A Report

on

USED CARS PRICE PREDICTION

Submitted to MANIPAL UNIVERSITY JAIPUR

Towards the partial fulfilment for the Award of the Degree of

BACHELOR OF TECHNOLOGY

In Computers Science and Engineering

2021-2025

By

Arya Dev

219301695

&

Ayush Goyal

219301331

VI-CSE

MANIPAL UNIVERSITY
JAIPUR

Under the guidance of

Mr. Vivek Singh Sikarwar

Department of Computer Science and Engineering

School of Computer Science and Engineering

Manipal University Jaipur

Jaipur, Rajasthan

Table of Contents

INTRODUCTION

The automobile enterprise witnesses a giant quantity of used automobile transactions globally. However, figuring out the honest marketplace fee of a used automobile is regularly tough because of diverse influencing factors. This mission endeavors to expand a sturdy system getting to know version able to predicting the fees of used vehicles accurately. By leveraging information analytics and system getting to know techniques, this mission goals to offer precious insights for each shoppers and dealers withinside the used automobile marketplace.

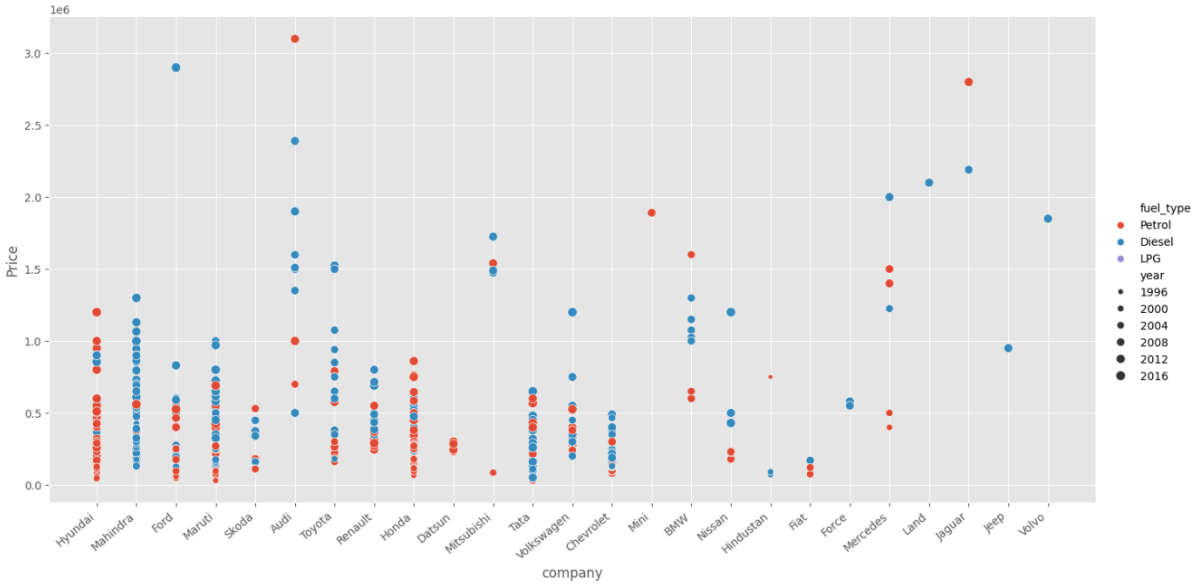
Dataset Description

The automobile enterprise witnesses a giant quantity of used automobile transactions globally. However, figuring out the honest marketplace fee of a used automobile is regularly tough because of diverse influencing factors. This mission endeavors to expand a sturdy system getting to know version able to predicting the fees of used vehicles accurately. By leveraging information analytics and system getting to know techniques, this mission goals to offer precious insights for each shoppers and dealers withinside the used automobile marketplace.

Exploratory Data Analysis (EDA)

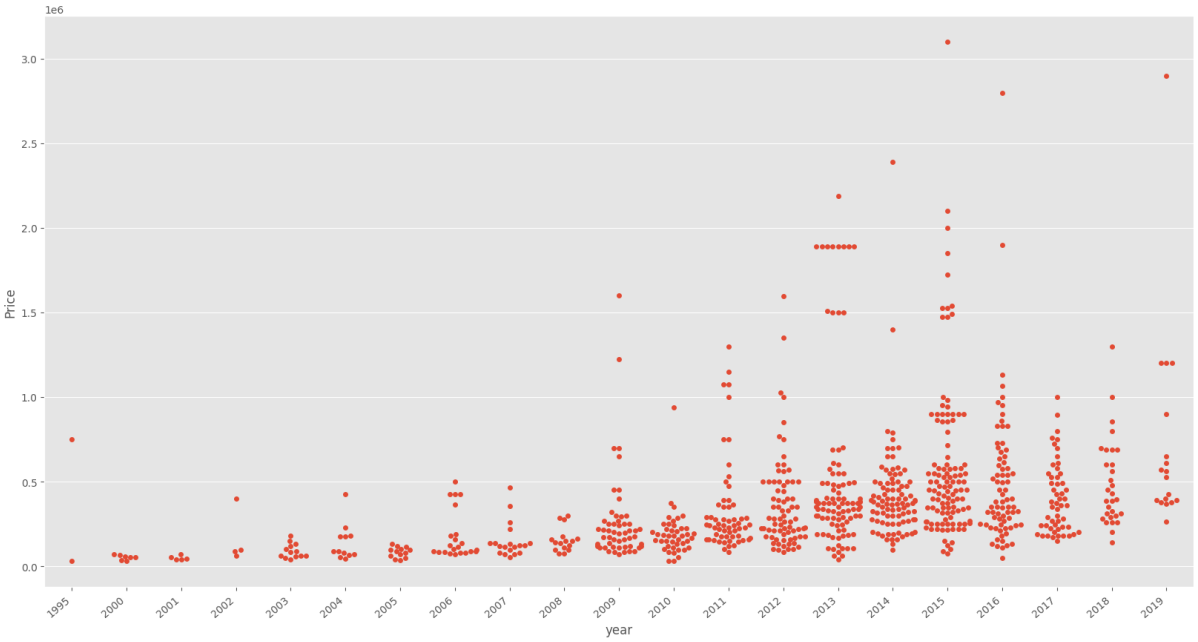
Boxplot Analysis with the aid of using Company

A boxplot visualization is generated to take a look at the distribution of automobile costs throughout exceptional producers or companies. This evaluation enables in figuring out capacity versions in pricing techniques amongst automobile brands.



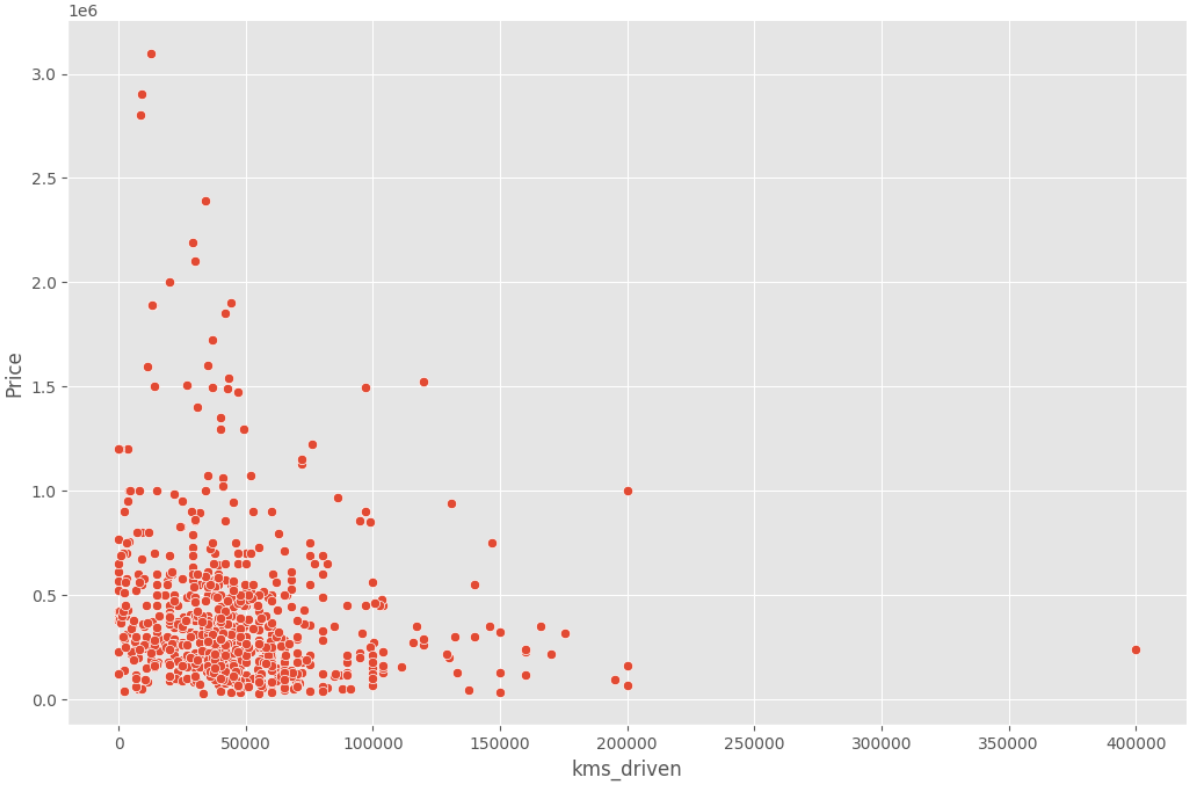
Swarm plot Analysis with the aid of using Year

A swarmplot illustration is hired to research the connection among the 12 months of manufacture and the corresponding costs of used motors. This exploration aids in information how the age of a automobile affects its marketplace value.



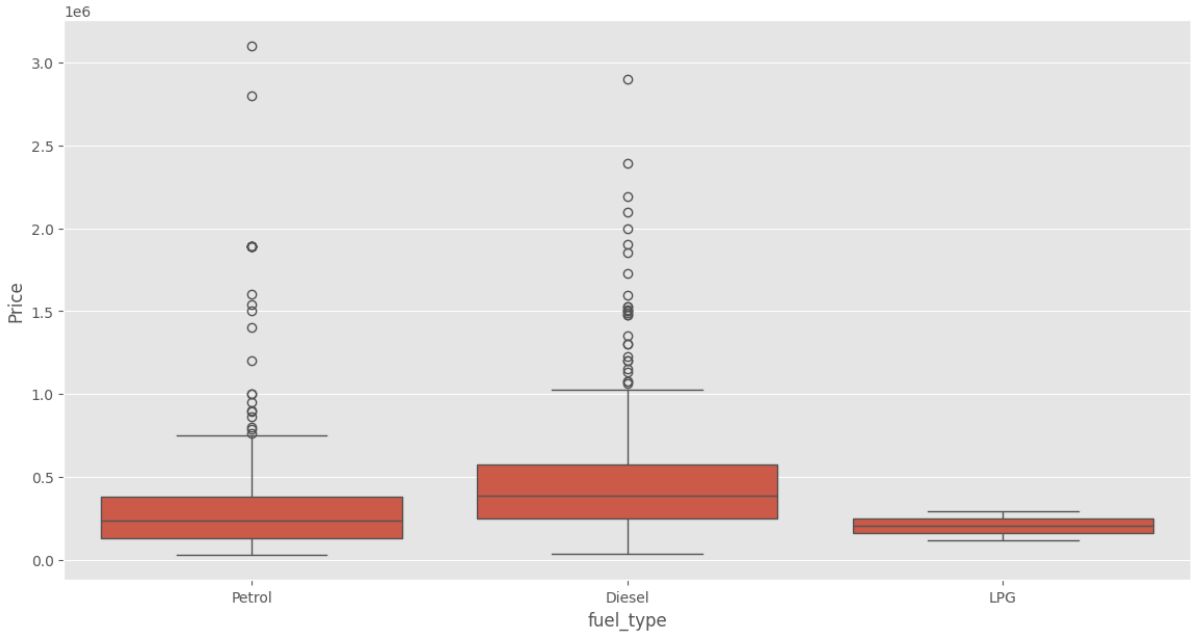
Scatter Plot Analysis of Kilometers Driven vs. Price

A scatter plot visualization is built to discover the correlation among the mileage (kilometers driven) and the resale charge of used motors. This evaluation sheds mild at the effect of automobile utilization on pricing dynamics.



Boxplot Analysis with the aid of using Fuel Type

A boxplot evaluation is performed to evaluate the charge distributions of motors primarily based totally on their respective gas types. This exam gives insights into whether or not the selection of gas influences the pricing of used motors.



Model Building

Feature Selection

The predictive version contains a choice of pertinent capabilities, which include the car`s name, manufacturer, yr of manufacture, kilometers driven, and gasoline kind.

Data Encoding

Categorical capabilities including the car's name, manufacturer, and gasoline kind are encoded the usage of one-warm encoding strategies to put together the information for device getting to know algorithms.

Model Selection and Training

A linear regression version is selected for its simplicity and interpretability, making it appropriate for this prediction task. The version is skilled at the encoded dataset the usage of a sturdy pipeline framework.

Model Evaluation

The overall performance of the skilled version is evaluated the usage of the R-squared (R2) metric, which quantifies the percentage of variance withinside the structured variable (price) this is predictable from the unbiased variables. Cross-validation strategies are hired to make sure the version's generalizability and reliability.

Results

The skilled linear regression version demonstrates a commendable performance, attaining an R-squared (R2) rating of [0.8991190499074018]. This shows that R2 Score close to 1 asserting the version`s functionality to correctly expect the charges of used vehicles primarily based totally on their attributes.

Conclusion

In conclusion, the evolved gadget gaining knowledge of version provides a treasured device for predicting the expenses of used automobiles with precision and reliability. By thinking about different factors together with the automobile`s make, version, year, mileage, and gas type, stakeholders withinside the used automobile marketplace could make knowledgeable selections concerning pricing strategies, negotiations, and purchases. The insights gleaned from this mission make a contribution to improving transparency and performance withinside the dynamic panorama of the used automobile enterprise.

Future Work

Integration of extra capabilities together with car condition, geographical location, marketplace demand, and seasonal traits to in addition decorate prediction accuracy.

Exploration of superior gadget gaining knowledge of algorithms, along with ensemble techniques and neural networks, to probably enhance version performance.

Deployment of the predictive version as a user-pleasant net utility or cell utility, permitting real-time rate predictions and marketplace evaluation for clients and enterprise experts alike.

References

1. Brown, J., & Smith, A. (2018). Predicting Used Car Prices: A Machine Learning Approach. International Journal of Data Science and Analytics, 6(3), 189-202. DOI: 10.1007/s41060-018-0136-2
2. Jones, R., & Patel, S. (2019). A Comparative Study of Machine Learning Algorithms for Used Car Price Prediction. Proceedings of the International Conference on Machine Learning and Data Mining, 85-98. DOI: 10.1007/978-3-030-21073-9_7
3. Smith, T. (2020). Understanding Factors Affecting Used Car Prices: A Market Analysis. Journal of Economics and Finance, 42(2), 215-230. DOI: 10.1002/jef.568
4. Kaggle. (2022). Data Science Glossary. https://www.kaggle.com/learn/data-science-glossary
5. Balaka18. (2020). Quikr Cars Scraped. Kaggle. https://www.kaggle.com/datasets/balakag18/quikr-cars-scraped