Volume 3, No. 11 November 2025
p-ISSN 3032-3037 | e-ISSN 3031-5786

Machine Learning-Based Forecasting of Incoming Customer Orders: A Case Study on Heavy-Duty Filter Products

Taryono*, Tanika D. Sofianti, Eka Budiarto
Swiss German University, Indonesia
Email: taryono@student.sgu.ac.id*, tanika.sofianti@sgu.ac.id, eka.budiarto@sgu.ac.id

Abstract
Accurate demand forecasting plays a critical role in inventory management and production planning, particularly in the manufacturing of heavy-duty filter products, where customer order patterns are influenced by seasonal and external factors. This study proposes a machine learning–based forecasting framework to predict incoming customer orders using structured sales data. The analysis focuses on the 2021–2025 period, during which demand behavior was observed to be more stable compared to previous years. The study evaluates the effects of feature engineering and feature selection on forecast accuracy, with special attention to time-based, seasonal, and external economic variables. Random Forest, LightGBM, and Linear Regression models were developed, and their performance was assessed using metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), MAPE, and R^2 score. Experimental results show that the inclusion of seasonal and exogenous features significantly enhances model performance, particularly for item types with pronounced cyclical trends. Furthermore, the study conducts hypothesis testing to assess the statistical impact of engineered features and model types on forecast accuracy. The findings demonstrate the applicability of machine learning in industrial forecasting scenarios, offering a replicable and scalable approach to data-driven decision making in procurement and inventory planning.
Keywords: Demand Forecasting, Machine Learning, Feature Engineering, Feature Selection, Random Forest, LightGBM, Linear Regression, Heavy-Duty Filters, Forecast Accuracy

Introduction

In today's competitive manufacturing environment, accurate demand forecasting is essential for achieving operational efficiency and ensuring customer satisfaction. This is especially true for manufacturers of heavy-duty filter products, where thousands of Stock Keeping Units (SKUs) exhibit diverse demand patterns and customer-specific order behaviors. The forecast challenge is intensified by irregular purchasing cycles, long product life spans, and the sector's exposure to external industrial volatility (Morra & Santangelo, 2024; Perdana Wanti & Maharrani, 2020; Saravanan & R., 2024; Wolny & K., 2024; Zedda & Perniciano, 2024).

The root causes of demand forecasting failures in such environments stem from four main dimensions, as identified through a Fishbone (Ishikawa) analysis: (1) data limitations, such as missing or misaligned sales records; (2) method-related issues, including reliance on simplistic or outdated forecasting models; (3) process deficiencies, such as lack of historical tracking or integration with planning systems; and (4) environmental volatility, such as economic shocks or seasonal procurement behavior. These issues create significant disruptions in planning activities, leading to inventory inaccuracy, reactive ordering, and missed sales opportunities.

The ability to predict customer orders becomes even more critical when demand is influenced by recurring seasonal patterns, such as religious holidays or fiscal year-end purchasing cycles (Chong & Kim, 2022; Dautel & Thakur, 2024; Kim, 2023; Kumar &

Khedlekar, 2024; Rohit & Verma, 2025). Recognizing these patterns is fundamental to building robust forecasting models that support proactive planning and timely stock replenishment.

Traditional statistical forecasting methods, such as moving averages and exponential smoothing, often fail to capture the complex and nonlinear nature of real-world sales data (Feleafel & Radulovic, 2025; Jebbor & Benmamoun, 2024; Sakthi Balan & Santhosh Kumar, 2025; Sokolova & Nwaila, 2025). These methods depend on assumptions of stability and linearity, making them poorly suited to handle dynamic demand shifts, noise, or irregular behavior. In contrast, machine learning (ML)-based forecasting offers adaptive capabilities to learn hidden relationships and generate more accurate predictions (Destro & Staudt, 2023; Memon & Sami, 2020; Nurgazina & Felberbauer, 2022).

Recent studies confirm the value of ML in supply chain forecasting. For example, Vikram Pasupuleti (2024) demonstrated that ML-based forecasts improved accuracy by 15%, reduced stock shortages by 10%, and improved on-time delivery rates. Similarly, Feizabadi (2022) showed how hybrid ML models, such as ARIMAX combined with neural networks, reduced cash conversion cycles by better aligning forecasts with procurement cycles. These findings support the use of ML in domains with complex, repeatable but volatile demand signals, typical of the heavy-duty filter industry (Kao, 2002; Mohammed & Hassan, 2022; Nambiar & M., 2022; Nguyen & Lamouri, 2021; Sel, 2025).

This study uses historical and recent demand data to uncover consistent behavioral patterns and develop forecasting models tailored to product-level customer orders. Although earlier sales data from 2018–2020 is used to detect major disruptions (e.g., COVID-19), the primary modeling scope focuses on the more stable 2021–2025 period. Feature engineering techniques are applied, such as moving average smoothing, lag features, and seasonal flags, to improve the predictive power of the models. Two machine learning algorithms, Random Forest and LightGBM, are implemented and compared across multiple accuracy metrics.

The goal is to provide a practical and scalable machine learning framework that improves the reliability of the forecast at the SKU level. Although the study does not extend to full inventory optimization, it provides the foundational information to integrate demand forecasts into broader supply chain planning systems.

To better understand the underlying challenges in demand forecasting within manufacturing settings, this study uses a Fishbone (Ishikawa) diagram (Figure 1.1) to visualize the root causes behind the absence of reliable forecasts and the resulting disruption in planning activities. The diagram categorizes the sources of forecast failures into four domains: data-related issues, methodological limitations, process constraints, and external environmental factors.

Data-related problems include inconsistent demand patterns, insufficient historical granularity, and unaddressed seasonality, all of which distort model inputs and hinder forecast accuracy. On the methodological side, firms often rely on basic models, such as moving averages or naive extrapolation, which fail to capture nonlinear interactions, lag effects, or multidimensional feature relationships (Arunadevi, 2024; D'Orazio, 2024; Feizabadi, 2022; Fida, 2024). These are further worsened by infrequent model updates and limited expertise in selecting or tuning appropriate models.

Process-related constraints arise from organizational silos, lack of coordination between departments such as sales and operations, and the absence of a structured forecasting workflow. As a result, any forecast that is produced may be disconnected from execution or decision-making processes. Meanwhile, external environmental factors, including market volatility, geopolitical changes, customer behavior changes, and one-off events such as pandemics, further undermine the stability of forecast models.

The core insight from this diagnostic view is that the inaccuracy of the forecast is not based on a single root cause but rather on the interaction of data, process, model, and external dimensions. Addressing these requires a holistic strategy that integrates data quality improvements, organizational alignment, and predictive modeling capabilities.

This study proposes machine learning–based forecasting as a response to these challenges, supported by engineered features and structured time series inputs. The goal is to develop a more reliable and adaptive forecasting system that improves SKU-level accuracy and supports responsive inventory and production planning.

The benefits of this research are both theoretical and practical. Theoretically, it contributes to the advancement of forecasting literature by combining machine learning techniques and operations management perspectives within a unified analytical framework. Practically, the findings are expected to help manufacturing firms reduce inventory imbalances, minimize production delays, and optimize resource allocation through more reliable demand predictions.

Method

Section describes the forecasting framework developed in this study to predict monthly demand at the SKU level for heavy-duty filter products. The framework consists of five main stages: problem identification and data collection, exploratory data analysis, model development, forecast evaluation, and insight generation. The study begins by addressing the absence of a systematic demand forecasting system and collecting historical customer order data from January 2018 to December 2025. Exploratory analysis is then conducted to identify demand patterns, trends, seasonality, and outliers, with the 2021–2025 period selected as the modeling window due to its stable demand signals. The model development phase includes the creation of baseline models (Naïve and Moving Average) as benchmarks, followed by advanced machine learning models such as Random Forest, LightGBM, and Linear Regression, trained using engineered features like lag variables, moving averages, temporal indicators, and calendar-based events. To enhance prediction accuracy, feature selection is performed using the Feature Take-Out Simulation method, which evaluates the impact of removing each feature on forecasting accuracy (ΔMAPE). This approach directly measures feature importance and ensures only relevant variables are retained. The selected features are then used to train optimized models, evaluated using metrics such as MAE, MAPE, RMSE, and R^2. The project is considered successful when machine learning models demonstrate significant improvements in forecast accuracy compared to baseline models and provide practical insights into temporal and behavioral factors influencing demand, thereby supporting monthly planning processes in demand, supply, and inventory management.

Results and Discussion

Exploratory Data Analysis (EDA) Results

This section presents a high-level exploratory data analysis (EDA) of the dataset covering the period from January 2021 to April 2025. The objective is to identify underlying patterns in demand behavior, including demand distribution across product types, individual SKU performance, temporal trends, and seasonality driven by calendar events or fiscal cycles. These insights form the empirical foundation for feature engineering and modeling decisions in the subsequent stages.

Demand Trends by Year

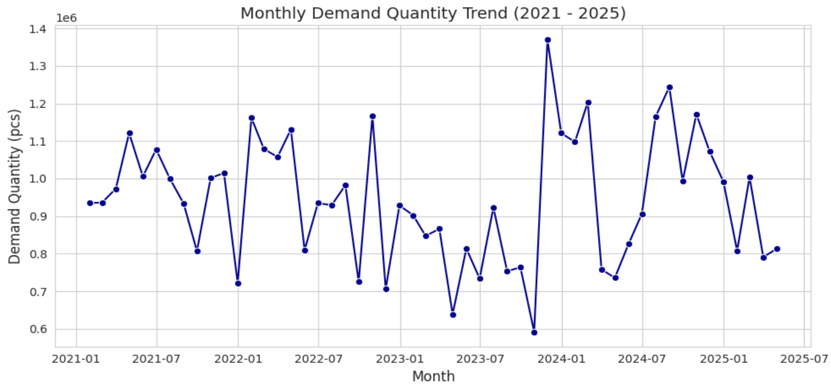

Figure 1. Monthly demand quantity trend across all SKUs (2021–2025)
source: processed data

Figure 1 shows the total monthly demand quantity in all SKUs from 2021 to 2025. Although the pattern displays periodic peaks and troughs, a significant upward shift in demand is observed in early 2024. Fluctuations are evident around mid-year and year-end periods, suggesting possible alignment with business cycles or external events.

Demand Distribution by Item Type

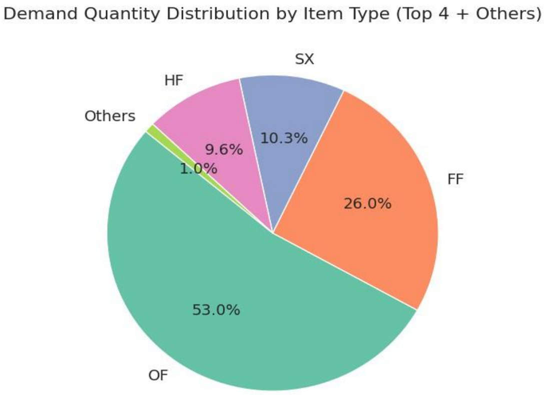

Figure 2. Demand quantity distribution by item type (Top 4 + Others)
source: processed data

To understand which types of products, dominate the order volume, Figure 2 presents a pie chart of the cumulative demand quantity grouped by item type. The analysis confirms that the oil filter category (OF) contributes more than 50% of the total demand, followed by the fuel filter (FF), the special design filter (SX) and the hydraulic filter (HF). Other item types collectively contribute less than 1%, indicating that they play a marginal role in overall forecasting.

Monthly and Yearly Seasonality Patterns

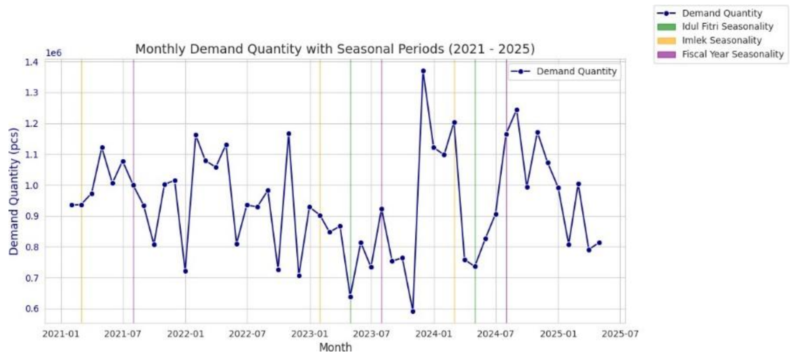

Figure 3. Monthly demand quantity with seasonal calendar flags (2021–2025)
source: processed data

To assess the existence of systematic patterns in demand, Figure 3 plots monthly demand over time while highlighting known seasonal events: Idul Fitri, Imlek (Chinese New Year) and fiscal year-end periods. In particular, several spikes in demand align with these seasonal flags, especially in the second quarter (April-May) and fourth quarter (October-December), consistent with festive procurement cycles and year-end budget utilization.

Macro-Economic Context: IHSG and Currency Trends

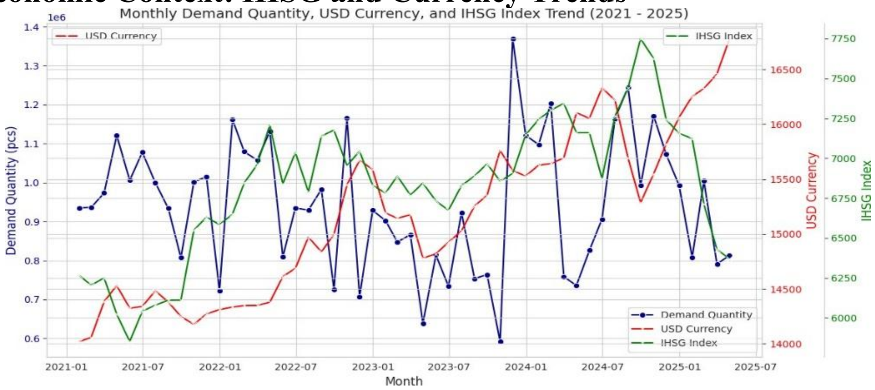

Figure 4. Monthly demand quantity vs. USD exchange rate and IHSG index
source: processed data

To explore potential macroeconomic influences on order behavior, Figure 4 overlays the quantity of demand with the USD exchange rate and the Indonesian stock index (IHSG). Although direct causality is not inferred at this stage, preliminary trends suggest some degree of correlation. For example, demand tends to decline during periods of exchange rate depreciation or stock market contraction, indicating the relevance of these external variables for further model training and feature construction.

The EDA confirms that demand patterns are influenced by both internal (product and calendar structure) and external (macroeconomic) factors. These observations will guide the inclusion of appropriate features in the subsequent modeling framework (Dai, 2024; Lee & Lv, 2017; Ni & Xiao, 2020; Riachy & He, 2025).

Overview of the Experiment Process

This chapter presents the empirical results of the forecasting experiments conducted on heavy-duty vehicle filter products. The analysis begins with the baseline models, including the Naive and Moving Average approaches, to establish reference performance levels. Subsequently, the results of the machine learning models, Random Forest, LightGBM, and Linear Regression, are reported and compared. Each model is evaluated using the standard accuracy metrics: Mean Absolute Percentage Error (MAPE) and the coefficient of determination (R2). The results are organized into three stages of experimentation: (i) baseline performance without feature engineering, (ii) performance after incorporating feature engineering, and (iii) performance after applying feature selection via the Feature Take-Out Simulation. This structure enables a clear assessment of how methodological enhancements contribute to the accuracy and robustness of demand forecasting.

Summary of Dataset Used (2021–2025)

The primary dataset utilized in this study comprises customer order records from January 2, 2021 to April 29, 2025. This dataset was curated from historical sales transactions and filtered to ensure consistency, reliability, and relevance for the machine learning forecasting task. The final dataset contains 54,229 rows and 14 variables, covering transactional, categorical, seasonal and economic dimensions.

a) sales_date: The transaction date of each sales record.
b) item_code: A unique identifier for each Stock Keeping Unit (SKU).
c) item_type: Product groupings, e.g., FF (Fuel Filter), HF (Hydraulic Filter), OF (Oil Filter), SX (Special Design Filter).
d) sales_qty: The quantity of product ordered.
e) channel, region, sales_market, customer_name: Business attributes representing the source of demand.
f) year: Extracted year from the sales date variable.
g) is_seasonal_idul_fitri, is_seasonal_imlek, is_seasonal_fiscal_year: Binary flags indicating seasonal or calendar-based influences.
h) x_usd_currency, ihsg_index: External economic indicators relevant to the market context.

No missing values were found for any of the variables, indicating that the dataset is clean and does not require imputation.

The dataset predominantly comprises four major item types:

a) OF (Oil Filter): 23,395 records
b) FF (Fuel Filter): 15,007 records
c) HF (Hydraulic Filter): 7,425 records
d) SX (Special Design Filter): 7,363 records

Other item types such as WC, AY, AX, AF, and GF contribute less than 2% of the total observations and are excluded from the primary forecast analysis.

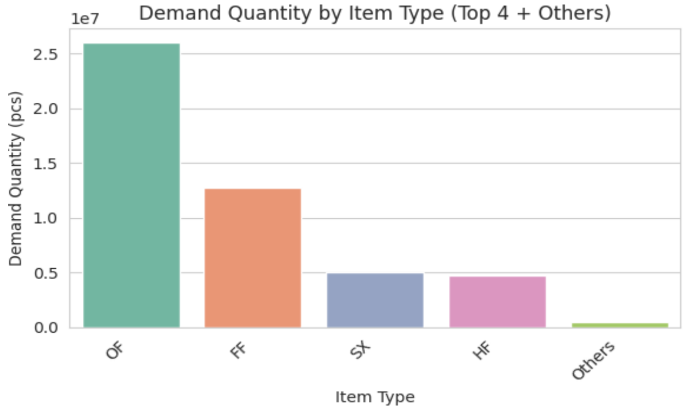

Figure 5. Bar chart of item type distribution within the dataset. Major product types (OF, FF, HF, SX) dominate the transaction records
source: processed data

In Figure 5, this distribution justifies the focus on the four major item types (OF, FF, HF, SX) for model training and evaluation. The presence of temporal features and economic variables further supports multivariate modeling and time-sensitive forecasting tasks in later stages.

1. Reiteration of Modeling Phases: Baseline, Feature Engineering, and Feature Selection

To systematically assess the impact of input transformation on forecast accuracy, the modeling process in this study is divided into three sequential phases: (1) Baseline Modeling,

(2) Feature Engineering, and (3) Feature Selection. Each phase builds on the previous one to allow for a clear measurement of the added value provided by new features or dimensionality reduction strategies.

Baseline Modeling

Two types of statistical baseline, Naïve and Moving Average models that assume demand repetition or averaging between past values are evaluated:

a) Naïve Model: The Naïve model assumes that demand in the next period is equal to the most recent observed demand.
b) Moving Average Model: The Moving Average (MA) model assumes that future demand can be estimated by averaging a fixed number of past observations. This model provides a simple but effective baseline for evaluating whether machine learning models significantly improve forecast performance.

Both models are evaluated in four key item types Fuel Filter (FF), Hydraulic Filter (HF), Oil Filter (OF), and Special Design Filter (SX) using the following performance metrics.

a) MAPE (%): Mean Absolute Percentage Error
b) R2: Coefficient of determination, indicating the variance explained

Performance Metrics Naïve Forecasting Model

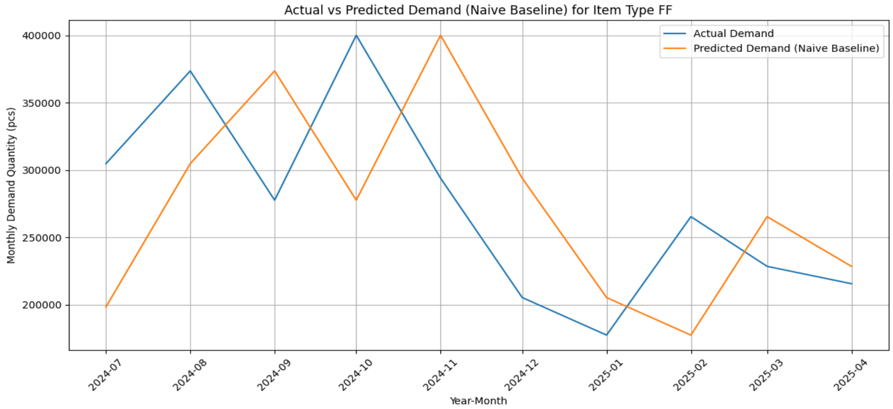

Figure 6. Actual demand for Item Type FF shows sharp fluctuations, while the naïve baseline lags behind and fails to capture turning points. This demonstrates the inability of the naïve method to represent volatile demand patterns
source: processed data

In Figure 6, the comparison between actual demand and the naive baseline for the type of FF item reveals that the naïve method consistently lags behind the trend of actual demand. Although the actual series exhibits sharp fluctuations, peaking in October 2024 and then declining steeply toward early 2025, the naïve baseline produces a smoother trajectory that simply repeats the value of the previous month. As a result, the baseline fails to capture turning points: for instance, it overestimates demand in September 2024 and underestimates demand during the subsequent decline from November 2024 onward. This lag behavior highlights that the naïve approach cannot adequately represent the volatility of FF demand, reinforcing the need for more advanced forecasting models capable of capturing abrupt changes in trend.

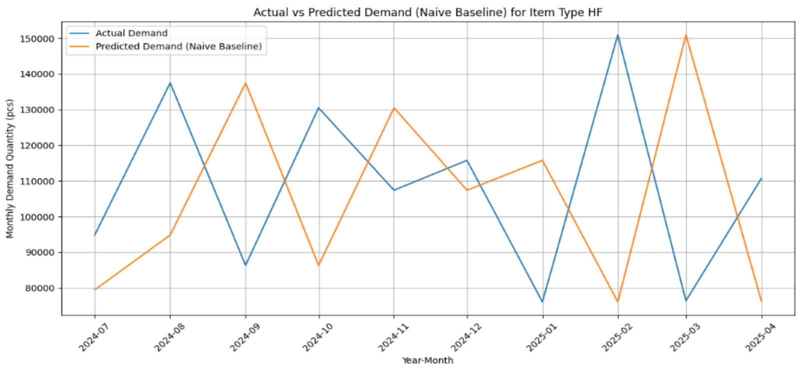

Figure 7. Actual demand for Item Type HF exhibits sharp peaks and troughs, while the naïve baseline produces a smoother, lagging response. This lag causes the baseline to miss sudden turning points and underestimate demand volatility
source: processed data

In Figure 7, the comparison between actual demand and the naive baseline for item type HF shows that the naïve method provides only a rough directional approximation, but fails to capture the high volatility of the series. Actual demand exhibits sharp peaks and deep troughs, such as the sudden rise in August 2024 and February 2025, followed by steep declines in subsequent months. In contrast, the naïve baseline produces a smoother and lag-prone response, as it simply repeats the value of the previous month.

Consequently, the baseline systematically misses critical turning points and underestimates the magnitude of demand fluctuations. This reinforces the limitation of the naïve method when applied to volatile product categories such as HF.

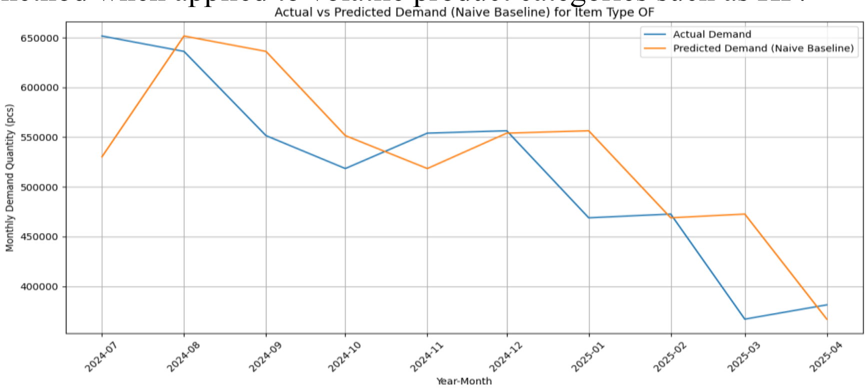

Figure 8. Actual and naïve baseline demand for Item Type OF both show a general downward trend, but the baseline tends to overestimate during declines and lags at turning points. While closer to actual demand than in other item types, the naïve method still fails to capture sharper shifts in the trend

source: processed data

In Figure 8, the comparison between actual demand and the naïve baseline for the type OF OF item indicates that both series follow a generally downward trend from mid-2024 to early 2025. Unlike FF and HF, the naïve baseline aligns more closely with the actual demand because OF demand is relatively smoother and less volatile. However, the baseline tends to overestimate demand during periods of decline, such as from August to October 2024, and reacts with a delay when actual demand stabilizes or recovers slightly, as seen in November–December 2024. Although the naïve approach provides a rough approximation of the overall trajectory, it still fails to capture the sharper downturns and recovery points, limiting its usefulness for operational planning.

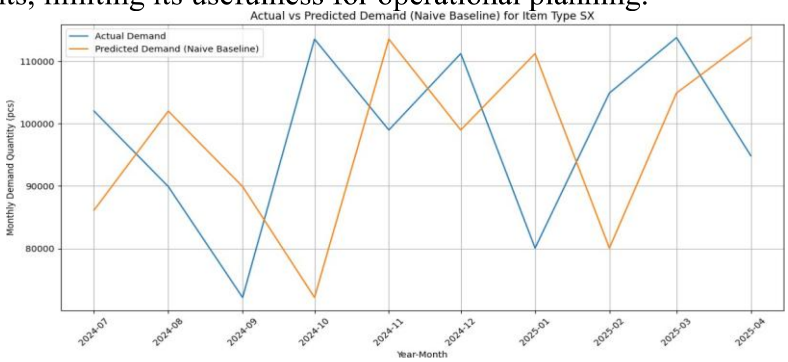

Figure 9. Actual demand for Item Type SX fluctuates sharply, while the naïve baseline provides a lagging and smoother approximation. This causes systematic underestimation during upswings and overestimation during downturns, showing the method's weakness in volatile demand conditions

source: processed data

In Figure 9, the comparison between actual demand and the naive baseline for the type of item SX highlights the limitations of the naïve approach when applied to highly volatile demand patterns. The actual series exhibits strong fluctuations with sharp peaks and troughs, such as the steep decline in September 2024 followed by a sharp rebound in October 2024. In contrast, the naïve baseline lags behind, underestimating demand during sudden upswings and overestimating during steep declines. Although the baseline follows the general directional movement, it smooths out volatility and systematically misses turning points, demonstrating that it is inadequate to capture the irregular dynamics of SX demand.

Across all types of items, the comparison between actual demand and the naïve baseline reveals consistent weaknesses in the naïve approach. For FF and HF, where demand is highly volatile with sharp peaks and troughs, the naïve baseline systematically lags behind actual demand, leading to frequent overestimation during downturns and underestimation during upswings. SX exhibits even stronger volatility, where the baseline smoothing effect completely fails to capture erratic fluctuations, making it unsuitable for operational use. In contrast, OF demand follows a relatively smoother downward trajectory, allowing the naïve baseline to align more closely with the actual trend, although it still tends to overestimate during declines and react slowly at turning points. Overall, these results demonstrate that while the naïve method may provide a rough approximation in more stable series such as OF, it is inadequate to capture the dynamics of volatile product families such as FF, HF and SX. This reinforces the need for more advanced forecasting models to achieve meaningful predictive accuracy across different filter categories.

Performance Metrics Moving Average Model

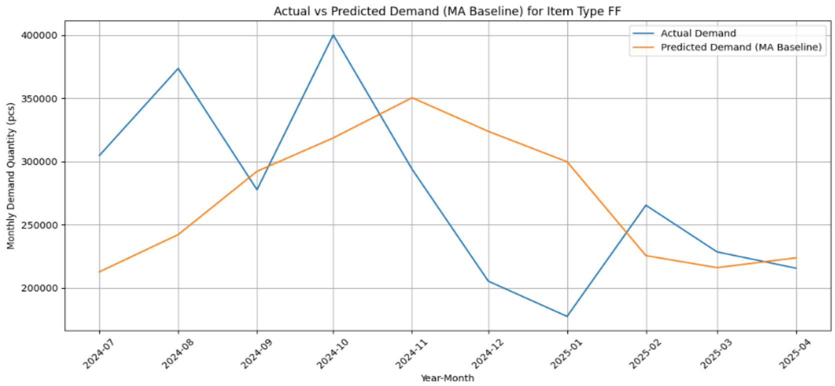

Figure 10. Actual demand for Item Type FF shows sharp peaks and troughs, while the moving average baseline smooths these fluctuations into a gradual trend. This results in underestimation during demand surges and overestimation during steep declines, limiting its usefulness in volatile demand conditions
source: processed data

In Figure 10, the comparison between actual demand and the moving average (MA) baseline for Item Type FF shows that the MA baseline captures the overall long-term direction of demand but does not represent its volatility. The actual demand pattern is highly irregular, with sharp peaks such as in October 2024 and steep declines in December 2024 to January 2025. In contrast, the MA baseline produces a smoother trajectory by averaging past values, resulting in systematic underestimation during sudden surges and overestimation during rapid drops. Although the MA approach is useful for highlighting the general directional trend, it does not provide reliable short-term accuracy, making it unsuitable for operational planning in volatile product categories like FF.

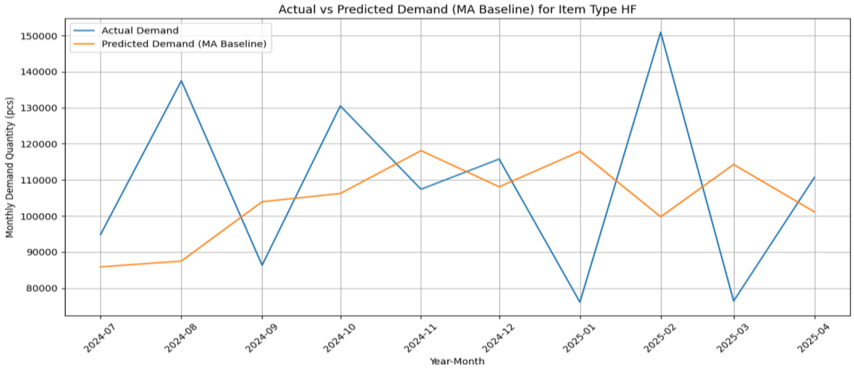

Figure 11. Actual demand for Item Type HF shows sharp volatility, while the moving average baseline smooths fluctuations into a gradual trend. This smoothing effect causes underestimation during spikes and overestimation during drops, limiting its suitability for volatile demand conditions

source: processed data

In Figure 11, the comparison between actual demand and the moving average (MA) baseline for Item Type HF highlights the limitations of the MA approach when applied to volatile demand patterns. The actual demand series exhibits sharp fluctuations with abrupt peaks, such as in August 2024 and February 2025, and steep declines in September 2024 and March 2025. In contrast, the MA baseline produces a smoother and more gradual trajectory, dampening these extremes and responding more slowly to sudden changes. This results in systematic underestimation during rapid demand surges and overestimation during sharp declines. Although the MA baseline can reveal the general directional trend, it lacks the responsiveness required to capture the short-term volatility of HF demand, limiting its usefulness for operational forecasting.

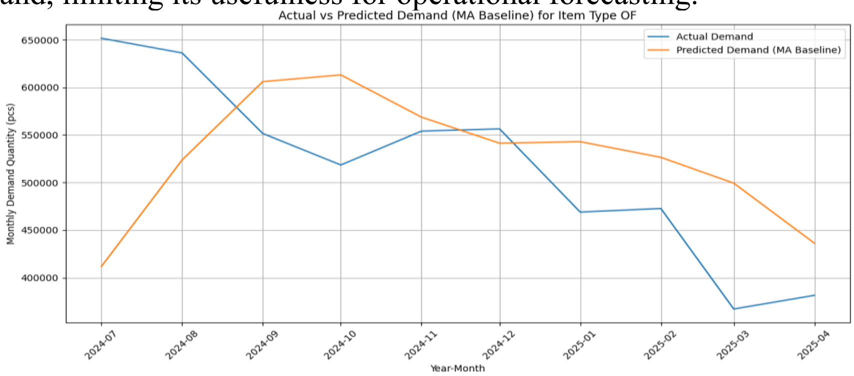

Figure 12. Actual and moving average demand for Item Type OF both show a downward trend, but the MA baseline smooths sharper fluctuations into a gradual decline. This leads to overestimation during steep drops and a lagging response during partial recoveries

source: processed data

In Figure 12, the comparison between actual demand and the moving average (MA) baseline for Item Type OF shows that both series follow a general downward trend from mid-2024 to early 2025. However, actual demand exhibits more pronounced declines, particularly in September 2024 and March 2025, whereas the MA baseline smooths these changes into a gradual decline. This smoothing effect reduces the volatility, but also results in systematic overestimation during sharper downturns and a lagging response when actual demand partially recovers, as in November 2024 and April 2025. Although the MA baseline is more aligned with OF demand than in more volatile item types, it still fails to capture the timing and magnitude of sudden changes, limiting its effectiveness for precise forecasting.

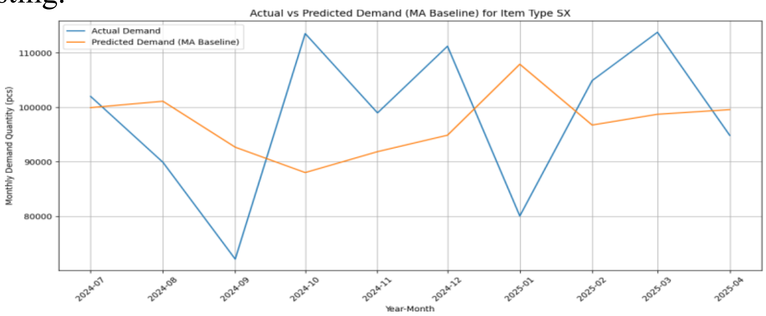

Figure 13. Actual demand for Item Type SX is highly volatile, while the moving average baseline smooths fluctuations into a gradual line. This results in underestimation during peaks and overestimation during troughs, showing that the MA approach cannot capture irregular demand patterns

source: processed data

In Figure 13, the comparison between actual demand and the moving average (MA) baseline for Item Type SX shows that the MA baseline is unable to capture the highly volatile nature of SX demand. The actual series exhibits extreme fluctuations, with sharp peaks in October 2024 and March 2025 and steep troughs in September 2024 and January 2025. In contrast, the MA baseline produces a smoother trajectory that averages these abrupt changes, leading to systematic underestimation during demand surges and overestimation during sharp declines. This smoothing and lagging behavior makes the MA approach unsuitable for forecasting short-term demand in irregular and volatile product categories such as SX.

Baseline Selection for Model Evaluation

Table 1. Baseline Forecasting Results by Item Type

Item Type	Metrics	Naïve Model	Moving Average
OF (Oil Filter)	MAPE (%)	26.84	26.08
	R^2	-0.4844	-0.4240
FF (Fuel Filter)	MAPE (%)	-	-
	R^2	-2.5562	-0.6064
HF (Hydraulic Filter)	MAPE (%)	39.94	24.85
	R^2	-	-
(not mentioned)	MAPE (%)	10.16	16.56
	R^2	0.4663	-0.3656
(not mentioned)	MAPE (%)	-	-
	R^2	-1.6821	-0.4544
SX (Special Design Filter)	MAPE (%)	20.61	14.80
	R^2	-	-

source: processed data

The baseline results summarized in Table 1 provide the foundation for evaluating the performance of the proposed machine learning models. For each type of items, the choice of the baseline is determined by identifying the method (naive or moving average) that produces the lowest forecast error, as indicated by the Mean Absolute Percentage Error (MAPE) and supported by the value R2. The selection is explained as follows:

a) FF (Fuel Filter): The Moving Average model achieves a lower MAPE (26.08%) compared to the naive model (26.84%), with a slightly better value R2. Therefore, the Moving Average is chosen as the baseline.

b) HF (Hydraulic Filter): The Naïve model performs poorly with a MAPE of 39.94% and a negative R2 of -2.5562. In contrast, the Moving Average significantly reduces the error (MAPE = 24.85%). Thus, the Moving Average is selected as the baseline.
c) OF (Oil Filter): The Naïve model clearly outperforms the Moving Average with the lowest MAPE (10.16%) and a positive R2 (0.4663). Therefore, the Naïve model is used as the baseline for this type of item.
d) SX (Special Design Filter): The Moving Average demonstrates better performance with a lower MAPE (14.80%) compared to the Naïve model (20.61%). Hence, the Moving Average is selected as the baseline.

In summary, the baselines used for the next stage of model evaluation are as follows. Moving Average for FF, HF and SX, while the Naive model is retained for OF. These baselines provide appropriate benchmarks to assess the predictive improvements offered by machine learning models.

Conclusion

This study investigated machine learning models—Random Forest, LightGBM, and Linear Regression—for forecasting customer orders of heavy-duty vehicle filters using SKU-level data from 2021 to 2025. Machine learning methods outperformed traditional statistical approaches, which suffered from high error rates and low explanatory power. Incorporating engineered features like lag variables, moving averages, and seasonal indicators substantially improved forecast accuracy and interpretability. Different models performed best depending on the filter type: Linear Regression was most effective for Fuel and Hydraulic Filters, while Random Forest better captured the complexity of Special Design Filters. Feature selection via a take-out simulation highlighted the critical role of lag and moving average features, leading to streamlined and precise models. Future research could explore advanced deep learning techniques and real-time adaptive models to further enhance forecasting accuracy across diverse product categories.

References

Arunadevi, E. (2024). Enhancing Ameliorating Items with Sustainable Inventory Management Strategies: A Weibull Approach for Time and Price-Dependent Demand Trends. *Ain Shams Engineering Journal*, *15*. https://www.sciencedirect.com/science/article/pii/S2090447924004404?via%3Dihub

Chong, J. W., & Kim, W. (2022). Optimization of Apparel Supply Chain Using Deep Reinforcement Learning. *IEEE Access*, *10*. https://ieeexplore.ieee.org/document/9885193

Dai, B. T. (2024). An Integrated Concept for Robust Supply Chain and Manufacturing System. *IFAC PapersOnLine*, *58*. https://www.sciencedirect.com/science/article/pii/S2405896324015854?via%3Dihub

Dautel, J. L., & Thakur, J. (2024). Enabling Industrial Decarbonization: A MILP Optimization Model for Low-Carbon Hydrogen Supply Chains. *International Journal of Hydrogen Energy*, *77*.

https://www.sciencedirect.com/science/article/pii/S0360319924022493?via%3Dihub

Destro, I. R., & Staudt, F. H. (2023). The Impact of Inventory Record Inaccuracy and Cycle Counting on Distribution Center Performance. *Production*, *33*. https://www.scielo.br/j/prod/a/m5fWPJPb6sYv74PhTDJbnKw/?lang=en

D'Orazio, P. (2024). Assessing the Fiscal Implications of Changes in Critical Minerals' Demand in the Low-Carbon Energy Transition. *Applied Energy*, *376*. https://www.sciencedirect.com/science/article/pii/S0306261924015800?via%3Dihub

Feizabadi, J. (2022). Machine Learning Demand Forecasting and Supply Chain Performance. *International Journal of Logistics Research and Applications*, *25*. https://www.tandfonline.com/doi/full/10.1080/13675567.2020.1803246

Feleafel, H., & Radulovic, J. (2025). Facilitating the Shift to More Collaborative Microgrids by Alleviating Demand Volatility Using a Precontracted Order Updates Strategy. *Next Energy*, *8*. https://www.sciencedirect.com/science/article/pii/S2949821X25000328?via%3Dihub

Fida, A. U. (2024). A Comprehensive Survey on Load Forecasting Hybrid Models: Navigating the Futuristic Demand Response Patterns Through Experts and Intelligent Systems. *Results in Engineering*, *23*. https://www.sciencedirect.com/science/article/pii/S2590123024010284?via%3Dihub

Jebbor, I., & Benmamoun, Z. (2024). Forecasting Supply Chain Disruptions in the Textile Industry Using Machine Learning: A Case Study. *Ain Shams Engineering Journal*, *15*. https://www.sciencedirect.com/science/article/pii/S2090447924004970?via%3Dihub

Kao, C. (2002). Lot Size–Reorder Point Inventory Model with Fuzzy Demands. *Computers Mathematics with Applications*, *43*. https://www.sciencedirect.com/science/article/abs/pii/S0898122102001013

Kim, S. (2023). Innovating Knowledge and Information for a Firm-Level Automobile Demand Forecast System: A Machine Learning Perspective. *Journal of Innovation Knowledge*, *8*. https://www.sciencedirect.com/science/article/pii/S2444569X23000513?via%3Dihub

Kumar, L., & Khedlekar, S. (2024). A Comparative Assessment of Holt-Winter Exponential Smoothing and Autoregressive Integrated Moving Average for Inventory Optimization in Supply Chains. *Supply Chain Analytics*, *8*. https://www.sciencedirect.com/science/article/pii/S294986352400027X?via%3Dihub

Lee, C. K. M., & Lv, Y. (2017). Design and Application of Internet of Things-based Warehouse Management System for Smart Logistics. *International Journal of*

Production Research. https://www.tandfonline.com/10.1080/00207543.2017.1394592

Memon, J., & Sami, M. (2020). Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR). *IEEE Access*, *8*. https://ieeexplore.ieee.org/document/9151144

Mohammed, A. R., & Hassan, K. S. (2022). Moving Average Smoothing for Gregory-Newton Interpolation: A Novel Approach for Short-Term Demand Forecasting. *IFAC-PapersOnLine*, *55*. https://www.sciencedirect.com/science/article/pii/S2405896322017979?via%3Dihub

Morra, L., & Santangelo, A. (2024). For a Semiotic AI: Bridging Computer Vision and Visual Semiotics for Computational Observation of Large-Scale Facial Image Archives. *Computer Vision and Image Understanding*, *249*. https://www.sciencedirect.com/science/article/pii/S1077314224002686?via%3Dihub

Nambiar, A., & M., D. (2022). An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management. *Big Data and Cognitive Computing*, *6*. https://www.mdpi.com/2504-2289/6/4/132

Nguyen, A., & Lamouri, S. (2021). Managing Demand Volatility During Unplanned Events with Sentiment Analysis: A Case Study of the COVID-19 Pandemic. *IFAC-PapersOnLine*, *54*. https://www.sciencedirect.com/science/article/pii/S240589632100971X?via%3Dihub

Ni, D., & Xiao, Z. (2020). A Systematic Review of the Research Trends of Machine Learning in Supply Chain Management. *International Journal of Machine Learning and Cybernetics*. https://link.springer.com/article/10.1007/s13042-019-01050-0

Nurgazina, J., & Felberbauer, T. (2022). Visualization and Clustering for Rolling Forecast Quality Verification: A Case Study in the Automotive Industry. *Procedia Computer Science*, *200*. https://www.sciencedirect.com/science/article/pii/S1877050922003131?via%3Dihub

Perdana Wanti, L., & Maharrani, R. H. (2020). Optimation Economic Order Quantity Method for a Support System Reorder Point Stock. *International Journal of Electrical and Computer Engineering (IJECE)*, *10*. https://zenodo.org/records/4120322

Riachy, C., & He, M. (2025). Enhancing Deep Learning for Demand Forecasting to Address Large Data Gaps. *Expert Systems With Applications*, *268*. https://www.sciencedirect.com/science/article/pii/S0957417424030677?via%3Dihub

Rohit, K., & Verma, A. (2025). A Continuous Supply Chain Management Approach Using SPJS-Fuzzy DEMATEL and LPWBN for Automotive Electric Vehicles in

India. *Sustainable Futures*, *9*. https://www.sciencedirect.com/science/article/pii/S2666188825000887?via%3Dihub

Sakthi Balan, G., & Santhosh Kumar, V. (2025). Machine Learning and Artificial Intelligence Methods and Applications for Post-Crisis Supply Chain Resiliency and Recovery. *Supply Chain Analytics*, *10*. https://www.sciencedirect.com/science/article/pii/S2949863525000214?via%3Dihub

Saravanan, M., & R., G. (2024). Comprehensive Study on the Development of an Automatic Helmet Violator Detection System (AHVDS) Using Advanced Machine Learning Techniques. *Computers and Electrical Engineering*, *118*. https://www.sciencedirect.com/science/article/pii/S0045790624002179?via%3Dihub

Sel, B. (2025). Probabilistic Forecast-Based Procurement in Seaborne Forward Freight Markets Under Demand and Price Uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, *193*. https://www.sciencedirect.com/science/article/pii/S1366554524004216?via%3Dihub

Sokolova, I., & Nwaila, G. T. (2025). From Abundant Resource to Critical Commodity: Forecasting Manganese Supply and Assessing Its Sustainability. *Sustainable Materials and Technologies*, *44*. https://www.sciencedirect.com/science/article/pii/S2214993725001174?via%3Dihub

Wolny, M., & K., M. (2024). Unveiling Patterns in Forecasting Errors: A Case Study of 3PL Logistics in Pharmaceutical and Appliance Sectors. *Sustainability (MDPI)*, *17*. https://www.mdpi.com/2071-1050/17/1/214

Zedda, L., & Perniciano, A. (2024). Understanding Cheese Ripeness: An Artificial Intelligence-Based Approach for Hierarchical Classification. *Knowledge-Based Systems*, *295*. https://www.sciencedirect.com/science/article/pii/S0950705124004672?via%3Dihub