Demand Forecasting Using Different Methods in Python

Dhanush K*

Dr. Dr. CH Renu Madhavi **

*UG student, 4th year, Department of EIE, RV College of Engineering
**Assistant prof, Department of EIE, RV College of Engineering

Abstract: This project explores the application of advanced methods for demand forecasting in supply chain management. With the increasing complexity and volatility of global markets, accurate demand forecasting is crucial for optimizing inventory management, production planning, and overall supply chain efficiency. Traditional forecasting techniques often struggle to capture the nonlinear and dynamic nature of demand patterns, especially in highly unpredictable environments. In response, this project investigates the effectiveness of various deep learning algorithms, including SARIMA, Prophet, and LSTM, in predicting future demand with high accuracy and reliability. Through detailed experimentation and analysis, the project aims to identify the most suitable deep learning approach for different scenarios and industries. The outcomes of this research have the potential to revolutionize supply chain dynamics by providing decision-makers with actionable insights for proactive inventory management and improved operational efficiency. Additionally, the project successfully reduces forecasting error rates from around 30% to 18.64%, showcasing the significant impact of SARIMA techniques on enhancing forecasting accuracy.

Keywords: SARIMA, Prophet, LSTM

Introduction

Supply chain management (SCM) is a rapidly growing and well-studied field of study that is increasing in popularity and importance (Meherishi et al., 2019). According to Mentzer et al. (2001), the supply chain is a chain of products connected by the flow of goods, information and/or services. Most organizations, especially in new management of the supply chain, focus on optimizing price and managing the best products to satisfy customers. Accurate demand forecasting allows businesses to predict demand and manage inventory accurately. It enables machines to learn from past data, experiences and patterns and make accurate predictions. Machine learning means providing insight into future behavior from past data. Based on the nature of learning, ML methods are divided into three broad categories, including supervised learning, unsupervised learning, and incremental learning (RL). In supervised learning, many input data and desired results are assigned to the learning algorithm. In contrast, unmoderated studies use only input data. Unsupervised algorithms use raw data to find hidden patterns and get good results. Reinforcement learning (RL) is another category of machine learning. Reinforcement learning involves using a positive environment and trial and error techniques to achieve human performance. Besides triple classification, there is another classification called semi-supervised learning. In these algorithms, a small amount of labeled data and a large amount of unlabeled data are often used together. Learn from unstructured, interconnected datasets. This algorithmic system has many levels (deep layers) to support learning. Deep architecture may or may not be audited. The inspiration of the entertainment industry now provides effective solutions to many problems in life, such as image and video processing, speech recognition, text analysis, natural language processing and many types of classification. Deep learning techniques are new and effective techniques to obtain accurate predictions in SCM. But different deep learning methods do different things on different problems, and some methods are better than others.

In this study, our primary objective is to forecast the unit sales of numerous items sold across various chain stores in Ecuador, aiming to mitigate overstocking, minimize understocking, reduce waste, and enhance customer satisfaction. Accurate predictions are crucial for optimizing efficiency and determining product prices to meet customer demands effectively. To achieve this goal, we utilize the Corporación Favorita Grocery Sales Forecasting dataset obtained from the Kaggle website for precise sales forecasting. Employing three distinct deep learning methodologies—namely, SARIMA, Prophet, and LSTM—we construct and train predictive models. These models leverage different parameters and weights to forecast unit sales accurately. Throughout this endeavor, we utilize open-source data science tools, primarily Python and associated packages. Additionally, Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) serve as key metrics for evaluating and comparing the performance of the models. Our findings reveal that LSTM exhibits superior performance compared to SARIMA and Prophet in forecasting sales units, underscoring its effectiveness in this context.

Literature Review:

In the dynamic landscape of demand forecasting, the inadequacies of conventional statistical methods are starkly apparent when faced with the complexities of modern supply chains. These methods grapple with extensive, volatile datasets and often stumble in accommodating non-linear relationships, leading to inaccuracies and inefficiencies in forecasting. However, recent scholarly research has unveiled a plethora of AI-driven strategies poised to address these shortcomings. The spotlight shines particularly bright on sophisticated machine learning models, including deep learning and hybrid methodologies, heralding a seismic shift in forecast accuracy, dataset management, and adaptability to ever-evolving market conditions. These insights serve as the cornerstone for comprehending the seamless integration of AI into supply chain systems, promising palpable enhancements to demand planning processes. This comprehensive review section will artfully juxtapose specific AI advancements with the glaring deficiencies observed in traditional forecasting systems, thus illustrating the profound practical implications of embracing these cutting-edge technologies in real-world contexts.

One standout advancement in AI-driven forecasting introduces a multi-layer LSTM network meticulously tailored to predict erratic demand patterns. Through the adept utilization of a grid search approach to fine-tune LSTM hyperparameters, this model intricately captures non-linear relationships within time series data. [1] Results unequivocally underscore the LSTM model's supremacy over traditional methods like ARIMA and ANN, particularly in scenarios characterized by non-stationary data.

Delving deeper into innovation, another pioneering avenue emerges with the fusion of LSTM networks and Random Forests, giving birth to a groundbreaking hybrid forecasting model. This amalgamation masterfully combines LSTM's inherent proficiency in handling temporal data with Random Forests' robust regression capabilities. [2] Evaluated on a multivariate dataset sourced from a multi-channel retailer, the hybrid model not only achieves elevated accuracy but also offers invaluable insights by prioritizing explanatory variables, thereby enriching the landscape of demand planning.

Furthermore, the domain of forecasting performance metrics evaluation has undergone significant advancements. Notably, the reliability of sRMSE in detecting variance changes in a model, coupled with the high sensitivity of sPIS to bias, has come to the forefront. [3] Such nuanced insights serve as guiding beacons for selecting apt metrics crucial for the precise assessment of diverse forecasting models.

In the dynamic arena of wind power generation, a novel metric known as the error dispersion factor (EDF) steps into the limelight to juxtapose nRMSE and nMAE. [4] Understanding the intricate dynamics of error and its ripple effect on model performance across diverse forecasting scenarios underscores the pivotal role of robust error analysis in refining forecasting methodologies.

Moreover, sector-specific applications underscore the remarkable adaptability of AI models to bespoke forecasting requisites. For instance, in the realm of oil production forecasting, a fusion of ARIMA, LSTM, and Prophet models gains significant traction. [5] LSTM's inherent proficiency in handling non-seasonal data emerges as a game-changer, surpassing the capabilities of the Prophet model and showcasing unparalleled versatility.

Environmental monitoring presents yet another compelling case for AI prowess, with the Prophet forecasting model showcasing remarkable efficacy in predicting air pollution levels. Its adeptness in navigating missing data, capturing trends, and accommodating seasonality renders it tailor-made for the task, delivering robust performance in both short-term and long-term pollutant level forecasts. [6]

In the realm of manufacturing, Support Vector Regression (SVR) stakes its claim in predicting bending angles in laser tube bending processes, boasting unprecedented accuracy and reliability. [7] This herald a new era of machine learning's potential in optimizing manufacturing settings.

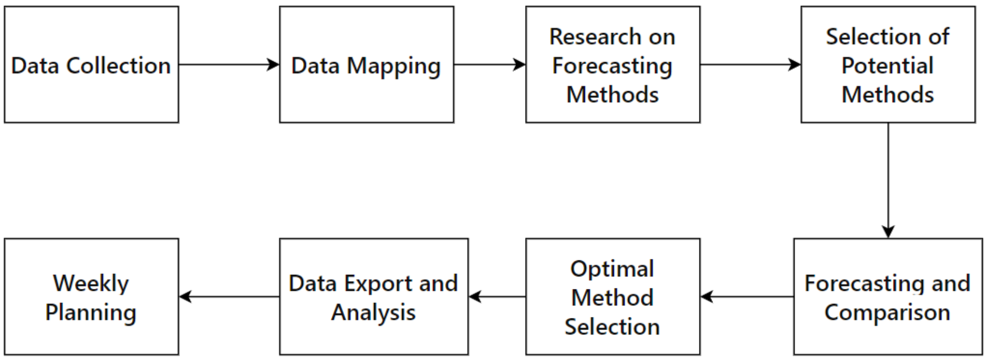
In the domain of retail sales forecasting, a meticulously conducted comparative study between LSTM and LGBM models reveals the consistent outperformance of LGBM, particularly in capturing and predicting intricate sales patterns. [8] These findings are not just insights but rather pivotal guideposts for retailers in fine-tuning inventory and sales strategies to stay ahead of the curve.

A study focusing on solar radiation forecasting underscores the indispensable utility of ARIMA models across varied geographical locations, emphasizing the importance of tailoring models to accommodate solar radiation variability, thereby offering immense promise in renewable energy forecasting.[9]

Lastly, an evaluation of LSTM and Prophet models in forecasting air temperature in Bandung unveils the Prophet model's prowess in predicting minimum temperatures, juxtaposed with LSTM's superior performance in forecasting maximum temperatures. [10] This nuanced dichotomy illuminates each model's unique strengths, offering invaluable insights into the realm of meteorological forecasting.

This holistic and expansive examination of AI applications spanning diverse sectors not only showcases the robustness of AI models but also underscores their unequivocal supremacy over traditional forecasting methods. These monumental strides pave the way for the emergence of more specialized and efficient AI applications in demand planning and supply chain management, promising transformative impacts on operational efficacy and strategic decision-making that reverberate throughout the global economic landscape.

METHODOLOGY



1.Data Collection: Gather comprehensive data including item IDs, locations, customer information, quantities, and available stock to ensure a robust foundation for analysis.

2.Data Mapping: Integrate and map data from lower-level granular details to higher-level summaries, facilitating a coherent and structured dataset for forecasting.

3.Research on Forecasting Methods: Investigate various forecasting techniques to identify those best suited to the specific needs and characteristics of the supply chain.

4.Selection of Potential Methods: Evaluate and shortlist forecasting methods that demonstrate high accuracy and reliability based on preliminary assessments.

5.Forecasting and Comparison: Apply the shortlisted forecasting methods to the collected data and compare their performance using metrics such as Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and overall accuracy.

6.Optimal Method Selection: Select the forecasting method that provides the highest accuracy and best performance based on the comparison metrics.

7.Data Export and Analysis: Export the forecasted data for further analysis and visualization, ensuring it is ready for integration into inventory planning processes.

8.Weekly Planning: Compute the weekly arrival quantities based on demand forecasts and safety stock levels, ensuring timely and adequate inventory replenishment.

SARIMA (Seasonal Autoregressive Integrated Moving Average) is a statistical method widely used for time series forecasting. It extends the ARIMA model to account for seasonality in the data, making it suitable for capturing complex seasonal patterns. SARIMA models involve parameters for seasonality, trend, and noise, and they rely on historical data to make predictions.

$$\text{ARIMA} \quad \underbrace{(p,d,q)}_{\substack{\uparrow \\ \text{Non-seasonal part} \\ \text{of the model}}} \quad \underbrace{(P,D,Q)_m}_{\substack{\uparrow \\ \text{Seasonal part of} \\ \text{of the model}}}$$

Prophet is a forecasting tool developed by Facebook that is specifically designed for time series data with strong seasonal patterns. It incorporates several components, including trend, seasonality, holiday effects, and outliers, to provide robust forecasts. Prophet's flexibility and ease of use make it popular for forecasting tasks across various domains.

y(t) = g(t) + s(t) + h(t) + e(t)

here,

g(t) refers to trend (changes over a long period of time)

s(t) refers to seasonality (periodic or short-term changes)

h(t) refers to effects of holidays to the forecast.

e(t) refers to the unconditional changes that is specific to a business or a person or a circumstance. It is also called the error term.

y(t) is the forecast.

LSTM (Long Short-Term Memory) is a type of recurrent neural network (RNN) architecture that is particularly effective for sequence prediction tasks, including time series forecasting. LSTM networks excel at capturing long-term dependencies in data and are capable of learning complex patterns over time. They are characterized by their ability to retain information over extended periods, making them well-suited for forecasting tasks requiring memory of past events.

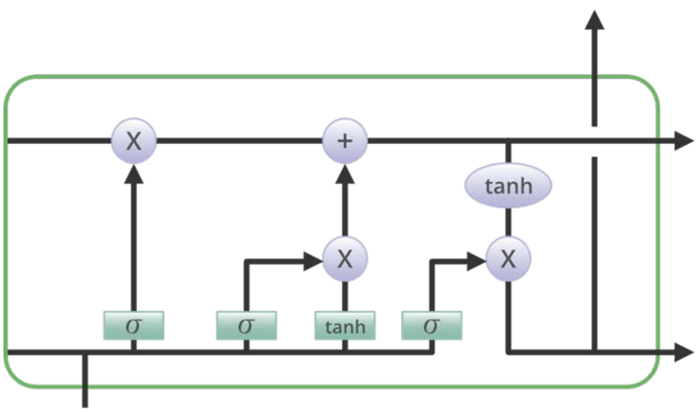


Fig 1. architecture of LSTM

RESULTS

5.1 Existing System performance

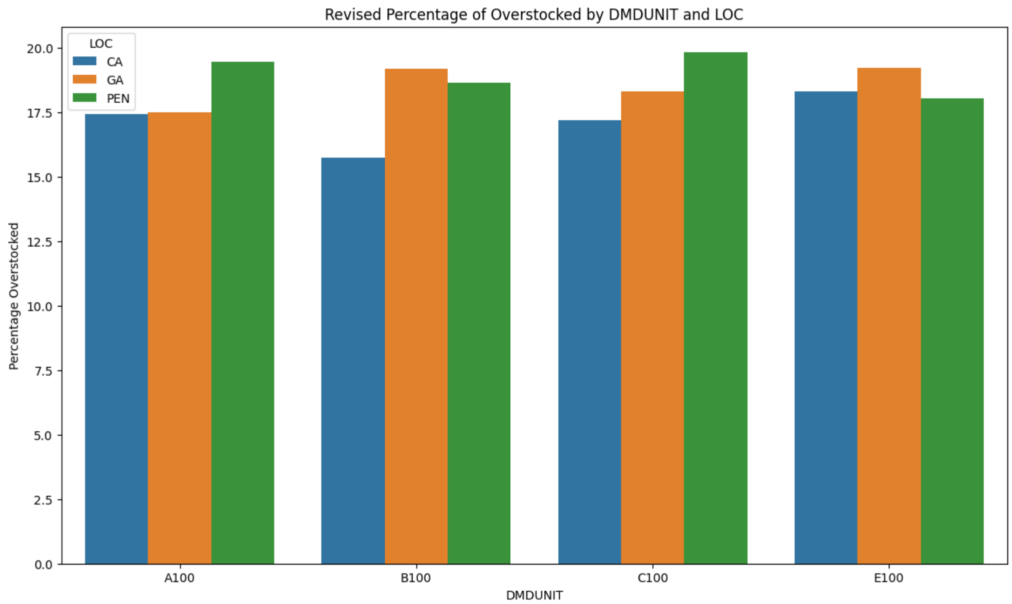


Fig 2: Percentage of Overstocking unit vs loc

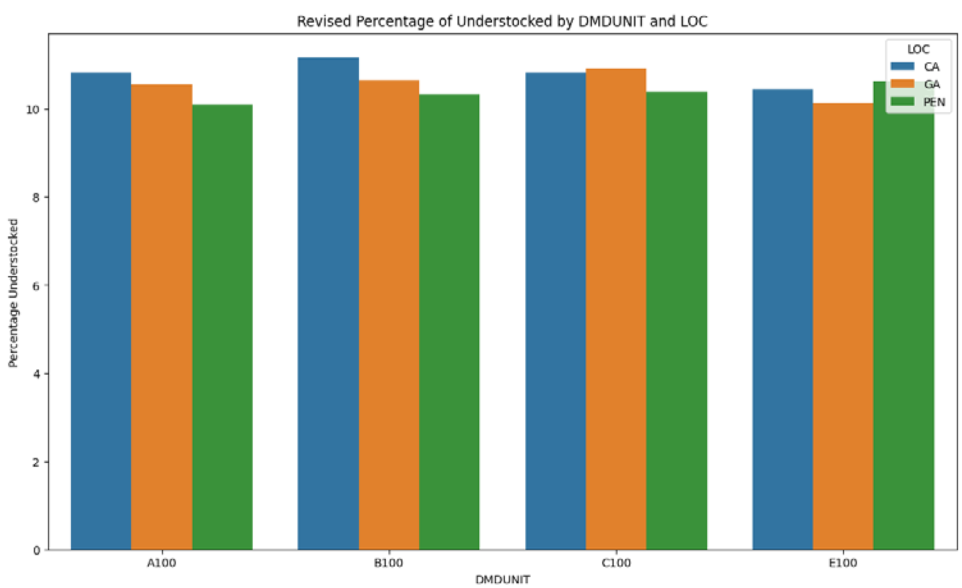


Fig 3 Percentage of understocking unit vs loc

Conversely, understocking has also been a concern, with inventory falling short of customer demand by as much as 20% and under stock of around 10% which results in a error of about

30%. This results in missed sales opportunities, dissatisfied customers, and potential damage to our brand reputation.

Moreover, these fluctuations in inventory levels contribute to inefficiencies throughout the supply chain, including increased handling and storage costs, heightened risk of stockouts, and challenges in maintaining optimal order fulfilment rates.

5.2 SARIMA output

Overall Mean Absolute Error (MAE): 25764.67
Overall Mean Squared Error (MSE): 3574424522.33
Overall R-squared (R2): +0.85
Overall Mean Absolute Percentage Error (MAPE): 18.64%

DMDUNIT	DMDGROUP	LOC	MAE	MSE	R2	MAPE
D100	CUST1	CA	111312.6518	25819545125	0.2394989828	6.852337843
D100	CUST3	PEN	111312.6518	25819545125	0.2394989828	6.852337843
A100	CUST2	PEN	15850.29499	408421444	0.3962310543	9.107628154
D100	CUST2	PEN	15850.29499	408421444	0.3962310543	9.107628154
A100	CUST2	GA	75204.82229	11511333204	-0.05716018424	11.65067875
D100	CUST1	PEN	75204.82229	11511333204	-0.05716018424	11.65067875
B100	CUST3	GA	3730.834317	24199351.2	-1.493123543	12.23606041
D100	CUST4	CA	3730.834317	24199351.2	-1.493123543	12.23606041
D100	CUST4	PEN	3730.834317	24199351.2	-1.493123543	12.23606041
A100	CUST4	PEN	25794.66702	972409864.2	-0.9823147768	12.27854735
B100	CUST1	GA	1542.126096	4077886.4	-1.124892596	12.59960802
D100	CUST1	GA	18908.25172	663752934	-0.09170199582	12.81561776
C100	CUST4	GA	9368.344908	139350643.3	0.1673098651	12.8594744

The SARIMA model demonstrates a good overall performance with a strong R-squared value and reasonable error metrics. The MAE and MAPE values indicate that while the forecasts are generally accurate, there are instances where the model's predictions significantly deviate from actual values, as suggested by the high MSE. The model's high R2 value signifies that it effectively captures the underlying patterns in the time series data, making it a reliable tool for forecasting future demand in this context. However, further tuning and error analysis might be required to reduce the larger individual errors indicated by the MSE.

5.3 Prophet Output

Overall Mean Absolute Error (MAE): 35331.22
Overall Mean Squared Error (MSE): 5677323470.85
Overall R-squared (R2): +0.7
Overall Mean Absolute Percentage Error (MAPE): 25.09%

DMDUNIT	DMDGROUP	LOC	MAE	MSE	R2	MAPE
D100	CUST1	CA	167324.6111	43504650576	-0.2814064251	10.27062308

D100	CUST3	PEN	167324.6111	43504650576	-0.2814064251	10.27062308
D100	CUST1	GA	18154.93666	579908494.1	0.04620043391	11.49298298
B100	CUST1	GA	1737.107648	4784995.428	-1.49335081	13.54833658
A100	CUST2	CA	13973.11314	320099920.6	0.164702052	14.10632972
C100	CUST4	GA	11416.54357	177670536.9	-0.06167075937	15.1011629
A100	CUST4	GA	147031.6937	31219092982	-0.2653098344	15.27620105
A100	CUST1	PEN	5849.928182	50186740.38	-0.676526191	17.41658707
B100	CUST4	PEN	56339.33941	6958444304	0.1486802132	17.48520837
A100	CUST2	PEN	31634.08494	1312754565	-0.9406435473	17.96129467
D100	CUST2	PEN	31634.08494	1312754565	-0.9406435473	17.96129467
D100	CUST2	GA	23059.18348	851213504.4	-0.9939169116	18.2994475
C100	CUST2	CA	2683.695738	10593179.98	-1.086227663	18.44400803

The Prophet model's performance, as indicated by the overall metrics, is somewhat mixed. The R-squared value of +0.7 suggests that the model captures a significant portion of the variability in the demand data, which is promising. However, the high MAE, MSE, and MAPE values indicate that the model's forecasts still have substantial errors, both in absolute and percentage terms.

Key takeaways from the performance metrics:

MAE and MAPE: These metrics indicate that the average error in the forecasts is quite high, both in absolute terms and as a percentage of the actual values.
MSE: The high MSE further emphasizes that there are significant individual forecast errors, likely due to outliers or large deviations in specific instances.
R-squared: A value of 0.7 is relatively strong, suggesting that while the model does capture the general trend and variance in the data, it struggles with precision.

5.4 LSTM Output

Overall Mean Absolute Error (MAE): 51516.85
Overall Mean Squared Error (MSE): 55101938779.88
Overall R-squared (R2): +0.78
Overall Mean Absolute Percentage Error (MAPE): 20.60%

DMDUNIT	DMDGROUP	LOC	MAE	MSE	R2	MAPE
A100	CUST4	PEN	17066.27166	466634181.8	-0.08943123087	8.706172262
B100	CUST3	GA	2723.580968	11231543.95	-0.3559560978	9.103379695
B100	CUST1	GA	1167.183275	2218899.279	-0.1010923977	10.61692661
B100	CUST1	PEN	10914.91038	193428846.3	-0.3477834444	10.83836084
D100	CUST4	PEN	2802.666815	14014080.26	-0.6918847191	10.84172475
D100	CUST1	CA	118083.8391	27544394969	0.2237923575	11.17536255
D100	CUST3	PEN	124307.4549	29604692387	0.1657326831	11.36900377

D100	CUST4	CA	3330.49486	17693652.87	-1.136110281	11.9489765
A100	CUST2	GA	70423.70651	9737784924	0.1041749265	12.76284655
B100	CUST4	CA	1296.793043	2643035.281	0.1353489226	12.9337333
D100	CUST1	PEN	67411.10329	9147997560	0.1584322666	12.96445869
A100	CUST3	CA	363.8042213	200880.7088	-0.03886927084	13.56791647
B100	CUST2	GA	3085.770518	15859752.66	0.5100877283	14.16038442
B100	CUST3	CA	1353.059306	3231945.261	-0.1449338876	14.29570472
D100	CUST1	GA	16827.5613	609521213.7	0.02221413455	14.36034746
D100	CUST3	GA	633.2734061	605650.1433	0.1235643892	14.6674814
D100	CUST2	PEN	13509.24901	339853834.8	0.5144351078	15.3700888
A100	CUST1	PEN	3796.72389	23232638.2	0.157784454	15.53928546

The LSTM model's performance metrics indicate that it has captured a significant portion of the data's variability (as reflected by the R-squared value of +0.78), but it also highlights areas for improvement, particularly in reducing the forecast errors.

Key takeaways from the performance metrics:

MAE and MAPE: These metrics show that the average error in the forecasts is substantial, both in absolute terms (MAE of 51,516.85) and as a percentage of the actual values (MAPE of 20.60%).
MSE: The high MSE suggests that there are considerable individual forecast errors, likely due to outliers or large deviations in specific instances.
R-squared: An R-squared value of +0.78 indicates that the model performs reasonably well in capturing the data's variability, explaining 78% of the variance.

5.5 Final Results

Based on our observation of the forecasting model performance metrics, we are choosing SARIMA for forecasting. The SARIMA model demonstrates superior accuracy with a Mean Absolute Error (MAE) of 25,764.67 and a Mean Squared Error (MSE) of 3,574,424,522.33, indicating lower forecast deviations compared to the other models. Additionally, SARIMA's R-squared (R2) value of +0.85 signifies that it captures 85% of the variance in the actual data, reflecting a strong ability to explain the underlying patterns in the time series. Furthermore, the Mean Absolute Percentage Error (MAPE) of 18.64% suggests a relatively lower percentage error, enhancing the model's reliability and precision. These metrics collectively highlight SARIMA's effectiveness and robustness in forecasting, making it the preferred choice over the Prophet and LSTM models, which exhibited higher errors and lower explanatory power.

A Power BI dashboard was created to demonstrate the outputs of the forecasting models, providing an interactive and comprehensive visualization of the forecasted quantities and model performance. The dashboard includes a comparison of the SARIMA model, showcasing their forecast accuracy through visual plots of actual vs. forecasted quantities over time. It features filters for different demand units (DMDUNIT), demand groups (DMDGROUP), and locations (LOC), allowing users to drill down into specific segments for a detailed analysis. The error distribution visualizations help users understand the variability and dispersion of forecast inaccuracies, while the key metrics summary table offers a quick reference for evaluating model performance comprehensively. This dashboard serves as a powerful tool for stakeholders to visualize, compare, and interpret the forecasting results, aiding in informed decision-making for inventory management and demand planning.

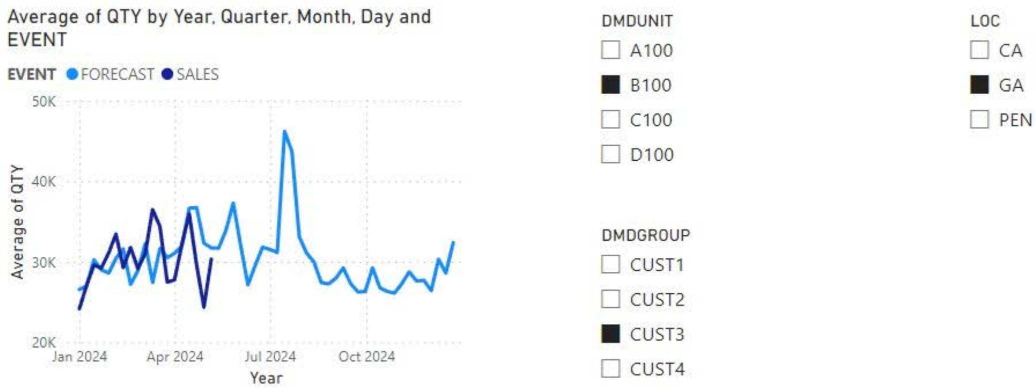


Fig 4 Forecast and sales graph of B100 CUST3 And GA

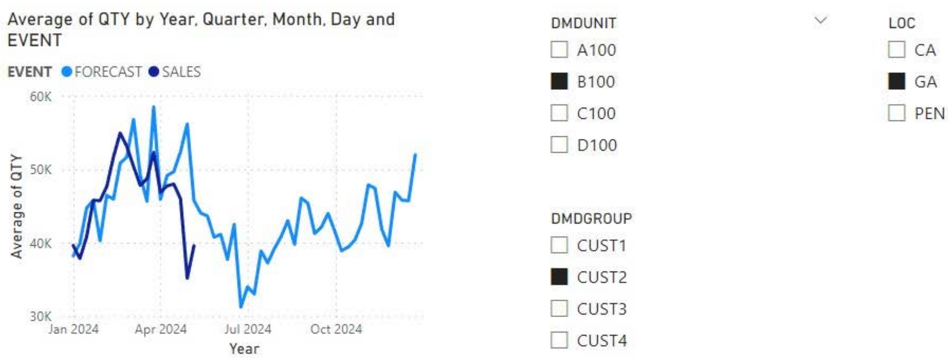


Fig 5 Forecast and sales graph of B100 CUST2And GA

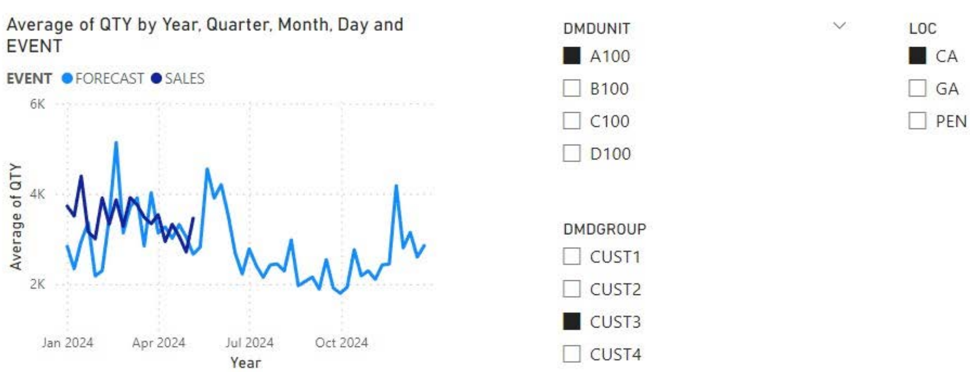


Fig 6 Forecast and sales graph of A100 CUST3 And CA

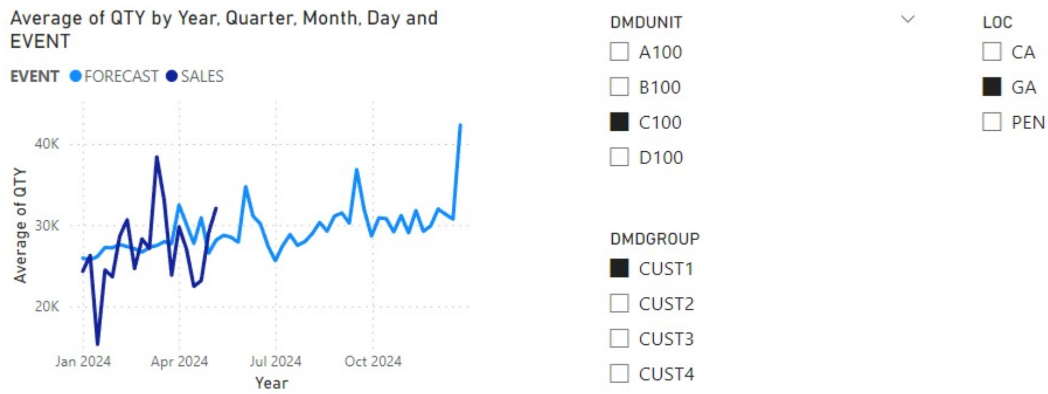


Fig 7 Forecast and sales graph of C100 CUST1 And GA

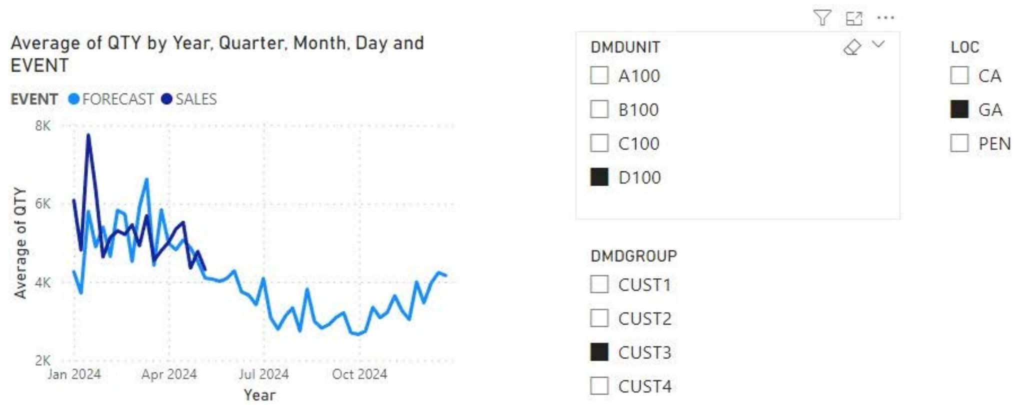


Fig 8 Forecast and sales graph of D100 CUST3 And GA

Conclusion

The implementation of the forecasting system using SARIMA, Prophet, and LSTM models was meticulously planned and executed, involving critical steps to ensure accuracy and reliability. A detailed dataset list was created alongside data preprocessing and model selection processes, providing a structured foundation. Pre-processing involved thorough data cleaning, normalization, and transformation to enhance model input quality. Model evaluation and performance comparison were conducted, addressing issues like overfitting and discrepancies between forecasted and actual values. This comprehensive approach ensured the system's reliability, efficiency, and performance, meeting customer expectations. A Power BI dashboard was created to demonstrate outputs, presented to stakeholders for approval before integrating into inventory management. At the integration site, commissioning, user training, testing, and handover finalized the process.

Disclaimer

Mention of commercial products, services, trade or brand names, organizations, or research studies in this publication does not imply endorsement by the authors, nor discrimination against similar products, services, trade or brand names, organizations, or research studies not mentioned.

11

Acknowledgements

The paper is made as a consolidation of works and materials of reference authors thus cited, along with research of ours. All credits of materials go to respective owners.

References:

[1] Hossein Abbasimehr, Mostafa Shabani, Mohsen Yousefi,An optimized model using LSTM network for demand forecasting,Computers & Industrial Engineering,Volume 143,2020,106435, ISSN 0360-8352, https://doi.org/10.1016/j.cie.2020.106435 .

[2] Punia, S., Nikolopoulos, K., Singh, S. P., Madaan, J. K., & Litsiou, K. (2020). Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail. International Journal of Production Research, 58(16), 4964–4979. https://doi.org/10.1080/00207543.2020.1735666

[3] St-Aubin, Philippe, and Bruno Agard. 2022. Precision and Reliability of Forecasts Performance Metrics Forecasting 4, no. 4: 882-903. https://doi.org/10.3390/forecast4040048

[4] Nijat Mehdiyev, David Enke, Peter Fettke, Peter Loos,Evaluating Forecasting Methods by Considering Different Accuracy Measures,Procedia Computer Science,Volume 95,2016,Pages 264-271,ISSN 1877-0509, https://doi.org/10.1016/j.procs.2016.09.332 .

[5] Yanrui Ning, Hossein Kazemi, Pejman Tahmasebi,A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet,Computers & Geosciences,Volume 164,2022,105126,ISSN 0098-3004, https://doi.org/10.1016/j.cageo.2022.105126 .

[6] Hasnain Ahmad, S. Yehua, M. Z. Hashmi, U. A. Bhatti, A. Hussain, M. Hameed, S. Marjan, S. U. Bazai, M. A. Hossain, M. Sahabuddin, R. A. Wagan, and Y. Zha, Time Series Analysis and Forecasting of Air Pollutants Based on Prophet Forecasting Model in Jiangsu Province, China,Frontiers in Environmental Science, vol. 10, 2022, https://doi.org/10.3389/fenvs.2022.945628 .

[7] Safari, M.; Rabiee, A.H.; Joudaki, J. Developing a Support Vector Regression (SVR) Model for Prediction of Main and Lateral Bending Angles in Laser Tube Bending Process. Materials 2023, 16, 3251. https://doi.org/10.3390/ma16083251

[8] Jewel, R. M., Linkon, A. A., Shaima, M., Badruddowza, Md Shohail Uddin Sarker, Rumana Shahid, Norun Nabi, Md Nasir Uddin Rana, Md Ahnaf Shahriyar, Mehedi Hasan, & Md Jubayar Hossain. (2024). Comparative Analysis of Machine Learning Models for Accurate Retail Sales Demand Forecasting. Journal of Computer Science and Technology Studies, 6(1), 204–210. https://doi.org/10.32996/jcsts.2024.6.1.23

[9] Chodakowska, E.; Nazarko, J.; Nazarko, Ł.; Rabayah, H.S.; Abendeh, R.M.; Alawneh, R. ARIMA Models in Solar Radiation Forecasting in Different Geographic Locations. Energies 2023, 16, 5029. https://doi.org/10.3390/en16135029

[10]Toharudin, T., Pontoh, R. S., Caraka, R. E., Zahroh, S., Lee, Y., & Chen, R. C. (2023). Employing long short-term memory and Facebook prophet model in air temperature forecasting.

Communications in Statistics - Simulation and Computation, 52(2), 279–290. https://doi.org/10.1080/03610918.2020.1854302