions.

II. LITERATURE REVIEW

Modern advancements in AI and ML have significantly enhanced demand forecasting accuracy and supply chain efficiency in the retail industry. Traditional methods such as exponential smoothing and moving averages fall short in capturing dynamic, nonlinear consumer behavior, leading researchers to adopt advanced ML-based methods.

Key Contributions

1. **Hobor et al. (2025)**
   Demonstrated that tree-based ensemble models such as XGBoost and LightGBM outperform deep learning models like N-BEATS and Transformers when handling irregular, localized retail demand patterns. The study emphasizes selecting an optimal model tailored to the data rather than relying on unnecessarily complex architectures.
2. **Nasseri et al. (2023)**
   Showed that ensemble models (Random Forest, XGBoost) achieve forecasting accuracy comparable to deep learning models like LSTM, while being computationally more efficient. The findings suggest that simpler, well-tuned ML models are often suitable and more practical for real-world forecasting tasks.
3. **Ray (2025)**
   Integrated AI-driven demand forecasting with supply chain optimization techniques, demonstrating improvements in inventory turnover, order planning precision, and stockout minimization. The study highlights the business value of combining forecasting with operational optimization.
4. **Venu Gopal Avula (2021)**
   Developed AI-powered forecasting solutions for retail operations focusing on demand planning, customer behavior analysis, and supply chain optimization. His work confirms the role of predictive intelligence in enhancing retail decision-making.

III. METHODOLOGY

The methodology for this study centers on developing an integrated predictive analytics framework for demand forecasting and inventory optimization in the retail domain. The process is divided into six primary components: data collection and preprocessing, demand forecasting model development, order optimization model formulation, visualization and decision-support design, model evaluation, and system architecture implementation.

A. Data Collection and Preprocessing

The study utilizes a comprehensive retail dataset consisting of two years of historical records across sales, product attributes, and inventory operations. The dataset includes:

- **Transactional Data:** Daily and weekly sales volumes across multiple product categories.
- **Product Data:** Product identifiers, categories, cost prices, and selling prices.
- **Inventory and Replenishment Data:** Stock levels, reorder quantities, safety stock values, and supplier lead times.
- **Temporal and External Factors:** Seasonality, promotional events, festive periods, and demand spikes.

Preprocessing operations were conducted to ensure data quality and readiness for modeling. Missing values were resolved using statistical imputation and interpolation techniques. Outliers were identified using the Interquartile Range (IQR) method and corrected where necessary. Feature engineering generated additional predictive attributes, such as lagged sales features, moving averages, demand volatility indicators, and promotional flags. All relevant features were normalized where appropriate, and the dataset was partitioned into training and testing subsets using time-based splitting to preserve temporal integrity.

B. Demand Forecasting Models

The first analytical layer of the framework focuses on predicting future demand using both classical time-series models and advanced machine-learning algorithms. Four primary models were implemented:

- **ARIMA (AutoRegressive Integrated Moving Average):** Used to capture linear patterns, trends, and seasonality in time-series data.
- **Prophet:** Applied for modeling complex seasonality patterns, holiday effects, and nonlinear trends.
- **XGBoost:** Gradient-boosting model capable of learning nonlinear interactions and handling irregular data patterns.
- **CatBoost:** A boosting algorithm optimized for categorical variables and suited for heterogeneous retail datasets.

Model hyperparameters were tuned using grid search and time-series cross-validation techniques. Historical sales data was used for model training, and unseen recent data was reserved for testing to evaluate predictive performance. Forecasts generated by each model served as inputs for downstream inventory optimization calculations.

C. Order Optimization Models

To translate demand forecasts into actionable inventory decisions, an optimization layer was incorporated. This module identifies optimal reorder points (ROP) and economic order quantities (EOQ) based on predicted demand patterns.

The following optimization approaches were implemented:

- **Deterministic Optimization:** Traditional calculations for safety stock and reorder points under relatively stable demand assumptions.
- **Cost-Based Optimization:** Minimization of total inventory cost, including holding, ordering, and stockout costs.
- **Constraint-Based Optimization:** Decision logic incorporating supplier capacity, lead time fluctuations, service-level constraints, and demand variability.

This ensures the system not only predicts demand accurately but also prescribes inventory replenishment decisions that balance cost efficiency and service-level performance.

D. Visualization and Decision Support

To improve interpretability and operational value, interactive dashboards were developed using Power BI and Tableau. These dashboards present:

- **Historical and Forecasted Demand Trends**
- **Inventory Turnover Analysis and Reorder Cycles**
- **Comparative Model Performance (ARIMA vs. ML vs. Prophet)**
- **Cost Optimization Metrics and Sustainability Indicators**

Visual storytelling empowers inventory managers to make informed, data-driven decisions while monitoring key performance indicators (KPIs) in real time.

E. Model Evaluation

Model and system performance were evaluated using quantitative and qualitative metrics:

- **Forecast Accuracy:**
    - Mean Absolute Error (MAE)
    - Mean Absolute Percentage Error (MAPE)
    - Root Mean Square Error (RMSE)
- **Inventory Metrics:**
    - Inventory Turnover Ratio

- Stockout Rate
- **Optimization Effectiveness:**
  - Total Inventory Cost Reduction (%)
  - Reduction in Overstock and Stockout Incidents
- **System Metrics:**
  - Dashboard Responsiveness
  - Usability Scores from User Evaluations

Cross-validation and temporal testing validated the framework's robustness and its ability to generalize to real-world retail operations.

F. System Architecture

The proposed architecture integrates authentication, data ingestion, processing, visualization, and reporting. It consists of three sequential layers:

1. Authentication and Entry Point

- **Login and Sign-Up Modules:**
  Users (e.g., retail managers) authenticate via login or register through the sign-up interface.
- Successful login transitions the user to the main dashboard.

2. Analysis and Processing Core

- **dash.html – CSV Upload Interface:**
  The secure dashboard allows users to upload transactional CSV files containing sales data.
- **Backend Processing (FastAPI + Python):**
  Upon file upload, the system performs:
  - Data preprocessing and feature engineering
  - Demand forecasting using ML/time-series models
  - Inventory optimization (ROP/EOQ calculations)
  - Plot generation (Matplotlib/Seaborn)
  - Storing results in MongoDB

3. Output and Reporting

- **Immediate Results:**
  dash.html displays interactive plots and analytics summaries returned from the backend.
- **Power BI Dashboard:**
  Stored metrics are exported to Power BI for long-term, enterprise-grade reporting, enabling comparative analysis, KPI monitoring, and strategic forecasting.

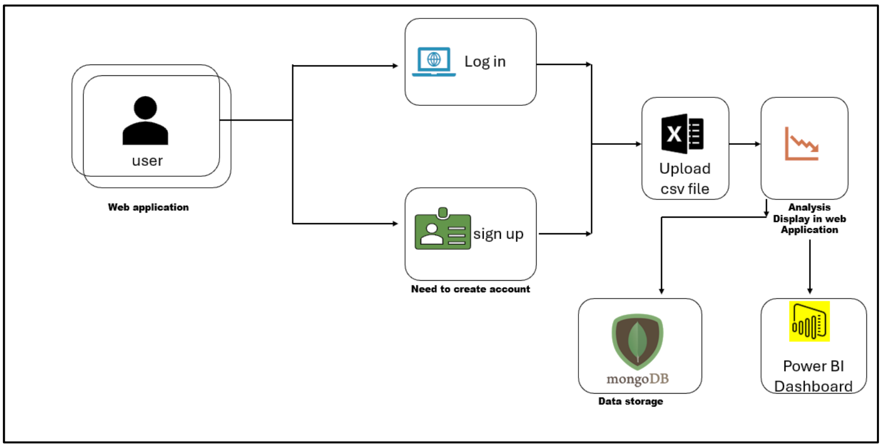

Fig 1: System Archieture

Scope International Journal of Science, Humanities, Management and Technology (SIJSHMT) - ISSN : 2455-068X
Vol.11 Issue 4 (2025) 58 - 67. Submitted 25/10/2025. Published 24/11/2025

IV. RESULTS AND ANALYSIS

A. Festival Sales Performance and Model Evaluation

1. Negative Change in Festival Sales by Category

The Year-on-Year (YoY) decline in festival-season sales (2024 → 2025) reveals significant reductions across several product categories. The largest decrease occurred in **Beauty & Hygiene**, which experienced a drop of approximately **–20,794 units**. This was followed by:

- **Foodgrains, Oil & Masala:** –14,544 units
- **Gourmet & World Food:** –12,263 units

The overall trend indicates a contraction in both the value and volume of festival-season sales, suggesting either declining consumer purchasing power or a shift in preferences within these segments. These reductions highlight the need for strategic adjustments in stock levels, targeted promotions, and category-specific sales interventions.

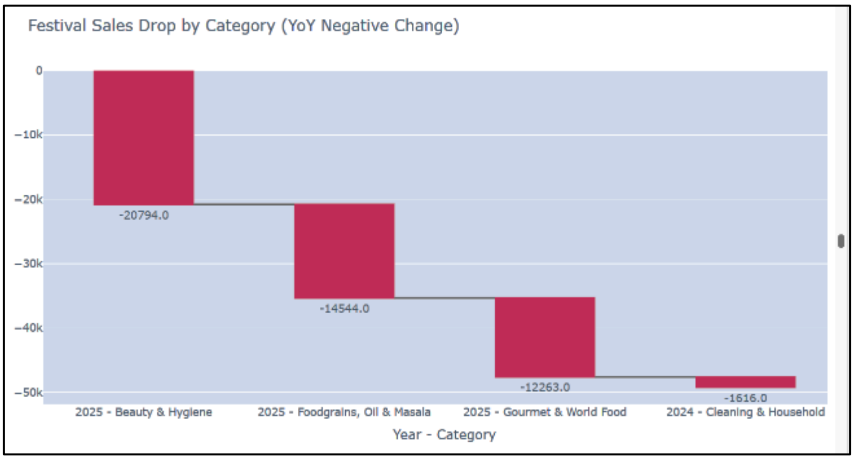

Fig. 2: Negative Change in Festival Sales by Category

2. Positive Change in Festival Sales by Category

Conversely, a subset of categories showed notable YoY improvements. The highest growth was observed in:

- **Beauty & Hygiene:** +25,032 units
- **Bakery, Cakes & Dairy:** +2,728 units
- **Cleaning & Household:** +415 units

The rise in these essential and hygiene-related product segments aligns with post-pandemic consumer behaviour, where increased preference for premium and health-oriented products has been widely observed.

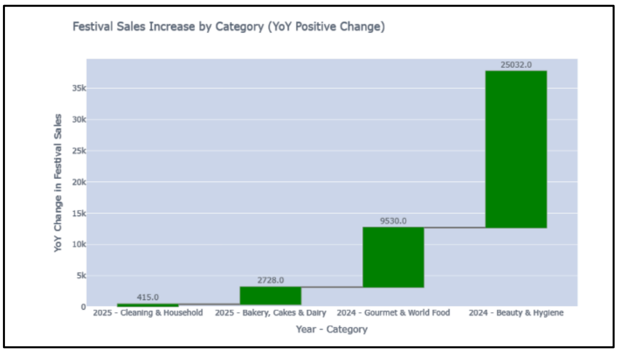

Fig. 3: Positive Change in Festival Sales by Category

Scope International Journal of Science, Humanities, Management and Technology (SIJSHMT) - ISSN : 2455-068X
Vol.11 Issue 4 (2025) 58 - 67. Submitted 25/10/2025. Published 24/11/2025

B. Total Weekly Sales Trend Across the Period (2023–2025)

Analysis of weekly sales trends over three years indicates a consistent upward trajectory in retail demand. The forecasted 2025 series reaches the highest weekly peaks, surpassing **500,000 units**, and demonstrates greater volatility than previous years.

This trend confirms the model's prediction of continued market expansion and increased sales variability. Such insights are critical for proactive inventory and supply-chain planning.

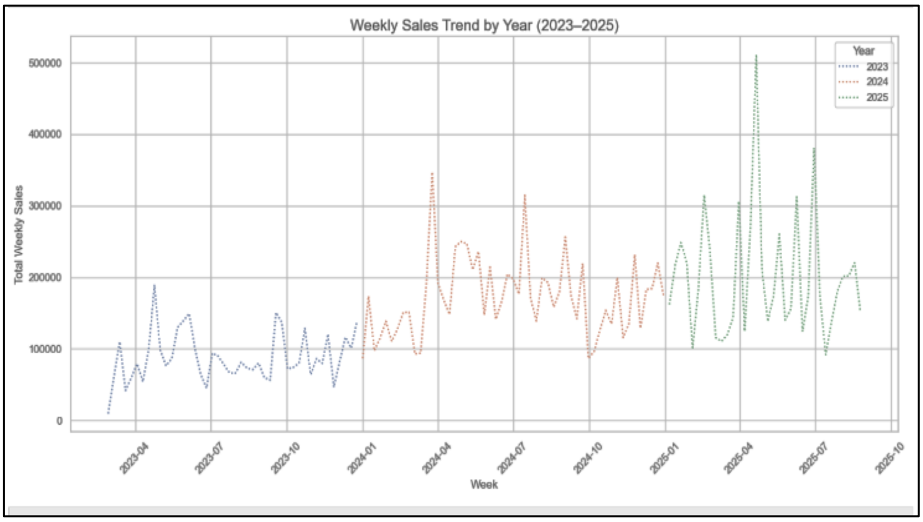

Fig. 4: Total Weekly Sales Trend Across the Period (2023–2025)

C. Model Accuracy by Category

1. Mean Absolute Error (MAE)

The predictive performance of the four forecasting models—XGBoost, CatBoost, ARIMA, and Prophet—was evaluated using MAE:

- **ARIMA and Prophet** showed superior accuracy in categories such as *Cleaning & Household* and *Foodgrains, Oil & Masala*.
- **CatBoost and XGBoost** performed competitively in broader categories like *Beauty & Hygiene* and *Kitchen, Garden & Pets*.

These outcomes suggest the value of adopting **category-specific hybrid forecasting strategies** that leverage strengths of both statistical and machine-learning methods.

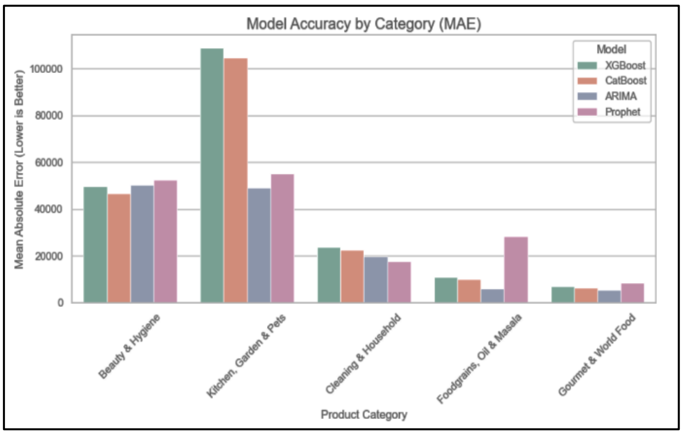

Fig. 5: Model Accuracy by Category (MAE)

Scope International Journal of Science, Humanities, Management and Technology (SIJSHMT) - ISSN : 2455-068X
Vol.11 Issue 4 (2025) 58 - 67. Submitted 25/10/2025. Published 24/11/2025

2. Root Mean Square Error (RMSE)
RMSE results further validate model performance:

- **ARIMA and Prophet** achieved the lowest RMSE in *Beauty & Hygiene* and *Kitchen, Garden & Pets*, demonstrating strong accuracy in high-volume categories.
- **CatBoost and XGBoost** produced the lowest RMSE for *Foodgrains, Oil & Masala* and *Gourmet & World Food.*

Collectively, the RMSE and MAE analyses confirm that **the optimal forecasting model differs by product category**, reinforcing the necessity for a hybrid model selection framework.

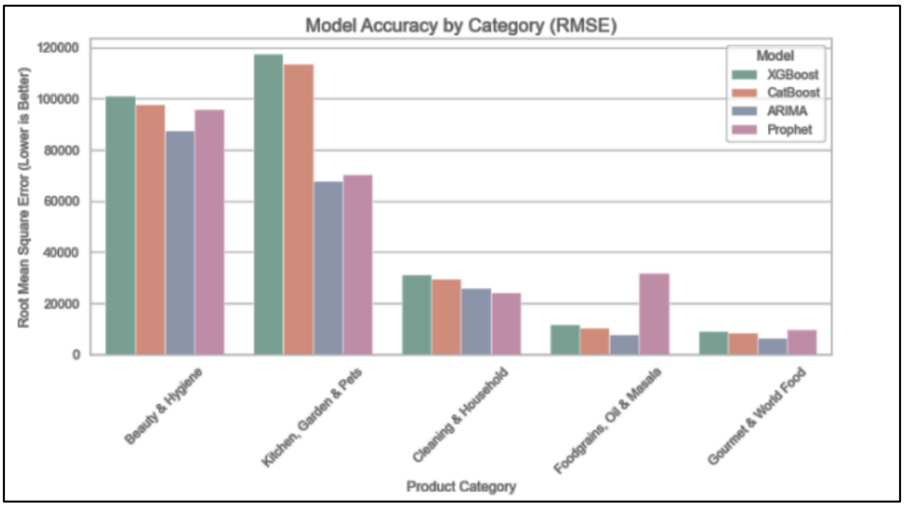

Fig. 6: Model Accuracy by Category (RMSE)

D. Inventory Optimization and Budget Allocation
The integrated forecasting and optimization engine generated reorder recommendations for five major product categories. The table below presents quantitative outputs without interpretation.

Table 1: Inventory Optimization and Budget Allocation

Category	Forecasted Sales ($)	Unit Cost ($)	Current Stock (Units)	Target Stock (Units)	Units to Reorder	Estimated Budget ($)
Beauty & Hygiene	82,185.20	200	50	453	403	80,600.00
Kitchen, Garden & Pets	76,546.80	350	10	241	231	80,850.00
Gourmet & World Food	12,015.75	75	30	178	148	11,100.00
Cleaning & Household	20,107.60	100	100	223	123	12,300.00
Foodgrains, Oil & Masala	6,319.50	50	250	140	0	0.00

Key Observations

- **Highest Reorder Volume:**

Beauty & Hygiene requires **403 units**, aligned with a high target stock level.

- **Largest Budget Requirement:**
  Kitchen, Garden & Pets requires **$80,850**, driven by a high unit cost of 350.
- **No Reorder Needed:**
  Foodgrains, Oil & Masala has current stock exceeding target stock; thus, reorder quantity is **zero.**

These results demonstrate the system's ability to translate predictive analytics into actionable procurement decisions.

E. Overall Project Success and Business Impact

The project's success is assessed based on operational improvements, financial impact, and enhanced decision-support capabilities.

1. Financial Control and Budget Efficiency

- **Targeted Procurement:**
  Using the Estimated Reorder Budget KPI, procurement teams can allocate funds efficiently—approximately **$184,850** for the projected cycle—reducing excess capital tied in slow-moving inventory.
- **Inventory Velocity:**
  An observed Inventory Turnover Rate exceeding **130x** reflects strong sell-through performance and efficient stock cycling.

2. Risk Mitigation and Strategic Foresight

- **Seasonal Preparedness:**
  The system detected a **66.10% uplift** during festival weeks, enabling precise pre-season stock planning.
- **Stock-Out Prevention:**
  The Inventory Health Matrix identifies high-demand, low-supply items in real time, reducing the risk of lost sales.

V. ACKNOWLEDGMENT

The superior performance of the **Prophet model (accuracy score: 0.96)** is a significant outcome of this research and justifies its selection as the primary forecasting engine and baseline for the final ensemble framework. Its ability to capture complex multi-seasonal patterns, trend shifts, holiday effects, and other external influences demonstrates Prophet's suitability for aggregated retail time-series data.

In comparison, the traditional **ARIMA model (0.55)** and the gradient-boosting model **CatBoost (0.69)** delivered moderate performance, indicating partial capability in modelling underlying demand structures but falling short in capturing the full variability and complexity inherent in retail demand cycles. Conversely, the **XGBoost model (0.00)** exhibited clear incompatibility with the dataset—likely due to architectural limitations, feature-scaling issues, or the static nature of the aggregated time series. This result highlights the importance of evaluating multiple modelling approaches and avoiding reliance on a single advanced algorithm that may fail under specific structural conditions.

The synthesized forecasting outputs were directly translated into operational decision layers. The inventory optimization subsystem provided precise guidance on reorder quantities and financial allocation. Notably, the **Kitchen, Garden & Pets** category exhibited the highest estimated reorder budget (**$80,850.00**), closely followed by **Beauty & Hygiene** (**$80,600.00**). These optimized projections enable more accurate procurement planning, reduce stockout risk, and improve inventory turnover, thereby supporting the overarching goal of strengthening operational efficiency in the retail environment.

A. Limitations and Future Research

While the current system demonstrates strong predictive capabilities and operational effectiveness, several areas offer potential for enhancement:

1. Explainable AI (XAI) for Inventory Decisions

Future work should incorporate SHAP, LIME, or similar XAI frameworks to improve the interpretability of models such as Prophet and CatBoost. Transparent rationalization of forecasts will support managerial trust and regulatory compliance in high-impact environments.

2. Adaptive and Robust Model Selection via AutoML

Implementing AutoML-driven model selection can help dynamically identify the best-performing forecasting algorithm for each product category. This will mitigate architectural failures, improve performance stability, and optimize hyperparameter tuning at scale.

3. Next-Generation Time-Series Architectures

Advanced deep-learning approaches—including **Diffusion Models**, **Mamba-based sequence models**, and hybrid attention-driven frameworks—should be evaluated for long-horizon forecasting. These models offer improved handling of distributional shifts and long-term dependencies, especially relevant for aggressive projected growth trends.

4. Dynamic, Risk-Aware Inventory Optimization

Transitioning from deterministic optimization to **Reinforcement Learning (RL)-based inventory control** will enable real-time decision-making under stochastic conditions such as fluctuating lead times, supply uncertainty, and volatile demand. RL integration can drive more adaptive stock and budget allocation policies.

B. Conclusion

This research successfully developed and validated an integrated AI-driven forecasting and optimization system designed to elevate operational efficiency across retail demand planning, model evaluation, and supply-chain decision-making. By converting raw transactional datasets into accurate multi-year demand forecasts and actionable inventory recommendations, the system achieves its core objective of supporting data-driven retail management.

Key Contributions and Achievements

1. High Predictive Accuracy

The Prophet model demonstrated the strongest performance with an accuracy score exceeding **0.96**, effectively capturing the non-linear, multi-seasonal dynamics of retail demand. Comparative model analysis confirmed that forecast accuracy differs significantly by product category, reinforcing the need for a **hybrid modelling strategy** to minimize overall error (MAE and RMSE).

2. Operationalization of Forecasts into Supply-Chain Decisions

A major innovation of the system is its ability to convert forecast outputs into concrete operational recommendations. Reorder budgets, target stock levels, and procurement priorities (as outlined in Table 1) provide immediate managerial value, enhancing both financial and supply-chain planning.

Overall, the developed system demonstrates strong potential for real-world deployment, offering a robust, scalable, and analytically transparent foundation for next-generation retail demand forecasting and inventory optimization.

REFERENCES

[1]. Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing: Introduction to the special issue on multi-channel retailing. *Journal of Retailing*, 91(2), 174–181.

[2]. Syndetons, A. A., Babai, Z., Boylan, J. E., Kolassa, S., & Nikolopoulos, K. (2016). Supply chain forecasting: Theory, practice, their gap, and the future. *European Journal of Operational Research*, 252(1), 1–26.

[3]. Ovechkin, B. N., Carpov, D., Chapados, N., & Bengio, Y. (2019). N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. *ICLR / arXiv*. https://arxiv.org/abs/1905.10437

[4]. Choi, T. M. (2018). A system of systems approach for global supply chain management in the big data era. *IEEE Engineering Management Review*, 46(1), 91–97.

[5]. Babai, M. Z., Ali, M. M., Boylan, J. E., & Syndetons, A. A. (2013). Forecasting and inventory performance in a two-stage supply chain with ARIMA (0,1,1) demand: Theory and empirical analysis. *International Journal of Production Economics*, 143(2), 463–471.

[6]. Syndetons, A. A., & Boylan, J. E. (2005). The accuracy of intermittent demand estimates. *International Journal of Forecasting*. https://www.sciencedirect.com/science/article/abs/pii/S0169207004000792

[7]. Huber, J., & Stuckenschmidt, H. (2020). Daily retail demand forecasting using machine learning with emphasis on calendric special days. *International Journal of Forecasting*, 36(4), 1420–1434.

[8]. 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.

[9]. Thomassen, S. (2010). Sales forecasts in clothing industry: The key success factor of supply chain management. *International Journal of Production Economics*, 128(2), 470–483.

[10]. Ali, Ö. G., Sayın, S., van Woensel, T., & Fransoo, J. (2009). SKU demand forecasting in the presence of promotions. *Expert Systems with Applications*, 36(10), 12340–12348.

[11]. Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*. https://facebook.github.io/prophet/static/prophet_paper_20170113.pdf

[12]. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. https://www.bioinf.jku.at/publications/older/2604.pdf

[13]. Rahman, K. O., Rezvi, R. I., Nasrullah, F., Islam, M. S., Hasan, M., Khanam, A., & Akash, A. H. (2025). Business model canvas: Business analytics on gas stations with C-stores in United States. *Journal of Business and Management Studies*, 7(1), 180–185.

[14]. Rahman, K. O., et al. (2025). Business model canvas: Business analytics on gas stations with C-stores in United States. *Journal of Business and Management Studies*, 7(1), 180–185.

[15]. Rahman, K. O., et al. (2025). AI-driven next-gen US retail: An empirical study on optimizing supply chains by leveraging artificial intelligence, business intelligence, and machine learning. *Journal of Computer Science and Technology Studies*, 7(1), 258–264.