Car Price Prediction using Machine Learning Algorithm

Prof.Bhapkar A.D[1], Jagtap Shivani[2], Deshmukh Prachi[3]

Assistant Professor, Computer Department, SPCOET SomeshwarnagarCollege, Baramati, India[1]

Student, Computer Department, SPCOET Someshwarnagar College, Baramati, India[2]

Student, Computer Department, SPCOET Someshwarnagar College, Baramati, India[3]

Abstract: A car price prediction has been a high-interest research area, as itrequires noticeable effort and knowledge of the field expert. Considerable numbers 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, we applied three machine learningtechniques (Artificial Neural Network, Support Vector Machine and Random Forest). However, the mentioned techniques were applied to work as an ensemble. The data used for the prediction was collected from the web portal autopijaca.ba using web scraper that was written in PHP programming language. Respective performances of different algorithms were then compared to find one that best suits the available data set. The final prediction model was integrated into Java application.

Keywords: Machine Learning, SVM Algorithm,Linear RegressionAlgorithm , Random forest regression,Decision Tree Regression, Jupiter, pycharm .

I.INTRODUCTION

The used car market is an ever-rising industry, which has almostdoubled its market value in the last few years. The emergence of online portals such as CarDheko, Quikr, Carwale, Cars24, and many others has facilitated the need for both the customer and the seller to be better informed about the trends and patterns that determine the value of the used car in the market. Machine Learning algorithms can be used to predict the retail value of a car, based on a certain set of features. Different websites have different algorithms to generate the retail price of the used cars, and hence there isn't a unified algorithm for determining the price. By training statistical models for predicting the prices, one can easily get a rough estimate of the price without actually entering the details into the desired website.

II.LITERATURE SURVEY

A literature review is a text of a scholarly paper, which includes thecurrent knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic. Literature reviews usesecondary sources, and do not report new or original experimental work.

1.Paper name: Car Price Prediction using SVM Techniques

Author: S.E.Viswapriya, DurbakaSaiSandeep Sharma, GandavarapuSathyakiran

The prediction of price for a car has been more popular in research area, and it needs predominant effort and information about the experts of this particular field. The number of different attributes is measured and also it has been considerable to predict the result in more reliable and accurate.

2.paper name: Predicting the Price of Cars Using Machine Learning and Data Science

Author: G. Kalpana1,Dr. A. Kanaka Durga,T. Anoop Reddy , Dr.G. Karuna

The Company Wishes To Develop An Algorithm To Predict The Price Of Pre-Owned Cars Based On Various Attributes Associated With The Car To Make A Sale Quickly, If The Price Is Reasonable And Satisfies Both The Seller And Buyer, By Comparing The Price Of Various Car

3.paper name: Car sales prediction using machine learning algorithmns

Author: K. Madhuvanthi, Nallakaruppan M-K, Senthilkumar N C, S. Siva Rama Krishnan

Sales prediction is the current numerotrend in which all the business companies thrive and it also aids the organization or concern in determining the future goals for it and its plan and procedure to achieve it. The data about car sales are derived from various sources sales of cars does not contain any independent variable since various factors suchas horse power; model, width, fuel type, height, price, city-mileage, highway-mileage and manufacturer are the various features that influence the sales.

III.PROPOSED SYSTEM

In this project we predict car price on the varying features and factors,and also with the help of experts knowledge the car price prediction has been done accurately. The most necessity ingredient for prediction is brand and model, period usage of car, mileage of car. The fuel type used in the car as well as fuel consumption per mile highly affect price of a car due to a frequent changes in the price of a fuel. Different features like exterior color, door number, type oftransmission, dimensions, safety, air condition, interior, whether it has navigation or not will also influence the car price. In this paper, we applied different methods and techniques in order to achieve higher precision of the used car price prediction..

IV.ALGORITHM

Linear Regression- is a machine learning algorithm based on supervised learning. It performs aregression task. Regression models a target prediction value based on independent variables. It is mostlyused for finding out the relationship between variables and forecasting. Different regression models differbased on – the kind of relationship between dependent and independent variables they are considering, andthe number of independent variables getting used.
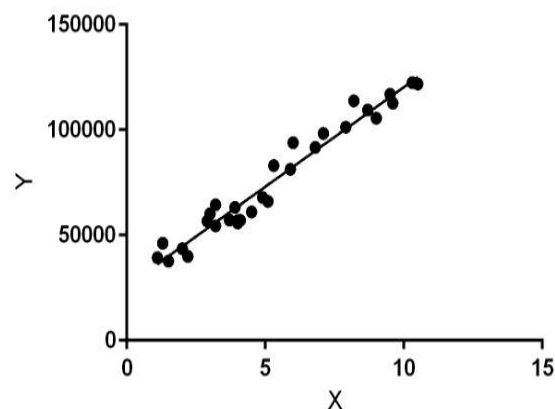

Random Forest Regression- Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap and Aggregation, commonly known as **bagging**. The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees. Random Forest has multiple decision trees as base learning models. We randomly perform row sampling and feature sampling from the dataset forming sample datasets for every model. This part is called Bootstrap.

Decision tree regression- Decision tree regressionobserves features of an object and trains a model in the structure of a tree to predict data in the future to produce meaningful continuous output. Continuous output means that the output/result is not discrete, i.e., it is not represented just by a discrete, known set of numbers or values

Discrete output example: A weather prediction model that predicts whether or not there'll be rain on a particular day.
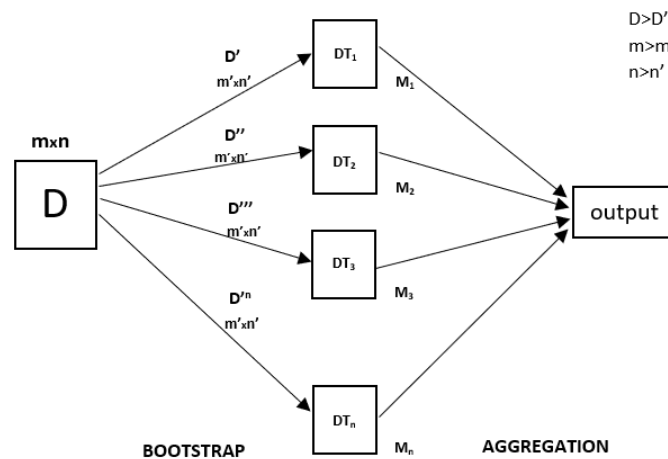
Continuous output example: A profit prediction model that states the probable profit that can be generated from the sale of a product.Here, continuous values are predicted with the help of a decision tree regression model.
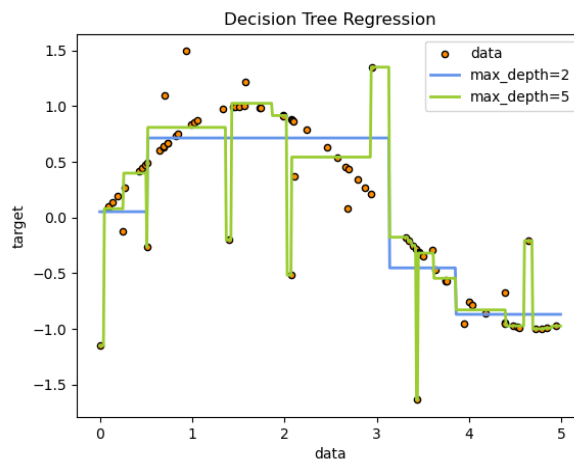
Support Vector Machine-Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.
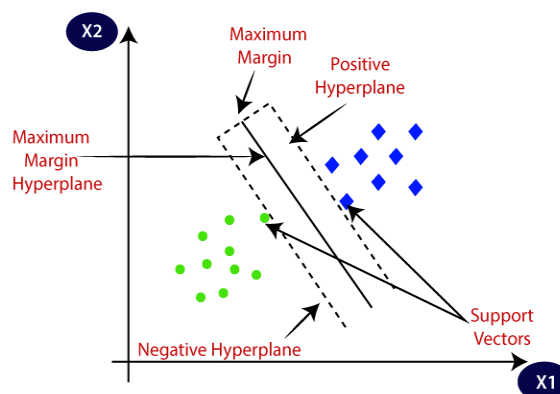
V.CONCLUSION

The increased prices of new cars and the financial incapability of thecustomers to buy them, Used Car sales are on a global increase. Therefore, there is an urgent need for a Used Car Price Prediction system which effectively determines the worthiness of the car using a variety of features. The proposed system will help to determine the accurateprice of used car price prediction. This system compares different algorithms for machine learning : Linear Regression,Decision treeRegression,SVM,Random forest.

REFERENCES

[1] SameerchandPudaruth, "Predicting the Price of Used Cars usingMachine Learning Techniques"; (IJICT 2014).

[2] Enisgegic, BecirIsakovic, Dino Keco, ZerinaMasetic, JasminKevric, "Car Price Prediction Using Machine Learning"; (TEMJournal 2019) .

[3] Ning sun, HongxiBai, YuxiaGeng, Huizhu Shi, "Price Evaluation Model In Second Hand Car System Based On BP Neural Network Theory"; (Hohai University Changzhou, China) .

[4] NitisMonburinon, PrajakChertchom, ThongchaiKaewkiriya, SuwatRungpheung, SabirBuya, PitchayakitBoonpou, "Prediction of Prices for Used Car by using Regression Models" (ICBIR 2018).

[5] Information regarding machine learning techniques and algorithmshttps: