Project Title: Car Resale value Prediction

Team Members

1913075 - Reshma.V

1913089 - Sri Hari.V

1913107 - Yavanika.V

1913507 - Soundhirarajan.C

**Project abstract**

Car resale value prediction system is made with the purpose of predicting the correct valuation of used cars that helps users to sell the car remotely with perfect valuation and without human intervention in the process to eliminate biased valuation.Due to the huge requirement of used cars and lack of experts who can determine the correct valuation, there is an utmost need of bridging this gap between sellers and buyers. This project focuses on building a system that can accurately predict a resale value of the car based on minimal features like kms driven, year of purchase etc. without manual or human interference and hence it remains unbiased In this project we have used different algorithms with different techniques for developing Car resale value prediction systems considering different features of the car. In a nutshell, car resale value prediction helps the user to predict the resale value of the car depending upon various features like kilometres driven, fuel type, etc

This resale value prediction system is made for general purpose to just predict the amount that can be roughly acquired by the user. We try to predict the amount of resale by best 70% accuracy so the user can get estimated value before he resales the car and doesn't make a deal in loss. The main idea of making a car resale value prediction system is to get hands-on practice for python using Data Science. Determining whether the listed price of a used car is a challenging task, due to the many factors that drive a used vehicle’s price on the market. The focus of this project is developing machine learning models that can accurately predict the price of a used car based on its features, in order to make informed purchases. We implement and evaluate various learning methods on a dataset consisting of the sale prices of different makes and models . We will compare the performance of various machine learning algorithms like Linear Regression, Ridge Regression, Lasso Regression, Elastic Net, Decision Tree Regressor and choose the best out of it. Depending on various parameters we will determine the price of the car. Regression Algorithms are used because they provide us with continuous value as an output and not a categorized value because of which it will be possible to predict the actual price a car rather than the price range of a car. User Interface has also been developed which acquires input from any user and displays the Price of a car according to user’s inputs.

**Literature Survey on the Car Resale value Prediction**

The first paper is Predicting the price of Used Car Using Machine Learning Techniques. In this paper, they investigate the application of supervised machine learning techniques to predict the price of used cars in Mauritius. The predictions are based on historical data collected from daily newspapers. Different techniques like multiple linear regression analysis, k-nearest neighbours, naïve bayes and decision trees have been used to make the predictions.

The Second paper is Car Price Prediction Using Machine Learning Techniques. Considerable number of distinct attributes are examined for the reliable and accurate prediction. To build a model for predicting the price of used cars in Bosnia and Herzegovina, they have applied three machine learning techniques (Artificial Neural Network, Support Vector Machine and Random Forest).

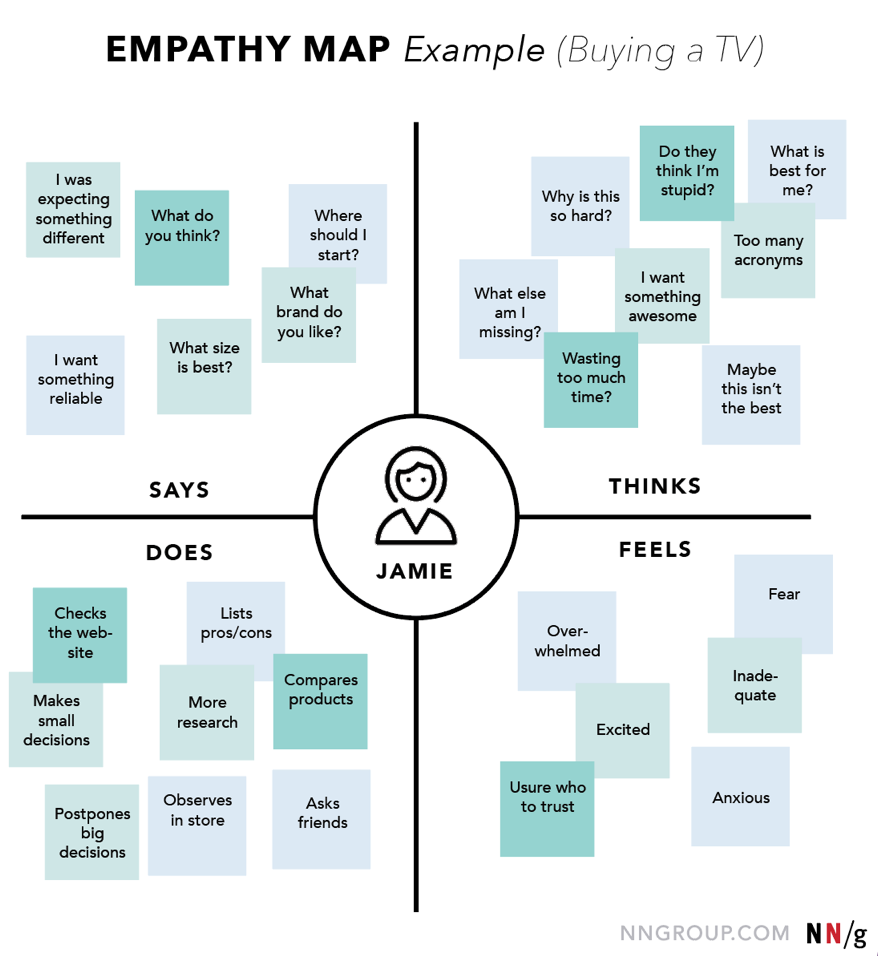
The Third paper is Price Evaluation model in second hand car system based on BP neural networks. In this paper, the price evaluation model based on big data analysis is proposed, which takes advantage of widely circulated vehicle data and a large number of vehicle transaction data to analyse the price data for each type of vehicles by using the optimized BP neural network algorithm. It aims to establish a second-hand car price evaluation model to get the price that best matches the car.

**Gathered Information**

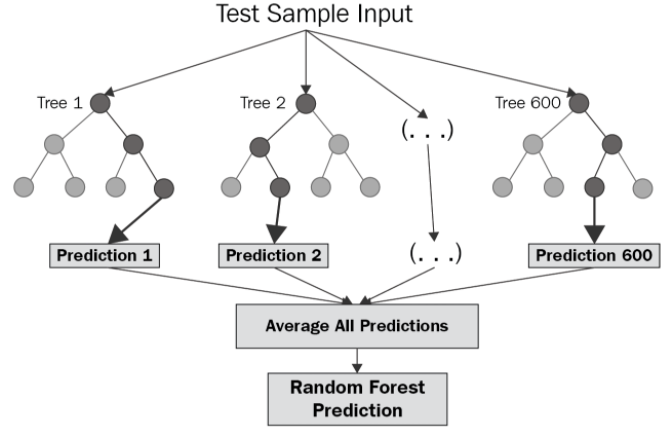
For accurate prediction and better model training, huge dataset of resale cars of Swift Dezire of 5 cities is gathered via web scraping cars24 website. This dataset contains data of 5 main features i.e., fuel type, kms driven, city, car purchase year and resale value. Here resale value becomes our target column whereas other columns served as features for our model. Data scraped consists of many unwanted characters like comma, whitespaces etc. which has to be removed as model can only understand numbers. Moreover, fuel type was converted into numerical codes via one-hot encoding. A one hot encoding is a representation of categorical variables as binary vectors. This requires that the categorical values be mapped to integer values. After data pre-processing, all 5 files, each representing each city has to be merged for model training. Various different machine learning algorithms were implemented on the dataset along with hyperparameter tuning using GRID SEARCH CV Reason behind GBR's good performance is because of its mathematical working.

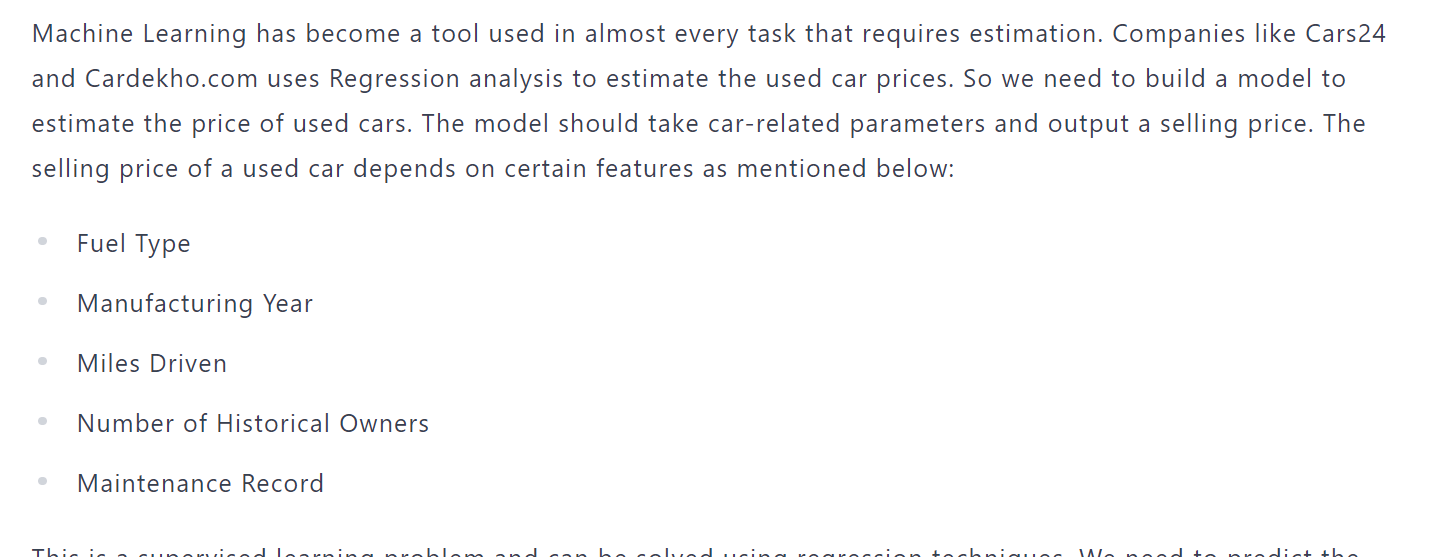
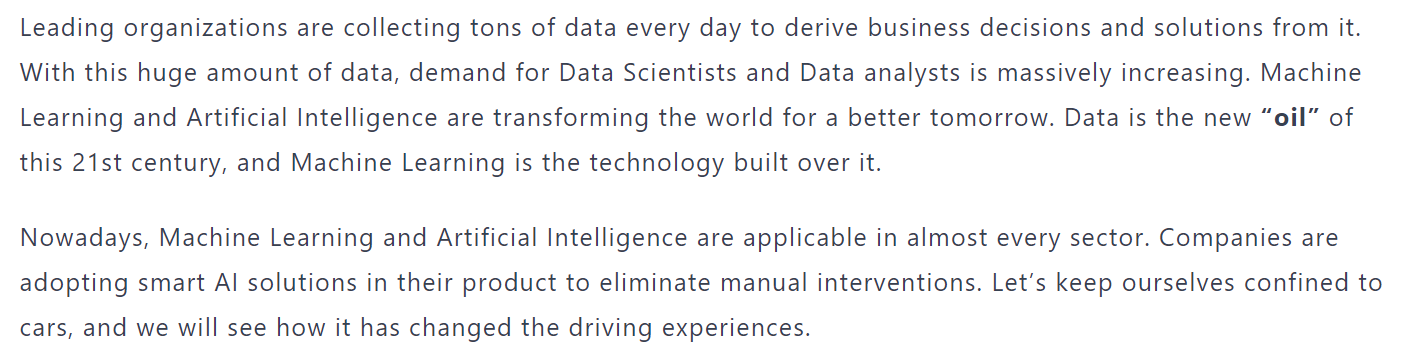
|  |  |
| --- | --- |
| Algorithms implemented | |
| Model Algorithm | RMSV |
| Support Vector Regression | 56000 |
| Logistic Regression | 86000 |
| Random Forest Regression | 78000 |
| Gradient Boosting Regression | 42000 |

**Empathy Map**



In order to predict the resale value of the car, we proposed an intelligent, flexible, and effective system that is based on using regression algorithms. Considering the main factors which would affect the resale value of a vehicle a regression model is to be built that would give the nearest resale value of the vehicle. We will be using various regression algorithms and algorithm with the best accuracy will be taken as a solution, then it will be integrated to the web-based application where the user is notified with the status of his product.



**Ideation******

