

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
YÖNETİCİLER İÇİN KARAR BİLİMLERİ (İNG) BİLİM DALI

PREDICTING THE PRICE OF CARS IN TURKIYE USING MACHINE LEARNING TECHNIQUES

Yüksek Lisans

YASİN ONUR ABAT

İSTANBUL, 2024

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
YÖNETİCİLER İÇİN KARAR BİLİMLERİ (İNG) BİLİM DALI

PREDICTING THE PRICE OF CARS IN TURKIYE USING MACHINE LEARNING TECHNIQUES

Yüksek Lisans

YASİN ONUR ABAT

Tez Savunma Jürisi

1. Tez Danışman : Doç. Dr. HÜSEYİN EKİZLER
2. Üye. : Prof. Dr. BERİL DURMUŞ
3. Üye : Dr. Öğr. Üyesi AYŞE NUR ADIGÜZEL TÜYLÜ

İSTANBUL, 2024

T.C.
MARMARA ÜNİVERSİTESİ
FACULTY OF BUSINESS ADMINISTRATION
DEPARTMENT OF BUSINESS ADMINISTRATION (IN ENGLISH)
DECISION SCIENCE FOR EXECUTIVES (IN ENGLISH)

TEZ ONAY BELGESİ

İşletme Anabilim Dalı Yöneticiler İçin Karar Bilimleri Bilim Dalı Yüksek Lisans öğrencisi YASİN ONUR ABAT'ın PREDICTING THE PRICE OF CARS IN TURKIYE USING MACHINE LEARNING TECHNIQUES adlı tez çalışması, Enstitümüz Yönetim Kurulunun 22.05.2024 tarih ve 2024-21/10 sayılı kararıyla oluşturulan jüri tarafından oy birliği / oy çokluğu ile Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 12 / 06 / 2024

Öğretim üyesi Adı Soyadı			**İmzası**
1	Tez Danışmanı	Doç.Dr. Hüseyin Ekizler	
2	Jüri Üyesi	Prof. Dr. Beril Durmuş	
3	Jüri Üyesi	Dr. Öğr. Üyesi Ayşe Nur Adıgüzel Tüylü	

ÖZET

MAKİNE ÖĞRENME TEKNİKLERİ KULLANARAK TÜRKİYE'DEKİ ARAÇ FİYATLARININ TAHMİN EDİLMESİ

Bu araştırmanın amacı, Türkiyedeki Otomobil ve SUV kategorilerinde bulunan araçların gelecekteki altı aylık satış fiyatını tahmin etmektir. Araştırma sahibinden.com'da yayınlanan ilanlardan elde edilen bilgilerle yapılmıştır. Araçların marka, model, paket tipi, model yılı, yakıt tipi, vites tipi, kasa tipi ve kilometre özellikleri kullanılarak, geçmiş dönem satış fiyatları üzerinden gelecekteki altı aylık satış fiyatları tahmin edilmiştir. Bu amaçla bu araştırmada altı farklı algoritma (ANN, ARIMA, SARIMA, XGBOOST, TBATS, HOLT WINTERS Forecasting) çalışılmıştır. Bu algoritmaların ürettiği sonuçların başarılarına göre en iyi algoritma olan HOLT WINTERS seçilmiştir. HOLT WINTERS'ın ürettiği sonuçlar araştımanın çıktısı olarak kabul edilmiştir. Bu araştırmanın sonucunda temel olarak şu soruya cevap verilmiştir; içinde bulunan günün tarihinden sonraki altı ay içinde satılmak/alınmak istenen araçların fiyatı ne olacaktır? Bu soruya cevap verilerek araç alım/satım işlemlerinde araçların değerinin daha doğru ve güvenilir bir şekilde belirlenmesine ortam oluşturulmuştur. Artık, satıcı ve alıcılar ekonomik olarak durumlarını daha iyi değerlendirebilecek ve işlemlerini daha iyi gerçekleştirebileceklerdir.

Anahtar Kelimeler: Araç Satış Fiyatı Belirleme, Otomobil Satış Fiyatı Tahminleme, SUV Satış Fiyatı Tahminleme, Araç Satış Fiyatı Tahminleme, Araç Satış Fiyatı Değerleme.

ABSTRACT

PREDICTING THE PRICE OF CARS IN TURKIYE USING MACHINE LEARNING TECHNIQUES

The aim of this research is to estimate the prospective six months sales price of vehicles in the Automobile and SUV categories in Turkiye. This research was done with the information obtained from the classifieds posted on sahibinden.com. By using the vehicles' brand, model, package type, model year, body type, gear type, fuel type and mileage characteristics for the future six months sales prices of the vehicle was estimated over the past selling prices of vehicles. For this regard, six different algorithms (ANN, ARIMA, SARIMA, XGBOOST, TBATS, HOLT WINTERS Forecasting) were studied in this research. According to the success of these algorithms, the best algorithm which is HOLT WINTERS was selected. The results of HOLT WINTERS were announced as output of the research. As a result of this research, basically the answer to the following question was sought; What will be the price of the vehicles to be sold/desired to be purchased within the time frame six months after the current date? By answering this question, it was created an environment to determine the value of the vehicles more accurately and reliably during vehicle purchase/sale transactions. Thus, sellers and buyers will be able to evaluate their situation economically better and be able to perform their transactions better.

Keywords: Vehicle Estimation Sales Price Determination, Automobile Sales Price Estimation, SUV Sales Price Estimation, Vehicle Sales Price Estimation,Vehicle Sales Price Valuation.

ACKNOWLEDGMENTS

Nowadays, acquiring a car is a challenging task for many and requiring extensive research prior to purchase.In addition to the difficulties of choosing the right car, many external factors such as low income, inflation, credit opportunities, car shortage can affect people's purchasing process. Therefore, the process of owning a car can take a long time for many people. So, we targeted to help people at their decision process with this study. The aim of this research is to estimate the prospective six months sales price of vehicles in the Automobile and SUV categories in Turkiye. This research was done with the information obtained from the classifieds posted on sahibinden.com. By using the vehicles' brand, model, package type, model year, body type ,gear type, fuel type and mileage characteristics for the future six months sales prices of the vehicle was estimated over the past selling prices of vehicles. As a result of this research, basically the answer to the following question was sought; What will be the price of the vehicles to be sold/desired to be purchased within the time frame six months after the current date? By answering this question, it was created an environment to determine the value of the vehicles more accurately and reliably during vehicle purchase/sale transactions. Thus, sellers and buyers will be able to evaluate their situation economically better and be able to perform their transactions better. Sellers will determine selling value of their vehicle better by seeing what the next six months price of their vehicle will be. And buyers will have the opportunity to decide faster when they see what the prices of the vehicles will be in the next six months.

I want to thanks to Assoc. Prof. HÜSEYİN EKİZLER from whose knowledge and experience i benefited during the realisation of this study, to my valuable colleague GÜRAY ATAMAN who supported me with his opinions at every stage of the research, always made me feel his sincerity and guided me correctly, and to my valuable managers GÜLCAN OGUNDUR, EMRE KESMEN and NAZIM ERDOĞAN who never withheld their support during my all study. I want also thank my dear wife DİDEM BİBER ABAT who has made an extraordinary effort to facilitate every stage of my life and has always supported me throughout my thesis study. Finally, I would thank my lovely family and all my friends who have always supported me in all my journeys, especially in this study, and who have never lost their trust in me.

YASİN ONUR ABAT

İstanbul, 2024

TABLE OF CONTENTS

ÖZET ..... *IV*
ABSTRACT ..... *V*
ACKNOWLEDGMENTS ..... *VI*
ABBREVIATIONS ..... *IX*
LIST OF TABLES ..... *XI*
LIST OF FIGURES ..... *XII*
1. INTRODUCTION *1*
1.1. The Aim of The Study **2**
1.2. The Scope of The Study **3**
2. AUTOMOBILE MARKET *4*
2.1. History of Automobile **4**
2.2. Worldwide Automobile Market **6**
2.3. Turkish Automobile Market **7**
2.4. Automobile History in Turkish Market **8**
2.5. Second Hand Automobile Market **9**
2.5.1. Worldwide Secondhand Automobile Market 9
2.5.2. Turkish Secondhand Automobile Market 10
2.6. Secondhand Automobile Trading Sites and Applications **11**
2.6.1. Worldwide Secondhand Automobile Trading Sites and Applications 12
2.6.2. Turkish Secondhand Automobile Trading Sites and Applications 14
2.7. Reasons of Prefering Secondhand Automobile **16**
3. LITERARURE REVIEW *17*
4. METHODOLOGY *20*
4.1. Machine Learning Algorithms **20**
4.1.1. ANN 20
4.1.1.1. Architecture of ANN 21
The architecture of ANN can be listed as follows: 21
4.1.1.2. Advantages of ANN 22
4.1.1.3. Disadvantages of ANN 23
4.1.1.4. Applications of ANN in Literature 23
4.1.2. XGBoost 25
4.1.2.1. Architecture of XGBoost 27
4.1.2.2. Advantages of XGBoost 28
4.1.2.3. Features Supported by XGBoost 29
4.1.2.4. Disadvantages of XGBoost 30
4.1.2.5. Applications of XGBoost In Literature 30
4.2. Forecasting Techniques **31**
4.2.1. Tbats 31
4.2.1.1. Architecture of Tbats 32
4.2.1.2. Advantages of Tbats 32
4.2.1.3. Disadvantages of Tbats 33
4.2.1.4. Applications of Tbats In Literature 34
4.2.2. Arima 34
4.2.2.1. Architecture of Arima 35

4.2.2.2. Advantages of Arima 37
4.2.2.3. Disadvantages of Arima 37
4.2.2.4. Applications of Arima in Literature 37
4.2.3. Sarima 38
4.2.3.1. Architecture of Sarima 38
4.2.3.2. Advantages of Sarima 39
4.2.3.3. Disadvantages of Sarima 39
4.2.3.4. Applications of Sarima In Literature 40
4.2.4. Holt Winters 40
4.2.4.1. Architecture of Holt Winters 41
4.2.4.2. Advantages of Holt Winters 42
4.2.4.3. Disadvantages of Holt Winters 43
4.2.4.4. Applications of Holt Winters In Literature 43
5. MODEL COMPARISON ***44***
5.1. Mean Squared Error (MSE) **44**
5.2. Mean Absolute Error (MAE) **44**
5.3. Mean Absolute Percentage Error (MAPE) **45**
5.4. Root Mean Squared Error (RMSE) **45**
6. DATA COLLECTION and DATA CLEANING ***46***
6.1. Data Collection **46**
6.2. Data Cleaning **47**
7. FINDINGS ***48***
7.1. ANN Results **48**
7.2. XGBoost Results **50**
7.3. Arima Results **55**
7.4. Sarima Results **56**
7.5. Tbats Results **58**
7.6. Holt Winters Results **60**
7.7. Model Comparison Results **62**
8. CONCLUSION & DISCUSSION ***63***
REFERENCES ***68***

ABBREVIATIONS

ANN	: Artificial Neural Networks
ARIMA	: Auto-regressive, integrated, and moving average
SARIMA	: Seasonal, Auto-regressive, integrated, and moving average.
XGBoost	: Extreme Gradient Boosting
TBATS	: Trigonometric seasonality, Box-Cox transformation, ARIMA errors, Trend, Seasonal components
SUV	: Sports Utility Vehicle
IQR	: Inter Quantile Range
USA	: United States of America
OECD	: Organisation for Economic Co-Operation and Development
EU	: European Union
US	: United States of America
e.g	: Exempli Gratia
TEMSA	: Termo-Mekanik Sanayi ve Ticaret A.Ş.
BMC	: British Motor Corporation
TOGG	: Türkiye'nin Otomobili Girişim Grubu
Otokar	: Otokar Otomotiv ve Savunma Sanayi A.Ş
R&D	: Research and Development
OEMs	: Original Equipment Manufacturers
CAGR	: Compound Annual Growth Rate
CV	: Commercial Vehicle
UK	: United Kingdom
MASFED	: Motorlu Araç Satıcıları Federasyonu
KNN	: K-Nearest Neighbors
MLP	: Multi-Layer Perceptron

GBDT : Gradient-Boosted Decision Trees

ARMA : Autoregressive Moving Average

AIC : Akaike Information Criterion

ML : Machine Learning

AI : Artificial Intelligence

NLP : Natural Language Processing

GBM : Gradient Boosting Machine

SES : Simple Exponential Smoothing

HES : Holt's Exponential Smoothing

WES : Winter's Exponential Smoothing

EWMA : Exponentially Weighted Moving Average

MSE : Mean Squared Error

RMSE : Root Mean Squared Error

MAPE : Mean Absolute Percentage Error

MAE : Mean Absolute Error

LIST OF TABLES

Table 1: An Example of Data Set 47
Table 2: Mean values of ANN in terms of MAE,MSE,RMSE and MAPE 49
Table 3: Median values of ANN in terms of MAE,MSE,RMSE and MAPE 50
Table 4: Mean values of XGBOOST in terms of MAE,MSE,RMSE and MAPE 54
Table 5: Median values of XGBOOST in terms of MAE,MSE,RMSE and MAPE 54
Table 6: Mean values of ARIMA in terms of MAE,MSE,RMSE and MAPE 56
Table 7: Median values of ARIMA in terms of MAE,MSE,RMSE and MAPE 56
Table 8: Mean values of SARIMA in terms of MAE,MSE,RMSE and MAPE 57
Table 9: Median values of SARIMA in terms of MAE,MSE,RMSE and MAPE 57
Table 10: Mean values of TBATS in terms of MAE,MSE,RMSE and MAPE 59
Table 11: Median values of TBATS in terms of MAE,MSE,RMSE and MAPE 59
Table 12: Mean values of HOLT WINTERS in terms of MAE,MSE,RMSE and MAPE 61
Table 13: Median values of HOLT WINTERS in terms of MAE,MSE,RMSE and MAPE 61
Table 14: Mean values of comparison indicators 62
Table 15: Median values of comparison indicators 62

LIST OF FIGURES

Figure 1: ANN Use Case Model 20
Figure 2: ANN Architecture 21
Figure 3: ANN activation function 21
Figure 4: XGBoost Features 25
Figure 5: A decision tree example 26
Figure 6: A Decision tree methodology 26
Figure 7: A generic unregularized XGBOOST algorithm 27
Figure 8: Features Supported by XGBOOST 29
Figure 9: The key features of TBATS 31
Figure 10: Model of TBATS 32
Figure 11: The stationarity of time series data in forecasting 36
Figure 12: Sarima model 38
Figure 13: Double exponential smoothing 41
Figure 14: Triple exponential smoothing 42
Figure 15: Sample data view 47
Figure 16: A few examples of ANN results 49
Figure 17: A few examples of XGBOOST results 54
Figure 18: A few examples of ARIMA results 55
Figure 19: A few examples of SARIMA results 57
Figure 20: A few examples of TBATS results 59
Figure 21: A few examples of HOLT WINTERS results 61

1. INTRODUCTION

It is evident that most individuals require the use of a motor vehicle for the purpose of transportation, which is essential for the completion of their daily routines. These routines may include commuting from one's place of residence to their place of employment, medical facilities, shopping centers, or social engagements with friends and family. Many people use public transportation for their transportation needs. On the other hand, there is a community whose economic situation allows them to buy a vehicle. However, the process of buying a vehicle is very difficult for many people. We can say the process of giving decision is much more difficult and time consuming for many people, because they are trying to buy a car with their savings over the years. At this stage, individuals are devoid of the opportunity for error, as a wrong decision could potentially result in the depletion of their savings. Consequently, individuals do extensive research before they buy or sell a car.

It can be said that the most important factor other than the durability and features of the car is the sales price of the car. People firstly make price research by looking the prices of both new one and second hands for their interested cars. At this point we can express when people compared the prices of new and second hands, many people tend to have a secondhand car instead of a new one because of the economic reasons. That’s why, the correct determination of the vehicle sales price is very important for both buyers and sellers. It is very important for buyers to determine the true value of the vehicle. Buyers need to be sure of the best price they can get. Likewise, it is very important for sellers to determine the right selling price. Sellers need to know what the best-selling price is.

At this point, we can say estimating the reselling price of vehicle is not a easy process. The prices of secondhand cars are affected by many factors. This makes prediction of sales prices of cars difficult. Because information about all factors is not always available. So, buyers must give the decision to purchase a car in terms of only a few factors. So, this situation causes a great mistake. For example, buyers pay too much for the car or sellers sell cheaper than market value of the car.

That's why, it is targeted to study with different methods to find a better application to learn the selling prices of secondhand vehicles in this research. As a result of this study, the selling price of secondhand cars can be determined a more accurate and reliable way. Sellers will determine selling value of their vehicle better by seeing what the next six months price of their vehicle will be. And buyers will have the opportunity to decide faster when they see what the prices of the vehicles will be in the next six months.

1.1. The Aim of The Study

Nowadays, acquiring a car can be a challenging task for many and requiring extensive research prior to purchase. Many external factors can impact the buying process of people such as low incomes, inflation, credit opportunities and car shortages etc. Therefore, the process of owning a car can take a long time for many people. And purchasing a car is a very expensive action for people in Turkiye. Many people must save money to buy a car. That's why, people don't want to lose their money by giving a wrong decision. So, it is quite normal making detailed research before buying a car. People firstly make price research by looking the prices of both new one and second hands for their interested cars. At this point we can express when people compared the prices of new and second hands, many people tend to have a secondhand car instead of a new one because of the economic reasons.

And it is well known fact that the sales sales number of secondhand vehicles in many countries is larger than the sales number of new vehicles. This is also evident in Turkiye where a vast market for secondhand vehicles exists. Turkiye is a developing country and car market is growing rapidly and gaining more importance with the various other projects. However, it is widely acknowledged that several significant factors, including high inflation, high taxes, and a shortage of cars, have an impact on consumer behaviors in Turkiye. As a result, people have had to resort to purchasing secondhand vehicles instead of new ones. Therefore, it has become crucial to forecast the resale value of secondhand cars for both purchasers and vendors.

That's why, the main goal of this research is to forecaste the prospective six months selling price of secondhand vehicles in the Automobile and SUV categories in Turkiye. With this study, basically the answer to the following question is sought; What will be the price of the vehicle to be sold/desired to be purchased within the time frame six months after the current date? By answering the blank question, it will be possible to determine the value of the vehicle more accurately and reliably during vehicle purchase/sale transactions. Sellers and buyers will be able to evaluate their situation economically and be able to perform their transactions better.

1.2. The Scope of The Study

In this study, we aimed to forecast the potential sales prices of vehicles in the automobile and SUV categories in Turkiye for six months period. This research was conductted using data gathered from classified posts on sahibinden.com. By using the vehicles' brand, model, package type, model year, body type , gear type, fuel type and mileage characteristics for the future six-month sales prices of the vehicle was estimated over the past prices of vehicles. For this regard, six different algorithms (ANN, ARIMA, SARIMA, XGBOOST, TBATS, HOLT WINTERS Forecasting) were studied in this research. According to the success of these algorithms, the best algorithm which is HOLT WINTERS was selected. The results produced by the selected algorithm (HOLT WINTERS) was announced as output.

The Data was collected from the classifieds in Automobile and SUV categories posted on sahibinden.com. The classifieds in the data set from January 2020 to December 2021 was scanned and information was gathered for vehicles. Then, some adjustments were done on the data set. Firstly, vehicles with thirty or more classifieds were included in the data set. Secondly, vehicles with data for at least 21 months in the last 24 months and vehicles with data data in the last month were included in the data set. Then, outliers were eliminated separately for each brand and model by using the Tukey outlier methodology (IQR). Finally, medians of classifieds were taken for the data set. After all these adjustments, we implemented our models with R Programing language to estimate the prospective six months sales price of vehicles in the Automobile and SUV categories in Turkiye.

At the end of this research, basically the answer to the question was sought; What will be the price of the vehicles to be sold/desired to be purchased within the time frame six months after the current date? By answering this question, it will be created an environment to determine the value of the vehicles more accurately and reliably during vehicle purchase/sale transactions. Sellers and buyers will be able to assess their economic situation more accurately, enabling them to conduct transactions more effectively.

2. AUTOMOBILE MARKET

The word "automotive" is derived from the Greek words "auto" (self) and "motivus" (movement). The meaning of the word proves this. When we looked at the meaning of automotive, we can see that it is a self-propelled vehicle. However, it is more than this meaning in our time. Even though it is just a car for some of us, it represents a way of life for some of us. It is also one of the fastest growing industries right now, thanks to evolving technologies and changing customer preferences. It is among the largest industries worldwide in terms of revenue. The automotive industry covers many business pillars such as designing and modification , development new technologies, production, marketing strategies, sales, repair of vehicles.

2.1. History of Automobile

It's amazing to think that the automotive industry has been around for over many years. It all started in the 15th century, when Leonardo da Vinci first designed transportation vehicles. These vehicles differed from modern-day automobiles in that they lacked seating for passengers. Despite this, they were exceptionally well designed.

The history of automobile production begins with the invention of the first steam-powered vehicle in 1672. It continues with the production of first steam vehicle in 1769 by Nicolas-Joseph Cugnot. Nicolas-Joseph Cugnot had designed the first self-propelled ground vehicle. It was a military tractor for the French army. It has only three wheels and no need for manpower. It could reach a maximum speed of four kilometres per hour. It had a driving capacity of fifteen minutes and required the boiler to be refilled after every twenty kilometres traveled. In addition, there is a known story about this model. The vehicle has crashed into a wall during the test and have destroyed the wall. As the vehicle is metal, there was no serious damage.

In 1789, Oliver Evans made history by becoming the first person in the United States to receive a patent for a self-propelled vehicle. This vehicle was capable of land and water travel via pedal wheels at the back. It weighed twenty-one tons. For the next eighty years, other inventors pursued similar projects. Most of the realized projects were successful. In the 18th century, steam cars were mainly produced due to the steam age. In 1859, Belgian engineer Ètienne Lenoir received patent with the name "Gas and expanded air engine". After that, he developed the first electrically ignited and water-cooled internal combustion engine in 1860. With the invention of the internal combustion engine, faster and long-distance cars were developed.

To mention further developments during this period, it is notable that the first motor vehicle designed for transporting people was invented in 1862. In addition to the vehicle with a speed of four km, a vehicle model with a speed of sixty-three km was also produced. This period was characterised by a significant advancement in automobile speed. In the following periods, more advanced vehicles with metal as well as leather seats and lighting began to come to the front. In 1885, German Karl Friedrich Benz produced the first gasoline car powered by an internal combustion engine.

The true rise of automobiles began with the emergence of internal combustion engines in the 19th century. The introduction of these engines allowed for higher quality cars to be produced. In the 1800s, an automobile brand embarked on a world tour, showcasing their own vehicle models. And technological advancements also increased during this period. For example, while the automobile industry consisted of vehicles produced by hand until 1924 (Pollard, 1995), Henry Ford tried to increase demand for automobiles by establishing a mobile assembly line for the first time (Berger, 2001). Ford focused its efforts on improving production methods to decrease costs. And the moving assembly line start its operation, thanks to lower cost of the Model T. It was a resounding success. in 1913-1914. This was a game-changer. It increased sales and potential customers. So, it can be said that during the 19th century, numerous automobile models were produced while new designs and brands emerged.

In the 20th century, significant advancements were made in the history of automobiles. Several automobile brands gained prominence and mass production enabled the production of a greater number of vehicles. Therefore, it can be stated that individuals easily were able to acquire automobiles of their own choosing during this period.

When examining the 21st century, electric vehicles emerge as a significant development in the automotive industry. Even though the history of electric cars dates back further, electric cars have taken on a more prominent role since the 2000s. Many countries are working towards transitioning to electric vehicles as a means of decreasing carbon emissions. Many automakers are taking steps to reduce carbon emissions and switch exclusively to electric vehicles. Over the next five years, it is thought that sales amount of electric vehicles will increase by at least four times. Consequently, over fifty million passenger cars and more than four million commercial vehicles will take to the roads. As prices decrease and more models become available, the ownership of electric vehicles will expand to a broader segment of vehicle owners. It is anticipated that by 2025, electric vehicles will be more than 10% of sales, with a potential for this figure to reach 20 to 30% until 2030.

In brief, when we look at the development process of automobiles, we can say the invention of the automobile is unlike any other invention. This difference arises from the fact that a certain person cannot say that I invented the automobile alone. Unlike other innovations, the automobile came into existence

through the consolidation of different inventions that emerged across the globe over a century. It was the ideas and contributions of multiple individuals that allowed the car to evolve into its advanced state during its development. It is estimated that the modern automobile emerged after the acquisition of around 100,000 patents. Therefore, we can say that consistent development and growth in the automotive industry over the last century has led to the present-day vehicle.

2.2. Worldwide Automobile Market

Today, there is a significant automobile market all over the world. The automotive industry has an significant position in the economies of many industrialized countries. The automotive sector is an important industry, with lots of jobs and a lot of money invested in it. It is a crucial factor in the economic growth of countries. So, it can be said that economically, the automotive sector is essential for all countries. When considering that cars make up around 70% of the industry's total motor vehicle production worldwide, it becomes clear that the automobile sector holds significant importance in the global economy. For instance, it is worth noting that in 2006, 87% of all motor vehicles produced in Europe were passenger cars (Onat, 2007). Furthermore, the manufacture of cars, along with its well-developed supply chain, also drives the production of other types of vehicles (Onat, 2007).

According to one of the latest reports for the automotive industry created by the OECD (Lejarraga, Kouzul-Wright, Primi, Toselli, & Wermelinger, 2016), according to their production volumes the production quantities of the top twenty countries, account for 90% of global production. The industry is distinguished by the presence of automotive hubs in various regions across different continents. It is worth noting that developed EU countries occupy a central position in the global market. Considering significant export countries within the sector, such as Germany, France, and the USA, it is apparent that most of them operate at 75-80% of their potential export capacity. However, east countries of Europe such as Turkiye, Poland, and the Czech Republic have experienced a decline in capacity ratios, now operating at only 35-40% (Trademap, 2021).

And also, when we looked at the statistic numbers, the automotive industry valued at around £4.5 trillion globally in 2019. It represents 5% of the global economy (Statista, 2020) and provides new job opportunities for around 80 million individuals. It is estimated that the worldwide automotive manufacturing market was valued at approximately £2.86 trillion as of 2021, with an estimated growth to £2.95 trillion in 2022. And also, in the fourth quarter of 2022, mergers and acquisitions deals in the automotive industry had a total value of three billion US dollars. The global total value of these deals varied between 2017 and 2022, with 2021 recording the highest overall total deal value in this period.

2.3. Turkish Automobile Market

The automotive market in Turkey is a very important, in a similar manner to other countries around the world. Its strong links to many other sectors (iron, plastics steel, chemicals, textiles, rubber, electronics, informatics) and export-oriented production strategies are worthy of note. 1,352,648 motor vehicles were produces in 2022 in Turkiye. This making Turkiye as the 13th important producer globally. The Turkish automotive market is an important part of the global automotive market. $20 billion value of vehicles and spare parts of vehicles are exported by Turkiye in a year. This is due to the presence of car manufacturers and suppliers in the country. It's so great to see the market of vans, buses and trucks in the world is leading by Turkish companies like BMC, Otokar . And it's even more impressive that Togg (Turkiye's Automobile Joint Venture Group Inc.) has been launched on the world market by the first all-electric vehicle company of Turkiye.

The automotive industry in Turkiye has seen some pretty big changes over the last 20 years! It's now a sector that's doing really well, with a trade surplus thanks to its high-number exports to developed countries. And also, it has enabled it to design and commission vehicles from scratch. And there’s more! Thanks to its export-oriented structure and its location right on the doorstep of both the European and Middle East/African markets, the Turkish automotive industry has become a real hub for vehicle production. It’s a great place for national and international companies to make their vehicles, in all kinds of segments.

Eight global Original Equipment Manufacturers (OEMs) has increased their production amount from 300,000 in 2002 to more than 1.3 million in 2021 in Turkey. Turkiye is an extremely advantageous location. Turkiye is offering a quality workforce and a developinh local market for manufacturers. Turkiye has already become very popular for manufacturing of commercial vehicles. Turkiye became the leading commercial vehicle (CV) producer in Europe In 2021. And also, the automotive market in Turkiye is seeking to enhance its R&D standards now.

It could be said that Turkey is home to a significant number of suppliers on a global scale. We are pleased to note that there are over 250 international suppliers operating from Turkey. And also, 30 of them within the first 50 global suppliers. And we can say many global brands that conduct their R&D solutions in Turkey. Ford, Fiat, Daimler are a few ones of the brands that are engaging in R&D solutions in Turkiye. Ford Otosan is one of three of the largest global R&D centers of Ford. R&D center of Fiat in Bursa is a distinctive facility for the Italian company, situated outside of Italy and serving the European market. In addition, Daimler's Istanbul-based R&D center serves manufacturing operations of truck and bus in Turkiye. And also, second R&D center of AVL Turkiye has been opened. It has targeted with second R&D office to progress hybrid vehicle technologies with autonomous driving capabilities. So, it can be said that Turkiye offers a favorable environment for vehicle production for manufacturers.

All these developments shows that Turkiye has a significant role in the vehicle production market. Automobile manufacturers are increasingly selecting Turkiye as a production center for their exports.

2.4. Automobile History in Turkish Market

When looking the history of the automotive industry in Turkiye, it can be said that the first automobile arrived in the country during the final period of the Ottoman Empire. Firstly, Sultan II. Abdülhamid ordered an electric car from the British Consulate, but it never actually hit the road. The first vehicle to be driven on Turkish roads was the Renault Landaulet model in Fenerbahçe in 1895.

Turkiye's initial production attempt in the automotive industry was the assembly plant founded by Ford Motor Company in 1929 to produce cars, tractors and lorries. However, this endeavor was unsuccessful due to the economic crisis of 1929. In the 1950s, these efforts to establish production continued in the sector. In 1954, Turkish Willys Overland Ltd. became the first Turkish partner to establish a factory for producing jeeps and pickup trucks for the army. Another factory, Türk Otomotiv Endüstrisi A.Ş. was established in 1955 to produce trucks. And Istanbul Bus Karoseri Sanayi A.Ş. began productions in 1963, investing in bus production (Taşıt Araçları Tedarik Sanayicileri Derneği, 2023). Then, in 1961, the Eskişehir State Railways factory produced the first completely domestic Turkish car, named Devrim. However, the production was limited to producing sample models due to various reasons, including insufficient demand (Uysal & Hatun Şimşek, 2022).

It is widely accepted that the primary automotive industry in Turkiye began with Otosan, which was founded in 1959 as a partnership between Ford and Koç. In 1966, Otosan entered into an agreement with the British company Reliant which led to the production of the first domestically produced car, the Anadol. OYAK-RENAULT and TOFAŞ were established in 1969 with French and Italian licenses respectively. These companies have been conducting large-scale production in the sector since 1971 (Güneş, 2012).

In the 1970s, studies were conducted on localizing component and part production. The sector initially displayed a domestic market-oriented approach until the 1980s when an outward-looking approach was adopted with the transition to a free market economy. The goal was to target a modern, quality oriented and internationally competitive sector. However, due to protectionist policies, a steady sectorial growth was impeded (Elmas, 2011). In the automotive sector (Köksal & Türedi, 2014), an export-driven was acquired a competitive edge with the influx of global brands in the 1990s.

Todays, Turkiye, a leading producer of commercial vehicles in Europe. It has important successes at the truck and bus segments as compared to other countries. In the first half of 2016, Turkiye's production of light commercial vehicles grew by 8%. This improvement made Turkiye as the most important commercial car producer between European countries and the second important producer in the world

(Pişkin, 2017). It should be pointed out that even with the rise in foreign investment focused manufacturing in the industry, over 50% of the automobile market is supplied by imports (Karbuz, Fahri, Silahçı & Çalışkan, 2017).

Mercedes Benz, Renault, Ford, Toyota, Hyundai, Fiat, Honda and Isuzu are some of the companies featured in the Forbes world's important automotive companies list, and they produce vehicles not only in global markets but also in Turkiye. Additionally, Turkish-based businesses like Hattat, Karsan, Otakar, Temsa, Türk Traktör and the German Man company also manufacture in Turkiye (Otomobil Sanayi Derneği, 2022). Since the 2000s, the industry has shifted its focus from expanding the production assembly process to technology, model development, and R&D studies suitable for the global competitive environment (Pişkin, 2017). Toyota, a leading brand ranked number one worldwide in terms of brand value and total sales, also manufactures Corolla, Verso, and Toyota C-HR models in Turkiye. Toyota exports 85% of its vehicles produced in Turkiye to over 90 countries worldwide every year (Otomobil Sanayi Derneği, 2022).

In addition, in the field of automobile production, the state's research on producing a domestic automobile brand has gained momentum and reached a certain point in recent years. TOGG, the first domestic car brand, was produced and launched by the Ministry of Technology and Science. TOGG will produce five different models by 2030. We are pleased to announce that the first of these Turkish vehicles has been launched in 2023.

2.5. Second Hand Automobile Market

After examining the global and Turkish new automobile markets, attention can be turned to the secondhand car market. It is evident that in most of countries, the sales sales number of secondhand vehicles has surpassed the new cars, highlighting the growing significance of the secondhand car market.

2.5.1. Worldwide Secondhand Automobile Market

The secondhand car market is one of the largest market in the US. It is larger than the new car market. And it continues to grow. (Lee, 2006). According to data from the US Department of Labour, prices for secondhand vehicles soared by 40.5 per cent in January 2022, year-on-year comparison of the same period last year. This increase in vehicle pricing also has a significant role in fueling inflation in the country to 7.5 per cent January 2022, the highest it has been in 40 years. The most substantial price increase was observed in the Dodge Grand Caravan, Nissan Versa and Toyota Prius models in the secondhand market of the United States of America. The Dodge Grand Caravan experienced a 69% price increase this year in January 2022 compared to the previous year, reaching a total of £19,017.

Secondhand car prices in the UK increased by almost 30% last year (2022), according to data from the British automobile association AA Cars. Over a two-year period, there was a 57% increase. In 2019, the average price of Ford Fiestas, one of the most popular cars aged between three and five years old in the country, was £7,448, which rose by 31% to £9,770 last year (2021). And the prices of Ford Focus cars rose by 43%, reaching 12,764 pounds, while Volkswagen Polos experienced a 34% increase, reaching 11,037 pounds. In Europe, Germany has 26.3% market value share. England has 23.7% market value share. France has 19.6% market value share. And Italy has 16.5% market value share. According to AutoScout24, Europe's biggest online car sales market, secondhand car prices in the European Union increased by 8% in the past year (2021). European car supply reduced by 11% in one year. Among EU countries, Austria saw the most significant rise in the average price of secondhand vehicles last year (2021) at 13%, followed by the Netherlands at 11%, Germany at 10%, Belgium at 9%, and Italy and Luxembourg at 7%. The Volkswagen Golf was the most popular secondhand across Europe in the preceding year. In 2020, the mean secondhand price of the Volkswagen Golf was €15,786, but this increased to €19,046 in 2021. The Audi A4, the second most preferred secondhand, saw an increment from €20,079 to €21,839. Notably, as of 2022, the minimum wage in Luxembourg is €2257, the Netherlands stand at €1725, in Germany it is €1624, and in Belgium it is €1658; meanwhile, Austria and Italy do not apply any minimum wage regulations.

Looking at these data, it can be easily said that the secondhand vehicle market is growing day by day globally. This suggests that the secondhand car market is worthy of our attention.

2.5.2. Turkish Secondhand Automobile Market

Turkiye has the fourth largest secondhand vehicle market in Europe, valued at $50 billion. (AutoScout24 ,2022) data indicates that secondhand car prices increased by 8 percent in the European Union last year (2021), while car supply declined by 11 percent. Nevertheless, secondhand vehicle prices in Turkiye increased by over 50 percent on average in the past year (2021) due to surging demand. Industry experts anticipate a further increase in the secondhand market in Turkiye due to the economic development measures being implemented.

According to the (Indicata report, 2022) on the secondhand market, the prices of used vehicles in Turkiye rose by an average of 63% last year. According to (Indicata report, 2022) for Turkiye, sales in the secondhand online market for passenger and light commercial vehicles increased by 10% in 2022, with 1.811.498 units sold, compared to 1.652.710 units sold in 2021. Looking at Turkiye, sales in the secondhand online market for passenger and light commercial vehicles increased by 10% in 2022, with 1.811.498 units sold, compared to 1.652.710 units sold in 2021.

Erkoç (2023) reviewed figures from the first quarter of 2023 and reported a 49.5% increase in pre-owned car sales in comparison to the previous year's (2022) same period. Aydin Erkoc observed that in the previous year (2022), the secondhand car market experienced a contraction, particularly in the second half of the year. There was a decline until November, followed by a recovery in December and a final increase of 8.5 percent by the end of the year. Furthermore, the market demonstrated a notable increase of nearly 50 percent compared to the previous year (2022). This is largely attributable to the positive momentum observed in January 2023. In January 2023, sales of secondhand goods reached 921,387 units, which represents a 97.2% increase compared to the previous year (2022). However, in February, the number of units sold decreased to 704,954, representing an increase of only 23.4%. This was due to the earthquakes that caused significant losses in 11 provinces of the country. In March, sales increased by 39.9%, surpassing one million units and reaching 1.066.442 units as Ramadan commenced. During the January-March period, the total sales of used vehicles amounted to 2.692.783 units, marking an increase of 891.717 units, or 49.5%, compared to the same duration of the prior year. It is important to note that these figures do not include data from the holiday season. Consequently, they provide an indication of a higher growth rate that will be reflected in the April data. Erkoç observed that the rise in automobile sales, the primary product in the vehicle market, significantly exceeds the general increase. In the first quarter, automobiles comprised the majority of secondhand vehicles sold, with 1,878,280 units in total sales. The amount of automobiles purchased increased by 59.4% compared to the previous year (2022).

2.6. Secondhand Automobile Trading Sites and Applications

Today, the secondhand car market holds a vital position in almost all countries globally. And we can say the secondhand car market is a substantial and strategically important market for also vehicle producers. Recognizing the potential of the secondhand car market, major car manufacturers have already entered or are ready to enter this market. Additionally, almost all car brands engage in secondhand sales, highlighting the commercial appeal of the secondhand car market. In addition to large corporations, car dealerships also hold a significant place in this market. They make significant profits through buying and selling vehicles. In the present day, with the advancement of the internet and technology, secondhand car dealers and buyers from all around the globe connect through the internet and engage in the commerce of vehicles with ease. At this point we can say, there are numerous websites and applications available where we can purchase and vend secondhand cars.

2.6.1. Worldwide Secondhand Automobile Trading Sites and Applications

AutoScout24 is one of the Europe's largest online car marketplaces with over 30 million monthly users. It also has more than 43,000 dealer partners and around 500 employees. It works in 18 countries, and it has more than 43,000 associated dealers in all important European markets. AutoScout24 provides a comprehensive platform for online car trading to consumers. Their wide-ranging marketplace offers new and secondhand vehicles including cars, motorcycles, caravans, and utility vehicles. The intelligent search function assists users in locating their ideal vehicle, and the price and car evaluation function enable full visibility of the price for buying or selling autos on the web.

Paruvendu.fr is the leading online marketplace for cars, with millions of vehicles from thousands of sellers. Users can easily filter their searches using a range of criteria including car condition, make and model, body type, colour, price range, and geographical location.

Gumtree is one of the largest British online classified sites, serving communities throughout the country. Gumtree is a platform where users could easily find everything they need. So, users can find a car in this platform easily filtering their searches.

Autotrader is a comprehensive automotive marketplace publishing multiple car listings for sale from numerous sellers. The website offers users the ability to sift through various car criteria, including new, secondhand, models, body style, colour, price, geography, and several other factors. The website also facilitates users in determining the value of their car and advertising it for sale.

CarGurus is an automotive platform providing an opportunity to compare car lists. The platform analyses new and secondhand vehicle listings every day to determine each car’s current market value. Based on prices, it ranks search results with 5 options. It facilitates buyers to obtain car values and sell their cars.

Cars.com helps users to make well-informed decisions about purchasing new or secondhand vehicles. It presents different search options to its customers.

Kelley Blue Book is renowned as a comprehensive automobile research website. KBB.com offers a plethora of valuable resources to assist buyers in conducting research, pricing, and shopping for a vehicle. Alongside the extensive search function for new and secondhand cars, the website also presents The Kelley Blue Book Price Advisor. This feature showcases the Fair Market Range, showing consumers what they should pay according to other purchases in the market.

AutoTempest.com collects results from numerous car shopping websites, enabling simple comparisons with each other. This website claims to have the most extensive collection of vehicle listings.

CarsDirect provides consumers with up-to-date car pricing information to assist them in obtaining a competitive deal on a new or secondhand vehicle. The latest pricing data is evaluated by CarsDirect's team of professionals to identify advantageous deals and pricing trends for the benefit of car buyers.

Craigslist is one of the platforms for classified adverts. Users can find a car in this platform easily filtering their searches.

EBay Motors is a popular auction website that presents cars and trucks classifieds. The eBay search engine offers comprehensive search parameters such as make, model, price, location, and delivery options. In addition, a market price analysis tool reveals how the asking prices stack up against other vehicles' prices on the site. And also, EBay Motors introduces seller ratings that assist potential buyers in making informed purchase decisions.

Edmunds is here to help find the perfect secondhand car. It helps to search for the right vehicle, gives an idea of what to expect to pay for a new car, and even lets to know when we think a used car is a great deal. This website presents a treasure trove of useful information for shoppers.

Carvana is a website that allows customers to browse through various car makes and models, trims, styles, and locations. Users can also trade or sell their vehicles, calculate payment amounts by using the loan calculator service. And also they can apply for credits from the platform.

Autolist helps it users to compare various new and secondhand car websites through a single user-friendly platform. It presents a comprehensive search engine, the website also features popular vehicle searches, and you can give filters according to price, body type, year, and city.

TrueCar is an online platform for buying and selling automobiles, offering over a million secondhand and new cars from certified dealerships nationwide. TrueCar has partnered with 250 firms to provide car-buying schemes. The user-friendly search tool facilitates the screening of automobiles by giving different filters.

Vroom tries to provide an improved car buying and selling experience. The website offers a step bystep guide for car buyers, enabling them to browse online, begin their purchase, upload documents and receive delivery at their doorstep. Shoppers can use a simple and user-friendly search tool to filter cars, trucks and SUVs, all of which are supplemented with Carfax vehicle history reports, and offered at fixed prices, without room for negotiation.

Facebook Marketplace is situated on the Facebook social media platform and functions as a classified advertising website. It allows users to give announcement to sell their cars.The website may not provide all necessary purchase details, but it does give potential buyers an idea of the local market.

ISeeCars.com offers an extensive search tool and provides information on listings with excellent, good and decent deals. It also provides data on vehicle prices in relation to market value, analyses of vehicle conditions, ratings of dealers, and much more. Furthermore, shoppers can browse through popular searches highlighted. Additionally, the website presents informative studies using its data about selling cars.

Cars24.com is an Indian start-up web platform that facilitates the sale of secondhand cars. Its user-friendly interface presents sellers to enter various parameters, such as car model, distance travelled, year of registration, and vehicle type, which enables certain algorithms to be run and predict the price.

CarTrade.com is a web and Android platform which enables users to conduct research on new cars available in India by looking car prices, mileage and reviews. The app also serves as a platform for the easy sale of secondhand cars to genuine buyers. This listing will be displayed to interested car buyers in their city. Potential improved version: Prospective car buyers can make use of both user and expert reviews, accompanied by images, to aid their decision-making process.

Get Vehicle Price is an Android application that utilizes a variety of parameters for predicting the prices of vehicles. These parameters include fiscal power, horsepower, and mileage. The application uses a machine learning method to accurately predicts the prices of cars, bicycles, electric vehicles, and hybrid vehicles. The algorithm of this application has been optimized for smart predictions of vehicle prices.

CarWale is a highly rated car application in India, designed for researching both new and secondhand vehicles. It presents precise secondhand car prices, expert reviews and an efficient car comparison tool. This application efficiently facilitates connection with nearby car dealers to provide the best offers available.

2.6.2. Turkish Secondhand Automobile Trading Sites and Applications

Sahibinden.com was established by the Aksoy Group in 2000. It is not only an e-commerce platform, but also a representative of the evolution of e-commerce in Turkiye, and currently one of the biggest advertising companies in the country. Sahibinden.com provides a variety of products and services across 10 categories to enhance the convenience of its users. These categories include Real Estate, Vehicles, Shopping, Construction Equipment, Spare Parts and Accessories, Services, Job Opportunities, Animal and Pets, Tutors, and Assistant Seekers. So, we can say Sahibinden.com is a platform where users could easily find everything they need. And users can find a secondhand car in this platform easily filtering their searches.

Arabam.com is a highly regarded online marketplace in Turkiye's automobile industry. It is founded in 2000. The Arabam.com aspires to be the foremost, innovative, and dependable technological platform for automotive trade and services. Arabam.com is an online platform for car listings that offers various services including Trink Sell, an instant car buying system, Arabam Kaç Para, a tool for market value inquiries, and My Car/Garage, which enables users to track their vehicle's market value and purchase car-related items at discounted prices.

İkinciyeni.com is Çelik Motor that is one of the most well-established firms within Anadolu Group. Çelik Motor established the ikinciyeni.com brand in 2013 by transferring its expertise and knowledge to the digital realm. The platform facilitates the buying and selling of secondhand vehicles through an auction method. Users can conveniently purchase and sell vehicles without facing any issues. And, it enables users to obtain vehicle price predictions within a few quick steps.

VavaCars aims to modernize the secondhand automotive industry and ensure the secure, effortless, and prompt buying and selling of vehicles. VavaCars is a subsidiary of Petrol Ofisi located in Turkiye and a member of the Vitol global conglomerate, founded over 50 years ago and dedicated to the energy industry.

Carvak is the Turkish brand of Kavak, one of the world's leading secondhand platforms. Kavak was founded in Mexico in 2016 and has managed to become a game-changing company in the secondhand car industry. It is the first investment outside Latin America of the company was in Turkiye in 2022. It aims to turn customers' car buying and selling process into a safe, transparent and fast experience.

Otoplus letgo has popularized secondhand shopping in Turkiye since 2016 by introducing "otoplus", an easy, fast, and safe method of buying and selling secondhands for everyone. They aim to break the mold in the secondhand car industry and offer a range of services under one roof that are necessary for buying and selling secondhands. As the largest and most dependable secondhand car platform in Turkiye, it aims to redefine the process of buying and selling secondhands by ensuring exceptional reliability, on par with purchasing a brand-new vehicle. If you are interested in purchasing a secondhand vehicle or selling your current one, otoplus provides a trustworthy end-to-end service.

Oto.net began operating in 2013, and it is an internet platform where over 1,500 automobiles trade every month by sale, purchase or barter services. Currently, it holds around 6,000 corporate members and almost 1,000,000 registered individual members. Oto.net is Turkiye's primary e-marketplace for used vehicles and has successfully supported more than 85,000 vehicle transactions. Bringing their unparalleled expertise in all aspects of buying and selling vehicles to the retail automotive market, Oto.net provides users with a new and innovative experience when it comes to purchasing, selling, or exchanging secondhand vehicles through their online platform.

Otonomi.com launched in 2016 as a platform for advertising new and secondhand motor vehicles. The platform aims to facilitate the buying and selling of secondhand vehicles. It provides users with a new and innovative experience when it comes to purchasing, selling, or exchanging secondhand vehicles through their online platform.

2.7. Reasons of Prefering Secondhand Automobile

There are several reasons why the secondhand car market is more extensive than that of new cars. Some of crucial causes are:

1. Cost (Secondhand vehicles can be purchased at a reduced expense according to their brand-new ones).
2. Taxation on new vehicles is higher compared to secondhand vehicles. (As a result, some people think it is more advantageable to purchase secondhand vehicles instead of brand-new ones. As a car's age increases in Turkiye, its tax decreases. Consequently, secondhand cars attract lower tax than new ones.)
3. Secondhand car resale (People who purchase new cars face an important reduction in price while resale, whereas there is little price differences for secondhand cars. So, someone can sell his/her car with greater profitability according to market conditions.)
4. The overall economic climate affects the demand for secondhand cars, with changes in both welfare levels and interest rates having an impact.
5. Access to diverse secondhand vehicle models; (Acquiring your desired brand-new vehicle model may pose a challenge due to limited availability. In addition, custom equipment installation and upgrade options may prolong the delivery process. Conversely, choosing to purchase a used vehicle grants you access to numerous models across various brands, allowing you to select freely.)

3. LITERARURE REVIEW

Many studies were done previously to estimate a secondhand car price using various algorithms and models. So, it can be said that there are many various algorithms to predict the selling price of a vehicle. As a result of our literature review, it is possible to summarize some of these studies as follows.

Murray and Sarantis (1999) have studied to forecaste the selling price of secondhand cars in the UK. They used different vehicle characteristics to understand the price differences among the different secondhand car models in UK. In this research, they aimed to create a price index for automobiles by using data on UK car characteristics.

Erdem and Senturk (2009) have studied with regression techniques to find the factors that influence secondhand car prices in Turkiye. Box-Cox transformation, log-linear and Semi-log methods are used in this study for 1074 secondhand data sets. At the end of the study, they found diesel engine, colors of the car, transmission type , sunroof, production location, production year and engine volume directly affects secondhand car prices.

Ecer (2013) has studied to find the factors that are affecting the price of secondhand cars. Therefore, she used two different algorithms. These are ANN and Hedonic model. By using hedonic model, the car features are determined that affects the prices of secondhand cars. According to hedonic model results, brand, model, engine power, model year, fuel type , mileage are important factors for the determination of a car selling price. And ANN was used as an alternative to hedonic model in order to estimate the selling price of secondhand cars.

Dastan (2016) has studied to find the factors that are affecting the secondhand car selling prices. Therefore, he used the data by getting from secondhand car advertisements on websites. He studied to analyze the effects of car characteristics on selling price of cars. He used the following car characteristics in his study; the brand, model, model year, front view camera, rear window defroster ,mileage, gear, fuel type, torque, fuel tank volume, ABS, panoramic glass roof, power steering, start/stop, sunroof, cooled torpedo.

Noor (2017) has tried to create a system to estimate vehicle price by using the supervised machine learning technique in their study. They used multiple linear regression method that provides 98% prediction precision. This study offers a system that predict the selling price of a secondhand car by using following factors such as vehicle brand, model, city, color , version, mileage, alloy wheels and power steering.

Pudaruth (2014) has studied to predict prices of secondhand cars in Mauritius. He used multiple linear regression , k-nearest neighbors , naive bayes tand decision tree to estimate the selling prices of secondhand cars. These models are compared to to find which presents the best performance. The study did not give precise results due to low number of data that have been used.

Gegic et al., (2019) have studied to predict selling prices of secondhand car in Bosnia by using three different machine learning techniques. ANN, Random Forest and Support Vector Machine models are used in the study. Data is taken from a local Bosnian website for 797 secondhands. Random Forest model has been seen the best model with 87.38% accuracy.

Monburinon et al., (2018) have studied to predict selling prices of secondhand by using regression models. The data is taken from a German e-commerce web site in this research. The purpose of this study is to develop a system to estimate the secondhand car price. They used different machine learning methods for comparison. And they tried to find the best model according to results.

Asghar et al., (2021) have studied to find an efficient way to help the sellers and buyers who estimate the best for their vehicle. In the study, one of the machine learning algorithm is regression was used. The study results show the model helps buyers and sellers to give a better decision about selling price of the secondhand car.

In the study of Özçalıcı (2017) the selling price of secondhand cars was tried to be estimated by using the data set obtained from the secondhand car market. The selling price of secondhand vehicles is affected by many factors. Which of these factors will successfully predict the selling price is determined by the genetic algorithm. The study showed that t the model year and engine power variables are more successful in estimating the price than other variables.

Çelik and Osmanoğlu (2019) have studied to develop a system that helps buyers and sellers to give a better decision about the selling price of their cars. They studied with linear regression and machine learning methods to estimate the selling price of cars in terms of some car characteristics. This study offers a system that forecaste the selling price of a secondhand vehicle by using following factors such as model, brand, mileage , package type, city, color, wheels and motor power.

Gajera et al., (2021) have studied with machine learning algorithms such as XG boost, KNN, linear-regression, Decision tree and Random Forest to build a statistical model to predict the selling price of a secondhand car. They used past data of consumers. They compared the results of models to find the best method for prediction of second hand car prices. They found Random Forest model is the best model with highest R- squared value: 0.93.

Bilen (2021) has targeted to find the best estimation model by using heuristic algorithms for prediction of a secondhand car. The data was taken from the advertisements from a classified website. The study was done with various methods and models on the dataset . As a result of study, the best succesful value observed with R2 was 0.8958 for prediction of a secondhand car.

Venkatasubbu and Ganesh (2019) have studied Multiple Regression algortihm , Lasso Regression algorthim, Regression tree algorthim in order to develop a system which will estimate the price of a secondhand vehicle, according to past consumer data. They compared the results of models to find the best method for prediction of second hand car prices.

AlShared (2021) has studied to estimate secondhand car prices by using some attributes. In this study, Linear Regression, Random Forest Regressor, Bagging Regressor have been used. the Random Forest Regressor had the best score at 95%. among all the models.

Samruddhi and Kumar (2020) have studied their model with secondhand cars data in order to forecaste the selling price of the secondhand cars.They have used K nearest Neighbor algorithm for prediction selling price of secondhand cars in this research. The results shows that the used model is predict the selling price of secondhand cars 85%.

Asilkan and Irmak (2009) have studied to predict the future prices of the secondhand cars by using ANN which is a very popular forecast technique in the recent years. The data of the study is taken from Europe-based secondhand automobiles web sites. The results obtained by ANN have been compared to other forecasting algorithms. As a conclusion, ANN was found successfully to forecast the future prices of the secondhand cars.

After detailed literature review it is showed that forecasting selling price of secondhand cars is a popular topic for many researchers in recent years. When the studies published in the literature are reviewed, it is seen that different models are studied to forecaste the selling price of secondhand cars. In this study, in addition to the algorithms in the literature, secondhand car sales price estimation will be performed with the help of the following algorithms: ANN, ARIMA, SARIMA, XGBOOST, TBATS, HOLT WINTERS Forecasting.

4. METHODOLOGY

In this section, we will look at the models which were used in order to predict the future six months sales prices of a secondhand car by using some factors such as the vehicle's brand, model, package type, model year, body type , fuel type, gear type and mileage. We will analyze six different algorithms (ANN, ARIMA, SARIMA, XGBOOST, TBATS, HOLT WINTERS Forecasting) was be studied in this research for this part.

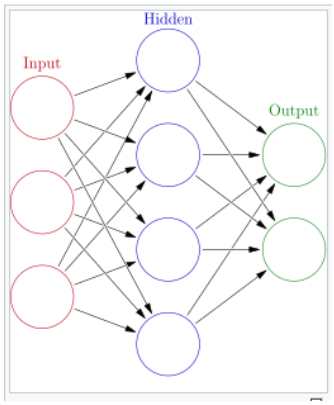
4.1. Machine Learning Algorithms

4.1.1. ANN

Artificial Neural Networks (ANN) are powerful algorithms that models complex samples and predict problems with great accuracy. It's based on the idea of biological neural networks in the human brain. ANN is a fascinating deep learning method that helps us understand how the brain works. While not identical, the operation of ANN is like that of biological neural networks. And it is significant to say that the ANN algorithm only accepts numerical and structured data.

An ANN is made up of little nodes, or artificial neurons, that try to do what our own neurons in our brains do. Each connection is like a synapse in our brains, sending signals to other neurons. An artificial neuron takes in the signals, processes them, and sends them on to other connected neurons. The signal at each connection is a number, and each neuron'output is calculated using a non-linear function of the sum of its inputs.Neurons are connected through edges, which can be evaluated during the learning process to increase/decrease the strength of the signal. Additionally, neurons may have a threshold that must be surpassed for a signal to be transmitted. In this diagram, each circle shows an artificial neuron, and each arrow shows a connection from the output of one neuron to the input of another.

Figure 1: ANN Use Case Model



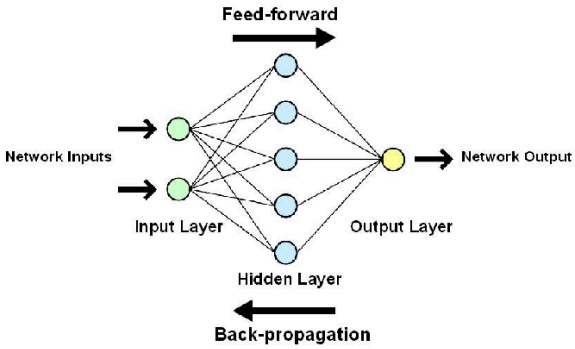
So, it can be stated that an ANN is a group of interconnected nodes that can find connections, make estimations, and solve various targets in machine learning and artificial intelligence. The connections' strength (weights) adjusts in training to take from the data, giving the network to operate with authority and expertise.

4.1.1.1. Architecture of ANN

The architecture of ANN can be listed as follows:

1. The network architecture is comprised of three layers: the input layer, the output layer and the hidden layer(s). This is a Multi-Layer Perceptron (MLP) architecture due to its multiple layers.

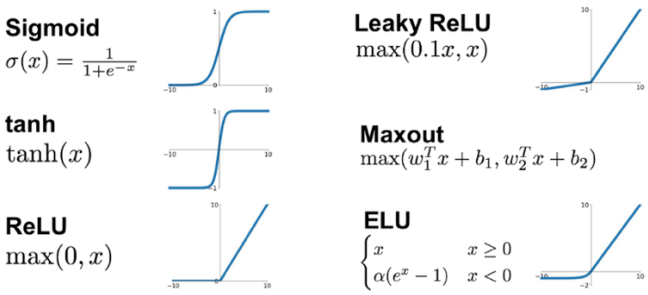
Figure 2: ANN Architecture



2. The hidden layer can be seen as a 'distillation layer' that takes the most connected insights from the inputs and sends them to the next level for detailed analysis.
3. The activation function plays a crucial role for two reasons.

- Firstly, it effectively understands the non-linear relationships between the inputs.
- Secondly, it efficiently transforms the input into a better output. There are a few activation functions.

Figure 3: ANN activation function



4. Finding the most suitable values of W weights is crucial for creating a successful model.

5. The optimization method uses a 'gradient descent' technique to calculate estimation errors. The optimal value for W is determined by making small adjustments and analyzing the effect on prediction error. The chosen W values are considered ideal as further changes do not reduce errors.

4.1.1.2. Advantages of ANN

ANN has a few key advantages that it suitable for specific problems and situations:

1. ANNs can learn and connect non-linear and complex interactions. This is significant because many of the connection between inputs and outputs are non-linear and complex. This makes ANN a very effective tool for non-linear statistical data analysis.
2. ANNs can applied in many areas. After learning from the original inputs, the model can be applied with other anonymous data.
3. ANN does not affect any constraints on the input variables (such as how they should be distributed), unlike many other forecasting approaches. In addition, many studies have shown that ANNs are better at simulating heteroskedasticity, high volatility data and nonconstant variance, due to their ability to find connections in the data without affecting any predetermined parts. This technique is highly effective in predicting financial analysis (such as stock prices) where data volatility is significant.
4. Problems are stated by using value pairs attribute.
5. While the target function can also be discrete or real, ANNs can output discrete or real features.
6. Problems in the training data does not affect the result of ANN learning techniques. The errors in the training samples are insignificant.
7. This method is ideal for situations where a quick evaluation is needed.
8. The time it takes to train ANNs can be affected by various factors.

4.1.1.3. Disadvantages of ANN

1. Constructing ANNs requires the use of two processors due to hardware dependency.
2. To understand how the network works, it's important to note that when an ANN responds, it doesn't provide an explanation for why or how it was selected. This can sometimes undermine trust in the network.
3. Network Structure Security. There is no rule for finding the structure of ANNs. With some experience, a suitable network structure can be developed.
4. Presenting problems to ANNs can be challenging as they require numerical data. Therefore, it is important to transform problems into numerical values before introducing them to the ANN. The method of representation chosen will directly impact the network's performance.
5. The lifetime of the network is uncertain. The training process concludes when error number of sample reaches a specific value.

4.1.1.4. Applications of ANN in Literature

ANNs are useful mathematical models that improve existing data analysis technologies. It is used in many areas such as spam detection in email, predictive analysis in business intelligence and natural language processing in chatbots. ANNs are also essential for Machine Learning (ML) and support Artificial Intelligence (AI) technology.

ANNs are being used more and more in different areas of modern life thanks to all the amazing technological advancements. They are used in so many different fields, including image processing, simulation, disease diagnosis, drug development, and market analysis. They are even employed in the military for predicting social events and in the automotive industry for price prediction and production methods. Additionally, they are used in robotics for circuit design and energy production for efficiency and policy. It's incredible to think that the use of ANNs is not limited to these areas. Simulation technologies are particularly important in space science, while stock value prediction is crucial in banking and finance. In the field of computer science, there are various applications, including energy production, energy efficiency, and energy policy. Other uses of stock value prediction include product price estimation, economic crisis prediction, management games, and exchange rate estimation. (Li et al., 2020)

ANNs are widely used in social media. For instance, Facebook's 'people you might know' feature recomends individuals you may know in real life.This is done by analyzing your profile byusing ANNs. Facial recognition is an interesting application of machine learning in social media. It is done by using ANN again.

E-commerce sites like Amazon and Zomato use personalized marketing to recommend products and restaurants based on users' browsing and order history. This practice is prevalent in various marketing segments, such as book sites, movie services, and hospitality sites. ANN identify the customer's preferences, past purchase history and other relevant information. This data customizes marketing campaigns to better suit the customer's needs.

ANNs are amazing tools used in oncology to improve algorithms that can determine cancerous tissue with the same accuracy as trained doctors. Additionally, facial analysis of patient photographs can be used to identify different rare diseases that manifest themselves in physical characteristics in their early stages. The use of ANNs for health services can be very beneficial for improving diagnostic processes.

These technologies have gained immense popularity and are now commonly found on mobilephones. For example, Siri, Alexa, and Cortana use speech recognition and natural language processing to connect with users and provide responses. Natural Language Processing (NLP) utilizes ANNs to manage language syntax, semantics, speech correction, and conversation. With NLP, you can handle many of the tasks of personal assistants with ease.

Asilkan and Irmak (2009) has planned to estimate the future prices of secondhand cars using ANNs. The input of the study has been taken from Europe based secondhand car websites. The results were compared with other time series forecasting methods. As a result, it has been shown that the future prices of secondhand cars can be successfully predicted by using ANNS.

Varshitha et al., (2022) has studied to create a model for secondhand car price estimation with low error value. They wanted to forecast the price of a secondhand car without any doubt for the buyer or seller. Therefore, they used ANNs.

Karakoçı et al., (2020) have designed an ANN for car price prediction and tested it using data from a a car classified website. They chose the ANN model because ANNs are often used for such calculations.

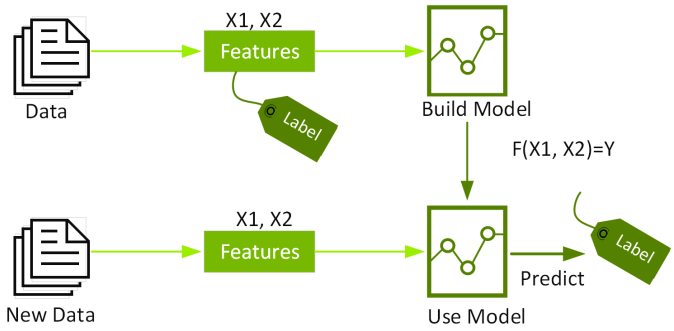
4.1.2. XGBoost

XGBoost is the most widely used and efficient gradient boosting algorithm in the field of data science. Developed by Tianqi Chen as an enhancement to the GBM algorithm, it is highly flexible and portable. The algorithm was designed with the goals of high efficiency, flexibility, and portability in mind, making it the top choice for data scientists. XGBoost outperforms other machine learning algorithms in various tasks, including classification, regression, and ranking.

XGBoost is a library for machine learning with regression, classification, and ranking challenges. It is a distributed and scalable library used for gradient-boosted decision trees (GBDT) in machine learning. XGBoost boosts trees in parallel, making it a highly efficient and effective tool for GBDT. A foundational understanding of the algorithms and concepts that form the base of XGBoost is necessary to comprehend it. These include decision trees, ensemble learning, gradient boosting and supervised machine learning.

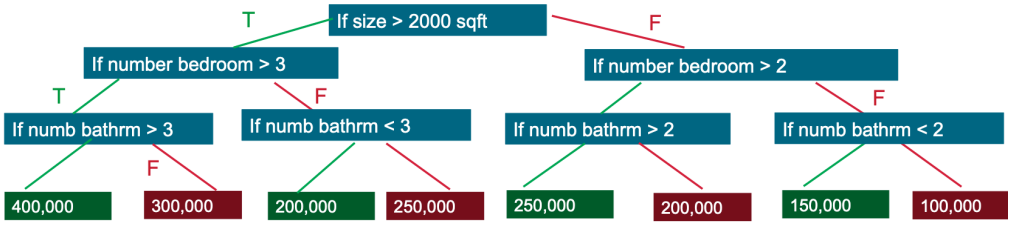
Supervised machine learning algorithms study a model to identify patterns in a labelled dataset's features, which enables the model to estimate new dataset labels based on their features with confidence.

Figure 4: XGBoost Features



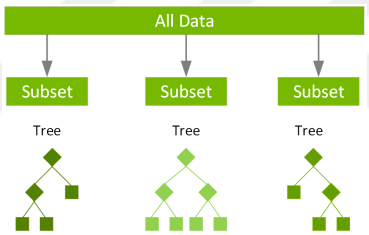
Decision trees predicts by evaluating a true/false feature tree of if-then-else questions. Decision trees are a powerful tool for category prediction through classification or continuous numeric value prediction through regression.

Figure 5: A decision tree example



Gradient Boosting Decision Trees (GBDT) is a powerful ensemble learning algorithm used for classification and regression, which outperforms any individual algorithm. It combines multiple machine learning algorithms to create a model that is highly accurate and reliable. While both random forest and GBDT construct a model using multiple decision trees, GBDT is known for its superior performance and is widely used in various applications. The distinction between the two arises solely from the methodology used to construct and integrate the trees.

Figure 6: A Decision tree methodology



The random forest algorithm employs a bagging technique to cnstruct multiple decision trees simultaneously. Each one is randomly coming from the dataset. The estimation is the average estimation made by all the decision trees.

Gradient boosting is an extension of the boosting technique that formalizes the process of creating inefficient models. The aim is to reduce errors by setting targeted results for the next model. The error gradient, hence, the name gradient boosting, determines the outcomes for each case in relation to the prediction.

GBDTs train decision trees iteratively, using the errors of the preceding model to set the following model. The final forecast is the weighted sum of all tree estimations. Thus, GBDT 'boosting' reduces bias and underfitting.

4.1.2.1. Architecture of XGBoost

XGBoost is a pretty clever algorithm that works by combining lots of little learners together to create one big, strong learner. A weak learner is a machine learning model that does a pretty good job of guessing, but not quite as well as it could. When weak learners are used together, they evolve into a strong learner that produces significantly more accurate results. XGBoost trains multiple decision trees, with each tree trained on a distinct subset of data. The estimations from each tree are combined to form the final forecast. It might be said that XGBoost represents an improvement over the GBM algorithm, as it employs a more regularised model to help prevent overfitting.

Figure 7: A generic unregularized XGBOOST algorithm

Input: training set $\{(x_i, y_i)\}_{i=1}^{N}$, a differentiable loss function $L(y, F(x))$, a number of weak learners M and a learning rate α.

Algorithm:

1. Initialize model with a constant value:

$$\hat{f}_{(0)}(x) = \arg\min_{\theta} \sum_{i=1}^{N} L(y_i, \theta).$$

2. For m = 1 to M:

   1. Compute the 'gradients' and 'hessians':

$$\hat{g}_m(x_i) = \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]_{f(x)=\hat{f}_{(m-1)}(x)}.$$

$$\hat{h}_m(x_i) = \left[\frac{\partial^2 L(y_i, f(x_i))}{\partial f(x_i)^2}\right]_{f(x)=\hat{f}_{(m-1)}(x)}.$$

   2. Fit a base learner (or weak learner, e.g. tree) using the training set $\left\{x_i, -\frac{\hat{g}_m(x_i)}{\hat{h}_m(x_i)}\right\}_{i=1}^{N}$ by solving the optimization problem below:

$$\hat{\phi}_m = \arg\min_{\phi \in \Phi} \sum_{i=1}^{N} \frac{1}{2} \hat{h}_m(x_i) \left[\phi(x_i) - \frac{\hat{g}_m(x_i)}{\hat{h}_m(x_i)}\right]^2.$$

$$\hat{f}_m(x) = \alpha \hat{\phi}_m(x).$$

   3. Update the model:

$$\hat{f}_{(m)}(x) = \hat{f}_{(m-1)}(x) + \hat{f}_m(x).$$

3. Output $\hat{f}(x) = \hat{f}_{(M)}(x) = \sum_{m=0}^{M} \hat{f}_m(x)$.

XGBoost has several adjustable parameters that can enhance the algorithm's performance. Here are the most significant parameters:

- n_estimators: Number of trees we want to grow sequentially.
- max_depth: The decision trees' maximum depth.
- η: The learning rate.
- γ: The minimum loss reduction.
- subsample: The fraction of the training data.

4.1.2.2. Advantages of XGBoost

1. XGBoost has gained immense popularity owing to its remarkable ability to help individuals and teams emerge victorious in data competition in recent years. In these competitions, participants compete to create the most effective models for forecasting and analyzing the data provided by various companies and researchers.
2. XGBoost was initially implemented in Python and R, but its popularity has led to package implementations for some programing language like Java, Scala, Julia, Perl etc. These abilities have made the XGBoost library more preferable to users.
3. The algorithm is precise and efficient with diverse data sets and issues. It is widely accepted as the standard for accurate predictive modelling results through machine learning due to its excellent documentation and convenience. This library is specifically designed for R, Python, and C++ and provides the fastest gradient-boosting technique with a high level of precision.
4. XGBoost has been seamlessly integrated into Python and R. Furthermore, it has been successfully integrated with distributed processing frameworks such as Apache Spark and Dask. Notably, XGBoost was awarded the prestigious Technology of the Year award by InfoWorld in 2019, cementing its position as a leading machine learning algorithm.
5. And the benefits of XGBoost are numerous and comprehensive, encompassing the following characteristics:
    1. XGBoost open-source development is supported by a large and growing global community of data scientists.
    2. It is used in many applications, including regression, classification, ranking, and user-defined prediction challenges.
    3. The library can run on OS X, Windows, and Linux platforms.
    4. It also supports cloud integration with AWS, Azure, Yarn clusters, and other ecosystems.
    5. The library is being used by many organizations across different industries.
    6. It was designed to be efficient, flexible, and portable.

4.1.2.3. Features Supported by XGBoost

Figure 8: Features Supported by XGBOOST



1. **Regularization:** XGBoost models can be quite complex and involve different forms of regularization, such as Lasso and Ridge, to impose penalties on the models.
2. **Handling sparse data:** XGBoost is great at handling sparse data, which means you don't have to worry about missing values.
3. **Block structures and parallel processing:** XGBoost stands out from other machine learning algorithms due to its ability to use multiple CPU cores simultaneously, thanks to its block structure in its system design. This unique feature allows XGBoost to work exceptionally quickly, making it a top choice for data scientists and machine learning practitioners.
4. **Cache awareness:** XGBoost was specifically designed with hardware optimization in mind. The algorithm allocates buffer memories at each level, utilizing the cache very well. Furthermore, for big data issues with very large datasets, the algorithm can compress this data into smaller versions, optimizing both the disk space and computational speed. This feature is commonly referred to as 'out of core' computation.
5. **Built-in cross-validation:** The XGBoost algorithm cross-validates models during development, reducing the risk of overfitting and maintaining a balance between bias and variance.
6. **Tree pruning:** XGBoost generates the maximum depth specified and prunes the tree by deleting splits that do not result in any positive gain. This approach prevents overfitting and ensures that XGBoost is not a greedy algorithm.

4.1.2.4. Disadvantages of XGBoost

1. **Complexity:** XGBoost is a complex algorithm that necessitates a certain degree of technical expertise to implement and optimise effectively. The configuration and tuning of the numerous hyperparameters involved can prove challenging, potentially leading to prolonged working times.
2. **Overfitting:** Although XGBoost incorporates regularisation techniques to prevent overfitting, it is still possible for the algorithm to overfit the training data. This can result in inaccurate predictions of new data.
3. **Memory usage:** XGBoost is known to be memory-intensive, particularly when dealing with large datasets. This can result in difficulties when attempting to run with less memory, which may lead to a reduction in performance.
4. **Less transparency:** XGBoost has frequently been regarded as a "black box" algorithm, which makes it challenging to interpret and comprehend the underlying rationale behind its predictions. This can make it challenging to identify and rectify any issues that may arise.

4.1.2.5. Applications of XGBoost In Literature

Ji, Wang et al., (2019) have studied a XGBoost forecasting model for the selling features of products. They have used different machine learning models to compare the results. The results showed that the XGBoost is outperformed than other models.

Gumus and Kiran (2017) studied to estimate the features by using XGBoost which are influcing the crude oil prices. They targeted to provide savings by predicting the crude oil price.

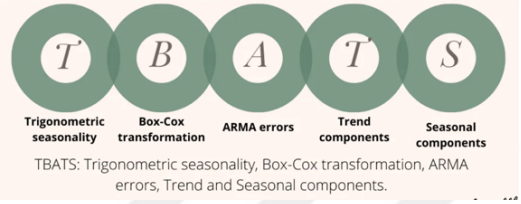
Niu (2020) has come up with a brilliant XGBoost sales prediction model to help solve Walmart's sales problem. He used sales data from Walmart supermarkets, which he got from the Kaggle competition. The results are in, and they show that this method really shines compared to other machine learning approaches.

4.2. Forecasting Techniques

4.2.1. Tbats

TBATS is one of the most popular forecasting methods in recent years. The aim of this method is to provide accurate forecasts through the implementation of exponential smoothing. The TBATS models are highly effective in managing data with numerous seasonal patterns, even when the seasonality varies over time. This model can accommodate intricate seasonality demonstrations without any restrictions, making it possible to generate elaborate long-term projections with confidence.

Figure 9: The key features of TBATS



The key features of the model are represented by TBATS:

- T: Trigonometric seasonality
- B: Box-Cox transformation
- A: ARIMA errors
- T: Trend
- S: Seasonal components

Time series forecasting is a crucial aspect of data science, applied statistics, and econometrics. Its primary target is to accurately forecaste future values of time series datasets, which are composed of regularly recorded variables and possess two essential properties.

- Trend - The long-term average direction of a variable.
- Seasonality- Fluctuation that occurs at specific time periods.

Forecasting models can be challenged by the complex seasonality that time series data can exhibit. Furthermore, the absence of detectable trends or seasonality

Moreover, you may encounter time series datasets that do not display any noticeable trend or seasonality, rendering them unsuitable for forecasting models that rely on these features. However, there is a solution to this problem: the TBATS forecasting model. This model is adept at handling intricate seasonality patterns and delivers unparalleled accuracy for such datasets.

The TBATS method applies the Box-Cox transformation to time series data and models it using an amalgamation of an exponentially smoothed trend, a seasonal component, and an autoregressive moving average (ARMA) component. Hyperparameter tuning is conducted using the AIC metric to eliminate any suboptimal components from the final model.

4.2.1.1. Architecture of Tbats

TBATS fits several models and considers various alternatives to determine the best fit. The final model will be created using Akaike information criterion (AIC). If what p and q values are suitable and residuals need modeling, ARMA is used to decide about final model.

Figure 10: Model of TBATS

Model:

$$y_t^{(\lambda)} = l_{t-1} + \phi\, b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t$$

$$l_t = l_{t-1} + \phi b_{t-1} + \alpha d_t$$

$$b_t = \phi\, b_{t-1} + \beta\, d_t$$

$$d_t = \sum_{i=1}^{p} \varphi_i d_{t-i} + \sum_{i=1}^{q} \theta_i e_{t-i} + e_t$$

Seasonal part:

$$s_t^{(i)} = \sum_{j=1}^{(k_i)} s_{j,t}^{(i)}$$

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos(\omega_i) + s_{j,t-1}^{*(i)} \sin(\omega_i) + \gamma_1^{(i)} d_t$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin(\omega_i) + s_{j,t-1}^{*(i)} \cos(\omega_i) + \gamma_2^{(i)} d_t$$

$$\omega_i = 2\pi j / m_i$$

Where:

$y_t^{(\lambda)}$ - time series at moment t (Box-Cox transformed)

$s_t^{(i)}$ - ith seasonal component

l_t - local level

b_t - trend with damping

d_t - ARMA(p,q) process for residuals

e_t - Gaussian white noise

Model parameters:

T - Amount of seasonalities

m_i - Length of ith seasonal period

k_i - Amount of harmonics for ith seasonal period

λ - Box-Cox transformation

α, β - Smoothing

ϕ - Trend damping

φ_i, θ_i - ARMA(p, q) coefficients

$\gamma_1^{(i)}, \gamma_2^{(i)}$ - Seasonal smoothing (two for each period)

4.2.1.2. Advantages of Tbats

1. **Multiple seasonality:** One of the principal advantages of TBATS models is that they can be employed to model datasets that encompass a multiplicity of disparate seasonal trends. This implies that they can be utilised to analyse datasets exhibiting both daily and hourly trends.
2. **Does not require stationary data:** It can be posited that TBATS models represent an extension of a basic exponential smoothing model. In a manner analogous to basic exponential smoothing models, TBATS models do not presume that the underlying data is stationary. This implies that they

can be employed even when the data is not fix and cannot be readily transformed into a fix state through the application of differencing.

3. **Does not require linear relationships in data:** Another advantage of TBATS models is that they do not necessitate the dependencies between different measurements to be linear. A variety of transformation techniques can be employed during the modelling process to accommodate scenarios where the relationships between successive values of a variable are not strictly linear.
4. **Highly flexible:** Another advantage of TBATS models is that they are highly flexible and can be extended to model a wide range of different types of time series data. A multitude of parameters can be modified to adapt the model framework to specific datasets.

4.2.1.3. Disadvantages of Tbats

1. **Can not incorporate covariates:** One disadvantage of TBATS models is that it is not straightforward to incorporate additional covariate information into TBATS models. Consequently, TBATS models are ill-suited to scenarios where the inclusion of additional covariates is necessary for the accurate prediction of the outcome variable.
2. **Not as well studied:** A further disadvantage of the TBATS models is that they are relatively niche and not well understood by practitioners who do not have experience working with time series data. Consequently, it is less straightforward for colleagues to provide feedback or contribute to projects that utilise TBATS models. Prior to their involvement, colleagues must first dedicate time to acquiring an understanding of the TBATS models.
3. **Difficult to explain:** TBATS models can be relatively difficult to explain. The model is susceptible to a multitude of transformations and modifications, rendering its explanation a challenging endeavour.
4. **Many hyperparameters:** Given the multitude of potential transformations and modifications that can be applied to TBATS models, it is evident that a multitude of hyperparameters and configuration options must be considered. This implies that additional time may be required for model optimisation and the identification of an appropriate model configuration for the given dataset.
5. **Computationally complex:** Given the multitude of hyperparameters and configurations that can be set, it is evident that arriving at a final model may prove to be a computationally intensive process. Many implementations of TBATS models automatically select the optimal model configuration for the user by attempting to identify the most effective model through a process of trial and error, whereby the performance of each model is evaluated. This process is inherently time-consuming and resource-intensive.

6. **Only appropriate for univariate time series:** The TBATS models are designed for univariate time series analysis, and thus are not suitable for joint modelling of multiple time series.
7. **Can be sensitive to outliers:** As with basic exponential smoothing models, TBATS models are susceptible to the influence of outliers. This is particularly the case if the outliers occurred relatively recently, prior to the generation of the forecast.
8. **Can be sensitive to mean shifts:** Similarly, TBATS models are susceptible to significant alterations in the mean value or disruption in the data. This is particularly the case if the mean shift or disruption in the data occurred relatively recently, prior to the generation of the forecast.
9. **Common implementations cannot handle missing values natively**: It can be challenging to address the issue of missing data. The majority of current implementations are unable to handle missing data in a robust manner.

4.2.1.4. Applications of Tbats In Literature

Fajar (2021) has tried to estimate the daily price of chili for controlling inflation by using TBATS. The data is taken from a website. The daily prices data of red curly chili, big red chili, red cayenne small chili, and green cayenne small chili are used in the study. The results showed that TBATS forecasting using chili price data is highly accurate.

Taş (2018) has analyzed hourly electricity prices in Turkiye by using TBATS. TBATS model was used to look the linear behavior of the electricity price.

4.2.2. Arima

The Autoregressive integrated moving average (ARIMA) is a widely used time series forecasting technique that utilizes past values to confidently predict future ones. ARIMA model is frequently applied in statistics and econometrics to accurately measure events occurring over a given period. Its purpose is to comprehensively analyze historical information or confidently anticipate future data within a given data series. It is the ideal choice when a measurement is registered at regular intervals, ranging from microseconds to daily, weekly, or monthly periods.

ARIMA model generates a linear equation that describes and predicts data in a time series. It consists of three components: auto-regressive, integrated, and moving average.

- AR (Auto regression): It is derived from historical data points.
- I (Integration): Based on the trend in the data.
- MA (Moving average): The error or noise derived from historical data.

ARIMA modelling combines the AR and MA elements from individual models. The addition of differencing (the 'I') enables both uncomplicated analyses. ARIMA is a powerful model which utilizes data to have a better insight of the dataset or predict future trends.

ARIMA models are the most popular class of models for predicting time series data. The model is represented as ARIMA (p,d,q), where p is the autoregressive model order, d is the degree of differencing, and q is the moving-average model order. It is important to define technical term acronyms when first used. ARIMA models differentiate to transform not fix time series into fix ones and then accurately forecast future values based on past data. These models predict future values by using auto correlations and moving averages of residual errors in the data.

The ARIMA model components are represented by standard notation, with each component functioning as a parameter. To indicate the type of ARIMA model used, the notation should be ARIMA with p, d, and q, where integer values replace the parameters. The parameters are defined as follows:

- p: the delay order.
- d: the level of differenced.
- q: the number of the moving average

4.2.2.1. Architecture of Arima

ARIMA(p,d,q) models are widely used in econometrics such as stock prices, demand predictions, and the spread of infectious diseases. This approach allows for accurate and reliable predictions, making ARIMA models a popular choice among econometricians. ARIMA would be applied to a differenced time sequences of order d, ensuring that the final differenced time series is stationary.

Forecasting a stationary time series is a breeze. It maintains a constant mean, variance, autocorrelation, and other statistical properties over time, which makes it a simple process. The statistical characteristics of the series are predicted to remain unchanged in the future, as compared to past observations. It is essential to comprehend the three components of an ARIMA model's name to gauge its functionality.

1. **AutoRegressive**: The AR(p) model is a regression model that confidently utilizes y values as predictors observation with previous data. The model comprises of a constant (c), parameters (φs), and white noise (ε) at time t.

$$\hat{y}_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \epsilon_t$$

2. **Integrated I(d):** It uses different times (d) until it accesses its stability. Integrated model of ARIMA is;

$By_t = y_{t-1}$ where B is called a backshift operator

Thus, a first order difference is written as

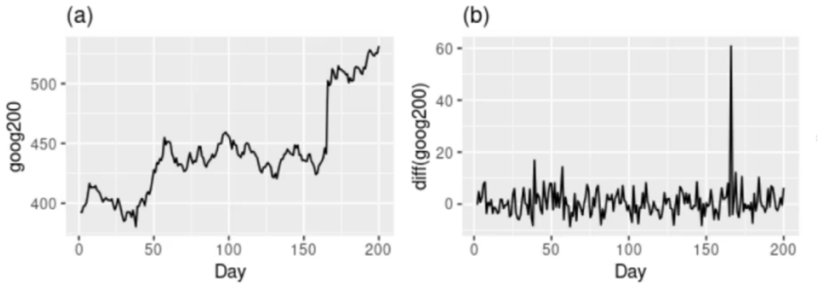
$y'_t = y_t - y_{t-1} = (1 - B)y_t$

In general, a *d* th-order difference can be written as

$y'_t = (1 - B)^d y_t$

When we look at two diagrams from Forecasting: Principles and Practice (2nd ed) by Rob J Hyndman and George Athanasopoulos. Figure A displays the non-stationary time series of Google's stock prices over 200 consecutive days. These diagrams provide clear evidence of the importance of understanding the stationarity of time series data in forecasting. Diagram A displays the non-stationary time series of Google's stock prices over 200 consecutive days. However, diagram B displays the stationary daily fluctuations in Google's stock price over the same 200 days, which are not affected by the time of observation. A first-order difference is necessary as the initial series is not stationary.

Figure 11: The stationarity of time series data in forecasting



3. **Moving average MA(q)**: A moving average model is a great way to use a regression-type model on past forecast errors. ε is white noise at time t, c is a constant and θs are parameters.

$$\hat{y}_t = c + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_q \epsilon_{t-q}$$

Combining all three types of these models gives use the ARIMA(p,d,q) model.

$$\hat{y}'_t = c + \phi_1 y'_{t-1} + \phi_2 y'_{t-2} + \ldots + \phi_p y'_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_q \epsilon_{t-q} + \epsilon_t$$

The ARIMA methodology is a statistical technique that helps us to accurately represent a time series by modelling the correlations in the data. ARIMA models are really clever because they use historical data to generate forecasts. This means they can be more accurate while still being simple to understand.

ARIMA models have several potential drawbacks, despite their parsimony. However, it is worth noting that the subjectivity involved in identifying the p and q parameters is a significant consideration. While

autocorrelation and partial autocorrelation are useful tools, the choice of p and q often depends on the skill and experience of the model developer. Additionally, ARIMA models are more complex than simple exponential smoothing and the Holt Winters method, which may result in less explanatory power.

ARIMA models are ideal for short-term estimation using time series data. The process of finding the optimal set of parameters for each application may require some experience and experimentation.

4.2.2.2. Advantages of Arima

1. **Flexibility:** ARIMA models are capable of accommodating a diverse array of temporal patterns, encompassing linear and nonlinear trends, seasonal patterns, and irregular fluctuations.
2. **Forecasting Accuracy:** When you use ARIMA models correctly, you can be sure you'll get accurate and reliable forecasts, especially for short- to medium-term predictions.
3. **Interpretability:** ARIMA models facilitate a transparent understanding of the influence of past observations and error terms on the present value of the time series.

4.2.2.3. Disadvantages of Arima

1. **Sensitivity to Parameter Selection:** The efficacy of ARIMA models is contingent upon the accurate selection of the parameters p, d, and q. Inappropriate parameter selection may result in suboptimal forecasts.
2. **Assumption of Linearity:** ARIMA models posit a linear relationship between the time series and its past observations. In the event that the relationship is non-linear, alternative models such as ARIMA-GARCH or nonlinear models may be more appropriate.

4.2.2.4. Applications of Arima in Literature

Autoregressive Integrated Moving Average (ARIMA) models have numerous applications across various industries. For example, ARIMA models are generally used in demand forecasting. ARIMA models are also effective in predicting future stock prices based on past prices.

Guha and Bandyopadhyay (2016) created an ARIMA model to predict the future price of gold in India, providing guidance to investors on when to buy or sell gold. The ARIMA model was chosen due to its suitability for time-series analysis in statistics.

Abonazel and Abd-Elftah (2019) have tried to create an ARIMA model for the Egyptian GDP data. Egypt’s annual GDP data for the years 1965 to 2016 is taken from the World Bank. Based on the fitted ARIMA model, Egypt's GDP for the next ten years was confidently forecasted.

4.2.3. Sarima

The data may have a seasonal element that can be modelled using SARIMA, which is an extension of ARIMA that incorporates seasonal components to process time series data with recurrent patterns at fixed intervals, ranging from daily, weekly to yearly.

SARIMA is a powerful extension of ARIMA that expertly captures seasonal patterns. It includes seasonal AR, seasonal MA, and seasonal I terms, in addition to the autoregressive (AR), moving average (MA), and integration (I) components. The seasonal AR terms establish a strong connection among the available data and its previous data at seasonal intervals, while the seasonal MA terms expertly reveal the dependence on the previous residuals at seasonal intervals. Seasonal integration's first component removes seasonal trends through seasonal differencing.

The seasonal ARIMA model order is conventionally represented as (p, d, q)(P, D, Q, m), where small letters indicate non-seasonal components and big letters indicate seasonal components.

4.2.3.1. Architecture of Sarima

SARIMA is a powerful extension of the ARIMA model that expertly handles data exhibiting seasonal patterns. While ARIMA models are commonly applied for forecasting and time series analysis, SARIMA models are generally designed to incorporate both seasonal and non-seasonal components, making them an ideal choice for complex data analysis.

Figure 12: Sarima model

where,

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

1. m = number of observations per year;
2. P = Number of seasonal AR terms;
3. D = Number of seasonal differences;
4. Q = Number of seasonal MA terms

We utilize capital letters for seasonal components of the model and small letters for the non-seasonal parts. Abbreviations for technical terms will always be explained upon their initial use.

SARIMA is a different type of ARIMA that expertly models seasonal effects in data. There are one more parameter that is S related for seasonal patterns. The s parameter indicates the duration of the seasonal cycle. A SARIMA (1,1,1) (1,1,1)12 model indicates that the current value depends on the previous value, the difference between the current and past values, the error term, the past seasonal

value, the difference between the current and past seasonal values, the seasonal error term, and the seasonality lasts for 12 periods. The seasonal component of SARIMA models comprises of three components;

1. Seasonal Autoregressive (P): This component is all about understanding the relationship between the current value of the series and its past values, especially at seasonal lags.
2. Seasonal Integrated (D): This component is similar to the non-seasonal differencing, but it's there to help us remove seasonality from the series.
3. Seasonal Moving Average (Q): This component is a great example of how the current value is related to the residual errors from previous predictions at seasonal lags.

Arima (1,0,1) (2,1,0) 12 model is mathematically expressed as;

$$y_t - y_{t-12} = W_t$$

$$W_t = \mu + \phi_1 W_{t-1} + \phi_2 W_{t-12} + \phi_3 W_{t-24} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

4.2.3.2. Advantages of Sarima

1. **Seasonal Modeling:** SARIMA represents an extension of ARIMA, incorporating seasonal components, which enables the model to more effectively capture and forecast seasonal patterns in time series data
2. **Improved Accuracy:** The incorporation of seasonal variations into SARIMA models enables the generation of more precise forecasts for datasets exhibiting pronounced seasonal fluctuations.
3. **Long-Term Forecasting:** SARIMA is an appropriate methodology for long-term forecasting tasks where seasonal trends play a significant role, such as sales forecasting for seasonal products.

4.2.3.3. Disadvantages of Sarima

1. **Complexity:** SARIMA models are a little more complex than traditional ARIMA models, thanks to the seasonal parameters that are included. This can mean that they require a little more expertise to implement and interpret accurately.
2. **Data Requirements:** SARIMA models necessitate the availability of sufficiently lengthy and comprehensive time series data to enable the estimation of seasonal parameters with a high level of accuracy. In the absence of such data, the reliability of the resulting forecasts may be compromised.
3. **Risk of Overfitting:** SARIMA models with lots of seasonal parameters can sometimes run the risk of overfitting, especially when we're working with datasets that don't have a lot of observations or noisy data.

4.2.3.4. Applications of Sarima In Literature

Perone (2021) has analyzed a few time series predicting models to estimate the spread of COVID-19 among the pandemic's second wave in Italy (after October 13, 2020). The data for the period 21 February 2020 to 13 October 2020 were taken from the website of the Italian Ministry of Health (www.salute.gov.it). She targeted to facilitate decision-making process of doctors, especially in the short-term with this study. This study emphasizes the broad applicability of ARIMA models in disease management. This research demonstrates the practical use cases of the ARIMA/SARIMA models in disease management. And also , we can say the ARIMA model in time series was successfully used by a Singaporean hospital during the SARS outbreak to accurately forecast the number of beds required in three days.

A Study was done to predict daily and monthly average global solar radiation in South Korea. The study forecasts solar radiation levels in South Korea using hourly data obtained over 37 years from the Korean Meteorological Administration. The SARIMA model is employed to achieve this prediction.

4.2.4. Holt Winters

Holt Winters Exponential Smoothing, a variant of this technique, is widely used in the industry. It allows for accurate modelling and prediction of value sequences over time. Holt Winters is very popular predictive technique for the data that displays trend and seasonal differentiation. This technique models the average value, trend slope, and cyclical repeating pattern (seasonality) of the time series with confidence.

Holt Winters is analyzing time sequences behavior, used in forecasting. When analyzing technical terms, abbreviations should be explained upon first use. This method models three components of the time series: an average level, a trend over time, and a repeating pattern of seasonality. This approach to modelling is crucial for accurate forecasting.

The Holt Winters method is a highly effective and widely used time series forecasting technique that expertly incorporates both trend and seasonality. This method is composed of three simpler smoothing methods that are expertly combined to create the powerful and reliable Holt Winters method.

- Simple Exponential Smoothing (SES): This method assumes that the time series remains constant in level. It is not suitable for use with series that display trends, seasonality, or both.
- Holt's Exponential Smoothing (HES): Holt's exponential smoothing is a powerful method that improves simple exponential smoothing by incorporating a trend component in time series data. However, this method may not be perfect choice for seasonal data.

- Winter's Exponential Smoothing (WES): It is a powerful extension of HES method which expertly incorporates seasonal analysis.

The Holt Winters model, also known as triple exponential smoothing, utilizes three smoothing techniques to define the model. There are three important parameters in this model. First one is alpha where is used for level of smoothing. Second, beta which is used for trend of smoothing. Third gamma where is used for seasonal of smoothing. Additionally, the parameter for the type of seasonality is specified as either additive or multiplicative. Additive seasonality displays a constant change in each season, while multiplicative seasonality involves a change in each season by a factor.

4.2.4.1. Architecture of Holt Winters

The Holt Winters technique applies exponential smoothing to encode past values and estimate future 'typical' values in an objective manner. Exponential smoothing uses an exponentially weighted moving average (EWMA) to smooth a time series. A smoothed version of the time series dataset, xt, called st, can be defined.

$$st=\alpha xt+(1-\alpha)st-1$$

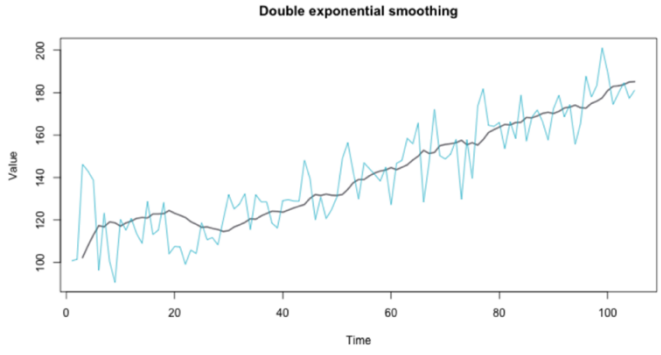
Charles Holt identified a problem with the basic EWMA model for time series with trends in the late 1950s. To address this issue, he confidently revised the simple exponential smoothing model to incorporate a linear trend. This modified method is referred to as Holt's exponential smoothing, which is a more complex but decisive approach. It incorporates two EWMAs, one for the smoothed values of xt and another for its slope, while also utilizing terms such as 'level' and 'trend.'

$$st=\alpha xt+(1-\alpha)(st-1+bt-1)$$

$$bt=\beta(st-st-1)+(1-\beta)bt-1$$

Double exponential smoothing noticeably improves the effectiveness of smoothed values in tracking the original time series, resulting in significantly improved forecasts.

Figure 13: Double exponential smoothing

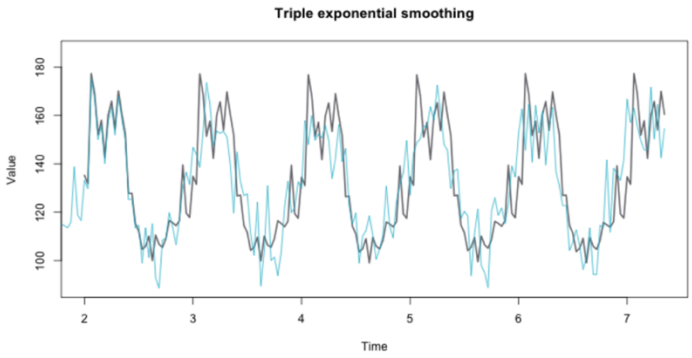


We should make an adjustment to forecast with this model. We should forecast the value in *m* time steps in the future. The formula for the *m*-step-ahead forecast, *Ft+m*, is:

$$Ft+m=st+mbt$$

When we look at the seasonality, we can say there is not a straight line. A different EWMA should be calculated. Peter Winters, a student of Holt, extended his teacher's model by adding a term that includes the variable L, which depends on the seasonality period and must be known beforehand.

Figure 14: Triple exponential smoothing



There are three characteristics namely level, trend, and seasonality that are represented through the exponential smoothing. So, it is called Holt Winters. The model forecasts present or upcoming values by calculating the cumulative impact of these three factors.

4.2.4.2. Advantages of Holt Winters

1. It is a very popular model for forecasting.
2. The model is recursive, which enables it to capture trends and seasonality in time series data. The model is capable of producing successful forecasting results for time series exhibiting a trend.
3. The software is capable of handling missing data by utilising modified Holt Winters recursive formulas.
4. Furthermore, the method has been demonstrated to yield highly effective results with minimal forecasting errors, particularly when seasonality is incorporated into the data.

4.2.4.3. Disadvantages of Holt Winters

1. One thing to be aware of with the Holt Winters method is that it does require the analyst to determine the initial values for level, trend, and seasonality. This can sometimes be tricky in practice, so it's good to be aware of that.
2. Another disadvantage of the original Holt method is that if the time series exhibits irregular time intervals or time-close observations, the method may produce wrong results.
3. Another thing to think about is that the original Holt method doesn't take the most recent inter-trends relation into account, which might affect the accuracy of the forecasts.

4.2.4.4. Applications of Holt Winters In Literature

Rahman et al., (2017) have tried to estimate the monthly revenue of the Bangabandhu bridge by using Holt Winters exponential smoothing model. The seasonal forecasting with trend issue was considerable importance. Therefore, the study has focus to forecast the monthly revenue using Holt–Winters.

Al-Hafid and Al-maamary (2012) targeted to develop a new model for predicting electricity load demand that minimises forecasting errors. This is very significant for electrical demand controllers to ensure sufficient electricity supply to cope with increasing demands. In the study, the Seasonal Holt Winters method was used to attempt to forecast the daily peak demand of electricity.

Yang. Et al., (2017) confidently studied the forecasting of aircraft failure rates using Holt Winters method. They assert that the Holt Winters seasonal model is a suitable method for predicting aircraft failure rates, based on their analysis of seasonal time series data.

When the studies published in the literature were examined for the algorithms which are targeted use for forecasting the selling price of secondhand cars; ANNS, ARIMA (SARIMA), XGBoost, TBATS, Holt Winters Forecasting, we can say they are the most suitable models for forecasting the selling price of secondhand cars. We will use all 5 models for our analysis. According to the success of these algorithms, the best algorithm will be selected. Then, the results produced by the selected algorithm will be announced as output.

5. MODEL COMPARISON

Forecast accuracy measures are essential in evaluating the effectiveness of forecasting models across various industries. These metrics provide significant results into the performance of predictions by assessing the differences between projected and actual values. They play a vital role in evaluating and improving the efficiency of forecasting models. After detailed reserch about comparing methods for forecasting studies, we decided to compare different forecasting models and machine learning algorithms in terms of MAE, MAPE, MSE and RMSE in this study. We used those 4-comparison metrics to measure the efficiency and accuracy of our 6 algorithms in this study.

5.1. Mean Squared Error (MSE)

Mean Squared Error is a useful tool that avoids the challenge of positive and negative forecast errors offsetting each other. It is calculated by:

- First, the square of the forecast error is calculated
- Then, the average of the squared forecast error is taken

The mathematical formula of MSE is;

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2$$

5.2. Mean Absolute Error (MAE)

This technique calculates the average values of the estimation errors. The main value is subtracted from estimated value, and the resulting error term is transformed into an absolute value. The mathematical formula of MAE is;

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

5.3. Mean Absolute Percentage Error (MAPE)

It is used to evaluate the results for a specific time period, such as comparing a monthly sale forecast to a weekly sale. Mean Absolute Percentage Error (MAPE) is dependent on the scale of the data. It is calculated by:

- First, the absolute forecast error is divided by the actual value
- Second, the average of individual absolute percentage error is calculated

The mathematical formula of MAPE is;

$$MAPE = mean(|e_i|/y_i) * 100$$

5.4. Root Mean Squared Error (RMSE)

Root Mean Squared Error is the square root of Mean Squared Error (MSE). It is a useful metric for calculating forecast accuracy. The mathematical formula of RMSE is;

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2}$$

This metric is scale dependent like MAE and cannot be used to make comparisons between series on different scales.

6. DATA COLLECTION and DATA CLEANING

6.1. Data Collection

The study collected data from 10,838 vehicles in the secondhand automobile and SUV categories posted on sahibinden.com between January 2020 and December 2021, out of approximately 700,000 monthly classifieds. This sample size provides a robust representation of the market trends during the specified time period. After a thorough analysis of the classifieds on sahibinden.com, we have identified the key attributes that strongly correlate with the selling price of cars. Our selection process ensures that we only consider the most significant factors in our price forecasting model. To estimate the prospective six months sales price of Automobile and SUV categories in Turkiye, we confidently select attributes such as the vehicle's brand, model, package type, model year, fuel type, gear type, body type, and mileage.

- **Vehicle's brand:** This attribute shows the brand of the car. For example, Toyoto, Mercedes, BMW etc.
- **Vehicle's model:** This attribute shows the model of the car. For example, Corollo, AMG, Egea etc.
- **Vehicle's package type:** This attribute shows the package type of the car. For example, 1.6 multi jet - Premium, Optimal etc.
- **Vehicle's model year:** This attribute shows the model year of the car. For example, 2023,2010 etc.
- **Vehicle's fuel type:** This attribute shows the fuel type of the car. For example, Diesel, Gasoline, Electric etc.
- **Vehicle's gear type:** This attribute shows the gear type of the car. For example, Automatic, Manuel etc.
- **Vehicle's body type:** This attribute shows the body type of the car. For example, Sedan, Hatch-back etc.
- **Vehicle's mileage:** This attribute shows the km of the car. For example, 45.000, 15.000 etc.

Table 1: An Example of Data Set

Brand	Model	Model Year	Package Type	Fuel Type	Gear Type	Body Type	Milege
Seat	Esco	2014	1.6 Style	Gasoline	Manuel	Sedan	17.644
Fiat	Linea	2005	1.6 Multi Jet-Dynamic	Diesel	Automatic	Hatchback	23.555
Kia	Rio	2007	1.5 CRDi-EX Advance	Electric	Manuel	Sedan	32.555
Toyoto	Corollo	2011	Elegant	Gasoline	Automatic	Hatchback	22.222
Mercedes	AMG	2020	Premium	Diesel	Manuel	Sedan	129.345
BMW	A4	2018	Sport	Electric	Automatic	Hatchback	258.744
Skodia	Octovia	2017	Optimal	Gasoline	Manuel	Sedan	12.233
Ctiroen	C3	2016	Premium	Diesel	Automatic	Hatchback	1.343.454

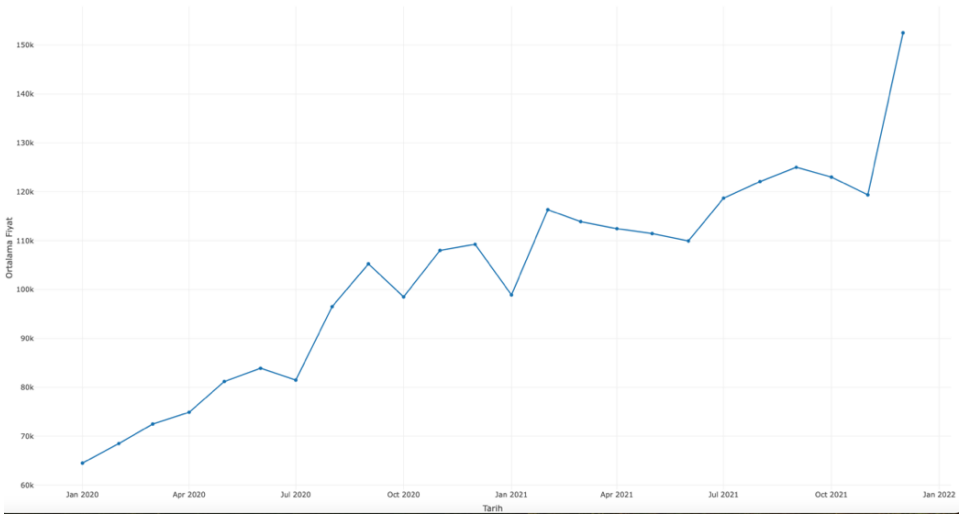
6.2. Data Cleaning

After we took the data from sahibinden.com, we made some adjustments on the data set.

1. Firstly, vehicles with 30 or more classifieds were included in the data set.
2. Secondly, vehicles with data for at least 21 months in the last 24 months and vehicles with data data in the last month were included in the data set.
3. Then, outliers were eliminated separately for each brand and model by using the Tukey outlier methodology (IQR).
4. Finally, medians of classifieds were taken for the data set.

After all these adjustments, we started to implement our models to estimate the prospective six months sales price of secondhand vehicles in the Automobile and SUV categories in Turkiye.

Figure 15: Sample data view



7. FINDINGS

This section confidently presents the findings of various forecasting techniques, including ARIMA, SARIMA, TBATS, HOLT WINTERS Forecasting, and machine learning algorithms such as ANN and XGBOOST. The ultimate goal is to accurately estimate the prospective six months sales price of vehicles in the Automobile and SUV categories in Turkiye. By using the vehicles' brand, model, package type, model year, body type , gear type, fuel type and mileage characteristics for the future six months sales prices of the vehicle was estimated over the past prices of vehicles.

7.1. ANN Results

After data cleaning, firstly, ANN was used to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye by using the vehicles' brand, model, package type, model year, body type, gear type, fuel type and mileage characteristics. We created ANN model with R programing language.

```
ann_model=nnetar(ts_data,decay=0.01, maxit=500, lambda = 'auto')

fitted_ann=forecast(ann_model,6)

ann_pred=fitted_ann$mean
```

Then we set parametric settings for the ANN model;

```
i=0

j=0

k=0

# result=NULL

for (i in c(seq(0,1,by=0.05))) {

for (j in c(seq(25,1000,by=25))) {

for (k in c(seq(0,1,by=0.05))) {

ann_model=nnetar(ts_data,decay=i, maxit=j, lambda = k)

ann_model=nnetar(ts_data)

fitted_ann=forecast(ann_model,6)
```

```
result_i=c(i,j,k,abs(fitted_ann$mean-test_data$real_price)/test_data$real_price,
mean(abs(fitted_ann$mean-test_data$real_price)/test_data$real_price))
result=rbind(result,result_i)
}}}
```

Then, we created the model comparision indicators for ANN model;

```
ann_ape=abs(fitted_ann$mean-test_data$real_price)/test_data$real_price
ann_mape=mean(abs(fitted_ann$mean-test_data$real_price)/test_data$real_price)
ann_mae=mean(abs(ann_pred$pred-test_data$real_price)) #mae
ann_mse=mean((ann_pred$pred-test_data$real_price)**2) #mse
ann_rmse=sqrt(mean((ann_pred$pred-test_data$real_price)**2)) #rmse
```

After all these calculations, ANN model worked and displayed the final results for use in order to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye.

Figure 16: A few examples of ANN results

arac	Jan_2022	Feb_2022	Mar_2022	Apr_2022	May_2022	Jun_2022
3686___2006___Manuel___Benzin___Coupe	₺179.441,31	₺179.404,25	₺179.676,64	₺179.596,00	₺179.640,14	₺179.766,87
3691___1997___Manuel___Benzin___Coupe	₺195.202,38	₺195.179,95	₺195.179,82	₺195.179,82	₺195.442,63	₺195.181,29
3712___1998___Manuel___Benzin & LPG___Hatchback 5 kapı	₺50.119,81	₺47.687,35	₺49.420,58	₺50.865,30	₺51.062,76	₺50.235,24
3713___1996___Manuel___Benzin & LPG___Hatchback 5 kapı	₺45.682,51	₺46.315,91	₺44.982,36	₺46.544,35	₺46.139,05	₺46.740,31
3718___2003___Manuel___Benzin & LPG___Hatchback 5 kapı	₺73.461,40	₺72.975,80	₺72.897,61	₺71.928,19	₺72.726,89	₺74.296,59
3718___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺73.499,48	₺72.589,79	₺72.266,09	₺73.494,14	₺74.280,17	₺73.725,80
3718___2005___Manuel___Benzin & LPG___Hatchback 5 kapı	₺72.055,14	₺77.522,70	₺78.310,82	₺74.929,90	₺80.095,27	₺80.192,36
3718___2006___Manuel___Benzin & LPG___Hatchback 5 kapı	₺96.243,18	₺98.780,72	₺99.489,59	₺100.217,57	₺100.551,84	₺100.473,22
3718___2007___Manuel___Benzin & LPG___Hatchback 5 kapı	₺83.809,17	₺82.873,77	₺85.107,91	₺85.314,00	₺84.893,18	₺85.187,95
3719___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺92.640,15	₺99.485,17	₺100.560,28	₺101.079,37	₺101.035,16	₺101.133,27
3726___2003___Manuel___Benzin & LPG___Sedan	₺70.180,27	₺69.945,11	₺69.684,87	₺69.575,89	₺70.034,75	₺70.436,28
3726___2004___Manuel___Benzin & LPG___Sedan	₺83.252,55	₺82.887,36	₺81.727,94	₺82.837,14	₺83.391,25	₺84.238,35
3726___2005___Manuel___Benzin___Sedan	₺85.232,82	₺116.218,13	₺83.833,68	₺82.547,43	₺104.013,45	₺98.582,49
3726___2005___Manuel___Benzin & LPG___Sedan	₺107.115,89	₺111.334,86	₺112.494,21	₺112.757,73	₺112.758,17	₺112.808,85
3726___2006___Manuel___Benzin & LPG___Sedan	₺107.007,60	₺101.736,72	₺96.369,23	₺95.437,84	₺103.703,81	₺98.796,53
3727___1998___Manuel___Benzin & LPG___Sedan	₺77.388,61	₺78.527,76	₺78.469,77	₺78.511,58	₺78.291,57	₺78.611,63

When we look at the model comparison indicators for ANN, mean value of ANN in terms of MAE, MSE, RMSE and MAPE is below.

Table 2: Mean values of ANN in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
ANN	62.987 TL	131.981.019.253	70.671	20%

When we look at the model comparison indicators for ANN, median value of ANN in terms of MAE, MSE, RMSE and MAPE is below.

Table 3: Median values of ANN in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
ANN	34.932 TL	1.503.076.236	38.769	20%

7.2. XGBoost Results

After data cleaning, secondly, XGBOOST was used to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye by using the vehicles' brand, model, package type, model year, body type , gear type, fuel type and mileage characteristics. We created XGBOOST model with R programing language.

```
train_data_xgb <- train_data %>% dplyr::select(active_date, real_price)

test_data_xgb <- test_data %>% dplyr::select(active_date, real_price)

date_train=seq (from = lubridate::as_date("2020-01-01"), by = "month", length.out = 24)

date_test=seq (from = lubridate::as_date("2022-01-01"), by = "month", length.out = 6)

extended_train_data <- train_data_xgb %>%

 dplyr::mutate(.,

              months = lubridate::month(date_train),

              years = lubridate::year(date_train))

extended_test_data <- test_data_xgb %>%

  dplyr::mutate(.,

              months = lubridate::month(date_test),

              years = lubridate::year(date_test))

train <- extended_train_data[1:nrow(train_data), ] # initial data

test <- extended_test_data[1:nrow(test_data), ]
```

```
pred <- extended_data_mod[(nrow(train_data) + 1):nrow(train_data), ] # extended time index


x_train <-  as.matrix(train %>%
                      dplyr::select(months, years))
x_pred <-  as.matrix(pred %>%
                     dplyr::select(months, years))
x_test <-  as.matrix(test %>%
                      dplyr::select(months, years))
y_train <- train$real_price


xgb_trcontrol <- caret::trainControl(
  method = "cv",
  number = 5,
  allowParallel = TRUE,
  verboseIter = FALSE,
  returnData = FALSE
)
xgb_grid <- base::expand.grid(
  list(
    nrounds = c(100, 200),
    max_depth = c(10, 15, 20), # maximum depth of a tree
    colsample_bytree = seq(0.5), # subsample ratio of columns when construction each tree
    eta = 0.1, # learning rate
    gamma = 0, # minimum loss reduction
    min_child_weight = 1,  # minimum sum of instance weight (hessian) needed ina child
    subsample = 1 # subsample ratio of the training instances
  ))
```

```
 xgb_model <- caret::train(
   x_train, y_train,
   trControl = xgb_trcontrol,
   tuneGrid = xgb_grid,
   method = "xgbTree",
   nthread = 1 )
 fitted_xgb <- xgb_model %>%
   stats::predict(x_test) %>%
   stats::ts(start = c(2022,1),frequency = 6)
```

Then we set parametric settings for the XGBOOST model;

```
i=0
j=0
k=0
result=NULL
for (i in c(seq(0.05,1,by=0.05))) {
for (j in c(seq(0.05,1,by=0.05))) {
for (k in c(seq(0.05,1,by=0.05))) {

xgb_trcontrol <- caret::trainControl(
   method = "cv",
   number = 5,
   allowParallel = TRUE,
   verboseIter = FALSE,
   returnData = FALSE  )
 xgb_grid <- base::expand.grid(
   list(
```

```
    nrounds = c(100, 200),
    max_depth = c(10, 15, 20), # maximum depth of a tree
    colsample_bytree = seq(0.5), # subsample ratio of columns when construction each tree
    eta = i, # learning rate
    gamma = j, # minimum loss reduction
    min_child_weight = k,  # minimum sum of instance weight (hessian) needed ina child
    subsample = 1 # subsample ratio of the training instances  ))
  xgb_model <- caret::train(
    x_train, y_train,
    trControl = xgb_trcontrol,
    tuneGrid = xgb_grid,
    method = "xgbTree",
    nthread = 1 )
  fitted_xgb <- xgb_model %>%
    stats::predict(x_test) %>%
    stats::ts(start = c(2022,1),frequency = 6)
result_i=c(i,j,k,abs(fitted_xgb$mean-test_data$real_price)/test_data$real_price,
mean(abs(fitted_xgb$mean-test_data$real_price)/test_data$real_price))
result=rbind(result,result_i) }}}
```

Then, we created the model comparision indicators for XGBOOST model;

```
  xgb_pred= as.numeric(fitted_xgb)
  xgb_ape=abs(xgb_pred-test_data$real_price)/test_data$real_price
  xgb_mape=mean (abs(xgb_pred-test_data$real_price)/test_data$real_price)
  xgb_mae=mean (abs(xgb_pred$pred-test_data$real_price)) #mae
  xgb_mse=mean ((xgb_pred$pred-test_data$real_price)**2) #mse
  xgb_rmse= sqrt(mean((xgb_pred$pred-test_data$real_price)**2)) #rmse
```

After all these calculations, XGBOOST model worked and displayed the final results for use in order to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye.

Figure 17: A few examples of XGBOOST results

arac	Jan_2022	Feb_2022	Mar_2022	Apr_2022	May_2022	Jun_2022
3686___2006___Manuel___Benzin___Coupe	₺126.067,95	₺105.131,33	₺117.286,80	₺112.591,14	₺114.908,56	₺124.212,86
3691___1997___Manuel___Benzin___Coupe	₺188.431,55	₺188.831,00	₺188.828,38	₺188.819,56	₺194.042,62	₺188.994,39
3712___1998___Manuel___Benzin & LPG___Hatchback 5 kapı	₺37.412,74	₺32.538,48	₺39.344,93	₺42.384,59	₺41.972,00	₺38.767,40
3713___1996___Manuel___Benzin & LPG___Hatchback 5 kapı	₺33.124,85	₺34.999,39	₺28.500,25	₺37.499,30	₺33.000,29	₺38.249,85
3718___2003___Manuel___Benzin & LPG___Hatchback 5 kapı	₺59.986,18	₺61.969,40	₺61.969,40	₺58.776,20	₺61.968,09	₺67.552,74
3718___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺57.393,01	₺56.899,31	₺56.899,31	₺60.970,63	₺62.157,01	₺59.637,91
3718___2005___Manuel___Benzin & LPG___Hatchback 5 kapı	₺64.482,92	₺65.464,97	₺68.863,71	₺65.053,52	₺69.929,84	₺72.251,95
3718___2006___Manuel___Benzin & LPG___Hatchback 5 kapı	₺68.472,27	₺71.793,12	₺68.943,49	₺72.959,56	₺73.866,55	₺72.022,84
3718___2007___Manuel___Benzin & LPG___Hatchback 5 kapı	₺67.354,59	₺68.466,38	₺78.262,93	₺76.452,14	₺74.548,72	₺76.220,99
3719___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺52.987,30	₺55.346,66	₺54.988,96	₺62.863,53	₺61.043,02	₺64.318,47
3726___2003___Manuel___Benzin & LPG___Sedan	₺62.628,46	₺59.952,54	₺57.892,20	₺57.518,03	₺62.472,23	₺66.359,91
3726___2004___Manuel___Benzin & LPG___Sedan	₺68.624,63	₺68.499,92	₺65.900,23	₺69.249,64	₺70.000,16	₺71.874,86
3726___2005___Manuel___Benzin___Sedan	₺71.927,48	₺94.502,26	₺72.482,48	₺67.589,20	₺78.467,84	₺78.467,84
3726___2005___Manuel___Benzin & LPG___Sedan	₺76.024,11	₺70.997,55	₺70.997,55	₺71.009,38	₺73.364,72	₺71.633,16
3726___2006___Manuel___Benzin & LPG___Sedan	₺82.147,67	₺77.921,48	₺73.119,82	₺71.537,38	₺77.384,47	₺75.282,91
3727___1998___Manuel___Benzin & LPG___Sedan	₺46.966,91	₺48.547,17	₺45.000,06	₺46.690,79	₺41.015,39	₺59.628,17

When we look at the model comparison indicators for XGBOOST, mean value of XGBOOST in terms of MAE, MSE, RMSE and MAPE is below.

Table 4: Mean values of XGBOOST in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
XGBOOST	149.571 TL	208.392.661.800	153.957	44%

When we look at the model comparison indicators for XGBOOST, median value of XGBOOST in terms of MAE, MSE, RMSE and MAPE is below.

Table 5: Median values of XGBOOST in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
XGBOOST	93.760 TL	9.101.650.322	95.403	45%

7.3. Arima Results

After data cleaning, thirdly, ARIMA was used to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye by using the vehicles' brand, model, package type, model year, body type ,gear type, fuel type and mileage characteristics. We created ARIMA model with R programing language.

```
ARIMAfit = arima(ts_data,order =c(0,0,1)) # arima Model adımı

arima_fit=predict(ARIMAfit, 6)  #tahmin adımı

arima_pred=arima_fit$pred
```

Then, we created the model comparision indicators for ARIMA model;

```
arima_ape=abs(arima_fit$pred-test_data$real_price)/test_data$real_price #hatasını oluşturuyor

arima_mape=mean(abs(arima_fit$pred-test_data$real_price)/test_data$real_price)  #ortalama hatayı buluyoruz

arima_mae=mean(abs(arima_fit$pred-test_data$real_price)) #mae

arima_mse=mean((arima_fit$pred-test_data$real_price)**2) #mse

arima_rmse=sqrt(mean((arima_fit$pred-test_data$real_price)**2)) #rmse
```

After all these calculations, ARIMA model worked and displayed the final results for use in order to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye.

Figure 18: A few examples of ARIMA results

arac	Jan_2022	Feb_2022	Mar_2022	Apr_2022	May_2022	Jun_2022
3686___2006___Manuel___Benzin___Coupe	₺127.271,72	₺98.955,85	₺98.955,85	₺98.955,85	₺98.955,85	₺98.955,85
3691___1997___Manuel___Benzin___Coupe	₺172.911,29	₺159.536,36	₺159.536,36	₺159.536,36	₺159.536,36	₺159.536,36
3712___1998___Manuel___Benzin & LPG___Hatchback 5 kapı	₺44.230,93	₺32.518,57	₺32.518,57	₺32.518,57	₺32.518,57	₺32.518,57
3713___1996___Manuel___Benzin & LPG___Hatchback 5 kapı	₺35.945,45	₺30.241,66	₺30.241,66	₺30.241,66	₺30.241,66	₺30.241,66
3718___2003___Manuel___Benzin & LPG___Hatchback 5 kapı	₺69.428,32	₺54.371,50	₺54.371,50	₺54.371,50	₺54.371,50	₺54.371,50
3718___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺62.295,93	₺53.682,55	₺53.682,55	₺53.682,55	₺53.682,55	₺53.682,55
3718___2005___Manuel___Benzin & LPG___Hatchback 5 kapı	₺81.473,87	₺60.327,27	₺60.327,27	₺60.327,27	₺60.327,27	₺60.327,27
3718___2006___Manuel___Benzin & LPG___Hatchback 5 kapı	₺81.545,92	₺63.802,68	₺63.802,68	₺63.802,68	₺63.802,68	₺63.802,68
3718___2007___Manuel___Benzin & LPG___Hatchback 5 kapı	₺76.478,84	₺67.376,12	₺67.376,12	₺67.376,12	₺67.376,12	₺67.376,12
3719___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺68.221,12	₺54.777,02	₺54.777,02	₺54.777,02	₺54.777,02	₺54.777,02
3726___2003___Manuel___Benzin & LPG___Sedan	₺61.971,77	₺54.760,48	₺54.760,48	₺54.760,48	₺54.760,48	₺54.760,48
3726___2004___Manuel___Benzin & LPG___Sedan	₺74.297,41	₺62.802,91	₺62.802,91	₺62.802,91	₺62.802,91	₺62.802,91
3726___2005___Manuel___Benzin___Sedan	₺84.559,09	₺69.436,03	₺69.436,03	₺69.436,03	₺69.436,03	₺69.436,03
3726___2005___Manuel___Benzin & LPG___Sedan	₺81.733,66	₺63.600,82	₺63.600,82	₺63.600,82	₺63.600,82	₺63.600,82
3726___2006___Manuel___Benzin & LPG___Sedan	₺89.328,81	₺70.613,52	₺70.613,52	₺70.613,52	₺70.613,52	₺70.613,52
3727___1998___Manuel___Benzin & LPG___Sedan	₺53.243,25	₺42.272,18	₺42.272,18	₺42.272,18	₺42.272,18	₺42.272,18

When we look at the model comparison indicators for ARIMA, mean value of ARIMA in terms of MAE, MSE, RMSE and MAPE is below.

Table 6: Mean values of ARIMA in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
ARIMA	152.574 TL	215.288.720.058	159.642	45%

When we look at the model comparison indicators for ARIMA, median value of ARIMA in terms of MAE, MSE, RMSE and MAPE is below.

Table 7: Median values of ARIMA in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
ARIMA	97.639 TL	10.146.560.424	100.730	46%

7.4. Sarima Results

After data cleaning, fourthly, SARIMA was used to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye by using the vehicles' brand, model, package type, model year, body type , gear type, fuel type and mileage characteristics. We created SARIMA model with R programing language.

```
sarima=arima(x = ts_data, order = c(0, 2, 1)) # sarima model adımı
sarima_fit=predict(sarima, 6)
sarima_pred=sarima_fit$pred
```

Then, we created the model comparision indicators for SARIMA model;

```
sarima_ape=abs(sarima_fit$pred-test_data$real_price)/test_data$real_price
sarima_mape=mean(abs(sarima_fit$pred-test_data$real_price)/test_data$real_price)
sarima_mae=mean(abs(sarima_fit$pred-test_data$real_price)) #mae
sarima_mse=mean((sarima_fit$pred-test_data$real_price)**2) #mse
sarima_rmse=sqrt(mean((sarima_fit$pred-test_data$real_price)**2)) #rmse
```

After all these calculations, SARIMA model worked and displayed the final results for use in order to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye.

Figure 19: A few examples of SARIMA results

arac	Jan_2022	Feb_2022	Mar_2022	Apr_2022	May_2022	Jun_2022
3686___2006___Manuel___Benzin___Coupe	₺169.760,88	₺174.521,75	₺179.282,63	₺184.043,50	₺188.804,38	₺193.565,25
3691___1997___Manuel___Benzin___Coupe	₺208.086,97	₺212.673,93	₺217.260,90	₺221.847,87	₺226.434,84	₺231.021,80
3712___1998___Manuel___Benzin & LPG___Hatchback 5 kapı	₺53.826,09	₺55.652,18	₺57.478,26	₺59.304,35	₺61.130,44	₺62.956,53
3713___1996___Manuel___Benzin & LPG___Hatchback 5 kapı	₺45.684,74	₺46.869,48	₺48.054,22	₺49.238,96	₺50.423,71	₺51.608,45
3718___2003___Manuel___Benzin & LPG___Hatchback 5 kapı	₺82.293,48	₺84.586,96	₺86.880,44	₺89.173,93	₺91.467,41	₺93.760,89
3718___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺77.500,00	₺79.375,01	₺81.250,01	₺83.125,02	₺85.000,02	₺86.875,02
3718___2005___Manuel___Benzin & LPG___Hatchback 5 kapı	₺88.891,31	₺91.282,62	₺93.673,92	₺96.065,23	₺98.456,54	₺100.847,85
3718___2006___Manuel___Benzin & LPG___Hatchback 5 kapı	₺95.365,22	₺97.980,44	₺100.595,66	₺103.210,88	₺105.826,11	₺108.441,33
3718___2007___Manuel___Benzin & LPG___Hatchback 5 kapı	₺89.744,57	₺91.989,14	₺94.233,71	₺96.478,28	₺98.722,85	₺100.967,42
3719___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺84.255,44	₺86.260,88	₺88.266,32	₺90.271,76	₺92.277,20	₺94.282,63
3726___2003___Manuel___Benzin & LPG___Sedan	₺71.208,70	₺72.217,40	₺73.226,10	₺74.234,81	₺75.243,51	₺76.252,21
3726___2004___Manuel___Benzin & LPG___Sedan	₺86.608,18	₺88.591,86	₺90.575,54	₺92.559,21	₺94.542,89	₺96.526,57
3726___2005___Manuel___Benzin___Sedan	₺105.130,44	₺107.760,88	₺110.391,32	₺113.021,76	₺115.652,20	₺118.282,64
3726___2005___Manuel___Benzin & LPG___Sedan	₺104.311,94	₺112.873,88	₺121.435,83	₺129.997,77	₺138.559,71	₺147.121,65
3726___2006___Manuel___Benzin & LPG___Sedan	₺111.782,61	₺114.815,23	₺117.847,84	₺120.880,45	₺123.913,07	₺126.945,68
3727___1998___Manuel___Benzin & LPG___Sedan	₺67.521,74	₺69.293,49	₺71.065,23	₺72.836,97	₺74.608,71	₺76.380,46

When we look at the model comparison indicators for SARIMA, mean value of SARIMA in terms of MAE, MSE, RMSE and MAPE is below.

Table 8: Mean values of SARIMA in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
SARIMA	128.925 TL	212.659.296.692	145.908	31%

When we look at the model comparison indicators for SARIMA, median value of SARIMA in terms of MAE, MSE, RMSE and MAPE is below.

Table 9: Median values of SARIMA in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
SARIMA	36.910 TL	1.606.396.992	40.080	19%

7.5. Tbats Results

After data cleaning, fifthly, TBATS was used to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye by using the vehicles' brand, model, package type, model year, body type , gear type, fuel type and mileage characteristics. We created TBATS model with R programing language.

```
tbats_model <- tbats(ts_data)

fitted_tbats=forecast(tbats_model,6)

tbats_pred=fitted_tbats$mean


tbats_pred_list2=c(data_limit_m$id[i],round(tbats_pred,2))

tbats_pred_list=rbind(tbats_pred_list,tbats_pred_list2)

tbats_pred_list=as.data.frame(tbats_pred_list)
```

Then, we created the model comparision indicators for TBATS model;

```
tbats_ape=abs(fitted_tbats$mean-test_data$real_price)/test_data$real_price

tbats_mape=mean(abs(fitted_tbats$mean-test_data$real_price)/test_data$real_price)

tbats_mae=mean(abs(tbats_pred$pred-test_data$real_price)) #mae

tbats_mse=mean((tbats_pred$pred-test_data$real_price)**2) #mse

tbats_rmse=sqrt(mean((tbats_pred$pred-test_data$real_price)**2)) #rmse
```

After all these calculations, TBATS model worked and displayed the final results for use in order to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye.

Figure 20: A few examples of TBATS results

arac	Jan_2022	Feb_2022	Mar_2022	Apr_2022	May_2022	Jun_2022
3686___2006___Manuel___Benzin___Coupe	₺153.730,62	₺158.167,70	₺159.898,92	₺159.376,69	₺157.657,22	₺156.117,53
3686___2006___Manuel___Benzin___Coupe	₺153.730,62	₺158.167,70	₺159.898,92	₺159.376,69	₺157.657,22	₺156.117,53
3691___1997___Manuel___Benzin___Coupe	₺202.581,86	₺202.581,86	₺202.581,86	₺202.581,86	₺202.581,86	₺202.581,86
3691___1997___Manuel___Benzin___Coupe	₺202.581,86	₺202.581,86	₺202.581,86	₺202.581,86	₺202.581,86	₺202.581,86
3712___1998___Manuel___Benzin & LPG___Hatchback 5 kapı	₺44.447,31	₺44.926,81	₺45.375,65	₺45.795,80	₺46.189,10	₺46.557,25
3712___1998___Manuel___Benzin & LPG___Hatchback 5 kapı	₺44.447,31	₺44.926,81	₺45.375,65	₺45.795,80	₺46.189,10	₺46.557,25
3713___1996___Manuel___Benzin & LPG___Hatchback 5 kapı	₺38.538,62	₺38.868,29	₺39.177,40	₺39.467,22	₺39.738,96	₺39.993,74
3713___1996___Manuel___Benzin & LPG___Hatchback 5 kapı	₺38.538,62	₺38.868,29	₺39.177,40	₺39.467,22	₺39.738,96	₺39.993,74
3718___2003___Manuel___Benzin & LPG___Hatchback 5 kapı	₺73.955,28	₺74.850,15	₺75.697,96	₺76.501,19	₺77.262,18	₺77.983,16
3718___2003___Manuel___Benzin & LPG___Hatchback 5 kapı	₺73.955,28	₺74.850,15	₺75.697,96	₺76.501,19	₺77.262,18	₺77.983,16
3718___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺78.783,53	₺81.691,74	₺84.701,39	₺87.815,80	₺91.038,40	₺94.372,73
3718___2004___Manuel___Benzin & LPG___Hatchback 5 kapı	₺78.783,53	₺81.691,74	₺84.701,39	₺87.815,80	₺91.038,40	₺94.372,73
3718___2005___Manuel___Benzin & LPG___Hatchback 5 kapı	₺77.408,04	₺78.121,26	₺78.790,59	₺79.418,72	₺80.008,19	₺80.561,39
3718___2005___Manuel___Benzin & LPG___Hatchback 5 kapı	₺77.408,04	₺78.121,26	₺78.790,59	₺79.418,72	₺80.008,19	₺80.561,39
3718___2006___Manuel___Benzin & LPG___Hatchback 5 kapı	₺94.311,91	₺95.618,04	₺96.872,22	₺98.076,50	₺99.232,87	₺100.343,25
3718___2006___Manuel___Benzin & LPG___Hatchback 5 kapı	₺94.311,91	₺95.618,04	₺96.872,22	₺98.076,50	₺99.232,87	₺100.343,25
3718___2007___Manuel___Benzin & LPG___Hatchback 5 kapı	₺87.300,36	₺87.388,77	₺87.469,48	₺87.543,16	₺87.610,43	₺87.671,85

When we look at the model comparison indicators for TBATS, mean value of TBATS in terms of MAE, MSE, RMSE and MAPE is below.

Table 10: Mean values of TBATS in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
TBATS	87.987 TL	2.348.941.947.010	107.555	24%

When we look at the model comparison indicators for TBATS, median value of TBATS in terms of MAE, MSE, RMSE and MAPE is below.

Table 11: Median values of TBATS in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
TBATS	35.323 TL	1.483.629.557	38.518	18%

7.6. Holt Winters Results

After data cleaning, sixthly, HOLT WINTERS was used to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye by using the vehicles' brand, model, package type, model year, body type, gear type, fuel type and mileage characteristics. We created HOLT WINTERS model with R programing language.

```
holt_winters=HoltWinters(ts_data,alpha = 0.95, beta = 0.15, gamma = 1)

hw_pred=predict(holt_winters, 6)
```

Then we set parametric settings for the HOLT WINTERS model;

```
i=0

j=0

k=0

result=NULL

for (i in c(seq(0.05,1,by=0.05))) {

for (j in c(seq(0.05,1,by=0.05))) {

for (k in c(seq(0.05,1,by=0.05))) {

holt_winters=HoltWinters(ts_data,alpha = i, beta = j, gamma = k)

abs(predict(holt_winters, 6)-test_data$real_price)/test_data$real_price

mean(abs(predict(holt_winters, 6)-test_data$real_price)/test_data$real_price)

result_i=c(i,j,k,abs(predict(holt_winters,

6)-test_data$real_price)/test_data$real_price,mean(abs(predict(holt_winters,

6)-test_data$real_price)/test_data$real_price))

result=rbind(result,result_i)}}}
```

Then, we created the model comparision indicators for HOLT WINTERS model;

```
hw_ape=abs(predict(holt_winters, 6)-test_data$real_price)/test_data$real_price

hw_mape=mean(abs(predict(holt_winters, 6)-test_data$real_price)/test_data$real_price)

hw_mae=mean(abs(hw_pred$pred-test_data$real_price)) #mae

hw_mse=mean((hw_pred$pred-test_data$real_price)**2) #mse

hw_rmse=sqrt(mean((hw_pred$pred-test_data$real_price)**2)) #rmse
```

After all these calculations, HOLT WINTERS model worked and displayed the final results for use in order to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye.

Figure 21: A few examples of HOLT WINTERS results

arac	Jan_2022	Feb_2022	Mar_2022	Apr_2022	May_2022	Jun_2022
3686__2006__Manuel__Benzin__Coupe	₺201.902,48	₺183.424,30	₺200.539,62	₺201.234,20	₺209.487,39	₺222.466,11
3691__1997__Manuel__Benzin__Coupe	₺199.403,27	₺195.214,62	₺198.059,22	₺201.964,68	₺210.863,72	₺207.929,50
3712__1998__Manuel__Benzin & LPG__Hatchback 5 kapı	₺56.947,51	₺52.819,20	₺61.146,30	₺66.171,31	₺67.703,35	₺66.006,24
3713__1996__Manuel__Benzin & LPG__Hatchback 5 kapı	₺43.960,98	₺45.918,94	₺40.344,91	₺50.477,11	₺46.960,42	₺52.946,23
3718__2003__Manuel__Benzin & LPG__Hatchback 5 kapı	₺84.748,72	₺87.528,53	₺89.493,51	₺88.314,75	₺93.778,08	₺101.341,54
3718__2004__Manuel__Benzin & LPG__Hatchback 5 kapı	₺74.723,88	₺73.608,42	₺73.923,41	₺78.945,29	₺81.279,00	₺79.413,09
3718__2005__Manuel__Benzin & LPG__Hatchback 5 kapı	₺83.858,21	₺85.419,92	₺90.091,08	₺87.287,48	₺93.997,61	₺98.045,30
3718__2006__Manuel__Benzin & LPG__Hatchback 5 kapı	₺94.331,36	₺98.397,63	₺97.183,26	₺103.209,72	₺106.244,59	₺105.896,19
3718__2007__Manuel__Benzin & LPG__Hatchback 5 kapı	₺80.019,43	₺80.381,57	₺91.015,23	₺90.561,20	₺89.751,29	₺92.218,40
3719__2004__Manuel__Benzin & LPG__Hatchback 5 kapı	₺71.841,63	₺75.032,49	₺76.173,86	₺86.420,00	₺87.065,84	₺92.395,93
3726__2003__Manuel__Benzin & LPG__Sedan	₺73.557,44	₺69.449,36	₺66.580,92	₺65.244,59	₺69.777,71	₺73.823,33
3726__2004__Manuel__Benzin & LPG__Sedan	₺83.476,73	₺82.683,57	₺80.397,56	₺84.210,48	₺85.490,90	₺87.863,37
3726__2005__Manuel__Benzin__Sedan	₺103.334,91	₺127.917,19	₺107.758,94	₺105.914,49	₺120.104,62	₺122.264,68
3726__2005__Manuel__Benzin & LPG__Sedan	₺100.408,52	₺95.616,55	₺97.296,42	₺99.259,56	₺103.934,05	₺103.844,87
3726__2006__Manuel__Benzin & LPG__Sedan	₺113.864,55	₺109.831,67	₺106.353,45	₺105.935,49	₺113.204,35	₺112.135,83
3727__1998__Manuel__Benzin & LPG__Sedan	₺73.703,72	₺76.041,10	₺73.927,74	₺77.543,71	₺73.398,48	₺93.455,06

When we look at the model comparison indicators for HOLT WINTERS, mean value of HOLT WINTERS in terms of MAE, MSE, RMSE and MAPE is below.

Table 12: Mean values of HOLT WINTERS in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
HOLT WINTERS	52.321 TL	125.400.096.121	58.527	17%

When we look at the model comparison indicators for HOLT WINTERS, median value of HOLT WINTERS in terms of MAE, MSE, RMSE and MAPE is below.

Table 13: Median values of HOLT WINTERS in terms of MAE,MSE,RMSE and MAPE

	MAE	MSE	RMSE	MAPE
HOLT WINTERS	29.321 TL	1.045.828.497	32.339	15%

7.7. Model Comparison Results

After all calculations are implemented for all models in the study, we checked the final results in terms of main 4 comparison indicators such as MAE, MSE, RMSE and MAPE in order to estimate the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye. These metrics represent how well the forecasting methods performed. Lower the values of these metrics, the more successful forecasting model is. So, we conclude the study by choosing the lower the values of these metrics for our forecasting models.

Table 14: Mean values of comparison indicators

	MAE	MSE	RMSE	MAPE
ARIMA	152.574 TL	215.288.720.058	159.642	45%
SARIMA	128.925 TL	212.659.296.692	145.908	31%
HOLT WINTERS	52.321 TL	125.400.096.121	58.527	17%
ANN	62.987 TL	131.981.019.253	70.671	20%
TBATS	87.987 TL	2.348.941.947.010	107.555	24%
XGBOOST	149.571 TL	208.392.661.800	153.957	44%

Table 15: Median values of comparison indicators

	MAE	MSE	RMSE	MAPE
ARIMA	97.639 TL	10.146.560.424	100.730	46%
SARIMA	36.910 TL	1.606.396.992	40.080	19%
HOLT WINTERS	29.321 TL	1.045.828.497	32.339	15%
ANN	34.932 TL	1.503.076.236	38.769	20%
TBATS	35.323 TL	1.483.629.557	38.518	18%
XGBOOST	93.760 TL	9.101.650.322	95.403	45%

According to results of these 4 indicators, the best estimation is performed by HOLT WINTERS algorithm for the selected dataset. HOLT WINTERS successfully passed all other forecasting methods in the study. HOLT WINTERS model resulted the best performance with mean value of MAE 52.321, MSE 125.400.096.121 error and R2 (RMSE) 58.527 and MAPE 17% performance value. When we look at the meadian value of MAE 29.321, MSE 1.045.828.497 error and R2 (RMSE) 32.339 and MAPE 15% performance value the best. HOLT WINTERS estimated the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye by using the vehicles' brand, model, package type, model year, body type, gear type, fuel type and mileage characteristics the best. That's why we will continue to study with HOLT WINTERS for our analysis in the future. And the results of HOLT WINTERS algorithm will be announced as output at the end of the study.

8. CONCLUSION & DISCUSSION

The aim of this research is to estimate the prospective six months selling price of secondhand vehicles in the Automobile and SUV categories in Turkiye. Because there are many applications in the market that evaluate the current price of cars. For example, “Otobid” service of sahibinden.com that predict the current selling price of a secondhand car. It uses light gbm and regression models. And also there is another service (“Araç Değerleme”) of sahibinden.com that predict the current selling and buying price of a secondhand car. It uses regression and random forest models for prediction. An it can be said that there are services like these that are presented by other companies such as Carvak, Arabam.com,Vavacars, Otoplus Let go in the market. On the contrary, there is not enough applications to predict future prices of secondhand cars. With this study, we basically targeted to answer to the following question is, “What will be the price of the vehicle to be sold/desired to be purchased within the time frame six months after the current date?”. By answering the blank question, it will be possible to determine the value of the vehicle more accurately and reliably during vehicle purchase/sale transactions. Sellers and buyers will be able to evaluate their situation economically and be able to perform their transactions better.

This research was done with the information obtained from the classifieds posted on sahibinden.com. By using the vehicles’ brand, model, package type, model year, body type, gear type, fuel type and mileage characteristics for the future six months sales prices of the vehicle was estimated over the past selling prices of vehicles. The data was collected from the classifieds in Automobile and SUV categories posted on sahibinden.com. The classifieds in the data set from February 2020 to December 2021 were scanned and information was gathered for vehicles. Then, some adjustments were done on the data set. Firstly, vehicles with 30 or more classifieds were included in the data set. Secondly, vehicles with data for at least 9 months in the last 12 months were included in the data set. Then, outliers were eliminated separately for each brand and model by using the Tukey outlier methodology (IQR). Finally, medians of classifieds were taken for the data set. After all these data cleaning adjustments are done, we implemented our models to estimate the prospective six months sales price of vehicles in the Automobile and SUV categories in Turkiye.We successfully implemented various forecasting techniques, including ARIMA, SARIMA, TBATS, HOLT WINTERS Forecasting, and machine learning algorithms such as ANN and XGBOOST using algorithms from libraries in R programing language. According to the success of these algorithms, the best algorithm was selected.

As a result of the estimation processes using various forecasting models, the HOLT WINTERS model resulted the most succesfull mean value of MAE 52.321, MSE 125.400.096.121 error and R2 (RMSE) 58.527 and MAPE 17% performance value. When we look at the meadian value of MAE 29.321, MSE 1.045.828.497 error and R2 (RMSE) 32.339 and MAPE 15% performance value the best. According to studies in literature, it can be said that the ability of Holt Winters forecasting is very good for predicting the price for the long period. When we looked at the past studies in literature, Rahman, et al., (2017) have tried to predict the monthly revenue of the Bangabandhu bridge by using Holt Winters model. The seasonal forecasting with trend issue was considerable importance. Therefore, the study has focus to forecast the monthly revenue using Holt–Winters. And also, Al-Hafid and Al-maamary (2012) aimed to develop a new model by using HOLT WINTERS method for forecasting electricity load demand that minimizes forecasting errors. This is important for demand controllers to ensure sufficient electricity supply to cope with increasing demands. These studies look like our study in terms of study target. The results showed that HOLT WINTERS are highly accurate like in our study. HOLT WINTERS' results are better than other methods in our study. The best estimation is resulted by HOLT WINTERS algorithm for the created dataset. HOLT WINTERS successfully passed all other forecasting methods in the study. HOLT WINTERS estimated the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye by using the vehicles' brand, model, package type, model year, body type, gear type, fuel type and mileage characteristics the best.

ANN performance with mean value of MAE 62.987, MSE 131.981.019.253 error and R2 (RMSE) 70.671 and MAPE 20% performance value. And also, when we look at ANN meadian value of MAE 34.932, MSE 1.503.076.236 error and R2 (RMSE) 38.769 and MAPE 20% performance value. When we looked at the past studies in literature Asilkan and Irmak (2009) has planned to estimate the future prices of secondhand cars using ANN. Varshitha et al., (2022) has studied with ANN to create a working model for secondhand car price prediction with low error value. And also, Karakoçı et al., (2020) have designed an ANN for car selling price estimation from a car selling website. They chose the ANN model because ANNs are often used for such calculations. According to those studies, it has been shown that the future prices of secondhand cars can be successfully predicted by using ANN. However, ANN results were not enough good in our study. ANN estimated the prices of secondhand cars, but HOLT WINTERS' results were better than ANN in our study.

TBATS performance with mean value of MAE 87.987, MSE 2.348.941.947.010 error and R2 (RMSE) 107.555 and MAPE 24% performance value. And also, when we look at TBATS meadian value of MAE 35.323, MSE 11.483.629.557 error and R2 (RMSE) 38.518 and MAPE 18% performance value. When we looked at the past studies in literature, Fajar (2021) has tried to predict the daily price of chili for controlling inflation by using TBATS. It looks like our study in terms of study target. The results

showed that TBATS forecasting using chili price data is highly accurate. However, TBATS results were not enough good in our study. TBATS estimated the prices of secondhand cars, but HOLT WINTERS' results were better than TBATS in our study.

XGBOOST performance with mean value of MAE 149.571, MSE 208.392.661.800 error and R2 (RMSE) 153.957 and MAPE 44% performance value. And also, when we look at XGBOOST meadian value of MAE 93.760, MSE 9.101.650.322 error and R2 (RMSE) 95.403 and MAPE 45% performance value. When we looked at the past studies in literature, Gajera et al., (2021) have used XGBOOST to build a statistical model to forecaste the selling price of a secondhand car. And XGBOOST did not peformed very well in their study like us. They chose Random Forest model is the mosts successfull model with highest R- squared value: 0.93 and we chose the HOLT WINTERS model resulted the most successful performance with mean value of MAE 52.321, MSE 125.400.096.121 error and R2 (RMSE) 58.527 and MAPE 17% performance value. When we look at the meadian value of MAE 29.321, MSE 1.045.828.497 error and R2 (RMSE) 32.339 and MAPE 15% performance value the best.

ARIMA performance with mean value of MAE 152.574, MSE 215.288.720.058 error and R2 (RMSE) 159.642 and MAPE 45% performance value. And also, when we look at ARIMA meadian value of MAE 97.639, MSE 10.146.560.424 error and R2 (RMSE) 100.730 and MAPE 46% performance value. When we looked at the past studies in literature, it is seen that the ARIMA model was chosen due to its suitability for time-series analysis in statistics. Guha and Bandyopadhyay (2016) created an ARIMA model to predict the future price of gold in India, providing guidance to investors on when to buy or sell gold. It looks like our study in terms of study target. The results showed that ARIMA forecasting using the future price of gold in India highly accurate. However, ARIMA results were not enough good in our study. ARIMA estimated the prices of secondhand cars, but HOLT WINTERS' results were better than ARIMA in our study.

SARIMA performance with mean value of MAE 128.925, MSE 212.659.296.692 error and R2 (RMSE) 145.908 and MAPE 31% performance value. And also, when we look at SARIMA meadian value of MAE 36.910, MSE 1.606.396.992 error and R2 (RMSE) 40.080 and MAPE 19% performance value.

When we looked at the past studies in literature, it is seen that the SARIMA model was chosen in order to analyze seasonal effects for time-series analysis in statistics. Perone (2021) has analyzed a few forecasting models to estimate the spread of COVID-19 during the pandemic's second wave in Italy (after October 13, 2020). It looks like our study in terms of study target. The results showed that SARIMA can predict the speed of spread of COVID-19. However, SARIMA results were not enough good in our study. SARIMA estimated the prices of secondhand cars, but HOLT WINTERS' results were better than SARIMA in our study.

According to mean and median results of these 4 indicators (MAE, MSE, RMSE, MAPE) the best forecasting is resulted by HOLT WINTERS model for the selected dataset. Generally,the mean values are taken into consideration in this kind of studies. However, in our study, we still checked both mean and median values. HOLT WINTERS successfully passed all other forecasting methods in the study. HOLT WINTERS estimated the future six months sales price of vehicles in the Automobile and SUV categories in Turkiye by using the vehicles' brand, model, package type, model year, body type , gear type, fuel type and mileage characteristics the best. That's why we will continue to study with HOLT WINTERS for our analysis in the future. And the results of HOLT WINTERS algorithm will be announced as output at the end of the study.

When the results obtained in the study findings are reviewed, it is understood that the most successful forecasting is resulted by the HOLT WINTERS model for the selected dataset. HOLT WINTERS successfully passed all other preprocessing methods in the study. However, it can be said there are some difficulties for this study. One of them is that it's really important to use a good feature selection/filtering method with the prediction algorithm to get the best performance. We chose and studied the most important features of a car for prediction of selling price of a car in this study. It can be said that more features can be selected and the study can be reapplied for better results. In addition, It's worth noting that the same algorithms can produce different results in different data sets. So we can reapply our study with different data sets to validate the results. And also, While market conditions and changes in car prices may cause the data to become outdated, our model remains reliable. To ensure continued accuracy, we recommend updating the price information at regular intervals when training models with the shared dataset.

On the other hand, we can mention there are some limitions for this study. The study yielded successful results despite the limitations. The date range selected for the analysis coincides with the covid pandemic period. Therefore, we can say that there is unusual enviroment for the second hand car market. Pirces of second hand cars has increased because of the many factors during the covid pandemic period. For example a worldwide crisis has born in automobile production during this period because of less of chip production. This sitation directly influenced the selling prices of secondhand cars and new cars. Prices of cars in the market has increased very much. Additionally, inflation has significantly increased due to economic conditions in Turkiye. That's why, prices of cars in the market has increased very much in Turkiye. So, we can say this is an usual sitiation for our study.The estimation of the prices of the cars has become difficult because of the uncertainity in the market. However, taking all these factors into consideration, predicting vehicle prices remains a challenging task.When we look at the results of the study, the study produced successful results despite some limitations. Our model accurately predicts the prospective six months sales price of vehicles in the Automobile and SUV categories in Turkiye.

This study shows that a number of points significant to researchers, automotive companies,automobile dealers and consumers. The results of this study can help for the companies and consumers who are related in the secondhand car market while making their purchasing and selling decision. The results will undoubtedly benefit users, especially researchers, who want to determine the price of their second-hand vehicles. From now on, researchers who want to study this subject can increase their contribution to the literature by taking into account a wider population, companies as well as consumers, and developing different models. And also companies can use the findings of this study to develop a new service for their costumers. For insance, sahibinden.com can develop a new service to show the future prices of a secondhand car by using the findings of the study. And other companies in the market can follow this development and they can develop new services about this area. Thus, people can see the selling prices of their vehicles for the next 6 months. And they will be able to assess their economic situation more accurately, enabling them to conduct selling/purchasing transactions more effectively.

After this study, Firstly, we will apply advanced machine learning techniques to validate the model with other various methods by enhancing its optimization and accuracy. We will measure the impact of changes in the secondhand market on the trained models and develop a method to tolerate such changes. Then, we will turn the trained model for predicting secondhand prices into a reliable service for buyers and sellers on sahibinden.com. Thus, all people will use this service safetly while they are purchasing or selling their vehicles. And also, we can create an automatic, interactive system that contains selling prices of secondhand cars. This system can allow users to know the selling price of a similar car using a recommendation engine.

REFERENCES

Abonazel, M. R., & Abd-Elftah, A.İ. (2019). Forecasting egyptian GDP using ARIMA models. *Reports on Economics and Finance*, 5(1), 35-47.

Alshared, A. (2021). Secondhands price prediction and valuation using data mining techniques. *Rochester Institute of Technology,* 1-37.

Asghar, M., Mehmood, K., Yasin, S., Khan, Z. M. (2021). Secondhands price prediction using machine learning with optimal features. *Pakistan Journal of Engineering and Technology*, 4(2), 113-119.

Asilkan, Ö., Irmak, S. (2009). İkinci el otomobillerin gelecekteki fiyatlarının yapay sinir ağları ile tahmin edilmesi. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi,* 14(2), 375-391.

Azcanlı, A. (2009). *Otomobilin Hikâyesi: Hayatımızı Değiştiren Mucize Araç.* İstanbul: Doğan Kitap Yayıncılık ve Yapımcılık

Berger, M. L. (2001). *The Automobile In American History And Culture: A Reference Guide, American Popular Culture.* London: Greenwood Press.

Bilen, M. (2021). Predicting secondhand prices with heuristic algorithms and creating a new dataset. *Journal of Multidisciplinary Developments.* 6(1), 29-43.

Bohr, H. (2020). Drug discovery and molecular modeling using artificial intelligence. *Artificial Intelligence in Healthcare*, 61-83.

Bourquin, J., Schmidli, H., Van Hoogevest, P., & Leuenberger, H. (1998). Advantages of artificial neural networks (ANNs) as alternative modelling technique for data sets showing non-linear relationships using data from a galenical study on a solid dosage form. *European Journal of Pharmaceutical Sciences*, 1(7), 5-16.

Cumhur. E., Senturk, I. (2009). A hedonic analysis of secondhand prices in Turkiye, *International Journal of Economic Perspectives*, 3(2), 141-149.

Çelik, O., Osmanoglu, O. (2019). Prediction of the prices of secondhand cars. *European Journal of Science and Technology*, (16), 77-83.

Dam, A. & Siegel, R.(2022,February 10). *Prices climbed 7.5% in January, compared with last year, continuing inflation's fastest pace in 40 years.*Washingtonpost.https://www.washingtonpost.com/business/2022/02/10/cpi-inflation-january/

Dastan, H. (2016). Türkiye'de ikinci el otomobil fiyatlarını etkileyen faktörlerin hedonik fiyat modeli ile belirlenmesi, *Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 18(1), 303-327.

Ecer, F. (2013). Türkiye'de 2. el otomobil fiyatının tahmini ve fiyat belirleyicilerinin tahmini. *Anadolu Üniversitesi Sosyal Bilimler Dergisi,* 13(4), 101-112.

Elmas, G. (2011). *Bursa Bölgesi'nde otomotiv lojistiği ve otomotiv terminallerinde kapasite analizi.* [Unpublished doctoral disertation]. İstanbul Üniversitesi.

Erkoç, A.(2023). The Motor Vehicle Dealers Federation (MASFED). https://masfed.com.tr/?page_id=1606

European Automobile Manufacturers' Association, (2023). www.acea.be

Fajar, M. (2021). Forecasting chili prices using TBATS. *International Journal of Scientific Research in Multidisciplinary Studies*, 7(2), 1-5.

Gajera, P., Gondaliya, A., Kavathiya, J. (2021). Old car price prediction with machine learning. *International Research Journal of Modernization in Engineering Technology and Science,* 3(3), 284-290.

Gegic, E., Becir, I., Dino, K., Zerina, M., Jasmin, K.(2019). Car price prediction using machine learning techniques. *TEM Journal,* 8(1), 113-118.

Guha, B., & Bandyopadhyay, G. (2016). Gold price forecasting using ARIMA model. *Journal of Advanced Management Science*, 4(2), 117-121.

Gumus, M., & Kiran, M.S. (2017, October 5-8). Crude oil price forecasting using XGBoost [Conference presentation]. *2017 International Conference on Computer Science and Engineering (UBMK)*, Antalya, Turkey.

Güneş, S. (2012). Türk toplumu ve otomobil. *SDÜ Fen Edebiyat Fakültesi Sosyal Bilimler Dergisi,* 25, 213-230.

Hafid, M.S. & Al-maamary, G. H. (2011). Short term electrical load forecasting using Holt-Winters method. *Al-Rafidain Engineering*, 20(6), 15-22.

Indicata Report (2022, Aralık), *Aralık 2022 Turkiye otomotiv 2.el online pazar analiz raporu*. Indicata.

https://www.indicata.com.tr/

Ji, S., Wang, X., Zhao, W., & Guo, D. (2019). An application of a three-stage XGBoost-Based Model to sales forecasting of a cross-border e-commerce enterprise. *Hindawi Mathematical Problems in Engineering,* 2019, 1-15.

Karbuz, F. & Çalışkan, E. (2017). Otomotiv Sektör Raporu. İstanbul Ticaret odası. https://www.ito.org.tr/tr

Köksal, Y. & Kahraman Türedi, M. (2014). Tüketici otomobil tercihinde etkili olan bilgi ve iletişim kanalları üzerine bir inceleme. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 17(32), 105-125.

Lejarraga, I., Kouzul-Wright, A., Primi, A., Toselli, M., and Wermelinger, M. (2016, November). Upgrading pathways in the automotive value chain. [Conference presentation]. *Background Document For the 7th Plenary Meeting of the OECD Initiative for Policy Dialogue On GVCs, Production Transformation and Upgrading, OECD*, Paris.

Lee, J. (2006). *Empirical analysis of wholesale secondhand auctions* [Unpublished master's thesis]. University of California.

Li B., Lee, Y., Yao, W., Lu, Y., & Fan, X. (2020). Development and application of ANN model for property prediction of supercritical kerosene. *Computers and Fluids*, 209, 1-18.

Monburinon, N., Chertchom, P., Kaewkiriya, T., Rungpheung, S., Buya, S., Boonpou, P. (2018). Prediction of prices for secondhand by using regression models. *5th International Conference on Business and Industrial* Research, 115-119.

Murray, J., Sarantis, N. (1999). Price-quality relations and hedonic price indexes for cars in the United Kingdom. *International Journal of the Economics of Business*, 6(1), 5-27.

Niu, Y. (2020, October 23-25). Walmart sales forecasting using XGBoost algorithm and feature engineering [Conference presentation]. *2020 International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE),* Chengdu, China.

Noor, K., Jan, S. (2017). Vehicle price prediction system using machine learning techniques. *International Journal of Computer Applications,* 167(9), 27-31.

Onat, M. G. (2007). *Otomotiv sektöründe oranlar yöntemi aracılığı ile finansal analiz* [Unpublished master's thesis]. Marmara Üniversitesi.

Otomotiv Sanayi Derneği, (2022). http://www.osd.org.tr/osd-uyeleri

Otomotiv Sanayi Derneği, (2022). *Toyato*. http://www.osd.org.tr/osd-uyeleri/toyota/

Özçalıcı, M. (2017). Predicting secondhand car sales price using decision trees and genetic algorithms. *Alphanumeric Journal*, 5(1), 103-114.

Özşahin, İ., & Özşahin, D. U. (2020). Neural network applications in medicine. *Biomedical Signal Processing and Artificial Intelligence in Healthcare*,[Academic Press].183 -206.

Perone, G. (2021). The determinants of COVID-19 case fatality rate (CFR) in the Italian regions and provinces: an analysis of environmental, demographic, and healthcare factors. *Science of the Total Enviroment,* 760, 144625

Pişkin, S. (2017). Türkiye Otomotiv Sanayii Rekabet Gücü ve Talep Dinamikleri Perspektifinde 2020 İç Pazar Beklentileri Otomotiv Sektör Raporu. Taşıt Araçları Tedarik Sanayicileri Derneği. https://www.taysad.org.tr/tr/bilgi-bankasi/otomotiv-raporlari.

Pollard, M. (1995). *Henry Ford ve Ford*. Ankara: İlk Kaynak Yayınları.

Presidency of the rebuplic of Turkiye investment office, (2023).Otomotiv. https://www.invest.gov.tr/en/sectors/pages/automotive

Pudaruth, S. (2014). Predicting the price of secondhands using machine learning techniques. *International Journal of Information & Computation Technology*, 4(7), 753- 764.

Rahman, H., Salma, U., Hossain, M., & Khan, T. F. (2016). Revenue forecasting using Holt–Winters exponential smoothing. Research & Reviews: *Journal of Statistics,* 5(3), 2278-2273

Samruddhi, K., Kumar, R. A. (2020). Secondhand price prediction using k-nearest neighbor based model. *International Journal of Innovative Research in Applied Sciences and Engineering, 4*(3), 686-689.

Smith, R. & Bruce, H. (2004). *Lean Maintenance: Reduce Costs, Improve Quality, And Increase Market Share*. USA: Elsevier

Statista, (2020). *Global automotive industry revenue between 2017 and 2030*, Statista GmbH. https://www.statista.com/

Taş, B. (2018). *Electricity price forecasting using hybrid time series models* [Unpublished master's thesis]. Middle East Technical University.

Taşıt Araçları Tedarik Sanayicileri Derneği, (2023). https://www.taysad.org.tr/tr

Toyata, (2023). http://www.toyotatr.com/?m=p&pid=8

Trademap, (2021). *Export Potential Map*, Trademap, Trade & Market Intelligence Section. https://exportpotential.intracen.org/en/markets/geomap?whatMarker=s&what=65&fromMarker=w&exporter=w&toMarker=j

Uysal, Ö. & Hatun Şimşek, G., (2022). *Yarım asırlık bir hayalden gerçeğe:Devrim arabaları.* http://devrimarabasi.com.tr/bir-basari-oykusu/

Varshitha, J., Jahnavi, K., Lakshmi, C. (2022, January 25-27). Prediction of secondhand prices using artificial neural networks and machine learning [Conference presentation], 2022 *International Conference on Computer Communication and Informatics (ICCCI*), Coimbatore, India.

Venkatasubbu, P., Ganesh, M. (2019). Secondhands price prediction using supervised learning techniques, *International Journal of Engineering and Advanced Technology*, 9(1S3), 2249-8958.

Yang, Y., Yu, H., & Sun, Z. (2017, April 28-30). Aircraft failure rate forecasting method based on Holt-Winters seasonal model [Conference presentation]. *2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA),* Chengdu, China.

Yılmaz, E. (2016). *Türkiye otomotiv sanayinin gelişimi ve ikinci el otomobil talep fiyatının belirleyicileri* [Unpublished master's thesis]. Gaziosmanpaşa Üniversitesi.