T.C.
BAHCESEHIR UNIVERSITY
GRADUATE SCHOOL
THE DEPARTMENT OF ARTIFICIAL INTELLIGENCE
(INTERDISCIPLINARY)
MASTER'S PROGRAM IN ARTIFICIAL INTELLIGENCE

COMPARISON OF THE SPARE PART DEMAND FORECASTING MODELS IN AUTOMOTIVE INDUSTRY

MASTER'S THESIS

ÇAĞATAY ÇİFTÇİ

ISTANBUL 2025

T.C.
BAHCESEHIR UNIVERSITY
GRADUATE SCHOOL
THE DEPARTMENT OF ARTIFICIAL INTELLIGENCE
(INTERDISCIPLINARY)
MASTER'S PROGRAM IN ARTIFICIAL INTELLIGENCE

COMPARISON OF THE SPARE PART DEMAND FORECASTING MODELS IN AUTOMOTIVE INDUSTRY

MASTER'S THESIS

ÇAĞATAY ÇİFTÇİ

THESIS ADVISOR

ASST. PROF. TAMER UÇAR

ISTANBUL 2025

T.C.

BAHCESEHIR UNIVERSITY

GRADUATE SCHOOL

…./…/…

MASTER THESIS APPROVAL FORM

Program Name:	Artificial Intelligence (English,Thesis)
Student's Name and Surname:	Çağatay ÇİFTÇİ
Name Of The Thesis:	Comparison of The Spare Part Demand Forecasting Models in Automotive Industry
Thesis Defense Date:	22/05/2025

This thesis has been approved by the Graduate School which has fulfilled the necessary conditions as Master thesis.

Assoc. Prof. Yücel Batu SALMAN

Institute Director

This thesis was read by us, quality and content as a Master's thesis has been seen and accepted as sufficient.

	Title/Name	**Institution**	**Signature**
Thesis Advisor's	Asst. Prof. Tamer UÇAR	Bahçeşehir University	
Member's	Asst. Prof. Derya BODUR	Bahçeşehir University	
Member's	Asst. Prof. Ümit ÖZTÜRK	Beykent University	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Çağatay Çiftçi

Signature:

ABSTRACT

COMPARISON OF THE SPARE PART DEMAND FORECASTING MODELS IN AUTOMOTIVE INDUSTRY

Çağatay Çiftçi
Master's Program in Artificial Intelligence
Supervisor: Asst. Prof. Tamer Uçar

May 2025, 72 pages

This thesis aims to compare the performance of different methods for spare parts demand forecasting, which is one of the most important parts of effective inventory management in the automotive industry. In contrast to traditional statistical models, DeepAR, Prophet, LSTM, and Attention-based LSTM models, which are widely used especially in the field of NLP, are used to examine the ability of these methods to predict spare parts order data in the automotive industry. The research classifies spare parts time series as stationary, non-stationary, seasonal, and non-seasonal using data from an automotive company in Turkey since 2018. Thus, the effectiveness of the analyzed models on different time series structures was evaluated in detail. One of the key contributions of this study is highlighting the effectiveness and benefits of the Attention mechanism, initially developed for NLP, in enhancing time series-based demand forecasting. The results demonstrate how advanced forecasting models can improve metrics that are used in inventory management in the automotive sector, leading to greater operational efficiency and reduced costs.

Key Words: Spare Parts Demand Forecasting, Inventory Management, Automotive Industry, Attention Mechanism

ÖZET

OTOMOTİV SEKTÖRÜNDE YEDEK PARÇA TALEP TAHMİN MODELLERİNİN KARŞILAŞTIRILMASI

Çağatay Çiftçi
Yapay Zeka Yüksek Lisans Programı
Tez Danışmanı: Dr. Öğr. Üyesi Tamer Uçar

Mayıs 2025, 72 sayfa

Bu tez, otomotiv sektöründe etkin envanter yönetiminin en önemli parçalarından biri olan yedek parça talep tahmini için farklı yöntemlerin performansını karşılaştırmayı amaçlamaktadır. Geleneksel istatistiksel modellerin aksine özellikle NLP alanında yaygın olarak kullanılan DeepAR, Prophet, LSTM ve Attention tabanlı LSTM modelleri kullanılarak bu yöntemlerin otomotiv sektöründeki yedek parça sipariş verilerini tahmin edebilme yetenekleri incelenmiştir. Araştırmada, Türkiye'deki bir otomotiv firmasının 2018 yılına ait verileri kullanılarak yedek parça zaman serileri durağan, durağan olmayan, mevsimsel ve mevsimsel olmayan olarak sınıflandırılmıştır. Böylece, analiz edilen modellerin farklı zaman serisi yapıları üzerindeki etkinliği ayrıntılı olarak değerlendirilmiştir. Bu çalışmanın en önemli katkılarından biri, başlangıçta NLP için geliştirilen Dikkat mekanizmasının zaman serisi tabanlı talep tahminini geliştirmedeki etkinliğini ve faydalarını vurgulamaktır. Sonuçlar, gelişmiş tahmin modellerinin otomotiv sektöründe envanter yönetiminde kullanılan ölçütleri nasıl iyileştirebileceğini ve böylece daha fazla operasyonel verimlilik ve daha düşük maliyetler sağlayabileceğini göstermektedir.

Anahtar Kelimeler: Yedek Parça Talep Tahmin, Envanter Yönetimi, Otomotiv Endüstrisi, Dikkat Mekanizması

To the Greatest Supporter of My Life, Dilara.

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my supervisor Asst. Prof. Dr. Tamer UÇAR for his guidance, advice, criticism, encouragements and insight throughout the research.

I would also like to thank my wife, Dilara SAVUR ÇİFTÇİ, for her endless love, patience, and motivation during this journey.

My heartfelt thanks also go to my parents, Sebahat and Mehmet ÇİFTÇİ, for their lifelong support and encouragement. Without them, none of this would have been possible.

TABLE OF CONTENTS

ETHICAL CONDUCT iii

ABSTRACT iv

ÖZET v

DEDICATION vi

ACKNOWLEDGEMENTS vii

LIST OF TABLES xi

LIST OF FIGURES xii

LIST OF ABBREVIATIONS xiii

Chapter 1 Introduction 1

1.1 Theoretical Overview 1

1.1.1 The importance of demand forecasting in the automotive industry ... 2

1.1.2 The effect of demand forecasting on stock planning 2

1.2 Statement of the Problem 2

1.3 Purpose of Study 3

1.4 Research Questions 3

1.5 Significance of the Study 3

1.6 Definitions 4

Chapter 2 Literature Review 5

2.1 Demand Forecasting 6

2.2 DeepAR Literature Review 9

2.3 Prophet Literature Review 13

2.4 LSTM Literature Review 17

2.5 Attention Based LSTM Literature Review 20

Chapter 3 Methodology 27

3.1 Data Preparation 27
3.1.1 Creation univariate time series 27
3.1.2 Removing outliers with hampel filter 28
3.1.3 Resampling 28
3.1.4 Time series classification 29
3.1.4.1 Stationary test 29
3.1.4.2 Seasonality test 30
3.1.5 Normalization 30
3.2 Hyperparameter Optimization 31
3.3 DeepAR Model 33
3.3.1 Data preprocessing 33
3.3.2 Model architecture and hyperparameter optimization 34
3.4 Prophet Model 37
3.4.1 Data preprocessing for prophet model 37
3.4.2 Model architecture and hyperparameter optimization 38
3.4.2.1 Trend component 38
3.4.2.2 Seasonality component 38
3.4.2.3 Holiday effects 39
3.4.2.4 Hyperparameter optimization 39
3.4.2.5 Advantages of prophet 40
3.5 Long Short-Term Memory Model 40
3.5.1 Data preprocessing for lstm model 40
3.5.2 Model architecture and hyperparameter optimization 41
3.5.2.1 Hyperparameter optimization 43
3.6 Attention Based LSTM 45
3.6.1 Multi head attention mechanism 45
3.6.1.1 Scaled dot-product attention 46
3.6.1.2 Multi-head attention in time series forecasting 46
3.6.2 Model architecture and hyperparameter optimization 47
3.7 Evaluation Metrics 49
3.7.1 Mean absolute error 49

3.7.2 Root mean squered error 50
3.7.3 Symmetric mean absolute perentage error 51

Chapter 4 Findings 53

4.1 Result by Time Series Datasets 53
4.1.1 Stationary time series 54
4.1.2 Non-stationary time series 57
4.1.3 Seasonal time series 60
4.1.4 Non-seasonal time series 63

4.2 Overall Performance Summary 66

Chapter 5 Discussions and Conclusions 70

REFERENCES 73

LIST OF TABLES

TABLES

Table 1 Sample From The Data Used For The Thesis 27
Table 2 Values For Which Dataset Is Prepared 34
Table 3 Best Performing Parameters For Deepar 36
Table 4 Best Performing Parameters For Lstm... 44
Table 5 Best Performing Parameters For Attention Based Lstm 49

LIST OF FIGURES

FIGURES

Figure 1 LSTM architecture 42
Figure 2 LSTM forget gate 42
Figure 3 LSTM input gate 43
Figure 4 Scaled dot-product and multi-head attention (Vaswani et al. (2017)) 46
Figure 5 Performance comparison for stationary time series 55
Figure 6 Error distribution for stationary time series parts 56
Figure 7 Performance comparison for non-stationary time series 59
Figure 8 Error distribution for non-stationary time series parts 59
Figure 9 Performance comparison for seasonal time series 62
Figure 10 Error distribution for seasonal time series parts 63
Figure 11 Performance comparison for non-seasonal time series 65
Figure 12 Error distribution for non-seasonal time series parts 66
Figure 13 Overall performance comparison across all time series types 68

LIST OF ABBREVIATIONS

LSTM	Long Short-Term Memory
RNN	Recurrent Neural Network
DNN	Deep Neural Network
SEQ2SEQ	Sequence-to-Sequence
NLP	Natural Language Processing
TPE	Tree-structured Parzen Estimator
ETS	Exponantial Smoothing
MHA	Multi-Head Attention
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Squared Error
SMAPE	Symmetric Mean Absolute Percentage Error
MSE	Mean Squered Error
ARIMA	AutoRegressive Integrated Moving Average
SARIMA	Seasonal AutoRegressive Integrated Moving Average
SKU	Stock Keeping Unit

Chapter 1

Introduction

Many industries have turned to statistical or artificial intelligence-based applications to manage their inventories. The most important of these applications is accurately predicting future demand. The automotive industry is undoubtedly one of the most important industries for demand forecasting. Accurate demand forecasting is a crucial phase for the efficient operation of supply chains, optimizes inventory levels, and minimizes costs related to overstocking or stockouts.

In recent years, significant developments in demand forecasting methodologies have left traditional and statistical models behind, with machine learning and deep learning techniques producing more effective outputs. However, many studies in the automotive sector have yet to fully realize the potential of emerging forecasting models such as DeepAR and Prophet. In addition, investigating the impact of attention mechanisms on demand forecasting in the automotive industry will be one of the innovative aspects of this study. This study seeks to address these gaps by comparing various forecasting models on datasets categorized into stationary, non-stationary, seasonal, and non-seasonal classifications subsets.

1.1 Theoretical Overview

This research focuses on time series forecasting, utilizing deep learning-based forecasting models. Machine learning models such as Prophet, LSTM, and DeepAR offer superior performance by learning temporal dependencies in the data.

An important novelty investigated in this study is the impact of the attention mechanism on demand forecasting. The attention mechanism has been widely applied in natural language processing, but relatively little was explored in time series forecasting. This study evaluates the effectiveness of Attention based LSTM models compared to other approaches.

1.1.1 The importance of demand forecasting in the automotive industry. Forecasting spare parts demand is critical for aftermarket services, especially for discontinued vehicles.

Inaccurate forecasting results in inefficient utilization of resources, elevated inventory costs, and disruptions within the supply chain. Conversely, effective forecasting enhances supply chain flexibility by optimizing supply planning, reducing lead times, and ensuring the availability of spare parts availability. The automotive industry faces unique challenges in demand forecasting due to fluctuations in part consumption, seasonal trends, and unexpected changes in market demand. Forecasting methodologies must meet these complexities to improve overall operational efficiency.

1.1.2 The effect of demand forecasting on stock planning. Stock planning constitutes an essential component of inventory management, wherein demand forecasting plays an important role in enhancing the accuracy of stock replenishment decisions.

Organizations devise various optimization strategies aimed at improving key performance indicators employed in inventory planning, including fill rate and stock turnover ratios, while also trying to minimize costs. The foremost element within these optimization strategies is demand forecasting. Companies can improve inventory planning accuracy using advanced forecasting techniques such as LSTM, Prophet, and DeepAR. Segmenting datasets based on statistical characteristics, as done here, helps identify the most appropriate forecasting models for different demand patterns.

1.2 Statement of the Problem

Despite the importance of accurate demand forecasting in the automotive industry, existing research has primarily focused on traditional statistical methods. Limited studies examine the performance of DeepAR and Prophet models in this context. Furthermore, the role of the attention mechanism in improving forecast accuracy has not been extensively investigated in the automotive industry.

1.3 Purpose of Study

The main objective of this study is to compare the performance of DeepAR, Prophet, LSTM, and Attention based LSTM models in automotive spare parts demand forecasting. The study aims to:

- To evaluate the applicability of DeepAR and Prophet models in the automotive industry,
- Evaluate the effect of the attention mechanism on prediction accuracy,
- Determine the most appropriate forecasting model for different time series data characteristics.

1.4 Research Questions

This research seeks to answer the following questions:

- How do DeepAR and Prophet models perform in forecasting automotive spare parts demand?
- How does the attention mechanism impact time series forecasting in the automotive sector?
- Which forecasting model performs best for stationary, non-stationary, seasonal, and non-seasonal data subsets?

1.5 Significance of the Study

This study contributes to the literature by comparing emerging forecasting techniques in the automotive industry. The key contributions include:

- Bridging the research gap on the application of DeepAR and Prophet in automotive spare parts demand forecasting,
- Evaluating the effectiveness of Attention based LSTM models in time series forecasting,

- Offering a segmentation-based performance analysis to identify the best forecasting approach for different data patterns.

1.6 Definitions

Demand Forecasting: The process of predicting future product demand using historical data and machine learning or deep learning models.

Attention Mechanism: A neural network component that selectively focuses on relevant parts of input data to improve predictive accuracy.

DeepAR: A deep learning-based probabilistic forecasting model designed for sequential data.

Prophet: A forecasting model developed by Facebook that accounts for seasonality and external factors.

LSTM (Long Short-Term Memory): A recurrent neural network (RNN) type designed to capture long-term dependencies in sequential data.

Stationary Time Series: A time series whose statistical properties, such as mean and variance, remain constant over time.

Non-Stationary Time Series: A time series with changing statistical properties over time.

Seasonality: Recurrent patterns in time series data that repeat at regular intervals.

Stock Planning: The process of determining optimal inventory levels to meet future demand while minimizing costs.

Chapter 2

Literature Review

This section presents detailed research on demand forecasting and model research used in demand forecasting. These researches are not only within the scope of the automotive industry, but the models and methods used are intended to shed light on this study in the automotive industry.

Research on demand forecasting models in the automotive industry has evolved from traditional statistical approaches to integrate machine learning and deep learning techniques. Classic statistical models, such as ARIMA and exponential smoothing, are favored for their clarity and success in stationary time series forecasting. However, the challenge of working with complex and non-linear models has led to the search for more advanced techniques. Recent research indicates that deep learning models are increasingly effective for forecasting time series data. Specifically, models such as DeepAR, Prophet, LSTM, and Attention-based LSTM have gained importance due to their ability to capture long-term dependencies, non-stationary trends, and seasonal patterns.

DeepAR, developed by Amazon, uses an autoregressive recurrent network that learns probabilistic forecasts across multiple time series. Studies have demonstrated its superior performance in handling sparse datasets and producing well-calibrated forecast intervals.

Prophet, a model developed by Facebook, is widely used for forecasting time series with strong seasonal components and missing data by exploiting an additive model structure.

In contrast, LSTMs as recurrent neural networks are particularly effective in capturing sequential dependencies in time series data. An emerging area of interest is the application of attentional mechanisms in time series forecasting. Originally developed for NLP tasks, attentional mechanisms improve prediction accuracy by enabling models to focus on relevant time steps in a sequence.

Attention-based LSTMs integrate attention mechanisms into LSTMs and effectively detect complex dependencies in time series data.

The effectiveness of this approach in forecasting spare parts demand is an ongoing research topic.

2.1 Demand Forecasting

The integration of machine learning techniques into inventory management has significantly advanced the field of demand forecasting, particularly in the automotive industry. Nazarenko (2024) presents a hybrid approach combining short-term demand forecasting with inventory refill strategies. The research shows the significant role of machine learning and deep learning techniques such as LSTM in effectively forecasting spare parts demand within automotive production. It shows that a systematically formulated supply approach, guided by machine learning insights, can improve inventory management by reducing stockouts and overstock inventory. Accuracy in demand forecasting is a crucial factor in inventory control. While traditional statistical techniques like ARIMA are frequently used for predicting spare parts demand, Nazarenko's research (2024) reveals that deep learning models, especially LSTM, are advanced to these methods in effectively detecting demand variability.

The study also emphasizes the need to integrate short-term forecasts into refilling strategies to adapt well to the variability in demand variations. The choice of forecasting model significantly impacts the success of inventory control policies. Nazarenko (2024) compares different machine learning approaches, including LSTM, ARIMA, and XGBoost, for predicting battery demand in a car production factory. The results show that LSTM-based models provide more accurate demand predictions, reducing stockout risks while maintaining optimal stock levels. This discovery reinforces the case for employing deep learning models in forecasting spare parts, mainly when demand trends show non-stationary behavior and seasonal variations. The use of machine learning in inventory management signifies a movement toward data-driven decision-making in the automotive industry.

According to Nazarenko (2024), hybrid systems that include inventory simulation and predictive modeling are more effective than classic refilling tactics that only use historical demand data. The study offers insightful information about how an optimal replenishment strategy coupled with short-term demand forecasts can result in more effective inventory management, which lowers expenses and raises service standards. These results add to the increasing amount of research supporting the application of deep learning methods to forecast demand for spare parts.

The impact of demand forecasting on inventory control is a very important factor to consider. In their 2015 study, Rego and Mesquita (2015) analyzed several forecasting and inventory control strategies using a large-scale simulation of spare parts demand forecasting and inventory management. Their study demonstrates how good forecasting methods may guarantee high service levels while lowering holding costs and stockouts. Using a dataset with over 10,000 SKUs, the study assesses many demand forecasting techniques, such as bootstrapping, the Syntetos-Boylan Approximation (SBA), and the simple moving average (SMA). The results indicate that advanced probabilistic forecasting techniques, such as compound Poisson distributions and bootstrapping, can establish more efficient inventory rules that reduce costs and improve fill rates. This study highlights the key need for better inventory management in the automotive spare parts industry. With powerful forecasting techniques such as DeepAR, Prophet, LSTM, and Attention-based LSTM, inventory management can be improved by increasing the forecasting accuracy on datasets with different characteristics.

Rožanec, Kažič, Škrjanc, Fortuna and Mladenić (2021) conducted a thorough comparative analysis of 21 forecasting models based on machine learning and statistics for predicting automotive OEM demand. Their research showed that machine learning techniques performed better than conventional statistical techniques, especially when dealing with unpredictable and non-stationary demand. The study also demonstrated the advantages of using global machine learning models, which reduce predicting mistakes by pooling product data according to historical demand size. The study also suggested a set of assessment criteria that offer a more in-depth understanding of the advantages and disadvantages of various forecasting models. These results support our research goals as they highlight how powerful deep learning models such as DeepAR, Prophet, LSTM, and Attention-based LSTM can improve the accuracy of demand forecasting in the automobile sector, as opposed to ML models.

The efficiency of deep learning models, specifically Deep Neural Networks (DNN) and Recurrent Neural Networks (RNN), in predicting demand for automobile accessories is further supported by the study conducted by Cadena, Albright, Wang and Ajmera (2023). According to their research, RNN models perform poorly when forecasting over long periods, but they are especially effective in short-term forecasting because they can capture sequential relationships. On the other hand, models that successfully model intricate time dependencies, such as LSTM and DeepAR, shows better performance for long-term predictions. The study also provides empirical evidence that choosing key time stages in the sequence allows attention-based mechanisms to enhance forecasting accuracy. Their findings revealed that the RNN model was effective for short-term forecasting, attaining an RMSE of 51,000 for a regional dataset and reducing the error rate to 0.08%. The importance of incorporating attention into time series forecasting models is further supported by the expectation that models with attention mechanisms, such as Attention-based LSTM, will produce more accurate predictions for longer time horizons. In the automobile sector, AI-driven demand forecasting models have shown notable gains in accuracy and inventory efficiency.

Omprakash (2024) compared AI-driven forecasting techniques (e.g., neural networks, deep learning, and reinforcement learning-based optimization) with traditional time series forecasting methods, demonstrating the positive impact of AI-based models on inventory management. The study found that AI-assisted forecasting systems diminished the Mean Absolute Percentage Error (MAPE) by 23%, resulting in a 15% reduction in excess inventory. Furthermore, manufacturers using AI-powered inventory management systems experienced a 30% decrease in stockout incidents, facilitating improved supply alignment demand.

These findings highlight the advantages of AI-driven approaches in forecasting spare parts demand in the automotive industry. The study supports the claim that AI-based time series forecasting models, such as DeepAR and Prophet, perform better in high-variability environments.

At the same time, deep learning models like Attention-based LSTM and LSTM demonstrate strong potential for capturing long-term dependencies and adapting to seasonal demand fluctuations.

Although attention-based LSTM was not explicitly compared to Prophet and DeepAR in this study, the findings primarily emphasize the general benefits of AI-driven forecasting models in improving inventory control and reducing forecasting uncertainty. A more thorough empirical comparison is necessary to reach definitive conclusions about the relative effectiveness of Attention-based LSTM and Prophet in forecasting spare parts demand.

This study reinforces the importance of AI-based forecasting techniques in spare parts inventory management. It highlights the need for further comparative analysis of LSTM, Attention-based LSTM, DeepAR, and Prophet better to understand their specific strengths and limitations in forecasting accuracy.

2.2 Deepar Literature Review

Building on these foundations, it is crucial to investigate how DeepAR, in particular, enhances and expands upon the approaches that have been covered thus far. The automobile industry's intrinsic non-stationarity, seasonality, and intermittency make forecasting spare component demand particularly difficult. Deep learning techniques have become popular due to the limitations of traditional statistical models, like exponential smoothing and ARIMA, in handling complex and large-scale time series. Amazon's DeepAR model, which uses autoregressive recurrent neural networks to increase probabilistic forecasting accuracy, is among the most critical developments in this area Salinas, Flunkert, Gasthaus and Januschowski (2020). DeepAR effectively recognizes shared patterns and relationships through a single model trained on several related time series.

In contrast to traditional methods that assess parameters individually for each series, this approach is particularly beneficial for predicting spare parts, given that the demand for different parts often exhibits comparable trends. DeepAR has demonstrated superior forecasting performance compared to traditional and alternative machine learning models. According to empirical evaluations, DeepAR provides more accurate forecasts than traditional methods such as Croston and ETS. For instance, DeepAR demonstrates approximately 33% lower error in the 0.5-risk metric compared to the Croston model and 54% lower error in the 0.9-risk metric. In tests conducted on electricity and traffic datasets, DeepAR outperforms the MatFact method by 56% in the ND (Normalized Deviation) metric and 13% in RMSE (Root Mean Squared Error) Salinas et al. (2020). DeepAR improves decision-making in uncertain environments by creating probabilistic forecasts using Monte Carlo sampling. Its flexibility in adapting to unpredictable demand variations makes it key for efficient spare parts inventory management. Additionally, its capability to learn from related time series allows for accurate forecasting even in cases where historical data is sparse. These findings support the growing interest in combining advanced models such as Prophet and Attention-based LSTM with deep learning-based probabilistic forecasting techniques like DeepAR to enhance demand forecasting accuracy in the automotive sector.

In the automotive industry, effective demand forecasting requires models tailored to address real-world data challenges, including anomalies and incomplete information. Traditional statistical models often struggle with incomplete datasets and unpredictable data changes, leading to unreliable forecasts. In contrast, DeepAR has demonstrated its ability to tackle these issues successfully. Bohlke-Schneider, Kapoor, and Januschowski (2020) point out that DeepAR effectively manages missing values by masking their influence on loss calculations and using a probabilistic imputation approach. This strategy eliminates manual imputation, minimizing potential biases and enhancing the model's reliability and robustness. DeepAR employs a model averaging technique to ensure stability in forecasts, which standardizes predictions throughout retraining cycles. Empirical evidence indicates that model averaging leads to a 42.2% improvement in forecast stability, reducing prediction variations caused by model instability retraining. Beyond missing data, anomalies—such as sudden spikes or drops in demand—can significantly distort conventional forecasting models. Bohlke-Schneider et al. (2020) describe an anomaly detection framework in DeepAR that identifies and mitigates extreme deviations before they impact future forecasts. The system assesses predicted values by comparing them with actual observations, replacing recognized anomalies with missing values that the model imputes automatically.

This approach has been highly effective in reducing the impact of rare events, leading to significant improvements in forecasting accuracy. In a particular case study, the Mean Absolute Error (MAE) decreased from 0.19 to 0.04 following anomaly detection applied.

This resilience is particularly valuable for spare parts forecasting, where abrupt demand fluctuations due to supply chain disruptions or market dynamics are common.

These findings support the claim that DeepAR provides a more reliable and flexible approach to demand forecasting in dynamic industrial settings when combined with models such as Prophet and Attention based LSTM.

Forecasting spare parts demand in the automotive sector requires models capable of handling highly unpredictable and non-stationary data. Conventional statistical techniques like ARIMA and exponential smoothing frequently struggle to accurately capture rapid fluctuations in demand, leading to unreliable forecasts. Conversely, recent studies indicate that probabilistic forecasting models utilizing deep learning, like DeepAR, excel in managing high variability and recognizing non-stationary trends Fan and Lixin (2024).

DeepAR, which leverages autoregressive recurrent networks, has been demonstrated to perform effectively under volatile conditions. Empirical results show that DeepAR reduces the Root Mean Square Error (RMSE) by 12.7% compared to traditional time series models during periods of high volatility. This makes DeepAR particularly well-suited for spare parts forecasting, where seasonality and uncertainty in demand play crucial roles in inventory management.

Recent research also highlights the benefits of integrating attention mechanisms into time series forecasting models, which enhances their ability to handle volatility Fan and Lixin (2024). Attention-oriented models, especially Transformers, excel at capturing long-range dependencies in sequential data, enhancing predictive accuracy. Research shows that Transformer models outperform architectures such as LSTM and GRU, achieving a Mean Absolute Percentage Error (MAPE) of 8.05%.

Despite demand fluctuations, DeepAR continues to perform well thanks to its probabilistic framework, which effectively manages missing data and minimizes anomalies.

Studies show that integrating DeepAR's probabilistic forecasting with attention-based models can significantly enhance the precision of spare parts demand forecasts. This approach is designed to improve inventory management reliability and optimize supply chain operations within the automotive industry. Additionally, DeepAR's automated systems for addressing missing values and identifying anomalies increase its adaptability amid considerable demand shifts, strengthening its effectiveness in dynamic forecasting environments. These insights indicate that a combined strategy utilizing DeepAR alongside attention-augmented architectures could yield more precise demand predictions, leading to improved supply chain strategies and enhanced inventory management in the automotive industry.

The irregular trends and recurring seasonal patterns in the automotive industry's spare parts usage make forecasting demand particularly challenging. Although SARIMA and other traditional statistical models have been widely used for time series forecasting, their limitations in handling complex seasonal dependencies have led to the growing adoption of deep learning-based methods. Recent studies indicate that DeepAR, particularly when optimized with hyperparameter tuning techniques like Optuna, significantly outperforms SARIMA in capturing seasonal patterns Zahid, Fitrianto, Silvianti, and Alamudi (2024). A comparative analysis of seasonal rice production forecasting demonstrates DeepAR's superior accuracy over SARIMA. DeepAR has a MAPE of 4.50% and a RMSE of 148,820, compared to SARIMA's MAPE of 5.44% and RMSE of 346,894. These results highlight DeepAR's ability to detect complex seasonal trends across various time series, leading to more precise predictions when significant seasonality exists.

In the automotive industry, seasonal demand variations greatly influence inventory management, making DeepAR a strong option for forecasting spare parts requirements. Unlike SARIMA, which necessitates manual parameter tuning for each dataset, DeepAR utilizes a recurrent neural network that automatically identifies seasonal trends. This capability is especially useful for managing spare parts, as demand often shows intricate yet predictable seasonal shifts. Additionally, the use of sophisticated deep learning algorithms in demand forecasting is justified due to their capacity to enhance stock levels and reduce holding costs.

By accurately detecting seasonal patterns, DeepAR improves inventory management and aids in sound supply chain decisions in the automotive sector sector. In many industries, especially within the automotive field, managing spare parts inventory effectively hinges on accurate demand forecasts.

Although traditional statistical methods such as exponential smoothing and ARIMA are popular for time series predictions, their deterministic characteristics restrict their capacity to handle uncertainty. On the other hand, DeepAR presents a probabilistic forecasting technique that delivers point estimates and appropriately calibrated confidence intervals.

A recent study by Wunsch, Kühnert, Wallner and Ziebarth (2024) found that DeepAR outperforms traditional methods in forecasting urban water demand. Specifically, DeepAR achieved a coefficient of determination (R^2) of up to 0.98 in certain district-metered areas (DMAs), while conventional models struggled with R^2 values below 0.80. DeepAR's Prediction Interval Coverage Probability (PICP) also consistently exceeded 0.80, indicating that most observed values fell within the expected confidence intervals. This feature is especially useful for forecasting the demand for spare parts, as changes in order volumes can greatly affect inventory management.

DeepAR enhances forecasting accuracy by incorporating external factors such as weather trends, seasonal shifts, and holidays, all crucial for demand prediction. Even during notable fluctuations, studies show that DeepAR's probabilistic intervals effectively capture demand variations, affirming its reliability as a forecasting tool in diverse fields like finance, industrial supply chains, and urban resource management.

DeepAR enables inventory managers to produce forecasts with confidence intervals, allowing them to make knowledgeable decisions that could reduce stockout risks and cut down on excess inventory costs. Its ability to model complex relationships and incorporate uncertainty makes DeepAR particularly effective in improving the accuracy of spare parts demand forecasting in dynamic and unpredictable environments.

2.3 Prophet Literature Review

Facebook created Prophet, a time series forecasting algorithm that can provide precise and scalable predictions even with complicated seasonality and erratic demand swings.

Prophet utilizes an additive model framework that integrates trend, seasonality, and external factors such as holidays, enabling more flexible and understandable forecasts than traditional methods like ARIMA. Key advantages include managing missing data, automatically detecting changepoints, and allowing the incorporation of domain expertise through user-defined parameters. Because of these characteristics, Prophet works especially well for corporate forecasting applications that need automated, large-scale projections Taylor and Letham (2017). Prophet has shown notable gains in accuracy in demand forecasting when working with high-dimensional data. The predictive effectiveness of the model is improved by its capacity to incorporate external regressors, such as promotional effects and stock management at the Stock Keeping Unit (SKU) level. According to Taylor and Letham (2017). Prophet outperforms conventional forecasting techniques regarding MAPE, especially in business applications where trend changes and numerous seasonalities are significant factors. Prophet's adaptability makes it a solid contender for estimating the demand for spare parts since it can more accurately simulate SKU-based inventory variations and outside factors. Furthermore, Prophet is a perfect model for optimizing inventory levels and enhancing forecasting dependability in the automobile industry since multi-stage variable selection techniques can further increase Prophet's accuracy by accounting for the influence of outside factors on demand for spare parts.

When dealing with high-dimensional data in demand forecasting, identifying the most relevant predictors while avoiding overfitting is a significant challenge. Ma, Fildes and Huang (2016) demonstrated that SKU-level demand forecasting accuracy can significantly improve by incorporating intra- and inter-category promotional information through a multi-stage LASSO regression approach. The analysis showed that intra-category data alone improved forecast accuracy by 12.6%, whereas adding inter-category information provided an extra 5% boost in accuracy.

These findings emphasize the importance of leveraging cross-category relationships in demand forecasting models. Similar methods can be utilized to assess the effects of promotions, seasonality, and SKU-level interactions when forecasting the demand for spare parts. A multi-step variable selection strategy, like the one employed by Ma et al. (2016), proves beneficial in handling high-dimensional data. It helps identify significant external demand factors while minimizing the risk of overfitting. Concentrating on the most relevant intra- and inter-category variables supports the optimization of inventory levels and helps prevent stock-out risks.

Additionally, the ability to manage large datasets and complex interactions makes such structured regression techniques well-suited for enhancing demand forecasting accuracy in dynamic environments. Future research could explore integrating deep learning-based approaches, such as Prophet, to further improve prediction performance by capturing complex seasonal dependencies and external influences in spare parts forecasting models.

Prophet has demonstrated strong capabilities in financial forecasting, particularly in handling complex time series data with seasonal patterns and trend fluctuations. Yurttabir and Kiymetli Sen (2021) applied the Prophet model to forecast the financial performance of 173 companies in the BIST Manufacturing Sector from 2009 to 2020. Their study evaluated the model's accuracy using MSE, RMSE, and MAPE. The results indicated that MSE values ranged from 0.0185 to 25.0147, RMSE values varied between 0.1361 and 5.0015, and MAPE values were measured between 0.1002 and 4.6634. While Prophet provided accurate forecasts for numerous companies, it displayed significant variations in other cases, highlighting inconsistencies in its predictive capability across different financial scenario conditions. These findings show that Prophet is a powerful financial forecasting tool that integrates external regressors and automatically recognizes structural changes in time series data. Its versatility suggests potential applications in various forecasting domains, including predicting spare parts demand. Since demand in the automotive sector is influenced by external shocks, SKU-level inventory fluctuations, and seasonal trends, Prophet's ability to incorporate multiple predictive variables may improve accuracy in the forecasting sector.

Additionally, evaluation metrics such as MSE, RMSE, and MAPE provide a structured approach for comparing Prophet with other forecasting models. In the context of spare parts demand forecasting, these metrics support model selection and enhancement, resulting in more accurate demand predictions and better inventory management. Prophet has shown strong performance in time series forecasting, especially in managing seasonal variations and external factors without extensive hyperparameter tuning. In a comparative analysis of oil production forecasting, Ning, Kazemi and Tahmasebi (2022) assessed the effectiveness of ARIMA, LSTM, and Prophet across various oil wells in the Denver-Julesburg Basin. Their findings reveal that Prophet effectively captured seasonal variations, such as reduced oil production in winter, a task that other models often struggled to recognize.

However, the study found that Prophet did not consistently outperform ARIMA and LSTM in all cases. While Prophet achieved a MAE of 1.34 and a RMSE of 2.32 in some scenarios, ARIMA and LSTM were often more robust for general oil production trends. The findings suggest that Prophet's ability to identify seasonality is particularly beneficial in cases where cyclical changes significantly influence demand.

Considering seasonality, SKU-specific inventory tactics, and promotions that influence spare parts demand forecasting, Prophet's ability to address these external factors is beneficial. Additionally, utilizing hybrid modeling methods such as merging Prophet with machine learning feature selection—can improve accuracy in forecasting. To properly assess Prophet alongside other forecasting models, using structured metrics like RMSE and MAE provides a reliable framework for selecting model improvement. Combining machine learning models, like Prophet and Long Short-Term Memory (LSTM) networks, has dramatically improved demand forecasting precision across multiple industries, including e-commerce. In their research, Sharma, Patel, and Gupta (2024) examined how LSTM and Prophet models perform in predicting online shopping demand trends, analyzing their respective strengths and weaknesses.

Their study found that LSTM excels at modeling short-term fluctuations by effectively capturing complex sequential dependencies, while Prophet performs well in handling long-term seasonality and external factors such as promotions and holidays.

The researchers proposed a hybrid ensemble model combining Prophet and LSTM, which reduced prediction errors by 15-20% compared to individual models. This improvement highlights the potential of leveraging deep learning-based sequence modeling and statistical time series decomposition to enhance forecasting accuracy.

These results indicate that analogous hybrid methods might enhance forecasting demand for automotive spare parts, as demand is affected by seasonality, SKU-specific inventory tactics, and external market influences. Prophet is adept at identifying seasonal trends and changes, which enhances its effectiveness in long-term inventory management. In contrast, LSTM is proficient in capturing short-term fluctuations, thus serving as a beneficial complement. Using structured performance evaluation metrics, such as RMSE and MAE, facilitates the optimization of hybrid models to produce results that are both more reliable and easier to interpret forecasts. Traditional forecasting methods often struggle to capture complex demand patterns, leading to inefficient replenishment strategies. Recent research highlights that LSTM networks can significantly improve demand forecasting accuracy, optimizing stock control strategies.

2.4 LSTM Literature Review

Nazarenko (2024) examined how LSTM can be utilized for short-term vehicle battery demand forecasting in an automotive manufacturing setting. The research assessed various machine learning models, such as LSTM, ARIMA, and XGBoost, for their demand prediction capabilities and influence on inventory management. LSTM showed enhanced predictive accuracy, especially in managing demand variations and lead time constraints.

Integrating LSTM-based demand forecasts into the inventory management system improved stock level optimization and refill strategies. This suggests that LSTM forecasting can significantly reduce stock imbalances, minimize excess inventory costs, and avoid stockouts, thereby aiding in enhancing supply chain resilience and reducing cost efficiency.

Effective inventory management is vital for matching supply with demand in the automotive sector. Integrating LSTM into inventory control systems can significantly enhance demand forecasting accuracy and guide replenishment choices. Considering the effects of fluctuating demand, lead times, and production schedules, machine learning-based forecasting solutions are promising for optimizing spare parts management.

LSTM-based predictive models have shown significant potential in improving the forecasting accuracy of spare parts inventory, particularly for large datasets with short replacement cycles. Liao, Ye, Yin, Yan, Ma and Zuo (2024) developed an LSTM-based inventory prediction method specifically designed for automotive spare parts with high demand variability and short replenishment periods. Their study demonstrated that adjusting the time step in the LSTM model affected the forecasting error rates for brake pad sales data, with error values ranging from 2.78% to 5.42%, depending on the chosen time step.

Compared to conventional statistical models, LSTM offers greater flexibility in recognizing intricate demand patterns and nonlinear dependencies, leading to more precise stock-level predictions. The research indicates that forecasting techniques utilizing LSTM are especially effective for automotive spare parts inventory management, where demand fluctuations and short replacement cycles pose challenges for traditional forecasting approaches.

Traditional statistical models like ARIMA have been extensively used for demand forecasting. However, their reliance on linear assumptions may limit their ability to capture complex temporal dependencies in spare parts demand effectively. In contrast, LSTM is well-suited for modeling long-term and nonlinear relationships, making it an essential tool for supply chain optimization. These benefits enhance inventory management, reduce surplus stock, and boost supply chain efficiency in the automotive industry.

Conventional forecasting models frequently fail to account for intricate demand changes in the automotive industry, resulting in ineffective spare parts inventory management. However, recent developments in deep learning forecasting techniques have shown enhanced accuracy over traditional statistical approach techniques.

Oukassi, Hasni and Layeb (2023) investigated the performance of LSTM networks in forecasting demand within the automotive manufacturing industry. They compared their effectiveness against the widely used AutoRegressive Integrated Moving Average (ARIMA) model. Their findings indicate that LSTM outperformed ARIMA in predicting spare parts demand, particularly in handling intermittent and fluctuating demand patterns.

Furthermore, the study implemented a sequence-to-sequence (Seq2Seq) architecture, which improved forecasting accuracy by better capturing long-term dependencies in demand data. The results suggest that LSTM-based forecasting techniques can be critical in optimizing inventory levels, reducing forecasting errors, and enhancing supply chain efficiency.

Considering the intricate demand fluctuations in the automotive industry, incorporating LSTM into inventory management systems may yield significant advantages. Its capacity to model nonlinear relationships and adjust to evolving demand patterns establishes it as an effective tool for enhancing replenishment strategies and reducing inventory issues and imbalances.

LSTM-based models have demonstrated high accuracy in short-term load forecasting, particularly in real-world energy systems. Abdulwahid and Haghifam (2024) tested LSTM for short-term load forecasting using actual data from the Wasit Thermal Power Plant in Iraq. The study evaluated the model's performance using various error metrics, demonstrating notably low forecasting errors. For Unit 1, the LSTM model yielded an RMSE of 0.0123, a MAPE of 2.749%, and an R^2 value of 0.7706. In contrast, Unit 2 produced an RMSE of 0.0079, a MAPE of 1.826%, and an R^2 value of 0.8470.

The results indicate that LSTM is a powerful forecasting method that effectively identifies intricate time-dependent patterns and reduces prediction errors. Its established effectiveness in forecasting energy loads implies that LSTM-based approaches could significantly enhance inventory management and demand forecasting in the automotive sector, where accurate short-term predictions are crucial due to fluctuating spare parts requirements—moreover, incorporating data preprocessing methods like feature selection and normalization further boost LSTM's predictive capabilities performance.

2.5 Attention Based LSTM Literature Review

Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gomez, Kaiser, and Polosukhin (2017) introduced the Transformer architecture, which replaced conventional RNNs with a self-attention mechanism in their seminal paper *"Attention Is All You Need."* This innovation revolutionized natural language processing (NLP) by enabling more effective parallelization, improved long-range dependency modeling, and enhanced sequence-to-sequence performance, particularly in machine translation tasks.

The core concept of self-attention allows the model to efficiently capture complex dependencies within sequential data, regardless of their spatial distance. Although initially designed for NLP applications, the attention mechanism has since been successfully applied to various domains, including time series forecasting. By enabling models to focus on the most relevant time steps in sequential data, integrating attention mechanisms into LSTM-based models has significantly improved forecasting performance.

Traditional LSTM models typically face challenges with long-term dependencies because of their sequential design, which can result in information loss as time progresses. In contrast, Attention-based LSTM models strategically allocate varying weights to past observations, enhancing the capture of significant patterns. This method has proven especially useful in demand forecasting, where balancing long-term trends with short-term fluctuations is crucial.

Attention-based LSTM models can significantly improve spare parts demand forecasting by focusing on the most pertinent historical data while minimizing the effects of irrelevant or minor inputs. Given the influence of supply chain disruptions, economic factors, and seasonal trends on the demand for automotive spare parts, integrating attention mechanisms allows forecasting models to adapt flexibly to these variations. Additionally, enhancing the accuracy of demand spike predictions enables attention layers to emphasize crucial timeframes, resulting in improved inventory management. This, in turn, reduces costs associated with overstocking and stockouts. This study investigates how attention-based LSTM models can further enhance demand forecasting methodologies in the automotive sector by combining the sequential learning capabilities of LSTMs with the adaptive weighting power of attention mechanisms.

LSTM networks and attention mechanisms have significantly improved demand forecasting accuracy across various industries, including ride-hailing services. Ye, Ye, Yan, Wang, Chen, and Li (2021) developed an attention-based LSTM model to predict short-term demand for online car-hailing services using temporal, geographical, and weather-related variables. Their study demonstrated that the LSTM + Attention model outperformed traditional approaches such as backpropagation neural networks (BPNN), gradient boosting decision trees (GBDT), RNN, and standard LSTM.

According to experimental results, the LSTM + Attention model achieved a Mean Squared Error (MSE) of 18.100 and a Mean Absolute Error (MAE) of 3.015 in 10-minute interval forecasting, outperforming LSTM (MSE: 21.841, MAE: 3.3026) and RNN (MSE: 22.298, MAE: 3.299). This shows that the model's capacity to adjust the significance of various temporal and spatial elements significantly improves its forecasting accuracy, establishing it as a dependable tool for demand prediction.

The LSTM + Attention model provides significant advantages for predicting spare parts demand in the automotive industry. It adeptly handles demand variations by focusing on critical time steps and spatial correlations, essential for inventory management. The demand for spare parts is affected by seasonal patterns, disruptions in the supply chain, and unforeseen failures. As a result, an LSTM model augmented with attention can minimize noise and emphasize crucial historical data points, leading to improved forecasting precision.

Additionally, incorporating external factors like weather, maintenance schedules, and market trends can significantly improve predictions, as evidenced by research on ride-hailing demand forecasting.

Assessing performance metrics such as MSE and MAE offers a solid benchmark for comparing the LSTM + Attention model with other forecasting approaches like Prophet and DeepAR, highlighting its efficiency in optimizing spare parts inventory management.

Integrating attention mechanisms with Long Short-Term Memory (LSTM) networks has significantly improved time series forecasting accuracy, mainly when dealing with complex, non-stationary data.

Abbasimehr and Paki (2021) proposed a hybrid model that combines LSTM with multi-head attention, enhancing prediction accuracy across multiple real-world time series datasets compared the proposed model against deep learning techniques such as standard LSTM and conventional forecasting models like ARIMA, ETS, and MLP.

The results demonstrated that in seven out of sixteen datasets, the attention-enhanced LSTM model achieved the lowest Symmetric Mean Absolute Percentage Error (SMAPE). Furthermore, when considering the overall ranking across all datasets, the Multi-Head Attention model achieved an Average Rank (AR) of 3.0625. In contrast, the proposed LSTM + Multi-Head Attention model improved the ranking to 2.375. In contrast, traditional statistical models such as ETS and ARIMA exhibited significantly poorer performance, with AR values of 6.875 and 6.78125, respectively. These improve the ability of attention-based LSTM models to allocate importance dynamically to critical time steps, enabling more accurate long-term forecasts.

Integrating attention mechanisms with LSTM models can substantially improve forecasting automotive spare parts demand.

Given the high volatility in spare parts consumption due to seasonality, supply chain disruptions, and fluctuating usage trends, an attention-enhanced LSTM model can effectively capture long-term dependencies while filtering out irrelevant variations. This feature is crucial for inventory management because it minimizes stock shortages and excess, resulting in more precise demand forecasting.

Additionally, structured evaluation metrics such as MAE and SMAPE provide a clear benchmark for comparing attention-based LSTM models with other forecasting techniques like Prophet and DeepAR. By applying these methodological enhancements, inventory managers can develop more adaptable and efficient stock management strategies, boosting supply chain resilience and optimizing expenses.

Attention mechanisms enhance prediction accuracy in time series forecasting by dynamically weighting essential time steps. By combining Dilated Convolutional Neural Networks (DCNN), Long Short-Term Memory (LSTM), Autoencoders (AE), and Attention Mechanisms (AM), Ji, Huang, Chen, Yin, Zuo, Chen and Bai (2023) presented a hybrid residential short-term load forecasting model. The study discovered that recognizing the oscillatory character of low-load data a problem that conventional algorithms frequently ignore increased forecasting accuracy.

Their model outperformed solo LSTM and traditional statistical techniques by achieving a MSE of 0.0041 and a MAE of 0.0333 at a 15-minute resolution. Similar advantages can be obtained for anticipating spare parts demand by including attention processes into LSTM. Supply chain interruptions, unforeseen malfunctions, and seasonal trends affect demand swings in the automobile sector. By removing noise and prioritizing the most significant historical data points, attention-based LSTM can produce more accurate forecasts. Additionally, integrating attention mechanisms with Seq2Seq structures might improve long-term dependence modeling, guaranteeing that abrupt changes in demand are appropriately represented. The research goals of this study are well-aligned with the findings of Ji, Huang, Chen, Yin, Zuo, Chen and Bai (2023), which indicate that such a method can be modified to improve inventory management and lower forecasting errors in spare parts demand prediction.

Time series forecasting has significantly improved with the advancement of deep learning models, particularly Transformer-based architectures and attention-enhanced recurrent networks.

In their comprehensive analysis of deep learning architectures for time series forecasting, Edirisooriya, Weliwatta, Amarasena and Ganesh (2024) evaluated multiple models, including LSTM, GRU, Transformer, and hybrid techniques. Their research demonstrates that attention-based models—such as Transformers and Attention-LSTM outperform traditional statistical methods like exponential smoothing and ARIMA in identifying complex, multi-faceted patterns and long-term dependencies.

The study's empirical results indicate that Transformer models reduce forecasting errors by approximately 18.7% compared to standard LSTMs. Moreover, hybrid models that combine LSTM and ARIMA enhance the MAPE by as much as 12.4% compared to LSTM alone. These improvements are significant in forecasting demand for spare parts, which is affected by seasonal changes, demand variability, and external influences that present considerable forecasting challenges.

By assigning varying weights to key time steps, attention-based LSTM models improve the Seq2Seq forecasting framework, enabling better capture of long-term dependencies and sudden demand changes.

Furthermore, hybrid models that merge statistical methods with deep learning like ARIMA-LSTM or Transformer-enhanced LSTM can enhance forecasting accuracy even more. These strategies support the goals of this study by refining inventory management and minimizing forecasting uncertainty in the automotive industry. Significant improvements in predictive accuracy have been observed when attention mechanisms are incorporated into time series forecasting, particularly in capturing complex feature dependencies. Ramachandran, Neergaard, Maier, and Bayer (2025) introduced a deep learning model integrating cross-attention blocks with Continuous Wavelet Transform (CWT) to enhance heat demand forecasting. Their study demonstrated that incorporating attention mechanisms improved the model's ability to capture critical temporal dependencies and enhance interpretability.

The model achieved a MAE of 0.105 ± 0.06 kWh and a MAPE of 5.4% ± 2.8%, outperforming the second-best model, which had an MAE of 0.10 ± 0.06 kWh and a MAPE of 5.6% ± 3%. Furthermore, their approach reduced the number of trainable parameters by 97% compared to conventional LSTM-based methods without sacrificing predictive accuracy.

These findings suggest that Attention-based LSTM models could improve forecasting spare parts demand by capturing critical fluctuations and filtering out extraneous variability. Given the influence of seasonality, supply chain disruptions, and unexpected failures in the automotive sector, employing advanced feature extraction techniques, such as cross-attention blocks and Continuous Wavelet Transform (CWT), could further refine demand trend analysis.

Furthermore, hybrid forecasting models that combine ARIMA, LSTM, and attention mechanisms can improve predictive accuracy by utilizing the flexibility of deep learning while preserving the statistical strength of traditional forecasting methods. These advanced forecasting methods provide a more dependable approach to managing inventory fluctuations, ultimately reducing costs and improving supply chain efficiency.

The accuracy of time series forecasting has significantly improved with the integration of attention mechanisms into deep learning models such as LSTM and Gated Recurrent Unit (GRU).

Yadav and Singh (2025) introduced an attention-based deep learning system for managing electric vehicle (EV) charging demand, utilizing LSTM and GRU architectures to capture time-dependent variations in demand. Their study found that attention-based LSTM outperformed traditional deep learning models, achieving an accuracy of 98.2%, compared to 90.2% for GRU and 96.4% for standard LSTM. The ability to dynamically assign importance to various input factors—such as location, time of day, and weather conditions—led to reduced forecasting errors and enhanced real-time load prediction findings are particularly relevant to spare parts demand forecasting, where external factors like supply chain disruptions, economic conditions, and seasonal trends significantly impact demand fluctuations. Attention-based models refine demand forecasting in EV charging networks by adapting to short-term and long-term fluctuations and enhance inventory planning for automotive spare parts. By integrating attention mechanisms into Seq2Seq LSTM models, the capacity to identify sudden demand changes and alleviate stock imbalances can be significantly improved.

Additionally, hybrid forecasting models that integrate ARIMA with LSTM or employ Transformer-based architectures can boost predictive accuracy by merging deep learning's flexibility with the statistical dependability of traditional forecasting techniques. These advanced approaches facilitate the development of more efficient inventory management systems, reducing stock shortages and preventing overstock in the automotive sector.

The increasing significance of deep learning methods, especially those that use attention processes, to handle the complexity and unpredictability of spare parts demand is highlighted by the corpus of research on demand forecasting for the automotive sector. In addition to better capturing non-linearities, seasonalities, and volatility than conventional statistical techniques, models like DeepAR, Prophet, LSTM, and Attention-based LSTM also show robustness to abrupt fluctuations and incomplete data. By methodically contrasting cutting-edge deep learning techniques with traditional statistical methods, this thesis seeks to expand on these findings. It focuses on how attention-based approaches, such as LSTM architecture, can increase predictive accuracy for non-stationary, seasonal, and volatile demand.

The focus on multi-dimensional and irregular data settings, characteristic of the automotive spare parts industry, and the investigation of hybrid methodologies that combine attention and probabilistic forecasting techniques make this work novel. This thesis aims to give practitioners practical insights into inventory planning optimization through a thorough experimental design and real-world data analysis, ultimately supporting more robust and cost-effective supply chain management.

Chapter 3

Methodology

3.1 Data Preparation

This study utilized spare parts order data from a Turkish automotive firm, which had been gathered since 2018. We carefully carried out the data preparation process to facilitate a fair comparison of time series forecasting models. The steps below outline the data preparation process.

Table 1

Sample From The Data Used For The Thesis

Part Number	Date	Order Quantity Daily	Part Definition
04L131547L	2018-01-03	1	CONTA
WHT009072	2024-08-24	7	SOMUN
3G5807417F GRU	2024-08-29	2	ARKA TAMPON
3V1941015C	2024-08-28	4	LED FAR
2Q0819669	2024-08-31	131	POLEN FİLTRESİ

3.1.1 Creation univariate time series. The date and order quantity daily variables from the dataset were used to construct the time series for each part:

$$Y_t = (y_1, y_2, \dots, y_t)$$

Where $\boldsymbol{Y_t}$ represents the order quantity at a given time point.

3.1.2 Removing outliers with hampel filter. Outliers can adversely affect the forecasting accuracy of time series models. We applied the Hampel filter to identify and adjust extreme values to mitigate this issue. This filter detects outliers by analyzing the median absolute deviation (MAD) Pearson (2002):

$$\text{MAD} = \text{median}(|Y_i - \text{median}(Y)|)$$

$$Y_i \text{ is an outlier if } |Y_i - \text{median}(Y)| > k \cdot \text{MAD}$$

Where **k** is typically set to 5. Outliers were replaced with the median value.

3.1.3 Resampling. Maintaining a consistent frequency in time series data is essential for forecasting models. Given that the raw data was collected daily, it was resampled to weekly and monthly intervals performed:

$$Y_t^{agg} = \sum_{i=1}^{n} Y_i$$

where Y_t^{agg} represents the aggregated values after resampling and Y_i denotes the original time series values. Resampling was done using mean and sum aggregation methods Hyndman and Athanasopoulos (2018).

In the context of the automotive spare parts industry, inventory management is often governed by a planning structure referred to as the "matrix approach". Under this concept, forecasts and procurement decisions are typically managed on a monthly or quarterly basis, as this cycle aligns better with operational and supply chain constraints.

Consequently, monthly resampling was deemed appropriate and 3-month ahead forecasts were produced to support this planning methodology.

This decision ensures that the developed forecasting models are not only methodologically sound but also practically applicable for real-world inventory management in the automotive sector.

3.1.4 Time series classification. To make model comparisons more meaningful, the time series were classified into four categories: stationary, non-stationary, seasonal, and non-seasonal.

3.1.4.1 Stationary test. The Augmented Dickey-Fuller (ADF) test was used to determine whether a time series is stationary Dickey and Fuller (1979). The test evaluates the null hypothesis H_0 that the time series has a unit root (non-stationary) against the alternative hypothesis H_1 that the series is stationary:

$$H_0\text{: The time series is non-stationary}$$

$$H_1\text{: The time series is stationary}$$

The test uses the following regression equation:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta Y_{t-i} + \varepsilon_t$$

Where γ must be statistically significant for the series to be considered stationary. Time series with a p-value below 0.05 were classified as stationary, while others were labeled non-stationary.

3.1.4.2 Seasonality test. We assessed the seasonal strength of each time series using seasonal decomposition Findley et al.,(1998). This approach separates a time series into its trend, seasonal, and residual parts. As a metric, we used the standard deviation of the seasonal component compared to the observed series:

$$S = \frac{\sigma_{\text{seasonal}}}{\sigma_{\text{observed}}}$$

If the seasonal strength S was greater than 0.2, the time series was classified as seasonal; otherwise, it was labeled non-seasonal.

Following these steps, each spare part's time series was classified into one of four categories: stationary, non-stationary, seasonal, or non-seasonal, making it ready for modeling.

3.1.5 Normalization. In order to guarantee that all time series data are on a comparable scale, we employed Min-Max scaling.

This transformation standardizes each time series within the range of 0 to 1, employing the subsequent formula:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

The normalized value is represented as the original value, while the minimum value and maximum value are defined within the series. This step guarantees that all models are provided with data within a standardized range, thereby enhancing model convergence and performance.

3.2 Hyperparameter Optimization

This study performed hyperparameter optimization on all models DeepAR, Prophet, LSTM, and Attention-based LSTM using Optuna, a state-of-the-art optimization framework. The effectiveness of time series forecasting models is heavily influenced by the chosen hyperparameters. Therefore, a proper tuning strategy is crucial for attaining the best predictive accuracy and stability across varying data types.

Optuna features a flexible optimization approach termed Define by Run, which allows the search space to be created dynamically and conditionally throughout the optimization process. This differs from traditional grid or random search techniques that necessitate pre-defined and rigid search structures. Optuna permits the adaptive adjustment of parameters and conditional logic based on previous choices made during the trial.

To assess each trial during optimization, a custom objective function was established for each model.

This function takes the sampled hyperparameters from the trial as input and produces a validation error, such as the Mean Squared Error, as output. Optuna aims to minimize this error over a user-defined number of trials trials.Formally, the optimization process can be expressed as:

$$\theta^* = \arg\min_{\theta \in \Theta} \mathcal{L}\left(f_\theta(X_{\text{train}}), Y_{\text{val}}\right)$$

where θ denotes the hyperparameter set, f_θ , which is the model parameterized by θ, and $\mathcal{L}$ and is the loss function calculated on the validation set.

Optuna employs the Tree-structured Parzen Estimator (TPE) algorithm to speed up convergence and lower computational expenses. This Bayesian optimization technique models the probability distributions of successful trials, allowing for intelligent guidance in future sampling.

Moreover, Optuna's pruning mechanism was employed to terminate unpromising trials early, based on intermediate validation scores. This early stopping strategy reduces the total computation time and allocates resources more effectively across promising configurations. The pruning condition can be formalized as:

$$\text{Stop trial } t \text{ if } \mathcal{L}_t^{(i)} > E\left[\mathcal{L}^{(i)} \mid \text{top-}k \text{ trials}\right], \quad \text{for some } i \in \{1, \dots, N_{\text{epochs}}\}$$

In addition to computational efficiency, Optuna also provides built-in visualization tools that assist in analyzing the sensitivity of model performance to individual hyperparameters and their interactions. This interpretability aspect not only improves model selection but also contributes to the transparency of the modeling process.

In summary, Optuna was selected for this study due to its:

- Efficiency: Faster convergence to optimal configurations using TPE and pruning,

- Flexibility: Dynamic and conditional search space construction suitable for diverse model architectures,
- Scalability: Capability to handle complex models such as Attention-based LSTM,

- Interpretability: Visualization and analysis support for understanding parameter effects,
- Reproducibility: Built-in mechanisms for saving and resuming optimization studies.

Integrating Optuna into the modeling pipeline allowed for the principled and automated fine-tuning of each forecasting model, ensuring a fair and optimized comparison of their predictive capabilities.

3.3 Deepar Model

The DeepAR model represents a probabilistic forecasting framework rooted in RNNs. It has been carefully crafted to effectively capture temporal dependencies and learn from a variety of interrelated time series. This model underwent training using historical data related to spare parts orders and was subsequently evaluated on an out-of-time (OOT) dataset for the purpose of assessing its forecasting capabilities and accuracy.

3.3.1 Data preprocessing for deepar model. The data was aggregated on a monthly basis to be consistent with the forecasting horizon. In automotive spare parts inventory management, planning is typically performed on a monthly basis, and forward-looking forecasts rarely exceed a 3-month window. Therefore, monthly aggregation was chosen to ensure that the forecasts align with real-world operational planning practices.

The dataset was organized as follows:

Table 2

Values For Which Dataset Is Prepared

Context Length	Prediction Length	Target Variable
14	3	Monthly aggregated order quantity

The time series for each spare part was transformed into a structured matrix format:

$$X_{\text{train}} = [Y_{t-14}, Y_{t-11}, \ldots, Y_{t-1}]$$

$$X_{\text{test}} = [Y_{t-14}, Y_{t-11}, \ldots, Y_{t-1}, Y_t, Y_{t+1}, Y_{t+2}]$$

Where X_{train} represents the training dataset with past observations and X_{test} includes both historical and future data points.

3.3.2 Model architecture and hyperparameter optimization. The DeepAR model consists of multiple RNN layers using LSTM units to capture sequential dependencies. In the context of time series, DeepAR models the conditional probability distribution as follows:

$$P\left(z_{i,t_0:T} \mid z_{i,1:t_0-1}, x_{i,1:T} \right) = \prod_{t=t_0}^{T} P\left(z_{i,t} \mid h_{i,t}, \theta_{i,t} \right)$$

where T represents the prediction length, $z_{i,t}$ denotes the observed data, and $h_{i,t}$ is the hidden state of the LSTM network.

DeepAR follows an autoregressive structure, where at each time step, the model estimates the probability distribution of the next value:

$$h_{i,t} = f\left(h_{i,t-1}, z_{i,t-1}, x_{i,t}; \Theta\right)$$

where $x_{i,t}$represents additional covariates, and Θ represents the learnable model parameters.

During training, the probability distribution used was either negative binomial or Gaussian, in line with the methodology described in the original paper. Unlike the Student-t distribution, which was previously considered, the negative binomial distribution was chosen due to its effectiveness in handling intermittent demand and heavy-tailed errors.

The loss function used for training was the negative log-likelihood (NLL)

$$\mathcal{L} = -\sum_{i=1}^{N} \sum_{t=t_0}^{T} \log P\left(z_{i,t} \mid h_{i,t}, \theta_{i,t} \right)$$

While the original DeepAR paper Salinas et al. (2020) employed grid search to optimize hyperparameters based on the negative log-likelihood on the validation dataset, this study adopted Optuna, benefiting from its advanced features such as dynamic search space construction, pruning, and TPE-based optimization, as detailed in Section 3.2.

The hyperparameter space was defined as follows:

Number of layers: $L \sim \{2, 3, 4, 5\}$

Number of cells per layer: $C \sim \{40, 60, 80, 100\}$

Dropout rate: $d \sim \mathcal{U}(0.1,\ 0.5)$

Learning rate: $\eta \sim \mathcal{U}(0.0005,\ 0.005)$

Batch size: $B \sim \{32, 64, 128\}$

Epoch: Fixed at 500

Optuna minimizes the objective function $f(\theta)$ using Tree-structured Parzen Estimator (TPE) optimization:

$$\theta^* = \arg\min_{\theta \in \Theta} f(\theta)$$

Where θ^* represents the optimal set of hyperparameters from the predefined search space Θ.

After running optuna for 100 trials**,** the best-performing configuration was identified as:

Table 3

Best Performing Parameters For Deepar

Number of Layers	Number of Cells per Layer	Drop out Rate	Learning Rate	Batch Size
4	42	0.2	0.004	64

The optimized model was then trained using these parameters, ensuring improved generalization and predictive performance compared to manual tuning.

3.4 Prophet Model

The Prophet model represents an additive regression-based time series forecasting framework developed by Facebook Taylor and Letham (2017). It is specifically designed to accommodate non-linear trends, including seasonality and holiday effects while exhibiting robustness to missing data and outliers. Prophet is constructed upon a generalized additive model (GAM) framework, which contributes to its high interpretability and flexibility for diverse forecasting applications scenarios.

3.4.1 Data preprocessing for prophet model

3.4.1 Data preprocessing for prophet model. The data was transformed into a format compatible with Prophet, where:

ds: Represents the date column.

y: Represents the target variable (monthly aggregated order quantity).

The dataset was structured as follows:

$$Y_t = T_t + S_t + H_t + \varepsilon_t$$

Where:

T_t represents the trend component, capturing long-term variations, and the seasonality component, capturing periodic fluctuations.

H_t denotes holiday effects, capturing special event-driven impacts.

ε_t is The error term accountinsor residual variance.

3.4.2 Model architecture and hyperparameter optimization. There are three main components for the Prophet model.

3.4.2.1 Trend component. Prophet models trend using a piecewise linear function with automatic changepoint detection. It assumes the trend follows a linear or logistic growth pattern:

Linear trend model:

$$T_t = kt + m$$

Where k is the growth rate, and is the initial offset.

Logistic growth model:

$$T_t = \frac{C}{1 + e^{-k(t-m)}}$$

Where C is the carrying capacity, k is the growth rate, and is the midpoint. Prophet automatically detects changepoints in trend by segmenting the time series and assigning different growth rates before and after the detected points.

3.4.2.2 Seasonality component. Prophet models seasonality using Fourier series expansion to capture periodic effects:

$$S_t = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi n t}{P}\right) + b_n \sin\left(\frac{2\pi n t}{P}\right) \right)$$

Where P is the period (e.g., 12 months for yearly seasonality), and a_n b_n are learned coefficients.

3.4.2.3 Holiday effects. Prophet facilitates the explicit modeling of holidays and special events by incorporating them as regressors within the model. Each holiday effect H_t is represented as an indicator function that adjusts the forecast accordingly.

$$H_t = \sum_i \kappa_i \cdot 1(t \in D_i)$$

where:

D_i is the set of dates associated with the holiday i.

K_i is the learned impact of the holiday on the forecast.

3.4.2.4 Hyperparameter optimization. To optimize Prophet's performance, hyperparameters were tuned using Optuna, an advanced hyperparameter optimization framework (Akiba et al., 2019). The search space included:

Changepoint prior scale: $CPS \sim \mathcal{U}0.01,\ 0.5$

Seasonality prior scale: $SPS \sim \mathcal{U}(0.01,\ 10)$

Interval width: $IW \sim \mathcal{U}(0.1,\ 0.9)$

The most effective configuration was selected after one hundred trials were executed using Optuna, ensuring maximum forecasting accuracy.

3.4.2.5 Advantages of prophet. Prophet offers several advantages for time series forecasting:

Automatic Changepoint Detection: Adjusts trend shifts dynamically.

Flexible Seasonality Modeling: Uses Fourier terms for different seasonal cycles.

Handles Missing Data: Works well with gaps in time series.

User-Defined Holiday Effects: Incorporates domain knowledge.

Interpretable Parameters: Easily explainable forecasting components.

3.5 Long Short-Term Memory Model

The Long Short-Term Memory (LSTM) network is a specific kind of RNN built to handle long-term dependencies in time series data Hochreiter and Schmidhuber (1997). LSTM effectively tackles the vanishing gradient problem faced by traditional RNNs, making it a strong choice for time series forecasting. In this research, we developed an LSTM-based model designed to forecast the monthly demand for spare parts.

3.5.1 Data preprocessing for lstm model. The dataset was preprocessed as follows:

Date column (ds): The timestamp of each observation.

Target variable (y): Monthly aggregated order quantity.

Time step selection: The model was trained using a rolling window of 14 months to predict the next month's demand.

The dataset was structured as follows:

$$\boldsymbol{X_{train}} = \left[Y_{\{t-14\}}, Y_{\{t-11\}}, \dots, Y_{\{t-1\}}\right]$$

$$\boldsymbol{X_{test}} = \left[Y_{\{t-14\}}, Y_{\{t-11\}}, \dots, Y_{\{t\}}\right]$$

Where X_{train} represents the training dataset with past observations, and X_{test} includes both historical and predicted data points.

3.5.2 Model architecture and hyperparameter optimization. The LSTM model consists of stacked LSTM layers**,** allowing it to capture temporal dependencies effectively. The architecture includes:

Input layer: Takes sequences of 14 months as input.
LSTM layers: Stacked LSTM units to capture long-term dependencies.
Fully connected layer: Outputs the final predicted order quantity.
Activation function: $tanh$ for hidden layers, linear for output.

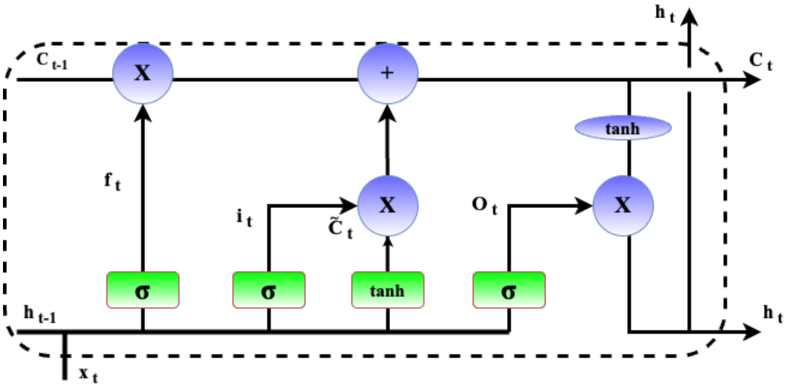

Figure 1. LSTM architecture.

The forward pass of an LSTM cell is formulated as follows:

Forget Gate:

$$f_t = \sigma\left(W_f \cdot \left[h_{\{t-1\}}, x_t\right] + b_f\right)$$

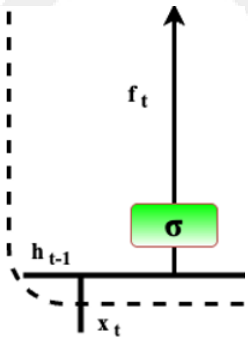

Figure 2. LSTM forget gate.

Input Gate:

$$i_t = \sigma\left(W_i \cdot \left[h_{\{t-1\}}, x_t\right] + b_i\right)$$

$$\widetilde{C}_t = tanh\left(W_C \cdot \left[h_{\{t-1\}}, x_t\right] + b_C\right)$$

$$C_t = f_t * C_{\{t-1\}} + i_t * \widetilde{C}_t$$

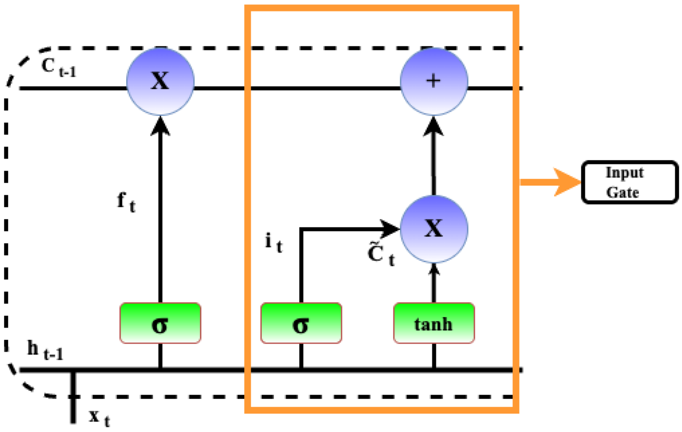

Figure 3. LSTM input gate.

Output Gate:

$$o_t = \sigma\left(W_o * \left[h_{\{t-1\}}, x_t\right] + b_o\right)$$
$$h_t = o_t * tanh(C_t)$$

Where:

X_t is the input at the time step .

h_t is the hidden state.

C_t is the cell state.

W_f, W_i, W_c, W_o and b_f, b_i, b_c, b_o are the trainable weight matrices and bias terms.

σ represents the sigmoid activation function.

$tanh$ is the hyperbolic tangent activation function.

3.5.2.1 Hyperparameter optimization. In this study, hyperparameter optimization was carried out using Optuna, a state-of-the-art automatic hyperparameter tuning framework.

Optuna, allowing dynamic construction of the search space and efficient exploration through advanced sampling algorithms such as Tree-structured Parzen Estimator (TPE).

Instead of manually testing predefined hyperparameter combinations, Optuna was employed to automate and optimize the search for the best model configuration based on validation performance. The objective function minimized the validation loss (e.g., Mean Absolute Error), and the search was time constrained, ensuring the optimization process remained practical for the experiment duration.

Optuna's efficient sampling strategy significantly reduced the time and computational cost of finding optimal parameters, while avoiding overfitting.

The best hyperparameter set discovered through this process is shown below:

Table 4

Best Performing Parameters For Lstm

Number of Layers	Number of Units per Layer	Dropout Rate	Learning Rate	Batch Size
2	128	0.3	0.001	32

The LSTM model trained with these parameters exhibited improved predictive accuracy and generalization on unseen data compared to models tuned via manual search. Optuna's integration contributed to faster development cycles and more reliable results, making it a robust choice for hyperparameter tuning in deep learning-based forecasting tasks.

3.6 Attention Based LSTM

The Attention-Based LSTM Model integrates Long Short-Term Memory (LSTM) networks with the Multi-Head Attention (MHA) mechanism to enhance forecasting accuracy in time series prediction. While LSTMs are effective at capturing sequential dependencies, they often struggle with long-range dependencies due to vanishing gradients Hochreiter and Schmidhuber (1997). The Attention Mechanism, originally introduced by Bahdanau, Cho, and Bengio (2015), enables models to dynamically focus on the most relevant time steps, overcoming this limitation.

3.6.1 Multi head attention mechanism. The Attention Mechanism has significantly improved deep learning applications by enabling models to selectively weigh different input time steps based on their relevance to the prediction task Bahdanau et al. (2015); Vaswani et al. (2017).

Several types of attention mechanisms exist:

Global Attention: Considers all prior time steps with varying importance.

Local Attention: Focuses on a subset of past observations.

Self-Attention: Evaluates dependencies within a sequence without requiring external inputs.

MHA, introduced by Vaswani et al. (2017) in the Transformer architecture, extends Self-Attention by employing multiple attention heads. This allows the model to capture diverse perspectives on the input sequence, improving representational power.

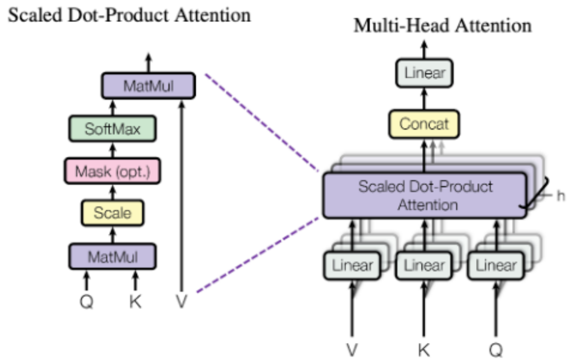

Figure 4. Scaled dot-product and mha (Vaswani et al. (2017)).

3.6.1.1 Scaled dot-product attention. The Scaled Dot-Product Attention mechanism, forming the foundation of MHA, computes attention scores using three learned matrices:

- Query (Q): Represents the embedding of the current time step.
- Key (K): Represents past time steps in the sequence.
- Value (V): Contains relevant information from past observations.

The attention scores are computed using the Scaled Dot-Product Attention formula:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)$$

where d_k is the dimensionality of the key vectors used for scaling, preventing excessively large dot-product values.

3.6.1.2 Multi-head attention in time series forecasting. Instead of relying on a single attention head, MHA projects Q, K, and V into multiple subspaces, each processed independently and subsequently concatenated:

$$MHA(Q,K,V) = Concat(head_1, head_2, \ldots, head_h)W^O$$

where each attention head is computed as:

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)$$

The final output is a linear transformation of the concatenated outputs from all heads.

3.6.2 Model architecture and hyperparameter optimization. The proposed model integrates a LSTM layer with MHA to capture both short-term and long-range dependencies in time series data.

The LSTM layer extracts temporal dependencies, while MHA improves forecasting accuracy by assigning different importance scores to different time steps.

The architecture consists of:

Input Layer: A time series window of 14 time steps is fed into the model.

LSTM Layer: A 128-unit LSTM processes sequential dependencies and outputs hidden states.

MHA Layer: 7 attention heads capture long-range dependencies.

Dense Layer: A fully connected layer refines the extracted features.

Output Layer: The final forecast is produced.

The final model formulation is:

$$H_t = \mathrm{LSTM}(X)$$

$$C_t = MHA(H_t, H_t, H_t)$$

$$\hat{y_t} = \sigma(WC_t + b)$$

where:

H_t represents LSTM hidden states.

C_t is the attention-weighted context vector.

W and b are learnable parameters.

σ is the activation function.

To enhance the model's performance, hyperparameter tuning was conducted using Optuna. This framework allows efficient exploration of the hyperparameter space via advanced sampling strategies while being dynamically adaptable. The optimization was performed under a time constraint, ensuring practical training durations and avoiding unnecessary computational overhead.

The parameters that give the best results are as follows:

Table 5

Best Performing Parameters For Attention Based Lstm

LSTM Units	Attention Heads	Sequence Length	Learning Rate
128	7	14	0.0001

LSTM Units	Batch Size	Epoch
128	16	500

To prevent overfitting, Early Stopping was applied during training. Thanks to Optuna's hyperparameter optimization capabilities, the final model demonstrated improved generalization across stationary, non-stationary, seasonal and non-seasonal time series patterns.

3.7 Evaluation Metrics

To evaluate the performance of the forecasting models, three commonly used error metrics are employed**:** MAE**,** Symmetric sMAPE**,** and RMSE**.**

3.7.1 Mean absolute error. MAE measures the average magnitude of errors in the predictions, providing a straightforward interpretation of model performance:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}|$$

where:

y_i is the actual value,

$\widehat{y_i}$ is the predicted value,

n is the number of observations.

MAE measures the average magnitude of the errors without considering their direction. In the context of this thesis, MAE is particularly important for evaluating forecasting models in a direct and interpretable way. Since the dataset includes various spare parts with different order quantities, MAE helps understand how much deviation can be expected from the actual demand in real units, which is essential for inventory planning in the automotive industry.

3.7.2 Root mean squered error. RMSE penalizes larger errors more heavily, making it a useful metric when large deviations are critical:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}$$

where:

y_i is the actual value,

$\widehat{y_i}$ is the predicted value.

RMSE is sensitive to significant errors, as it penalizes them more heavily than MAE. This characteristic makes RMSE a valuable metric in this study, especially for identifying whether a model makes critical mistakes in forecasting peak or unexpected demand values.

In the spare parts context, sudden spikes in demand can have a significant operational impact, and RMSE helps assess how well models handle such situations.

3.7.3 Symmetric mean absolute perentage error. sMAPE is a scale-independent error metric that provides a percentage-based error measure:

$$\text{sMAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2} \times 100$$

where:

y_i is the actual value,

$\hat{y}_i$ is the predicted value.

This metric is particularly useful when dealing with datasets that contain a mix of large and small values.

sMAPE is a scale-independent error metric that expresses error as a percentage, making it especially suitable for comparing models across different spare parts with varying demand scales.

In this thesis, where the time series data contains both low-volume and high-volume parts, sMAPE ensures that models are not biased toward parts with larger quantities. It is also valuable for evaluating performance on non-stationary and seasonal data, where relative fluctuations matter more than absolute values.

These metrics together provide a comprehensive evaluation of model performance, balancing absolute error, relative percentage-based error, and penalization of large deviations.

Chapter 4

Findings

This chapter provides an in-depth evaluation of four advanced forecasting models DeepAR, Prophet, LSTM, and Attention-based LSTM utilized on real-world spare part demand data from a prominent Turkish automotive company. The dataset covers multiple years and captures diverse demand patterns affected by seasonality, market fluctuations, operational shifts, and unpredictable procurement behaviors. To ensure a thorough and context-aware comparison, the time series were initially preprocessed and then classified into four distinct structural categories: seasonal, non-seasonal, stationary, and non-stationary. This classification facilitated a more detailed analysis of each model's ability to adapt to different statistical characteristics, from predictable cyclical variations to erratic, high-variance patterns.

The forecasting ability of each model was evaluated using three commonly used error metrics: MAE, which shows the average size of errors in forecasts on a scale; RMSE, which disproportionately penalizes larger errors and is especially responsive to outliers; and sMAPE, a normalized measure that facilitates comparisons across series with varied scales by presenting forecast error as a relative percentage of the average absolute values of actual and predicted outcomes demand.

This chapter analyzes model performance across various time series categories and metrics to identify the model that produces the most accurate results. It also explores the reasons behind the superior performance of certain models under specific data conditions. This analysis offers valuable insights into the effectiveness of classical statistical models compared to deep learning architectures in the automotive aftersales sector. The findings form a basis for further discussions on hybrid modeling strategies, operational integration, and future development in demand forecasting systems.

4.1 Result by Time Series Datasets

4.1.1 Stationary time series. Stationary time series are characterized by a constant mean and variance over time, representing spare parts with stable and predictable demand patterns, such as components required for routine maintenance activities. In such series, forecasting models are primarily expected to capture consistent underlying trends while minimizing sensitivity to random noise or minor fluctuations.

In this context, Prophet demonstrated strong performance, achieving the lowest MAE (12.58) and a competitive sMAPE (24.15). Prophet's additive model structure, which combines stable trend components with fixed seasonalities modeled via Fourier series, is particularly well-suited to stationary series where the underlying data-generating process exhibits minimal structural change. Moreover, its ability to model changepoints only when necessary prevents overfitting to random fluctuations, preserving the stability of the forecasts. These characteristics explain why Prophet performed well in this setting, providing accurate baseline predictions without introducing unnecessary complexity. DeepAR, while inherently powerful in dynamic environments through its recurrent LSTM encoder and probabilistic forecasting structure, showed slightly inferior performance in stationary series (sMAPE of 22.83, RMSE of 30.77). The model's strength in modeling conditional probability distributions across related series introduced broader predictive intervals, which, while beneficial for heterogeneous or volatile data, led to less precise forecasts in a context where stability predominated. These findings address the first research question (*How do DeepAR and Prophet models perform in forecasting automotive spare parts demand?*), indicating that Prophet's structural simplicity and resistance to overfitting offered a better fit for stationary demand patterns compared to DeepAR's uncertainty-emphasizing architecture.

Attention-based LSTM achieved the best overall performance among all models for stationary time series, recording the lowest RMSE (16.30) and a competitive sMAPE (22.16). Its architecture integrates a MultiHead Self-Attention mechanism over the standard LSTM layers, enabling the model to selectively focus on stable, repetitive input patterns while ignoring irrelevant historical noise.

In stationary series, where a few critical historical windows dominate future behavior, this selective attention mechanism allowed the model to concentrate its predictive power on consistently informative periods, enhancing forecast stability and reducing error. In contrast, the standard LSTM model, which processes historical sequences uniformly without differentiated weighting, exhibited the weakest performance with a sMAPE of 33.96 and RMSE of 23.78. Without an explicit mechanism to prioritize relevant over irrelevant inputs, the LSTM model tended to overfit transient, low-importance fluctuations, diluting its ability to model the truly stable baseline behavior inherent to stationary series. This comparative analysis between the attention-augmented and standard LSTM architectures directly addresses the second research question (How does the attention mechanism impact time series forecasting in the automotive sector?), demonstrating that attention significantly enhances a model's ability to isolate and reinforce meaningful stability in demand patterns.

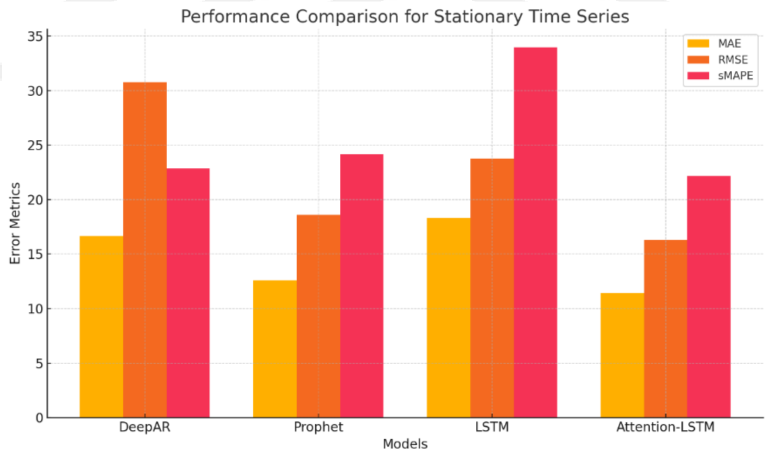

Figure 5. Performance comparison for stationary time series.

The bar plot illustrates a comparison of forecasting models using three error metrics: MAE, RMSE, and sMAPE, aimed specifically at stationary time series. The Attention-based LSTM model demonstrated the best overall performance, achieving the lowest values for both MAE and RMSE, which highlights its effectiveness in stable demand conditions due to its focus on the most relevant past observations.

Additionally, Prophet, known for its robust trend modeling abilities, performed well in terms of MAE. Meanwhile, while the DeepAR and LSTM models are powerful for dynamic series, they showed comparatively higher error rates in stationary contexts, emphasizing the critical role of model selection based on the time series's statistical characteristics.

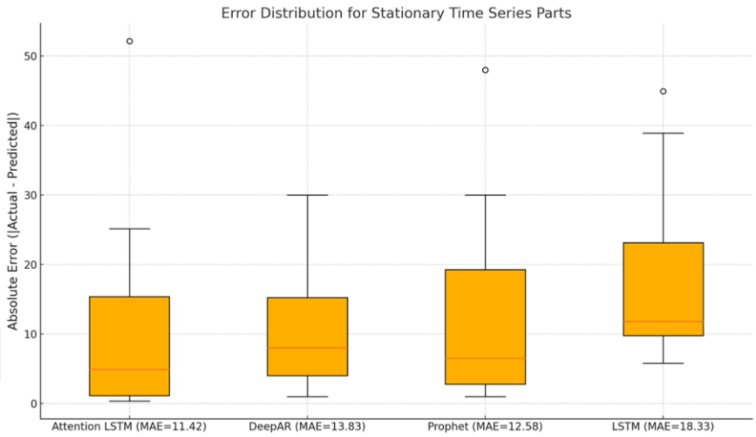

Figure 6. Error distribution for stationary time series parts.

The box plots show that the Attention LSTM model achieved the lowest median error and maintained the narrowest interquartile range, indicating more consistent and stable forecasts across various stationary series. In comparison, Prophet and DeepAR exhibited slightly wider distributions, suggesting a moderate level of variability in their predictions. Conversely, the LSTM model resulted in a wider error spread, highlighting greater inconsistency and a higher likelihood of significant deviations from actual values.

The analysis of outliers indicates that although all models sometimes generated extreme errors, Attention LSTM and Prophet exhibited significantly fewer and less severe outliers. In contrast, the LSTM model showed larger outliers, revealing its susceptibility to considerable forecasting errors, even in stationary time series.

This trend emphasizes the significance of advanced architectures, such as the attention mechanism, in attaining greater accuracy and enhanced performance in time series forecasting.

4.1.2 Non-stationary time series. Non-stationary time series are characterized by evolving mean levels, changing variances, or structural breaks over time. In the context of automotive spare parts demand, such series often reflect new model introductions, regulatory shifts, supply chain disruptions, or changing customer behaviors. Forecasting non-stationary demand requires models adapting dynamically to these structural changes rather than assuming long-term stability.

Within this group, DeepAR achieved the lowest sMAPE (13.08) among all models, indicating strong relative forecasting accuracy. DeepAR's autoregressive recurrent structure, built upon LSTM encoders, models the conditional distribution of future values given past observations. Its key advantage in non-stationary series arises from its global training approach, where it simultaneously learns from multiple related time series, allowing the model to generalize underlying patterns and absorb varying trends across series. This global perspective enhances its ability to adapt to trend shifts and regime changes, common in non-stationary data. However, while DeepAR captured relative patterns well, its RMSE was relatively high (190.19), suggesting that large absolute deviations still occurred, particularly around abrupt, unobserved shifts. Prophet also performed competitively, achieving a sMAPE of 13.14, but exhibited higher MAE (155.67) and RMSE (307.55) compared to DeepAR. Prophet's modeling of trends through piecewise linear or logistic growth with automatic changepoint detection enables it to handle gradual trend shifts, but its static changepoint structure can lag behind sudden and nonlinear trend evolutions typical of non-stationary environments.

These observations directly address the first research question (How do DeepAR and Prophet models perform in forecasting automotive spare parts demand?), indicating that DeepAR's global sequence modeling architecture offers superior adaptability in dynamic conditions, whereas Prophet's performance is constrained by its fixed changepoint assumptions.

The Attention based LSTM model recorded a moderate sMAPE (17.15) but suffered from high absolute errors (MAE: 217.33, RMSE: 480.80) in non-stationary series.

Its MultiHead Self-Attention mechanism, which selectively emphasizes relevant past observations, offers advantages in moderately dynamic environments; however, when faced with sharp structural breaks not adequately represented during training, the model's ability to focus can become misdirected. Without explicit changepoint modeling or external trend adjustment, the attention mechanism may continue emphasizing outdated historical patterns, resulting in substantial forecast deviations. Conversely, the classical LSTM model exhibited the weakest performance, with a sMAPE of 30.32 and RMSE of 494.02. Traditional LSTM models treat all past inputs with equal importance across the sequence, making them inherently less responsive to sudden shifts or emerging patterns. In highly non-stationary series, where historical relevance fluctuates significantly, the absence of dynamic attention or structural trend detection mechanisms critically limits LSTM's adaptability. This comparative analysis addresses the second research question (How does the attention mechanism impact time series forecasting in the automotive sector?), demonstrating that while attention mechanisms can enhance sequential models' responsiveness to moderate variations, their effectiveness diminishes in environments dominated by abrupt, large-scale structural changes unless complemented with explicit trend adaptation strategies.

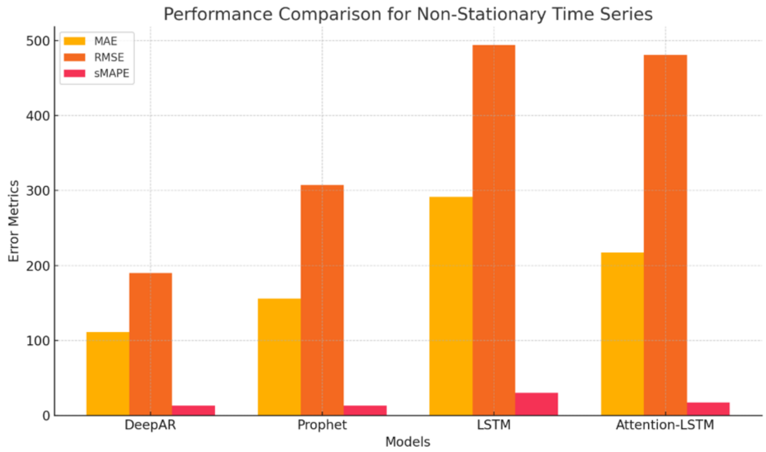

Figure 7. Performance comparison for non-stationary time series.

The bar plot presents a comparison of forecasting models' performance for non-stationary time series using MAE, RMSE, and sMAPE metrics. DeepAR recorded the lowest sMAPE, showcasing its superior adaptability to changing demand patterns. In contrast, Prophet displayed a similar relative error but had greater absolute deviations. Both LSTM and Attention based LSTM models experienced significant performance degradation, emphasizing the difficulties recurrent architectures face even when bolstered by attention mechanisms in accurately predicting demand that shifts structurally.

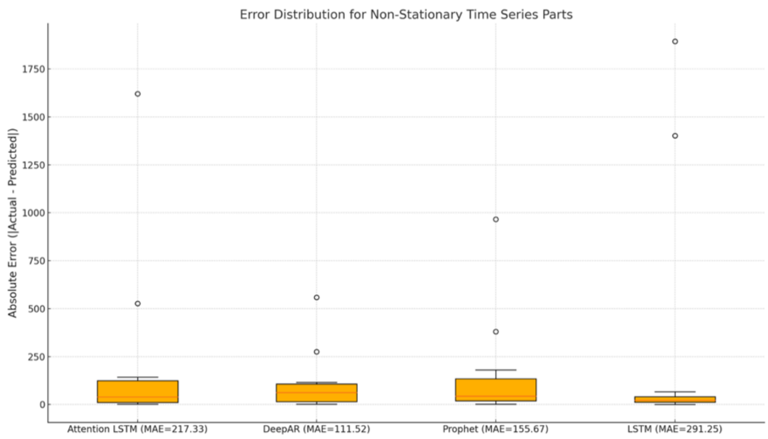

Figure 8. Error distribution for non-stationary time series parts.

The box plot shows that DeepAR achieved the lowest median errors and a narrower interquartile range (IQR), reflecting its strong and consistent predictions, even with the complexity of non-stationary series. Prophet maintained fair stability, though its error spread was wider relative to DeepAR. Conversely, the Attention LSTM and particularly the standard LSTM models demonstrated much greater variability, with LSTM displaying the highest degree of variability and the highest frequency of extreme outliers.

These results highlight the greater challenge that non-stationarity poses to traditional recurrent models like LSTM. Without mechanisms such as attention layers or probabilistic forecasting structures, standard LSTM struggles to capture dynamic trends and abrupt shifts typical of non-stationary series. Meanwhile, models like DeepAR, which are explicitly designed to handle probabilistic patterns and changing dynamics, outperform their deterministic counterparts. This comparison underscores the critical importance of model architecture choice when dealing with highly volatile and non-stationary demand patterns.

4.1.3 Seasonal time series. Seasonal time series are characterized by regular, repeating patterns or cycles at fixed intervals, typically driven by external factors such as weather conditions, service schedules, or customer behavior aligned with calendar periods. In automotive spare parts demand, seasonal patterns often emerge from factors like scheduled maintenance cycles, tire changes due to weather, or periodic service promotions. Effective forecasting in such series requires models that can accurately capture and align with these recurring structures.

Within this group, Prophet demonstrated competitive performance with a sMAPE of 18.60 and the lowest RMSE (6.525) among non-deep learning models. Prophet's architecture is inherently designed to handle seasonality through the decomposition of time series into trend, seasonality, and holiday effects. It models seasonality using Fourier series expansions, allowing it to flexibly fit periodic patterns with varying amplitudes and frequencies.

This explicit modeling of multiple seasonal components gives Prophet a natural advantage in environments where recurring behavior dominates, enabling it to achieve high accuracy in the seasonal spare parts demand data. DeepAR, despite being a global probabilistic model, achieved a higher sMAPE (19.82) compared to Prophet. While DeepAR's recurrent structure allows it to learn implicit seasonality across multiple series, the absence of an explicit seasonality component can lead to diminished precision when seasonal effects vary subtly between different series.

These observations contribute to addressing the first research question (*How do DeepAR and Prophet models perform in forecasting automotive spare parts demand?*), showing that explicit seasonality modeling in Prophet provides a distinct advantage over DeepAR's implicit pattern recognition when dealing with strongly seasonal time series.

Attention based LSTM achieved the best overall performance among all models for seasonal time series, with a sMAPE of 13.38, MAE of 6.166, and RMSE of 12.443.

By integrating a Multi Head Self-Attention mechanism on top of LSTM layers, the model dynamically identifies and prioritizes the most informative periods from past cycles, effectively learning the temporal structure of repeated seasonal patterns.

This selective focus allows the Attention-based LSTM to align its forecasting more closely with recurring demand peaks and troughs, thereby enhancing prediction accuracy beyond what can be achieved by standard recurrent architectures. In contrast, the classical LSTM model achieved a sMAPE of 16.80 and a RMSE of 14.733. Although capable of learning long-term dependencies through gated memory mechanisms, standard LSTM treats all historical inputs uniformly, which reduces its efficiency in emphasizing key points from prior seasonal cycles. Without an explicit attention mechanism, LSTM may underemphasize critical recurring patterns, leading to less precise forecasts.

This analysis addresses the second research question (How does the attention mechanism impact time series forecasting in the automotive sector?), illustrating that attention mechanisms significantly enhance a model's ability to learn, reinforce, and align with cyclical behaviors essential for accurately forecasting seasonal spare parts demand.

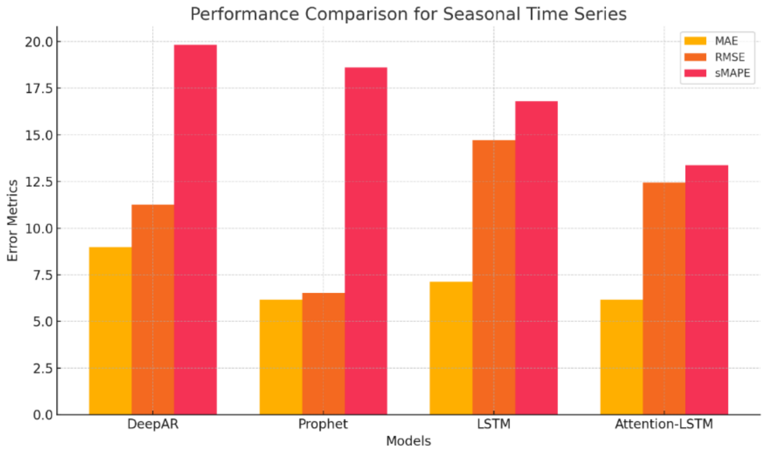

Figure 9. Performance comparison for seasonal time series.

The bar plot displays the performance comparison of forecasting models for seasonal time series using MAE, RMSE, and sMAPE metrics. The Attention-based LSTM model attained the highest overall accuracy by adeptly aligning with repeated seasonal patterns through dynamic attention mechanisms. Prophet, which utilizes explicit seasonality modeling via Fourier series, also showed strong results, especially in reducing absolute error. DeepAR and LSTM, although capable, showed shortcomings in completely capturing the cyclical patterns characteristic of seasonal spare parts demand.

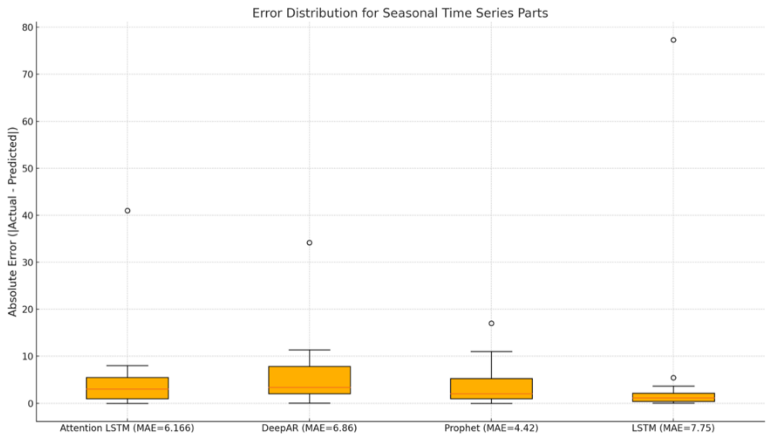

Figure 10. Error distribution for seasonal time series parts.

The box plot indicates that Prophet not only had the smallest median error but also displayed the narrowest interquartile range, underscoring its strong capacity to capture the data's seasonal patterns. Attention-based LSTM and DeepAR models performed reasonably well as well, with similar median errors and comparable variability, although their distributions were slightly wider than Prophet's. In comparison, the LSTM model showed a larger error spread, indicating higher variability and more moderate deviations from the actual values.

These findings indicate that models tailored for temporal structures, especially those that include trend and seasonality features like Prophet, are more effective in forecasting seasonal time series. Recurrent architectures such as Attention based LSTM and DeepAR also benefit from their sequence modeling capabilities, but are slightly less precise than Prophet. The comparatively weaker performance of LSTM highlights the limitations of relying solely on memory-based learning without explicit mechanisms to model periodic fluctuations.

4.1.4 Non-seasonal time series. Non-seasonal time series are characterized by the absence of recurring patterns or predictable cyclicality, often dominated by irregular demand fluctuations, one-off events, or sporadic procurement behaviors.

In the automotive spare parts sector, non-seasonal demand can arise from policy-driven purchasing decisions, unexpected repairs, or erratic consumer behavior. Forecasting such series is inherently challenging, as models must infer meaningful signals from data that exhibits low temporal regularity and high noise levels.

Within this group, Prophet achieved the lowest sMAPE (25.82) among all models, indicating relatively strong performance despite the lack of structured seasonality. Prophet's robustness stems from its ability to flexibly model trends using piecewise linear or logistic growth functions without relying heavily on seasonal components. In the absence of periodic structure, Prophet's focus on trend dynamics allowed it to provide a reasonable approximation of underlying demand movements without overfitting to noise. DeepAR, in contrast, produced a higher sMAPE (39.73), reflecting difficulty in generalizing across sparsely patterned series. While DeepAR's global sequence modeling and probabilistic framework are advantageous when latent structures exist across multiple related series, in highly irregular, individualistic demand profiles, its reliance on learned inter-series similarities becomes less effective. These findings contribute to answering the first research question (*How do DeepAR and Prophet models perform in forecasting automotive spare parts demand?*), demonstrating that Prophet's trend-centered modeling approach offers resilience against noise-dominated, non-seasonal data, whereas DeepAR's dependency on structural similarities limits its accuracy when such consistencies are absent.

Attention-based LSTM outperformed other models in terms of balancing relative and absolute errors, achieving a sMAPE of 30.00, MAE of 21.83, and RMSE of 31.46. The model's MultiHead Self-Attention mechanism enabled selective emphasis on informative past observations, even when clear cyclical structures were lacking. By learning to prioritize sporadic but meaningful historical signals while suppressing background noise, the Attention-based LSTM achieved greater forecast stability in the face of irregular demand fluctuations.

On the other hand, the standard LSTM model exhibited the weakest performance, with a sMAPE of 43.98 and a substantially higher RMSE.

Without a mechanism to differentiate between relevant and irrelevant past information, LSTM struggled to extract stable predictive signals from highly variable input sequences, leading to overfitting to random fluctuations and degraded forecast accuracy. This comparative analysis addresses the second research question (How does the attention mechanism impact time series forecasting in the automotive sector?), illustrating that attention mechanisms significantly enhance a model's resilience to noisy and irregular patterns, enabling more robust forecasting even when the data lacks strong temporal coherence.

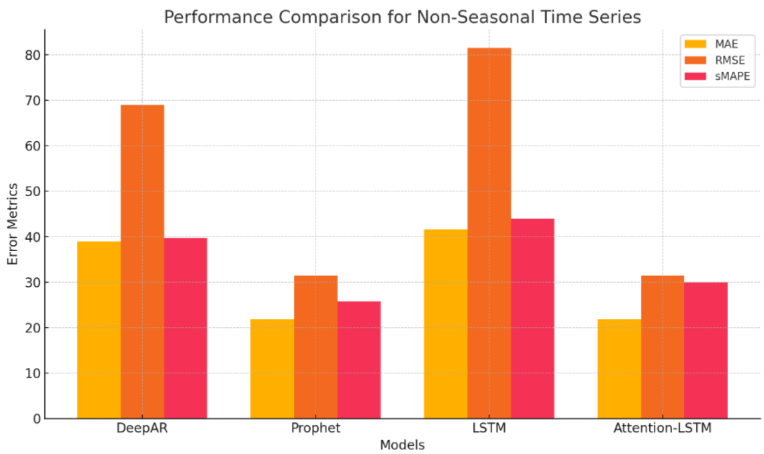

Figure 11. Performance comparison for non-seasonal time series.

The bar plot shows the forecasting performance of various models for non-seasonal time series, evaluated using MAE, RMSE, and sMAPE metrics. Prophet outperformed others by adeptly capturing trend dynamics, even without recurring seasonal patterns. The attention-based LSTM, by concentrating on key historical signals, achieved competitive forecasting accuracy. In contrast, DeepAR and LSTM faced significant performance declines due to their dependence on implicit pattern learning in contexts of highly irregular demand.

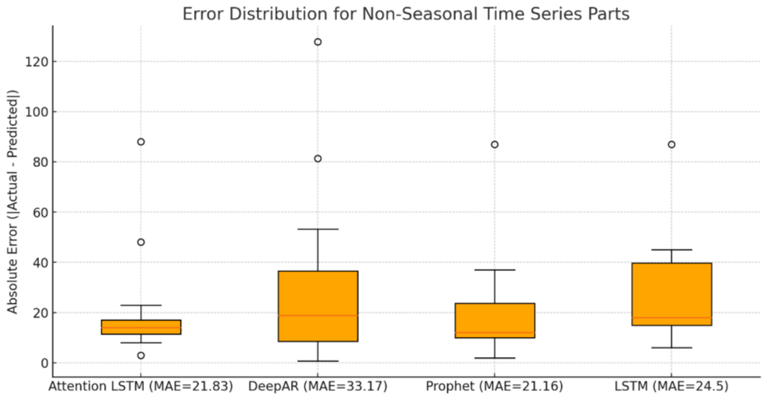

Figure 12. Error distribution for non-seasonal time series parts.

The boxplot shows the absolute error distributions of non-seasonal time series components across four models: Attention LSTM, DeepAR, Prophet, and LSTM. Attention LSTM records the lowest MAE of 21.83, closely followed by Prophet at 21.16, DeepAR at 33.17, and LSTM at 24.50. Although all models show some degree of outliers, the majority of the error values—represented by the interquartile range—are relatively compact, indicating stable predictions for most components. Prophet reveals a slightly more condensed distribution, implying it achieved more consistent predictive performance with non-seasonal data. Conversely, DeepAR shows a broader spread and more extreme outliers, highlighting greater variability in its predictions for certain components. This distributional analysis highlights that for non-seasonal patterns, Prophet and Attention LSTM models are better suited due to their lower central errors and narrower variability, whereas DeepAR exhibits higher error volatility across parts.

4.2 Overall Performance Summary

This section consolidates the forecasting performances of DeepAR, Prophet, LSTM, and Attention-based LSTM across all time series categories — stationary, non-stationary, seasonal, and non-seasonal — to determine which model consistently delivers the best results under varying demand patterns.

The Attention-based LSTM model consistently achieved superior performance across most categories, particularly excelling in stationary, seasonal, and non-seasonal series. By incorporating a MultiHead Self-Attention mechanism atop traditional LSTM layers, the model dynamically prioritized the most informative temporal segments, thereby enhancing its ability to capture stable, recurring, or sparse structures depending on the series characteristics. Its flexibility in emphasizing relevant past observations allowed it to deliver the lowest RMSE and competitive sMAPE values in three out of the four categories, demonstrating strong adaptability across diverse forecasting contexts.

Prophet also demonstrated robust performance, particularly in stationary and seasonal time series. Prophet's explicit modeling of trend and seasonality components via additive structures and Fourier expansions allowed it to align well with stable and cyclic demand patterns. Its strength was most apparent in series exhibiting clear seasonality or stable long-term behavior, where its changepoint detection and trend extrapolation mechanisms provided consistent forecasting accuracy with relatively low computational complexity.

DeepAR showed remarkable capability in forecasting non-stationary time series, leveraging its autoregressive LSTM encoder structure and probabilistic modeling framework. Its ability to generalize across multiple related series enabled it to adapt to evolving demand dynamics, although its broader prediction intervals and reliance on structural similarities across series limited its precision in stationary and highly irregular non-seasonal series.

The classical LSTM model, while conceptually powerful for sequence modeling, consistently exhibited the weakest performance across all categories. Without the support of attention mechanisms or explicit trend modeling,

LSTM struggled to distinguish between informative and irrelevant historical patterns,

leading to higher forecast errors, particularly in non-stationary and non-seasonal series where volatility and irregularity dominated.

These overall findings directly address the third research question (Which forecasting model performs best for stationary, non-stationary, seasonal, and non-seasonal data subsets?). The results clearly indicate that Attention-based LSTM, with its dynamic attention capabilities, is the most effective model across a majority of demand types. Prophet follows closely in environments characterized by stability and regularity, such as stationary and seasonal series. DeepAR excels in handling dynamically evolving series, but its advantage diminishes when faced with highly stable or noise-dominated data. LSTM, despite its recurrent learning capabilities, consistently underperformed relative to models enhanced with attention mechanisms or explicit structural modeling.

The comparative analysis suggests that model selection for spare parts demand forecasting should be tailored to the statistical nature of the series being modeled. Attention-based architectures provide the most versatile and accurate forecasting capabilities across varied demand environments, offering significant operational advantages in inventory management and strategic planning within the automotive sector.

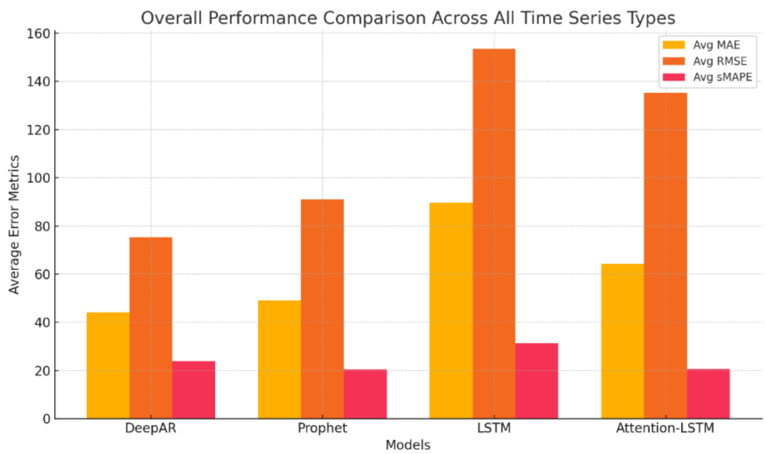

Figure 13. Overall performance comparison across all time series types.

The bar plot summarizes the overall forecasting performance of the models across stationary, non-stationary, seasonal, and non-seasonal time series. Attention-based LSTM consistently outperformed other models, achieving the lowest average error across MAE, RMSE, and sMAPE metrics. Prophet maintained strong performance, particularly in terms of minimizing relative errors. DeepAR showed moderate success but struggled with higher variance in absolute errors, while the classical LSTM model exhibited the weakest performance, highlighting the importance of incorporating attention mechanisms or explicit structural modeling for robust forecasting in diverse demand environments.

Chapter 5

Discussions and Conclusions

This study evaluated the performance of four different forecasting models DeepAR, Prophet, LSTM, and Attention-based LSTM on spare-part demand time series from a Turkish automotive company. The analysis was based on three standard error metrics (MAE, RMSE, and sMAPE), and the time series were categorized into four groups: seasonal, non-seasonal, stationary, and non-stationary. The results reveal that the performance of forecasting models varies significantly across different time series characteristics and that no single model dominates in all dimensions. However, Attention-based LSTM consistently outperformed others in most categories, particularly in terms of error stability and adaptability to complex patterns.

The results demonstrate that Attention-based LSTM achieved the lowest MAE and RMSE in seasonal and stationary series, where the model was able to focus on relevant temporal patterns and eliminate noise through the attention mechanism. Its performance in non-stationary series was also competitive, producing one of the lowest sMAPE values among all models. DeepAR, while demonstrating relatively low sMAPE in the non-stationary category, recorded much higher absolute errors (MAE and RMSE), indicating that although it captures overall trends well, it may still struggle with extreme spikes or outliers. Prophet offered strong interpretability and low error rates in seasonal and stationary series but exhibited performance degradation in more irregular or volatile patterns, especially in the non-stationary category. LSTM, although flexible and capable of learning long-term dependencies, produced the highest errors in both sMAPE and RMSE in the non-seasonal and non-stationary categories, suggesting that classical sequence learning alone may be insufficient in the face of erratic demand without the assistance of attention-based mechanisms.

From an operational standpoint, the results highlight the need for model-to-data alignment.

Rather than deploying a single forecasting model across all product lines, practitioners can benefit from a segmented modeling strategy, in which the time series are first classified based on statistical properties (e.g., seasonality, stationarity), and then the most suitable forecasting method is applied. For example, simpler and more interpretable models such as Prophet may be sufficient for seasonal, stable parts, whereas complex models like Attention-based LSTM are more suitable for unpredictable or high-variance series. This hybrid modeling framework allows for both efficiency and accuracy, optimizing the balance between computational cost and forecast quality.

The analysis also suggests that forecasting accuracy alone may not always reflect practical effectiveness. In some cases, models like DeepAR provided low relative error (sMAPE) but high absolute deviations (MAE), which may translate into critical stock shortages or excess inventory in real-world scenarios. Therefore, integrating these forecasting outputs with inventory cost models (e.g., shortage penalties, holding costs) would allow for more meaningful and actionable planning decisions.

While the results are encouraging, the study has its limitations. One key limitation is the evaluation of the models based solely on historical demand data, without considering external factors like promotions, macroeconomic indicators, weather influences, or supply-side constraints. Including these variables could enhance forecasting accuracy, particularly for non-stationary and non-seasonal data, where internal trends alone may not adequately account for changes in demand. Additionally, although the attention mechanism was beneficial in most cases, the transformer architecture—an advanced form of attention-based models—was not examined due to its greater data and computational needs. Future research could investigate the application of Transformer-based models or Temporal Fusion Transformers (TFT), which integrate attention, static covariates, and variable selection into a singular framework architecture.

Future research could explore hybrid or ensemble forecasting systems that integrate various models (such as statistical and deep learning) through techniques like model averaging, stacking, or meta-learning.

These systems can adaptively adjust the weights of each model's contribution based on historical performance or data characteristics, enhancing robustness across different types of time series. Additionally, utilizing automated model selection pipelines informed by meta-features like entropy, trend strength, or intermittency could offer scalable solutions for extensive inventories of spare parts without necessitating manual intervention intervention.In conclusion, this study confirms the utility of advanced deep learning methods, particularly Attention-based LSTM for forecasting demand in the spare-part domain, where traditional assumptions of regularity and seasonality often break down. The ability of attention mechanisms to identify and emphasize relevant historical segments improves forecast precision, especially in complex and high-noise environments. Combining these models with intelligent selection strategies, cost-sensitive loss functions, and operational constraints can further enhance their value in real-world supply chain settings. With continued development in interpretable AI and hybrid forecasting architectures, the automotive industry is poised to achieve more accurate, responsive, and cost-effective demand planning for its aftermarket operations.

REFERENCES

Abbasimehr, H., ve Paki, R. (2021). Improving time series forecasting using LSTM and attention models. Journal of Ambient Intelligence and Humanized Computing, 13, 673–691.

Abdulwahid, S., ve Haghifam, M.-R. (2024). Short-term load prediction based on LSTM network for Iraqi thermal power plant. International Journal of Electrical and Electronics Research, 12(4), 1461–1465.

Bohlke-Schneider, M., Kapoor, S., ve Januschowski, T. (2020). Resilient neural forecasting systems. International Workshop on Data Management for End-to-End Machine Learning (DEEM'20).

Cadena, E., Albright, K., Wang, H., ve Ajmera, S. (2023). Forecasting accessory demand in the automotive industry.

Dickey, D. A., ve Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association.

Edirisooriya, A., Weliwatta, R., Amarasena, H., ve Ganesh, A. (2024). Deep learning architectures for time series forecasting: A comprehensive survey. ResearchGate.

Fan, Y., ve Lixin, S. (2024). Comparative analysis of deep learning algorithms for predicting financial market time series. ISRG Journal of Arts, Humanities & Social Sciences, 2(4).

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., ve Chen, B. C. (1998). New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. Journal of Business & Economic Statistics.

Hyndman, R. J., ve Athanasopoulos, G. (2018). Forecasting: Principles and practice.

Ji, X., Huang, H., Chen, D., Yin, K., Zuo, Y., Chen, Z., ve Bai, R. (2023). A hybrid residential short-term load forecasting method using attention mechanism and deep learning. Buildings, 13(72).

Liao, W., Ye, G., Yin, Y., Yan, W., Ma, Y., ve Zuo, D. (2024). The inventory of spare parts prediction based on LSTM for large data and a short replacement cycle. IEEE International Conference on Computer Systems and Applications.

Ma, S., Fildes, R., ve Huang, T. (2016). Demand forecasting with high-dimensional data: The case of SKU retail sales forecasting with intra- and inter-category promotional information. European Journal of Operational Research, 249(1), 245–257.

Nazarenko, A. (2024). Inventory management system combining short-term demand forecasting and stock control policy: An automotive industry use-case. SSRN.

Ning, Y., Kazemi, H., ve Tahmasebi, P. (2022). A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet. Computers & Geosciences, 164, 105126.

Omprakash, M. K. (2024). Optimizing demand forecasting and inventory management with AI in automotive industry. Lappeenranta–Lahti University of Technology LUT.

Oukassi, H., Hasni, M., ve Layeb, S. B. (2023). Long short-term memory networks for forecasting demand in the case of automotive manufacturing industry. IEEE International Conference on Advanced Systems and Emergent Technologies.

Pearson, R. (2002). Outlier detection using Hampel identifier. Journal of Data Science.

Ramachandran, A., Neergaard, T. F. B., Maier, A., ve Bayer, S. (2025). Advancing heat demand forecasting with attention mechanisms: Opportunities and challenges. Friedrich-Alexander-Universität Erlangen-Nürnberg Preprint.

Rego, J. R., ve Mesquita, M. A. (2015). Demand forecasting and inventory control: A simulation study on automotive spare parts. International Journal of Production Economics, 161, 1–16.

Rožanec, J. M., Kažič, B., Škrjanc, M., Fortuna, B., ve Mladenić, D. (2021). Automotive OEM demand forecasting: A comparative study of forecasting algorithms and strategies. Applied Sciences, 11(6787).

Salinas, D., Flunkert, V., Gasthaus, J., ve Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. International Journal of Forecasting, 36(4), 1181–1191.

Sharma, A., Patel, N., ve Gupta, R. (2024). Leveraging LSTM and Prophet models for enhanced AI-driven demand prediction in e-commerce. Journal of Machine Learning Applications, 12(3), 45–62.

Taylor, S. J., ve Letham, B. (2017). Forecasting at scale. PeerJ Preprints.

Wunsch, A., Kühnert, C., Wallner, S., ve Ziebarth, M. (2024). Urban water demand forecasting using DeepAR-models as part of the Battle of Water Demand Forecasting (BWDF).

Yadav, K., ve Singh, M. (2025). A novel energy management of public charging stations using attention-based deep learning model. Electric

Power Systems Research, 238, 111090.

Ye, X., Ye, Q., Yan, X., Wang, T., Chen, J., ve Li, S. (2021). Demand forecasting of online car-hailing with combining LSTM + attention approaches. Electronics, 10(2480).

Yurttabir, A., ve Kiymetli Sen, I. (2021). Prophet model in financial performance forecast: Implementation in manufacturing sector.

Zahid, M. F., Fitrianto, A., Silvianti, P., ve Alamudi, A. (2024). Comparison between SARIMA and DeepAR with Optuna hyperparameter optimization for estimating rice production data in Indonesia. Indonesian Journal of Statistics and Its Applications, 8(2), 95–111.