



AUTO-गाड़ी

Analysis of the Automotive Industry

Paras Gupta



Table of contents

01



Introduction

02



Literature Survey

03



Research Gap

04



Problem Statement

Table of contents

05

1 3 5
T H
2 4 R

Methodology

06



Implementation

07



Result

08



Conclusion



01

INTRODUCTION



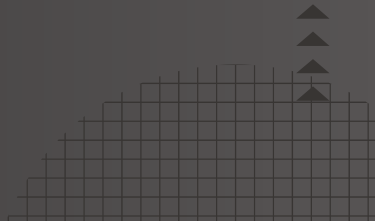
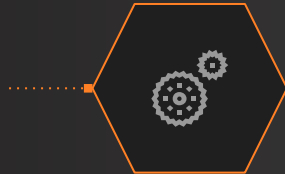
INTRODUCTION



The term "Automotive Industries" refers to all businesses and endeavours concerned with the production of motor vehicles, including the majority of its parts, such as their engines and bodywork, but excludes their tyres, batteries, and fuel.

The automotive industry continues to face a growing number of challenges and pressures. Cost pressure, competition, globalization, market shifts, and volatility are all increasing.

The background of the "AUTO-GAADI" project would involve data analysis and machine learning techniques to predict the price of a car and analyze the trends in car sales.



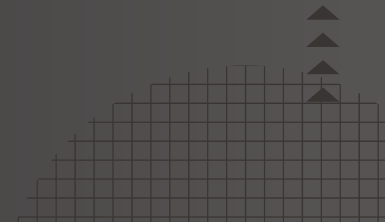
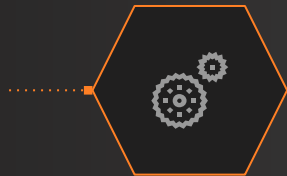
INTRODUCTION



To carry out this project, you would need access to a large dataset of car sales information, including factors such as car price, model, year, mileage, condition, location, and price. This data can be collected from various sources such as online car sales platforms, car dealerships, and public records.

The project would involve data cleaning and preparation to ensure that the data is accurate and consistent. This would include removing any missing or duplicate, converting data into a standardized format, and identifying any outliers or anomalies.

Overall, the project would require a strong understanding of data analysis and machine learning techniques, as well as a background in programming and statistics.

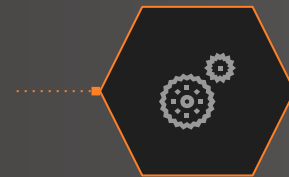


INTRODUCTION



The goal of this web app is to understand key performance indicators for automotive industry in the US and to predict the price of the Car with respect to Car Age, Fuel Type, Transmission Type and others .

We have trained our model with different datasets taken from “Data.world “ and The data consisted of 48 car brands and in total 915 different car models. This Price Prediction in our project is going to help the user to buy and sell a car with an unbiased price easily and without any loss .





02

LITERATURE SURVEY



LITERATURE SURVEY

To conduct a literature survey on the innovation of an AUTO-GAADI, we explored the existing systems of several company to gain more insight into the topic. The results of our survey are presented below.

There is an online platform **TrueCar** that provides pricing transparency and market insights to help consumers and dealers make informed car buying and selling decisions.

CarGurus - an online automotive marketplace that provides pricing insights, dealer reviews, and car valuation tools to help buyers and sellers make informed decisions.





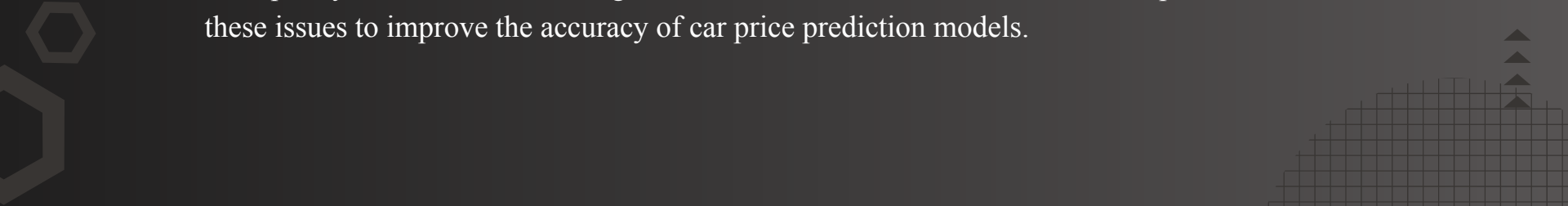
03

RESEARCH GAP





RESEARCH GAP

1. Incorporating non-linear relationships: Many existing car price prediction models assume linear relationships between variables. However, there may be non-linear relationships that could be important for predicting car prices. Further research could explore the use of non-linear models to improve accuracy.
 2. Incorporating unstructured data: Many car price prediction models rely on structured data such as car make, model, year, and mileage. However, unstructured data such as user reviews or social media sentiment could also be useful for predicting car prices. Further research could explore how to incorporate unstructured data into car price prediction models.
 3. Addressing data quality issues: The accuracy of car price prediction models can be impacted by data quality issues such as missing or incorrect data. Further research could explore how to address these issues to improve the accuracy of car price prediction models.
- 



04

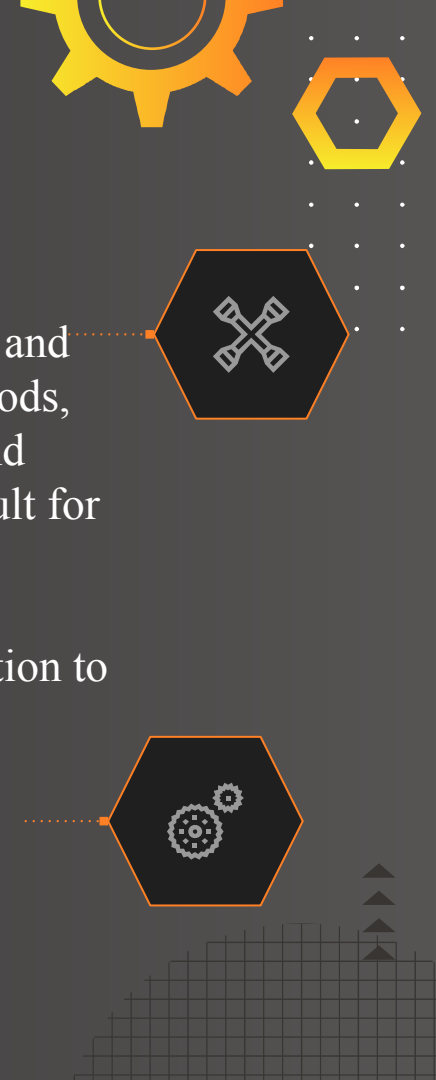
PROBLEM STATEMENT



PROBLEM STATEMENT

The current car sales industry faces a challenge in terms of transparency and consistency in pricing. There is a lack of standardization in pricing methods, and prices can vary significantly across different regions, dealerships, and sales channels. This can create confusion for buyers, and it can be difficult for sellers to determine the fair market value of a vehicle.

AUTO-GAADI can provide more reliable and accurate pricing information to help buyers and sellers make informed decisions.





05

METHODOLOGY



Snippet of Dataset Used:

	A	B	C	D	E	F	G	H	I	J
1	car	price	body	mileage	engV	engType	registratio	year	model	drive
2	Ford	15500	crossover	68	2.5	Gas	yes	2010	Kuga	full
3	Mercedes-	20500	sedan	173	1.8	Gas	yes	2011	E-Class	rear
4	Mercedes-	35000	other	135	5.5	Petrol	yes	2008	CL 550	rear
5	Mercedes-	17800	van	162	1.8	Diesel	yes	2012	B 180	front
6	Mercedes-	33000	vagon	91	NA	Other	yes	2013	E-Class	
7	Nissan	16600	crossover	83	2	Petrol	yes	2013	X-Trail	full
8	Honda	6500	sedan	199	2	Petrol	yes	2003	Accord	front
9	Renault	10500	vagon	185	1.5	Diesel	yes	2011	Megane	front
10	Mercedes-	21500	sedan	146	1.8	Gas	yes	2012	E-Class	rear
11	Mercedes-	22700	sedan	125	2.2	Diesel	yes	2010	E-Class	rear

ALGORITHM USED:

Random Forest Regression is a type of machine learning algorithm that is used for predicting continuous values (regression). It is an ensemble learning method that combines multiple decision trees to improve the accuracy and robustness of the prediction.

The basic idea behind the random forest regression algorithm is to construct multiple decision trees on randomly selected subsets of the training data, and then aggregate the predictions of all the trees to arrive at the final prediction. Each decision tree is constructed based on a random subset of features, and at each node of the tree, the best split is chosen from a random subset of features.

The final prediction of the random forest regression model is the average of the predictions of all the decision trees. This averaging process helps to reduce the variance of the predictions and improve the overall accuracy of the model.



Various Python libraries are used in this project like:

Numpy

Pickle

Pandas

Sklearn

Json

Plotly

Plotly.express , plotly.graph_objs

Random Forest Regressor



CODE SNIPPET

Training and Testing :

```
X = data.drop('Price', axis=1)
y = data['Price']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
                                                    random_state=42)
```

Random Forest Regression model and fit it to the Training Data:

```
rf = RandomForestRegressor(n_estimators=100, random_state=42)
rf.fit(X_train, y_train)
```

CODE SNIPPET

Testing data and evaluate the performance of the model:

```
y_pred = rf.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
print("Mean Squared Error: ", mse)
```

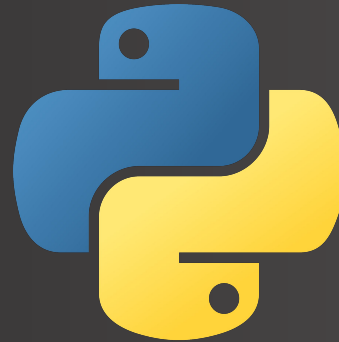
TECH-STACK used



HTML



CSS





06

IMPLEMENTATION



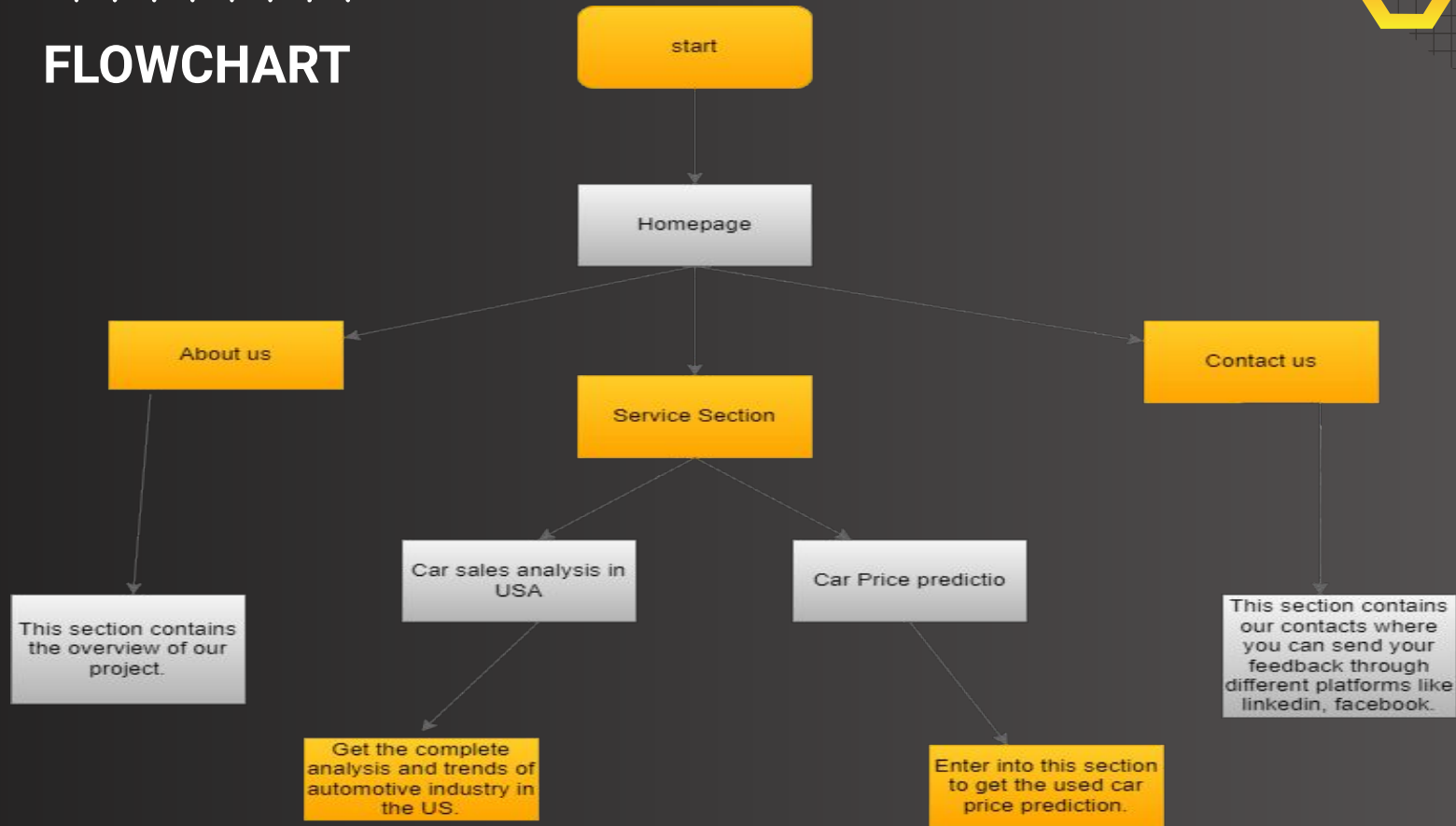
STRUCTURE OF THE PROJECT

- The home page consists of an about section, a services section and a contact section where in the services section you will find two options i.e, Car Sales Analysis in US and Car Price Prediction.
- The First Section consists of Car Sales Analysis in US where different queries are answered in the form of visualization related to Car Sales in US. The user get to know about the different trends in automotive industry.
- Using the above analysis, we can get a complete understanding about the automotive industry.

STRUCTURE OF THE PROJECT

- The Last Section consists of Used Car price prediction where the Machine Learning Model uses the Random Forest Regression to predict the price of the car.
- The prediction section consists of forms which take few different feature as input from the user in a given specified range. The most relevant features are taken into consideration for prediction also these features
- So, with this the User can get a complete analysis of the current trends related to automotive industry and Car Sales.

FLOWCHART



Automotive Industry - Analysis of Cars

Welcome to Auto-गाड़ी. Here you can get a complete analysis and recent trends about all types of Car in Automotive Industry. You can demonstrate how the Automotive Industry could harness data to take informed decisions.

GET STARTED

WHO WE ARE



Analysis of Automotive Industry

Automotive industries are all those companies and activities involved in the manufacture of motor vehicles, including most components, such as engines and bodies, but excluding tires, batteries, and fuel. The goal of this web app is to understand key performance indicators for automotive industry in the US. The data consisted of 48 car brands and in total 915 different car models. The automotive industry continues to face a growing number of challenges and pressures. Cost pressure, competition, globalization, market shifts, and volatility are all increasing. We understand the problem and created a web application where data analytics will allow you to see the unseen and make smart decisions and derive value from the vast amounts of data and to maintain market position and profits in the air of forever shrinking margins. To

make the automotive industry more understandable, we have used Data Analysis to make statistics and trends clearly visible to the users with Visualization. This web application help an user to get a clear and easy understanding and how the Automotive Industry could harness data to take informed decisions. Along with it we also predict an used car price which would help user to buy and sell a car by using the price prediction section.

SERVICES

Welcome to the Services section of AUTO-गाड़ी. Get your queries answered by the Car Sales Analysis in the US. Also get to know about the price of car by entering into the second section.

01

Car Sales Analysis in US

Enter into this section to know about the recent trends and analysis regarding the Car in US and get clear understanding.

Enter

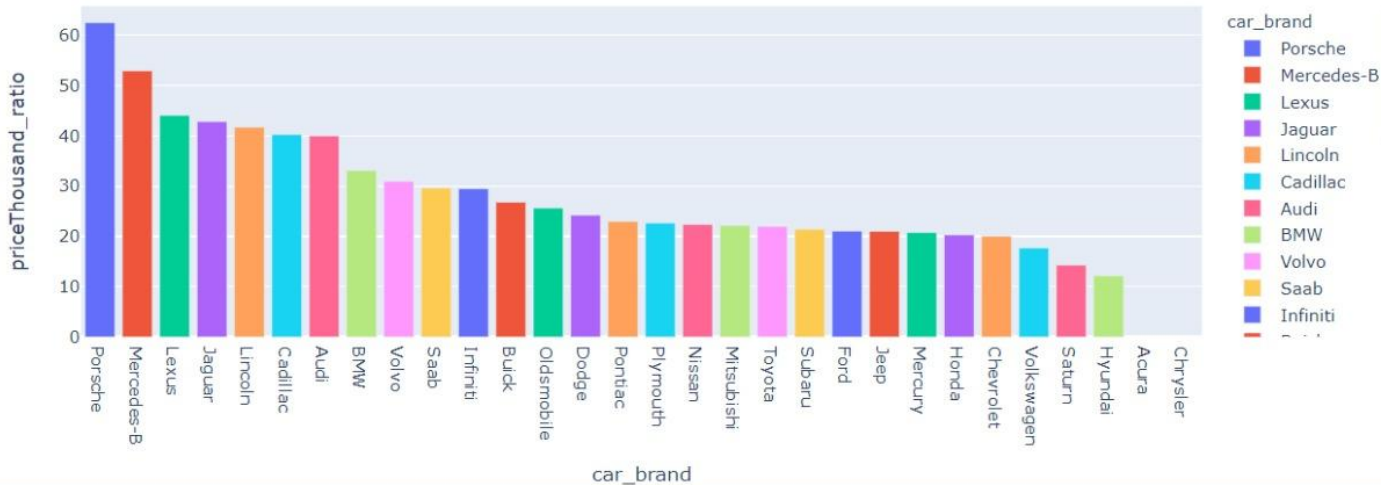
02

Predict Used Car Price

Enter into this section which will helps you to predict the price of the used car. By using this you can sell or buy a car by knowing the actual price.

Enter

Car brand price thousand ratio



Which car brand has the most Price Ratio?

After analyzing this bar chart plotted between Price in thousand ratio versus car brand.

1. Clearly the highest price in thousand ratio is Porsche.
2. The second highest is Mercedes-Benz.
3. The lowest is Hyundai whereas Acura and Chrysler have no price ratio.

Car Price Prediction

Enter the Year:

Enter the Showroom Price(In lakhs):

Enter the no of Kilometers Drived:

Enter the no of owners:

Enter the Fuel type:

Enter the seller Type:

Enter the Transmission type

Predict the price

You Can Sell The Car at 10.51 lakh



07

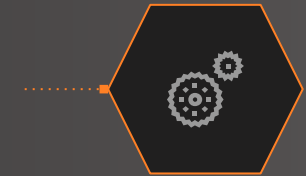
RESULT



RESULT

The AUTO-GAADI system should provide tangible benefits such as cost savings, increased sales, and improved customer satisfaction. It should also help reduce the risk of overpricing or underpricing cars, leading to more balanced supply and demand in the market.

Finally, the system should be able to keep up with changing market trends and customer preferences, ensuring its continued relevance and usefulness over time and patterns in car sales data, helping consumers to make informed decisions about pricing, inventory management, and sales strategies .





08

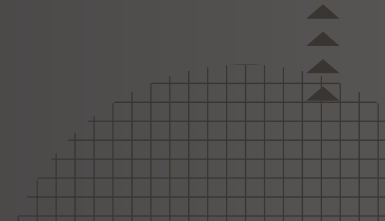
CONCLUSION



CONCLUSION

The AUTO-GAADI project is a promising initiative that has the potential to revolutionize the car sales industry by providing buyers and sellers with more accurate and reliable information about car prices and market trends.

However, to ensure the success of the project and maximize its social and environmental impact, it is important to conduct a thorough feasibility study, perform rigorous risk analysis, and engage with stakeholders throughout the design and implementation process.



REFERENCES

- “Predicting the Price of Used Cars using Machine Learning Techniques”, Sameer Chand Pudaruth, Computer Science and Engineering Department, University of Mauritius, Reunion, MAURITIUS
- “US Auto Production and Price Prediction in the Context of Multiple Regression Analysis”, Qilin Li, Department of Economics, University of California, Irvine, the U.S.
- “The Framework of Car Price Prediction and Damage Detection Technique”, Muhammad Kashif Shaikh, Department of Software Engineering, Sir Syed University of Engineering & Technology, Karachi, Pakistan.
- Web Development Tools: W3Schools.
- Stack Overflow
- Python tools: W3Schools.
- Youtube channels: Code with harry



Thanks!

