Optimizing Demand Forecasting in Supply Chain Management With Artificial Intelligence

A Technical Report
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

By
Ronit Reddy

Spring 2025

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Prof. Briana Morrison, Department of Computer Science
Prof. Rosanne Vrugtman, Department of Computer Science

Optimizing Demand Forecasting in Supply Chain Management With Artificial Intelligence

CS4991 Capstone Report, 2025

Ronit Reddy
Computer Science
The University of Virginia
School of Engineering and Applied Science
Charlottesville, Virginia USA
kus4ws@virginia.edu

ABSTRACT

o9 Solutions, a Dallas, Texas-based business planning company, found its proprietary demand forecasting workflow to be unreliable and inefficient, often necessitating manual intervention from demand planners. To streamline the forecasting system, I integrated machine learning (ML) algorithms with the input checks framework to ensure that, in the clients' historical product sales data, recurring spikes/dips were preserved and reflected in subsequent forecasts, and random spikes/dips were programmatically corrected and filtered. Through this work, I was not only able to catalyze general accuracy improvements in the company's forecasting models, but also observe the real-world applicability of artificial intelligence (AI) in supply chain analysis. Moving forward, greater efforts could be made in stratifying these methodologies, such as refining the algorithm selection for distinct product categories and sectors.

1. INTRODUCTION

In today's rapidly evolving global marketplace, accurate demand forecasting is more than just a competitive advantage for supply chain management companies. It is, rather, a critical necessity, allowing them to efficiently balance inventory, optimize production, and reduce operational costs. Despite the numerous contemporary advances in big data analytics, many organizations continue to rely on and struggle with unreliable forecasting systems.

During my summer 2023 internship at o9 Solutions, a leading American supply chain planning solutions company, I witnessed firsthand how inefficiencies in demand forecasting, particularly in anticipating demand trends and their catalysts can disrupt decision-making capabilities and operational workflows. Accordingly, I sought to support a more automated and intelligent approach to demand forecasting—one that could identify meaningful patterns and random fluctuations in sales data and act appropriately. To achieve this goal of heightened forecasting reasonability, I worked on employing ML algorithms that could enhance the input checks process and its linkage to the history correction module.

2. RELATED WORKS

While my work at o9 Solutions was largely informed by the company's platform and codebase, given the confidential nature of client relationships and the associated data, recent academic literature and research efforts detail several alternative techniques to address similar challenges in time series forecasting. Notably, Dantas and Oliveira (2018) presented their devised Bagged.Cluster.ETS approach, which combines various ML methods, including bagging, exponential smoothing, and clustering to generate ensembles of forecasts and reduce the aggregate covariance effect. Although this technique relies on making corrections post-forecast generation, whereas my work corrected the foundational logistics behind forecasting models before demand projections were made, it still reflects the importance of mitigating the effects of random anomalies in any data refinement procedures.

More recently, Zhang, et. al. (2024) proposed a method to enhance demand forecasting in their Hybrid Attention-based Long Short-Term Memory (HA-LSTM) network, which combines multi-head self-attention modules with LSTM layers to capture both local and global temporal dependencies in product demand data, effectively handling complex time series instances with multiple seasonalities, while also providing interpretable attention weights. In this, their work designates a clear emphasis on distinguishing between meaningful patterns and arbitrary noise in demand forecasting, emulating the principle of interpretability in accurately capturing recurring patterns and seasonal effects that were central to my work.

Although these studies did not directly influence the work I conducted at o9 Solutions, they embody the same foundational pillars of effective demand forecasting I explored and upheld during my internship. As such, irrespective of the different technological foundations in these approaches and mine, it is no surprise they all ultimately arrived at the same endpoint: enhancing demand forecasting through intelligent pattern preservation and data refinement.

3. PROJECT DESIGN

This section provides an overview of the design approach to the enhanced, AI-integrated forecasting workflow.

3.1 Review of System Architecture

The original demand forecasting architecture at o9 Solutions followed a two-step flow designed to clean and prepare historical sales data before generating final forecasts. First, actuals—the raw product sales data—were ingested and passed to the outlier correction plugin. From there, the outlier correction plugin attempted to identify and adjust anomalous spikes/dips, producing a set of corrected actuals that could then be fed into the stat forecasting module to remove any observable noise in the data and guarantee the final forecasting model(s) could focus on reflecting logical patterns and trends.

3.2 Requirements

This subsection reviews the key client needs and system limitations shaping the project solution.

3.2.1 Client Needs

o9 Solutions supports a vast, globally distributed partner network that spans system integrators and consultants, cloud providers, and technology innovators, all of whom require accurate and robust demand forecasts. In response, the core objective of this project was to identify the top problem areas within the company's existing forecasting workflow, and, armed with these insights, facilitate forecast reasonability improvements for different clients across different fields, thereby reinforcing its central mission to deliver transformative planning solutions that accelerate customer success and growth.

3.2.2 System Limitations

Despite the well-outlined functionality behind the original demand forecasting workflow, several shortcomings manifested to impede accurate demand forecasts. With regard to known demand drivers, the outlier correction plugin often failed to correct the effects of volatile events when required, namely in the case of stockouts, and mistakenly corrected the effects of predictable events when such needed to be retained, as in the case of holidays. More importantly, however, the original workflow generally failed to even identify recurring spikes/dips, such as those spurred by product launches and promotional events, and at the same time, rarely filtered out random spikes/dips. Because of these oversights, this workflow resulted in valuable information about repeated behaviors either being lost or treated as anomalies, as well as sporadic surges and drops in the data polluting forecasts' integrity.

3.3 Key Components

This subsection outlines the primary modules and functionalities present in the new system.

3.3.1 Specifications

For the Missing Drivers Identification task, which focused on capturing repeating events not fully recognized by the current system—including certain recurring holidays and geographic seasonal patterns—the specifications required a thorough classification methodology that could detect recurring spikes/dips within the historical product sales data and correspondingly link them to the respective missing demand drivers at the item/location/customer levels. This enhancement ultimately would help demand planners to incorporate critical demand signals into the forecasting workflow and in a greater sense, preserve the overall data maturity.

In parallel, the Random Anomalies Correction task, which aimed to handle sporadic and one-off anomalies in the historical data not tied to any known events, necessitated an explicit identification mechanism that could determine the exact random anomalies already processed by the outlier correction plugin. Then, any remaining anomalies could be programmatically passed through the more comprehensive and powerful history correction module, enabling demand planners to ensure the final dataset better represents true demand. These guidelines, when amalgamated, establish an optimal strategy to minimize demand planners' manual labor and maximize forecasting accuracy.

3.3.2 Challenges

One significant obstacle with the Missing Drivers Identification task was determining whether anomalies in the historical sales data truly reflected recurring demand patterns year-over-year. For instance, many regional and seasonal events do not recur at the same time each year within the given timeframe, nor do they always appear every year within the historical window; yet, the absence of a perfect annual repetition for a particular anomaly does not preclude the presence of a linked demand driver. This challenge demanded a more flexible detection logic in capturing patterns that recur approximately (rather than precisely) on a cyclical basis, without dismissing them as random noise.

On the other hand, the primary hurdle in the Random Anomalies Correction task was tracking processed anomalies in a manner that was both efficient and integrated with the existing architecture. Because the plugin lacked a clear mechanism for tracking processed anomalies, the devised workflow needed to not only correctly identify which spikes/dips had already been processed and which remained to be addressed, but also route any uncorrected anomalies through the history correction module, all while ensuring corrected entries were not redundantly flagged or modified. Consequently, the ensuing solution would require close coordination with the plugin's backend structure and the broader input checks system.

3.3.3 Solutions

As shown in Figure 1, the original input checks workflow treated all unknown anomalies as generic spikes/dips without differentiation.

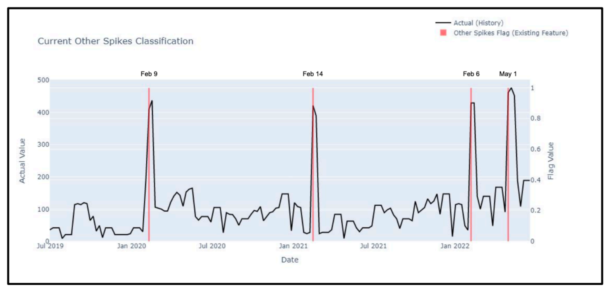

Figure 1: Anomaly Classification Outlook Prior to Missing Drivers Identification Task

To mitigate the numerous challenges of this shortcoming and, instead, distinguish recurring anomalies from truly random events with the Missing Drivers Identification task, I implemented a classification-based pipeline in Python. First, I leveraged Python's date-time libraries to create time-based features—such as day-of-year offsets and flexible annual windows—to detect whether an anomaly recurred around the same period in consecutive years. Building on this, I introduced an underlying ML model to classify anomalies as either "repeating" or "random." Logistically, the model calculates how frequently a given anomaly appears within the historical timeframe and weighs this frequency against a configurable threshold. If an anomaly's recurrence rate exceeds the threshold, the pipeline flags it as "repeating" and correlates it with a relevant demand driver. Conversely, if the recurrence rate falls below the threshold, the anomaly is deemed random. This refined approach is depicted in Figure 2**,** where anomalies now fall into distinct categories, allowing demand planners to preserve critical recurring patterns while systematically filtering out genuine noise.

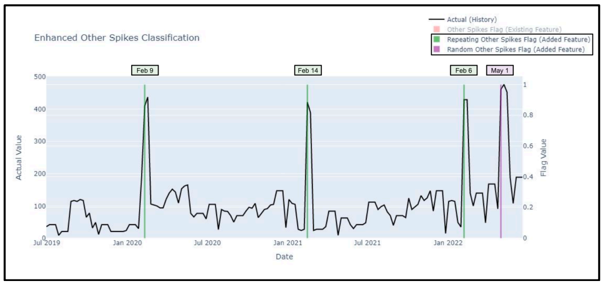

Figure 2: Anomaly Classification Outlook After Missing Drivers Identification Task

Nonetheless, further work was required. As illustrated in Figure 3, although random spikes/dips were now being fully identified, many of these sporadic, one-off anomalies were still escaping the outlier correction plugin, resulting in significant noise that undermined respective forecasts.

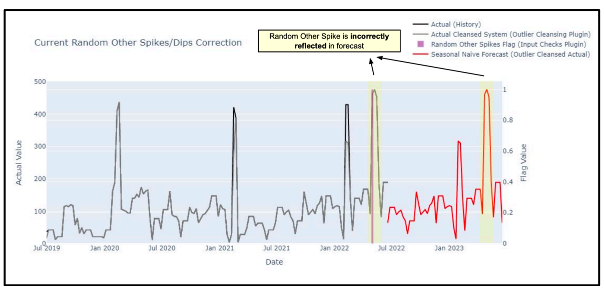

Figure 3: Demand Forecasting Outlook Prior to Random Anomalies Correction Task

To address this issue with the Random Anomaly Correction task, I configured ML algorithms that could quickly compare raw and corrected demand datasets and determine, in any case, which anomalies were addressed by the outlier correction plugin. From there, for these particular events, I introduced a specialized "Corrected Random Anomalies" flag to serve as an indicator the anomaly has been corrected, and passed these flags, along with the "Random Anomalies" flags, to the enhanced history correction module. This module would systematically correct any remaining random anomalies, ensuring the final demand forecast is free from distortions caused by isolated data irregularities. The impact of this adjustment is evident in Figure 4, which shows a substantially smoother demand forecasting outlook.

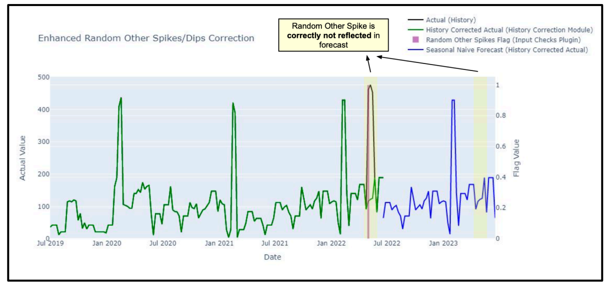

Figure 4: Demand Forecasting Outlook After Random Anomalies Correction Task

4. RESULTS

Overall, the enhanced forecasting workflow produced by these tasks yielded models that more accurately capture and reflect historical demand patterns. By integrating ML techniques into the input checks and linking them with the history correction module, the refined system can now preserve critical recurring demand drivers while filtering out random noise. As a result, the final forecasting models demonstrate a closer alignment with true demand trends, enabling more reliable and actionable predictions. In many test cases, these improvements translated to an increase in forecasting accuracy by upwards of 5%, underscoring the practical benefits of the solution in reducing manual interventions and enhancing overall forecast reliability.

Furthermore, the system now has the capabilities to provide demand planners and clients with granular analytics at key intersections of data processing. Through an intuitive tabular dashboard view, individuals can now filter and analyze areas where classifications or corrections have been applied, and access detailed metrics such as the number of corrected anomalies, total adjustment values, and percentage adjustments. This metadata can not only help anticipate systematic patterns and outliers, but also pinpoint specific data-driven refinement areas to the forecasting process, which undoubtedly improves inventory planning and supply chain responsiveness.

5. CONCLUSION

This project demonstrates the vast potential of AI in supply chain management contexts to revolutionize anomaly detection and data refinement processes. To this end, as it pertains to o9 Solutions, the company's improved forecasting workflow both reduces the manual intervention required from its demand planners and boosts the accuracy across its various forecasting models, ultimately enabling it to anticipate market fluctuations with greater efficiency and confidence. Beyond these immediate, tangible benefits, this work showcases how context-tailored AI modeling can shape more resilient supply chain networks, as data becomes an engine for strategic foresight and streamlined operations. On a personal level, this endeavor has sharpened my algorithmic development skills and highlighted the profound impact that ML algorithms can have in tackling complex, real-world challenges.

6. FUTURE WORK

While the updated anomaly classification and correction mechanisms substantially enhance o9 Solutions' procedures, several avenues remain for further refinement and expansion. First, incorporating more sophisticated deep learning architectures, including transformers and hybrid attention-based models, into the system software based on alignment with distinct client scenarios would enable the system to better capture sector-specific trends, ensuring forecasts remain relevant to each client's unique operational landscape. Beyond this front, adapting the system to consider external demand-shaping variables, such as weather information, could contextualize subtler patterns within large-scale demand data. Finally, implementing real-time data ingestion and continuous model retraining features, coupled with structured feedback from demand planners could guide iterative improvements in both usability and system performance. Taken together, these efforts promise greater forecasting precision, and, more importantly, deeper insights into the multifaceted nature of supply chain dynamics across diverse industries.

REFERENCES

Dantas, T. M., & Cyrino Oliveira, F. L. (2018). Improving time series forecasting: An approach combining bootstrap aggregation, clusters and exponential smoothing. *International Journal of Forecasting, 34*(4), 748–761. https://doi.org/10.1016/j.ijforecast.2018.05.006

Zhang, X., Li, P., Han, X., Yang, Y., & Cui, Y. (2024). Enhancing time series product demand forecasting with hybrid attention-based deep learning models. *IEEE Access, 12*, 190079–190091. https://doi.org/10.1109/access.2024.3516697