

Chapter 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This study investigated the application of machine learning models—Random Forest, LightGBM, and Linear Regression—to forecast incoming customer orders of heavy-duty vehicle filter products. The analysis covered the period 2021 to 2025 using historical sales data from the SKU level, with a focus on evaluating the effectiveness of feature engineering and selection in improving forecast accuracy.

The key conclusions drawn from this research are as follows:

- 1. Machine Learning Outperforms Statistical Baselines:** Compared to naïve and moving average models, machine learning models showed substantial improvements in predictive performance. Baseline models resulted in high MAPE (exceeding 20% for most item types) and negative or near zero R^2 scores, highlighting the limitations of traditional methods under volatile demand conditions.
- 2. Feature Engineering Enhances Accuracy and Interpretability:** The addition of engineered features, especially lag variables, moving averages (MA-1 to MA-6) and seasonal flags, significantly improved model performance. For example, the R^2 for Hydraulic Filters increased from negative values to as much as 0.60 using linear regression with engineered features.
- 3. Item-Type Specific Model Preferences Emerge:** Optimal models varied by item type:
 - For **FF (Fuel Filter)** and **HF (Hydraulic Filter)**, Linear Regression with feature engineering achieved the lowest MAPE and highest R^2 .
 - For **SX (Special Design Filter)**, Random Forest with feature engineering delivered the best performance, indicating its strength in handling complex patterns.
- 4. Feature Selection Yields Compact and Effective Models:** Through take-out simulation and R^2 /MAPE degradation tracking, it was found that lag and moving average features

were consistently the most important across all models. Removing less relevant features improved the generalizability and interpretability of the model without sacrificing accuracy.

5.2 Recommendations

Based on the findings, the following recommendations are proposed:

- **Operational Deployment:** The selected ML models should be integrated into a monthly forecasting routine within planning and procurement systems. Visualization dashboards can be used to communicate model outputs to relevant departments.
- **Expand to Inventory Integration:** While this study focused on forecasting incoming customer orders, future work should connect forecast outputs with inventory optimization modules such as reorder point and safety stock policies.
- **Incorporate Additional External Variables:** Although this study tested macroeconomic indicators (e.g., IHSG, USD), the results were mixed. Further exploration with more granular and leading external signals may enhance long-range forecasting capabilities.
- **Develop Multistep Forecasting Capabilities:** This research emphasized forecasts one month ahead. Future research could extend to multiperiod forecasting (e.g. 3- or 6-month horizons), which is valuable for strategic sourcing and capacity planning.
- **Adopt Feature Take-Out Simulation in Forecasting Workflows:** This technique proved useful in selecting high-contributing variables and should be embedded in future forecast pipelines for model refinement and validation.
- **Broader Product Application:** The modeling framework can be applied to other product families with similar demand volatility and high complexity of SKUs, ensuring adaptability in planning scenarios in different business units.

These recommendations provide a foundation for transitioning the forecasting framework from an analytical prototype to a deployable planning tool that supports more data-driven, agile, and resilient decision-making in manufacturing supply chains.

5.3 Implications

5.3.1 Academic Implications

- **Feature Engineering Impact:** Across all product families, feature engineering and selection reduced MAPE by 29% – 56% compared to baselines, with R^2 improvements ranging from +1.20 (HF) to +0.87 (FF). For example, Fuel Filters (FF) improved from 26.08% MAPE (baseline) to 16.32% MAPE (best model) and from $R^2 = -0.4240$ to 0.4387. These results provide robust empirical evidence that the incorporation of lags, moving averages, seasonal dummies, and macroeconomic indicators materially enhances the accuracy of forecasts in manufacturing environments lacking downstream consumption data.
- **Context-Dependent Model Allocation:** The research validates a model allocation logic that performs well for volatile demand profiles such as Special Design Filters (SX), improving MAPE from 14.80% to 9.71% and R^2 from -0.4544 to 0.2489. Linear regression is competitive or superior for stable seasonal categories such as Oil Filters (OF), improving MAPE from 10.16% to 7.67% with a high R^2 of 0.5262.

5.3.2 Methodological Implications

- **Replicable Constrained-Data Framework:** The Baseline $\rightarrow$ Feature Engineering $\rightarrow$ Feature Take-Out approach demonstrates that even with limited data, feature refinement can yield substantial gains. In Hydraulic Filters (HF), MAPE dropped from 24.85% to 10.82% (-56.5% error reduction), allowing linear regression to rival Random Forest accuracy within 0.5% MAPE for certain features.
- **Quantified Feature Importance:** The “Feature Take-Out” simulation revealed the most impactful features per category: SX: Removing lag increased MAPE from 10.44% to 12.15% (+1.71%) and decreased R^2 by -0.0447. FF: Removing ma increased MAPE from 16.36% to 20.8% (+4.02%) and decreased R^2 by -0.6981. HF: Removing is-seasonal-fiscal-year increased MAPE from 10.82% to 13.15% (+2.33%) and reduced R^2 by -0.5807 .

5.3.3 Managerial / Practical Implications

- **Operational KPI Improvements:** the best-performing models could reduce forecast error by up to 56.5%, which, according to the organization, can set the KPI target to: 20 - 25% reduction in excess inventory 10 - 20% fewer stock runs 5 - 10% improvement in on-time delivery rates. For example, improving the HF forecast from 24.85% to 10.82% MAPE could materially reduce the safety stock levels without sacrificing service quality.
- **ERP/MRP Integration Readiness:** Model outputs are in a time-series compatible structure, allowing direct integration into ERP/MRP demand planning modules to automate forward-looking Material Requirements Planning (MRP) runs, replacing reactive scheduling with proactive replenishment.

5.3.4 Implications of Feature Importance for Business Understanding

The analysis of feature importance provides critical business insights that go beyond technical modeling performance. By identifying which variables most strongly influence forecast accuracy, feature importance transforms machine learning output into actionable guidance for managers and decision makers. The implications can be summarized as follows.

1. **Turning technical insights into business levers:** The importance of the feature highlights the factors driving demand variability (e.g. sales history, seasonality, customer order cycles, macroeconomic conditions). This transparency allows forecasting to be placed as a decision support tool aligned with business realities.
2. **Prioritizing data quality and investments:** By classifying variables, the company can prioritize improving the accuracy, timeliness, and granularity of the most influential data streams. This ensures that IT and operations focus their investments where they will have the greatest impact.
3. **Aligning forecasting with operational strategy:** High-ranking features point to operational levers. For example, if seasonality is important, production and procurement must adapt to cyclical patterns. If customer order cycles dominate, customer relationship management (CRM) processes need to be strengthened. If macroeconomic factors are influential, external indicators should be integrated into Sales and Operations Planning (S&OP).

4. **Supporting change management and cross-functional buy-in:** Feature importance provides a clear rationale for departmental alignment. If forecast accuracy is heavily dependent on timely sales data, the sales team can see the direct impact of their reporting discipline. If supplier lead times emerge as influential, procurement gains a business case to stabilize supply performance.
5. **Risk mitigation and scenario planning:** By identifying sensitive drivers, feature importance highlights where risks are concentrated. If external variables such as oil prices or exchange rates are influential, managers can build contingency plans to reduce exposure to volatility in these areas.
6. **Guiding strategic recommendations:** Finally, feature importance provides an analytical justification for managerial recommendations. For example, if historical sales patterns are the strongest predictor, the company should invest in strengthening the data integration pipelines between the ERP and forecasting systems. If seasonality has a moderate influence, production planning must explicitly incorporate cyclical demand factors in capacity allocation.

In summary, feature importance serves as a bridge between technical forecasting models and business decision-making. It ensures that recommendations for process improvements, IT investments, and operational strategies are based on evidence, thus increasing their credibility and likelihood of adoption by management.

5.4 Roadmap for Implementation and Managerial Implications

Based on the findings and the advisors' input, a step-by-step roadmap is proposed to translate technical results into actionable business levers and to ensure scalability for other factories.

5.4.1 Implementation Roadmap

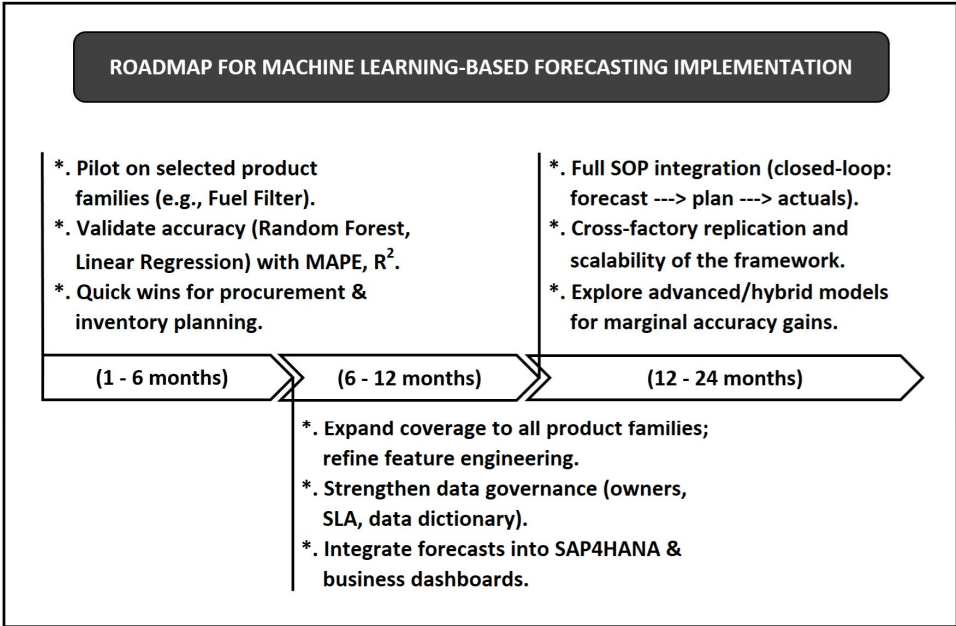


Figure 5.1: Roadmap Machine Learning-Based Forecasting Implementation.

In Figure 5.1, provides a structured roadmap for the staged implementation of machine learning-based forecasting within the company. The table divides the plan into three distinct phases, short-term, medium-term, and long-term, to make the adoption process realistic and actionable. In the **short-term** (1–6 months), the focus is on piloting forecasting for selected product families and demonstrating immediate business value through quick wins in procurement and inventory. The **medium-term** (6–12 months) emphasizes scaling to all product families, improving data governance practices, and integrating forecast results into the enterprise system (SAP S/4HANA) and dashboards. Finally, the **long-term** (12–24 months) aims for full integration into Sales and Operations Planning (S&OP), replication of the framework across factories and exploration of advanced or hybrid models for further accuracy improvements.

This roadmap provides a clear view of **which actions should be prioritized, when they should be executed, and how each stage contributes to the value of the business**. Not only does it connect the technical results of this study with managerial decision making, but also it demonstrates how the proposed forecasting framework can evolve into a scalable and sustainable solution for the organization.

Table 5.1: Examples of SLA Targets for Data Governance

Dimension	Example SLA Target	Relevance for Forecasting
Timeliness	Sales order data must be updated in ERP by the 3rd business day of each month.	Ensures ML models are trained on the most recent and complete data.
Completeness	99% of sales records must include product code, customer code, and quantity.	Prevents bias or gaps in training data that reduce forecast accuracy.
Accuracy	Data entry errors (e.g., incorrect item codes, duplicate orders) must be corrected within 24 hours.	Maintains data reliability and avoids misleading patterns in forecasting.
Consistency	Definitions of variables (e.g., fiscal year, sales_qty) must follow the data dictionary.	Aligns understanding across departments and improves reproducibility.

Table 5.1 illustrates practical examples of Service Level Agreements (SLAs) that can be applied to strengthen data governance within the organization. Each dimension, timeliness, completeness, accuracy, and consistency is translated into measurable targets to ensure that the input data used for machine learning-based forecasting are reliable and fit for purpose.

For example, timeliness guarantees that forecasting models are always trained on the most recent data, while completeness ensures that critical attributes such as product and customer codes are fully recorded. Accuracy prevents errors or duplications from distorting demand patterns, and consistency enforces common definitions between departments through the use of a data dictionary.

By presenting SLAs in this way, the reader gets a clear understanding of how data governance practices directly support the quality of forecasting outcomes. The table highlights that without disciplined management of these dimensions, the effectiveness of advanced forecasting models would be undermined by poor data quality.

5.4.2 Feature Insights to Business Actions

Table 5.2: Translating Feature Importance into Business Actions

Feature Category	Operational Lever	Business Value
Seasonality	Production and procurement adjustments	Align production with cyclical demand patterns
Customer Order Cycles	Strengthening CRM processes	Improved responsiveness to customer demand
Macroeconomic Indicators	Integrating external signals into S&OP	Reduced exposure to volatility in prices, currency, and indices
Historical Sales Patterns	Data-driven planning and pipeline integration	Higher accuracy in forecasting and planning reliability

Table 5.2 shows how the results of the analysis of feature importance can be translated into concrete managerial actions. Each feature category is connected with an operational lever that managers can influence, and then further linked to the corresponding business value. This structure bridges the gap between technical insights from machine learning models and practical decision making within the organization.

For example, identifying **seasonality** as a key driver can inform production and procurement adjustments, ensuring that manufacturing aligns with cyclical demand patterns. Insights into **customer order cycles** highlight the need to strengthen CRM processes, enabling the company to respond more effectively to client requirements. Likewise, the role of **macroeconomic indicators** emphasizes the importance of integrating external signals into Sales and Operations Planning (S&OP) to mitigate risks from market fluctuations. Finally, recognizing the importance of historical sales patterns supports the move toward more data-driven planning, increasing the accuracy and reliability of forecasts.

Through this mapping, the reader can understand how technical findings on feature importance can be operationalized into actions that generate tangible business benefits. It reinforces the practical relevance of the forecasting framework beyond its statistical performance, showing its value as a decision support tool for management.

5.4.3 Managerial Implications

Six managerial implications are outlined below.

1. **Turning technical insights into business levers:** Forecasting output becomes a decision support tool aligned with business realities.
2. **Prioritizing data quality and investments:** Focus resources on the most influential data streams to improve timeliness and accuracy.
3. **Aligning forecasting with operational strategy:** Ensure that the production, procurement, and CRM processes adapt to key drivers (seasonality, order cycles).
4. **Supporting change management and cross-functional buy-in:** Feature importance justifies departmental alignment and accountability.
5. **Risk mitigation and scenario planning:** Identify sensitive drivers (e.g., coal prices, oil prices) for contingency strategies.
6. **Guiding strategic recommendations:** Provide analytic justification for managerial actions, such as the integration of ERP forecasting.

5.4.4 Consulting and Scalability Opportunities

The proposed roadmap is **replicable and scalable** in other manufacturing contexts. By demonstrating the applicability of machine learning, this research offers a practical foundation for consulting engagements with factories facing similar demand forecasting challenges. Such engagements would not only improve forecast accuracy but also create tangible business value through better inventory planning, operational stability, and long-term competitiveness.