भारतीय सूचना प्रौद्योगिकी संस्थान, नागपुर
Indian Institute of Information Technology, Nagpur
Institute of National Importance by the Act of Parliament
The Indian Institute of Information Technology (PPP) Act, 2017

ITW

PROJECT REPORT ON USED CAR PRICE PREDICTION

PRESENTED TO
DR SHISHUPAL KUMAR

PRESENTED BY
VICKY KUMAR
BT22ECE051

Table of Contents

01 Abstract

02 Introduction

03 Dataset

04 Code and Result

05 Future work

06 Conclusion

Abstract

This report explores the development and evaluation of a prediction model for estimating the selling price of used cars. The dataset used is a CSV database containing various features such as car name, buying year, kilometer driven, fuel type, and present car price. Two regression algorithms, Linear Regression and Random Forest Regression, are employed to build predictive models. The report covers data preprocessing, model training, and evaluation. Performance metrics including Mean Absolute Error (MAE), Mean Squared Error (MSE), are utilised to assess the models' effectiveness. Additionally, feature importance analysis is conducted for the Random Forest model to identify significant predictors. The findings of this study provide insights into the predictive capabilities of the models and offer recommendations for future improvements.

Introduction

The used car market presents a complex and dynamic environment where various factors influence the selling price of a vehicle. In this report, we delve into the development and evaluation of a prediction model aimed at estimating the selling price of used cars. Leveraging a CSV database containing essential attributes such as car name, buying year, kilometer driven, fuel type, and present car price, we explore the effectiveness of two regression algorithms: Linear Regression and Random Forest Regression.

The primary objective of this report is to analyze the performance of these prediction models and provide insights into their predictive capabilities. By employing techniques in data preprocessing, exploratory data analysis, model training, and evaluation, we aim to understand the factors driving the selling price of used cars and to assess the accuracy of our predictive models.

This introduction outlines the structure of the report, which includes sections covering data preprocessing, EDA, model building, results analysis, and recommendations for future improvements. Through this study, we seek to contribute to the understanding of the used car market and provide valuable insights for stakeholders in the automotive industry.

Future Work

In the future, I plan to integrate the prediction model into a Flutter app for easy estimation of used car prices. To increase accuracy, I will:

1. Expand the dataset with additional relevant features.
2. Explore advanced regression algorithms and ensemble methods.
3. Fine-tune model hyperparameters.
4. Implement cross-validation techniques.
5. Analyze prediction errors and gather user feedback for continuous improvement.

Dataset

Car_Name	Year	Selling_Price	Present_Price	Driven_kms	Fuel_Type	Selling_type	Transmission	Owner
ritz	2014	3.35	5.59	27000	Petrol	Dealer	Manual	0
sx4	2013	4.75	9.54	43000	Diesel	Dealer	Manual	0
ciaz	2017	7.25	9.85	6900	Petrol	Dealer	Manual	0
wagon r	2011	2.85	4.15	5200	Petrol	Dealer	Manual	0
swift	2014	4.6	6.87	42450	Diesel	Dealer	Manual	0
vitara brezza	2018	9.25	9.83	2071	Diesel	Dealer	Manual	0
ciaz	2015	6.75	8.12	18796	Petrol	Dealer	Manual	0
s cross	2015	6.5	8.61	33429	Diesel	Dealer	Manual	0
ciaz	2016	8.75	8.89	20273	Diesel	Dealer	Manual	0
ciaz	2015	7.45	8.92	42367	Diesel	Dealer	Manual	0
alto 800	2017	2.85	3.6	2135	Petrol	Dealer	Manual	0
ciaz	2015	6.85	10.38	51000	Diesel	Dealer	Manual	0
ciaz	2015	7.5	9.94	15000	Petrol	Dealer	Automatic	0
ertiga	2015	6.1	7.71	26000	Petrol	Dealer	Manual	0
dzire	2009	2.25	7.21	77427	Petrol	Dealer	Manual	0
ertiga	2016	7.75	10.79	43000	Diesel	Dealer	Manual	0
ertiga	2015	7.25	10.79	41678	Diesel	Dealer	Manual	0
ertiga	2016	7.75	10.79	43000	Diesel	Dealer	Manual	0
wagon r	2015	3.25	5.09	35500	CNG	Dealer	Manual	0
sx4	2010	2.65	7.98	41442	Petrol	Dealer	Manual	0
alto k10	2016	2.85	3.95	25000	Petrol	Dealer	Manual	0
ignis	2017	4.9	5.71	2400	Petrol	Dealer	Manual	0
sx4	2011	4.4	8.01	50000	Petrol	Dealer	Automatic	0
alto k10	2014	2.5	3.46	45280	Petrol	Dealer	Manual	0
wagon r	2013	2.9	4.41	56879	Petrol	Dealer	Manual	0
swift	2011	3	4.99	20000	Petrol	Dealer	Manual	0
swift	2013	4.15	5.87	55138	Petrol	Dealer	Manual	0
swift	2017	6	6.49	16200	Petrol	Individual	Manual	0
alto k10	2010	1.95	3.95	44542	Petrol	Dealer	Manual	0
ciaz	2015	7.45	10.38	45000	Diesel	Dealer	Manual	0
ritz	2012	3.1	5.98	51439	Diesel	Dealer	Manual	0

conclusion

Our analysis indicates that the Random Forest model outperforms the Linear Regression model in predicting used car prices. It offers more accurate predictions and a better overall model usefulness score. Therefore, we recommend the Random Forest model for practical use in predicting used car prices due to its superior performance.

:

I extend my heartfelt gratitude to Dr. Shishupal Kumar for his invaluable guidance and support throughout this project. His expertise and mentorship were crucial in navigating through the complexities of machine learning and enhancing the quality of this work.

Thank you,
Dr. Shishupal Kumar, for your continuous support and encouragement.

Vicky kumar

CODE