Domain: Applied Data Science
Project: Car Resale Value
PredictionTeam Size :5
Team Leader: Vasipalli Nithin Bharath Reddy
Team Members: Venkata Sahith Andey ,Rekkala Gowtham Reddy ,PSS Rohit Varma ,V Pavan Varma

Project Design Phase-I
Proposed Solution Template

Proposed Solution Template:

Project team shall fill the following information in proposed solution template.

S.No.	Parameter	Description
1.	Problem Statement (Problem tobesolved)	Machine learning has become a tool used in almost every task that requires estimation. Companies like cars24 and cardekho. Com usesRegression analysis to estimate the usedcar prices. So we need to build a model to estimatethe price of cars. The model should take car- related parameters and output a selling price. The selling price of a used car dependsoncertain features as mentioned below • Fuel Type • Manufacturing year • Miles Driven • Number of Historical Owners • Maintenance Record This is a supervised learning problem and can be solved using regression techniques. We need to predict the selling price of a car basedon the given car's features. SupervisedRegression problems require labeled data where our target or dependent variable is the selling price of a car. All other featuresareindependent variables.
2.	Idea / Solution description	This project aims to deliver price prediction models to the public, to help guide the individuals looking to buy or sell cars and to give them a better insight into the automotive sector. Baying a used car from a dealer can be afrustrating and an unsatisfyingexperience as some dealers are known to deploy deceitful Dictactics to close a deal. Therefore, to help consumers avoid falling victims to mach tactics,this study hopes to equip consumers with right tools to guide them in their shopping experience. Another goal of the project is to explore newmethods to evaluate used cars prices and to compare t their accuracies. Considering thisisan interesting research topic in the research

		community, and in confirong heir footsteps,wehope to achieve significant results wing more advanced methods of previous work
3.	Novelty / Uniqueness	As there are so many ongoing experiments thatuse statistical approaches and some traditionalmethods to focus on predicting item sales. Most researches have experimented by takingasingle algorithm to predict sales. In this thesis Machine Learning algorithms such as Simple Linear Regression, Support Vector Regression, Gradient Boosting algorithm, andRandom Forest Regression are considered forpredictionand the most effective metrics suchas accuracy, mean absolute error, and max error are considered for measuring algorithm efficiency.This method will be very beneficial in the future for advanced item sales forecasting
4.	Social Impact / Customer Satisfaction	In the study, the variables having significant effects on the price of the second hand car were determined. A prediction model was established with these variables. The coefficientof determination (R2) of this modelwas calculated as 89.1%. The variables included in the estimation model are Brand, Model, Model Year, Fuel Type, Horse Power, Kilometer, Manual Air Conditioning, Fog Lights, Seat Air Cushion, Leather Steering Wheel, Wheel Rim, Automatic Air Conditioning, Start Stop, Rain Sensor, Sunroof,Electric Folding Mirrors, XenonHeadlight, Knee Airbag, Upholstery Leather, Memory Seat, 4X4, Parking Assistant, Vacuum Door.
5.	Business Model (Revenue Model)	Deciding whether a used car is worth the posted price when you see listings online canbe difficult. Several factors, including mileage,make, model, year, etc. can influence the actualworth of a car. From the perspective ofa seller,it is also a dilemma to price a used carappropriately[2-3]. Based on existing data, the aim is to use machine learning algorithmsto develop models for predicting used car prices.

6.	Scalability of the Solution	We started with understanding the use caseof machine learning in the Automotive industry and how machine learning has transformed thedriving experience. Moving on, we looked at the various factors that affect the resale value of a used car and performed exploratory data analysis (EDA). Further, we build a Random Forest Regression model to predict the resale valueof a used car. Finally, we evaluated the performance of the model using the R squared score and Residual Plot.
		We could have also used simpler regression algorithms like Linear Regression and Lasso Regression. Still, we need to make sure thereare no outliers in the dataset before implementing them. Pair plots and scatter plots help visualize the outliers.