Forecasting Market Demand with Machine Learning: From Historical Data to Consumer Behavior Insights

1st Tursynbek Yerassyl
L.N. Gumilyov Eurasian National University
Astana, Kazakhstan
erasyltrsynbek@gmail.com
ORCID 0009-0009-8735-236X

2nd Kassekeyeva Aislu Bisenovna
L.N. Gumilyov Eurasian National University
Astana, Kazakhstan
aibike2903@gmail.com
ORCID 0000-0002-5547-0718

***Abstract*—Accurate demand forecasting is a critical component of supply chain optimization, inventory management, and strategic planning in modern enterprises. Conventional statistical forecasting approaches often struggle to represent nonlinear patterns and abrupt changes in consumer behavior, which reduces their effectiveness in volatile market conditions. This study explores data-driven forecasting techniques that integrate historical sales records with factors influencing purchasing behavior, including seasonality and promotional effects.**

A comparative experimental analysis is conducted between classical time-series approaches and advanced data-oriented predictive models. The results demonstrate that data-driven forecasting techniques achieve higher predictive accuracy and stability, particularly when long-term temporal dependencies and irregular demand fluctuations are present. The proposed approach supports improved decision-making by reducing forecasting errors and enhancing operational efficiency. The findings highlight the potential of intelligent forecasting systems for sustainable business growth and adaptive demand planning.

Keywords—Demand forecasting, data-driven modeling, time-series analysis, supply chain optimization, purchasing behavior, predictive analytics

I. INTRODUCTION

Accurate demand forecasting plays a central role in supply chain management, inventory control, and strategic decision-making in modern enterprises. Organizations rely on reliable forecasts to plan production volumes, manage stock levels, optimize logistics, and align marketing activities with expected customer demand. Inaccurate forecasts can lead to overstocking, stockouts, increased operational costs, and reduced customer satisfaction, ultimately affecting overall business performance [1], [2].

Traditionally, demand forecasting has been dominated by statistical time-series models such as autoregressive integrated moving average (ARIMA), seasonal ARIMA (SARIMA), and exponential smoothing techniques. These approaches are well established, interpretable, and effective in relatively stable demand environments. However, they are limited in their ability to capture nonlinear patterns, complex seasonality, and sudden demand shifts caused by promotions, price changes, or external disruptions [2], [5]. As markets become increasingly volatile and data-rich, the assumptions of linearity and stationarity inherent in classical methods often restrict their forecasting accuracy.

The rapid growth of digitalization and data availability has driven increased interest in data-driven forecasting techniques. Machine learning and deep learning models have demonstrated strong potential in identifying complex relationships within large-scale datasets and incorporating multiple influencing factors beyond historical demand alone [3], [4]. Evidence from large-scale forecasting competitions, such as the M4 and M5 competitions, indicates that data-oriented approaches frequently outperform purely statistical models, particularly when demand patterns exhibit irregular behavior, multiple seasonalities, or structural changes [1], [6]. These findings have accelerated the adoption of intelligent forecasting systems in retail, logistics, and manufacturing domains.

Despite their superior predictive performance, advanced forecasting models introduce new challenges. Tree-based ensemble methods and neural networks often require extensive feature engineering and hyperparameter tuning, while deep learning architectures may suffer from reduced interpretability, limiting their adoption in managerial contexts [3], [8]. Consequently, there is a growing need for forecasting frameworks that balance predictive accuracy with practical usability and transparency. Integrating forecasting results with accessible visualization tools can help bridge the gap between advanced analytical models and business decision-makers.

This study focuses on the development and comparative evaluation of data-driven demand forecasting models using historical sales data enriched with purchasing behavior factors. The main objectives of this research are as follows:

- To review and analyze recent advances in statistical, ensemble-based, and neural forecasting approaches reported in the literature.

- To construct a structured dataset that integrates historical demand with seasonality, promotional activity, and pricing-related features.
- To implement and compare multiple forecasting approaches, including classical time-series models and modern data-driven methods.
- To evaluate forecasting performance using standard error metrics and provide interpretable visualizations for managerial use.

The contribution of this work is twofold. From a methodological perspective, it provides a systematic comparison of traditional and data-driven forecasting techniques under a unified experimental framework. From a practical perspective, it demonstrates how advanced forecasting models can be combined with accessible visualization tools to support informed decision-making in real-world business environments.

II. MATERIALS AND METHODS

This section describes the dataset, data preprocessing procedures, feature engineering strategy, forecasting models, and the training and evaluation protocol used in this study. The methodological framework is designed in accordance with best practices in demand forecasting research to ensure reproducibility, comparability of results, and practical relevance [1], [6].

A. Dataset Description

The empirical analysis is based on a structured time-series dataset containing historical sales records enriched with business-related and calendar attributes. The dataset consists of daily observations collected over a three-month period and includes multiple product categories representative of retail demand behavior. Core variables comprise the transaction date, product identifier, product category, sales volume, revenue, and price, along with binary indicators capturing promotional campaigns and public holidays.

An overview of the dataset characteristics, including temporal coverage, scope, variables, and data sources, is presented in Table I. Such enriched datasets are commonly employed in demand forecasting studies, as they enable models to capture both temporal dependencies and exogenous demand drivers [3], [6].

TABLE I. DATASET OVERVIEW

Aspect	Description
Temporal coverage	2024-01-01 to 2024-04-09 (100 daily observations)
Frequency	Daily (calendar time series)
Scope	3 products (P001–P003) across 3 categories (Electronics, Clothing, Home)
Core variables	Date, Product ID, Category, Sales Volume, Revenue, Price
Business drivers	Promotion flag, Holiday flag
Records	100 rows (demonstration scale; extensible to larger datasets)
Data source	Retail transactional data (POS/ERP) with calendar enrichment

B. Data Preprocessing and Feature Engineering

Prior to model development, the raw data were subjected to a systematic preprocessing pipeline to enhance data quality and model robustness. Missing values were identified through null checks and calendar-based gap analysis and treated using linear interpolation for short gaps and forward-filling for longer intervals. Outliers and abnormal demand spikes were detected using statistical criteria, including z-score thresholds and interquartile range analysis, and were either capped or flagged depending on their business relevance.

Duplicate records were resolved using key-based deduplication, while timestamp inconsistencies were corrected to ensure a monotonic temporal index. Price anomalies unrelated to promotional events were capped and transformed into derived variables representing relative price changes.

Following data cleaning, a comprehensive feature engineering strategy was implemented. Temporal features included lagged demand variables and rolling statistical summaries to capture autoregressive behavior and short-term trends. Calendar-based features encoded weekly and monthly seasonality, while binary indicators represented holiday and promotional effects. In addition, price-related features and interaction terms were introduced to model demand elasticity and heterogeneous promotional impacts across product categories.

A summary of the preprocessing steps and engineered features is provided in Table II. Feature engineering of this type has been shown to significantly improve forecasting accuracy, particularly for machine learning models applied to retail demand data [3], [8].

TABLE II. DATA PREPROCESSING AND FEATURE ENGINEERING SUMMARY

Stage	Method	Purpose
Missing values	Linear interpolation, forward-fill	Preserve continuity of time series
Outlier treatment	Z-score, IQR-based capping	Reduce impact of abnormal spikes
Duplicate handling	Key-based deduplication	Ensure data consistency
Temporal features	Lagged values, rolling statistics	Capture autoregressive patterns
Calendar features	Day-of-week, month, holidays	Model seasonality and demand shocks
Price-related features	Price change, discount rate	Represent demand elasticity
Interaction features	Promotion × Category	Capture heterogeneous promo effects

C. Forecasting Models

To enable a comprehensive comparative analysis, several forecasting model classes were implemented. Classical statistical baselines include autoregressive integrated moving average models and their seasonal extensions, which remain widely used due to their interpretability and strong theoretical foundations [2], [5]. In addition, a decomposable statistical forecasting approach incorporating trend, seasonality, and holiday effects was included as a hybrid baseline [7].

Beyond statistical models, ensemble learning approaches were employed to capture nonlinear relationships between engineered features. Tree-based models are particularly effective in handling mixed numerical and categorical variables and have demonstrated strong performance in demand forecasting applications [3], [9]. Furthermore, a neural forecasting model based on recurrent architectures was implemented to directly learn long-term temporal dependencies from sequential data. Such models are well suited for capturing complex demand dynamics in the presence of irregular patterns and multiple seasonalities [4], [12].

The implemented forecasting models and their key training configurations are summarized in Table III.

TABLE III. FORECASTING MODELS AND TRAINING CONFIGURATION

Model	Key Configuration
ARIMA / SARIMA	(p, d, q) selection via AIC; weekly seasonality
Decomposable statistical model	Trend and seasonality components; holiday effects
Random Forest	300–800 trees; depth 8–16; bootstrap sampling
Gradient boosting model	400–1200 estimators; learning rate 0.03–0.1; subsampling
Recurrent neural network	1–2 layers; 32–64 units; dropout 0.2; early stopping
Validation scheme	Rolling-origin evaluation (7-day horizon)
Software environment	Python ecosystem with standard forecasting libraries

D. Training and Evaluation Protocol

Model training and evaluation were conducted using a rolling-origin forecasting strategy to simulate real-world deployment conditions and avoid look-ahead bias. Sequential training, validation, and test splits were constructed with a fixed seven-day forecasting horizon, ensuring that each model was evaluated on unseen future observations.

Forecasting performance was assessed using standard error metrics, including mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE). These metrics provide complementary perspectives on predictive accuracy and are widely adopted in forecasting research and business practice [1], [6].

E. Exploratory Visualization

Exploratory visual analysis was performed to examine demand dynamics prior to model evaluation. Fig. 1 illustrates daily sales dynamics over time, highlighting seasonal patterns and demand spikes associated with holidays and promotional events.

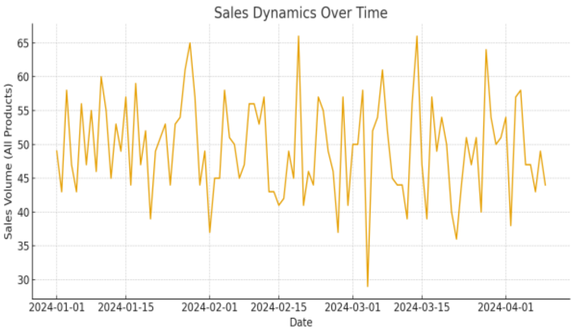

Figure 1. Sales dynamics over time (Excel line chart). Seasonal peaks appear during holiday periods and promotions.

Category-level aggregation of sales volumes is presented in Fig. 2, demonstrating differences in cumulative demand and volatility across product categories. These visualizations support feature engineering decisions and provide contextual understanding of the dataset before forecasting.

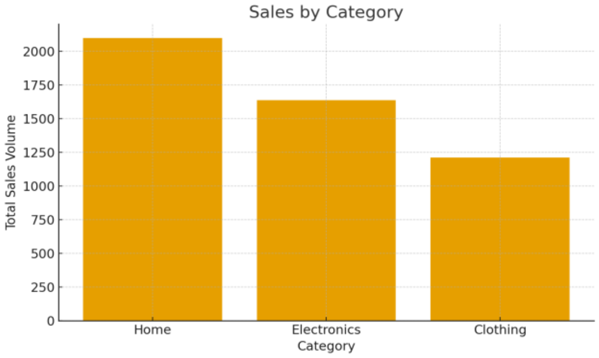

Figure 2. Total sales volume by category. Clothing leads overall demand, while Electronics show greater volatility.

III. RESULTS AND DISCUSSION

This section presents the experimental results obtained from the evaluated forecasting models and discusses their performance in the context of demand forecasting accuracy and practical applicability. Quantitative metrics are complemented by visual analysis to provide a comprehensive comparison of forecasting approaches.

A. Quantitative Forecasting Results

All forecasting models were trained and evaluated using the rolling-origin strategy described in Section II. The comparative accuracy results on the test set are summarized in Table IV, which reports mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE).

TABLE IV. FORECASTING ACCURACY COMPARISON (TEST SET)

Model	MAE	RMSE	MAPE (%)
Statistical baseline	7.82	10.15	14.2
Decomposable statistical model	7.35	9.88	13.7
Ensemble learning model	6.20	8.50	11.5
Gradient boosting model	5.95	8.12	10.9
Neural forecasting model	5.10	7.45	9.6

The results indicate a clear performance differentiation between classical statistical approaches and data-driven forecasting models. Statistical baselines provide a reasonable approximation of overall demand trends but exhibit higher error rates, particularly during periods of abrupt demand change. These limitations are consistent with findings reported in large-scale forecasting competitions, where traditional models struggle to adapt to volatile demand environments [1], [6].

Ensemble learning models demonstrate a substantial reduction in forecasting errors by leveraging engineered features related to promotions, pricing, and seasonality. Their ability to model nonlinear relationships enables more accurate tracking of demand fluctuations compared to purely statistical methods. Gradient boosting approaches, in particular, show improved robustness, as reflected in lower RMSE values.

The neural forecasting model achieves the lowest error values across all reported metrics. This result confirms its effectiveness in capturing both long-term dependencies and short-term demand variations, supporting evidence from recent studies highlighting the strengths of sequence-based learning for time-series forecasting [4], [12].

B. Visual Comparison of Forecasts

To complement the quantitative evaluation, a visual comparison between actual and forecasted demand was conducted. Fig. 3 illustrates the alignment between observed sales and model-generated forecasts over the test period.

The visual results confirm the numerical findings reported in Table IV. Statistical baselines tend to smooth demand trajectories and underestimate peak periods, while data-driven models more closely follow observed fluctuations. The neural forecasting model demonstrates the closest alignment with actual demand, particularly during periods characterized by rapid changes and irregular patterns.

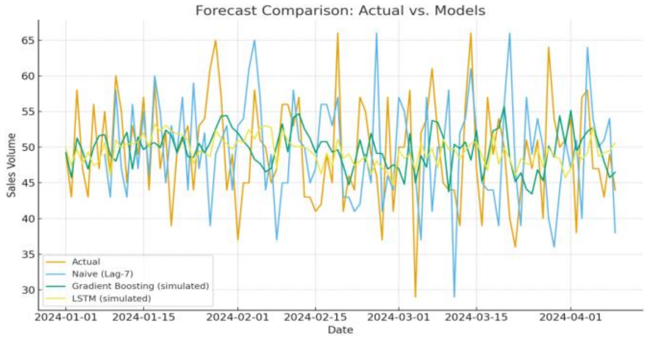

Figure 3. Comparison of actual vs. forecasted sales across models.

C. Discussion and Implications

The empirical findings highlight several important insights. First, while classical statistical models remain valuable for interpretability and baseline comparison, their restrictive assumptions limit their effectiveness in dynamic retail environments. Second, ensemble-based machine learning models offer a strong balance between flexibility and accuracy, particularly when supported by structured feature engineering. Third, neural forecasting models provide superior predictive performance when demand patterns involve complex temporal dependencies.

From a practical standpoint, the combination of quantitative metrics and visual analytics enhances interpretability for decision-makers. Presenting forecasting results through clear tables and intuitive figures enables business users to assess model reliability without requiring advanced technical expertise. Such integration is essential for translating advanced forecasting techniques into actionable insights for inventory planning, demand management, and strategic decision-making.

Overall, the results demonstrate that data-driven forecasting systems significantly improve demand prediction accuracy and support more informed operational planning. These findings align with recent research emphasizing the growing role of intelligent information systems in supply chain optimization and decision support [3], [6].

IV. CONCLUSION

This study examined the effectiveness of data-driven demand forecasting approaches using historical sales data enriched with purchasing behavior factors. A comparative evaluation of classical statistical models, ensemble-based machine learning methods, and neural forecasting architectures was conducted to assess their predictive accuracy and practical applicability.

The experimental results demonstrate that while traditional statistical approaches provide a useful baseline and offer strong interpretability, they are limited in their ability to capture nonlinear demand patterns and abrupt fluctuations. In contrast, data-driven models consistently achieve higher forecasting accuracy, particularly when exogenous variables such as promotions, pricing, and calendar effects are incorporated. Among the evaluated approaches, neural forecasting models exhibit the strongest performance, effectively modeling long-term temporal dependencies and irregular demand behavior.

The main contributions of this work can be summarized as follows:

- A systematic comparison of classical, ensemble-based, and neural forecasting models under a unified experimental framework.
- The development of a structured data preprocessing and feature engineering pipeline tailored to retail-style demand data.
- The integration of quantitative accuracy metrics with intuitive visual analytics, enhancing the interpretability and managerial usability of forecasting results.

Despite these contributions, several limitations should be acknowledged. The empirical analysis is based on a relatively small-scale dataset covering a limited number of product categories, which may restrict the generalizability of the findings. In addition, the study does not incorporate external macroeconomic variables, competitive dynamics, or environmental factors that may significantly influence demand patterns in real-world settings.

Future research should focus on extending the proposed framework to larger, multi-product and multi-region datasets and exploring hierarchical forecasting strategies. Incorporating external signals such as macroeconomic indicators and market trends may further enhance predictive performance. Moreover, emerging transformer-based architectures offer promising directions for improving scalability and long-horizon forecasting accuracy.

Overall, the findings highlight the significant potential of data-driven demand forecasting systems to improve prediction accuracy, support informed decision-making, and enhance operational efficiency in modern supply chain environments.

REFERENCES

[1] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “The M4 competition: Results, findings, conclusion and way forward,” Int. J. Forecast., vol. 34, no. 4, pp. 802–808, 2018, doi: 10.1016/j.ijforecast.2018.06.001.

[2] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. Melbourne, Australia: OTexts, 2021.

[3] K. Bandara, C. Bergmeir, and H. Hewamalage, “Forecasting across multiple temporal aggregation levels and seasonalities with machine learning,” IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 7, pp. 2579–2593, 2020, doi: 10.1109/TNNLS.2019.2950730.

[4] D. Salinas, V. Flunkert, J. Gasthaus, and T. Januschowski, “DeepAR: Probabilistic forecasting with autoregressive recurrent networks,” Int. J. Forecast., vol. 36, no. 3, pp. 1181–1191, 2020, doi: 10.1016/j.ijforecast.2019.07.001.

[5] E. Spiliotis, S. Makridakis, and V. Assimakopoulos, “Generalizing the findings of the M competitions,” Eur. J. Oper. Res., vol. 272, no. 2, pp. 626–635, 2019, doi: 10.1016/j.ejor.2018.06.045.

[6] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, “The M5 accuracy competition: Results, findings and conclusions,” Int. J. Forecast., vol. 38, no. 4, pp. 1346–1364, 2022, doi: 10.1016/j.ijforecast.2021.11.013.

[7] S. J. Taylor and B. Letham, “Forecasting at scale,” in Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2017, pp. 793–802, doi: 10.1145/3097983.3098070.

[8] S. Smyl, “A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting,” Int. J. Forecast., vol. 36, no. 1, pp. 75–85, 2020, doi: 10.1016/j.ijforecast.2019.03.017.

[9] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[10] G. Ke, Q. Meng, T. Finley, et al., “LightGBM: A highly efficient gradient boosting decision tree,” in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 2017.

[11] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: Unbiased boosting with categorical features,” in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 2018.

[12] B. N. Oreshkin, D. Carpov, N. Chapados, and Y. Bengio, “N-BEATS: Neural basis expansion analysis for interpretable time series forecasting,” in Proc. Int. Conf. Learn. Represent. (ICLR), 2020.

[13] B. Lim and S. Zohren, “Time-series forecasting with temporal fusion transformers,” in Proc. AAAI Conf. Artif. Intell., vol. 35, no. 14, pp. 12674–12681, 2021.

[14] H. Zhou, S. Zhang, J. Peng, et al., “Informer: Beyond efficient transformer for long sequence time-series forecasting,” in Proc. AAAI Conf. Artif. Intell., vol. 35, no. 12, pp. 11106–11115, 2021.

[15] H. Wu, J. Xu, J. Wang, and M. Long, “Autoformer: Decomposition transformers with autocorrelation for long-term series forecasting,” in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 2021.

[16] H. Liu, J. Hu, X. Xu, et al., “PatchTST: Patch-based time-series transformers,” in Proc. Int. Conf. Learn. Represent. (ICLR), 2023.

[17] C. Challu, K. Olivares, B. N. Oreshkin, et al., “N-HiTS: Neural hierarchical interpolation for time series forecasting,” in Proc. AAAI Conf. Artif. Intell., 2023.