CHAPTER 1

INTRODUCTION

**INTRODUCTION**

Machine Learning is the field of study that gives computers the capability to learn without being explicitly programmed. ML is one of the most exciting technologies that one would have ever come across. As it is evident from the name, it gives the computer that makes it more similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

achine Learning is the field of study that gives computers the capability to learn without

being explicitly programmed. ML is one of the most exciting technologies that one would have

ever come across. As it is evident from the name, it gives the computer that makes it more

similar to humans: The ability to learn. Machine learning is actively being used today, perhaps in

many more places than one would expect.

**1.1 Motivation**

A car price prediction has been a highinterest research area, as it requires noticeable effort and knowledge of the field expert. 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, we applied three machine learning techniques (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. Furthermore, the model was evaluated using test data and the accuracy of 87.38% was obtained.

**1.2 Problem Statement**

To build a model for predicting the price of used cars in, we applied one of the machine learning techniques i.e., Linear Regression. Using linear regression, there are multiple independent variables, but one and only one dependent variable whose actual and predicted values are compared to find precision of results. Our paper proposes a system where price is dependent variable which is predicted, and this price is derived from factors like kilometers driven, car purchase year, Car Company, car model, and the fuel type. Problem Statement - It is easy for any company to price their new cars based on the manufacturing and marketing cost it involves. But when it comes to a used car it is quite difficult to define a price because it involves it is influenced by various parameters like car brand, manufactured year and etc. The goal of our project is to predict the best price for a pre-owned car in the Indian market based on the previous data related to sold cars using Linear Regression.

**1.3 Purpose**

Purpose of Project The used car market is an ever-rising industry, which has almost doubled 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. The purpose of this project is to provide Car price prediction using machine learning without any human interference. In our day to day lives everyone buys and sells a car every day. Now there are limited facilities and applications to get an appropriate price for one’s car. Now we use.

**1.4 Objective and Scope**

Objective Of the Project - The goal of this project is to create an efficient and effective model that will be able to predict the price of a used car by using the Brand or Type of the car one prefers like Ford, Hyundai

•Linear Regression algorithm with better accuracy. Model of the car namely Ford Figo, Hyundai Creta

• Year of manufacturing like 2020, 2021

• Type of fuel namely Petrol, Diesel

• Number of kilometers car has travelled

Unsupervised machine learning analyzes and clusters unlabeled datasets using machine learning algorithms. These algorithms find hidden patterns and data without any human intervention, i.e., we don’t give output to our model. The training model has only input parameter values and discovers the groups or patterns on its own. Data-set in Figure A is Mall data that contains information about its clients that subscribe to them. Once subscribed they are provided a membership card and the mall has complete information about the customer and his/her every purchase. Now using this data and unsupervised learning techniques, the mall can easily group clients based.

CHAPTER 2

LITERATURE SURVEY

[1] Agencija za statistiku BiH. (n.d.), retrieved from: http://www.bhas.ba . [accessed July 18, 2018.]

[2] Listiani, M. (2009). Support vector regression analysis for price prediction in a car leasing application (Doctoral dissertation, Master thesis, TU Hamburg-Harburg).

[3] Richardson, M. S. (2009). Determinants of used car resale value. Retrieved from: https://digitalcc.coloradocollege.edu/islandora/object /coccc%3A1346 [accessed: August 1, 2018.]

[4] Wu, J. D., Hsu, C. C., & Chen, H. C. (2009). An expert system of price forecasting for used cars using adaptive neuro-fuzzy inference. Expert Systems with Applications, 36(4), 7809-7817.

[5] Du, J., Xie, L., & Schroeder, S. (2009). Practice Prize Paper—PIN Optimal Distribution of Auction Vehicles System: Applying Price Forecasting, Elasticity Estimation, and Genetic Algorithms to Used-Vehicle Distribution. Marketing Science, 28(4), 637-644.

[6] Gongqi, S., Yansong, W., & Qiang, Z. (2011, January). New Model for Residual Value Prediction of the Used Car Based on BP Neural Network and Nonlinear Curve Fit. In Measuring Technology and Mechatronics Automation (ICMTMA), 2011 Third International Conference on (Vol. 2, pp. 682-685). IEEE.

[7] Pudaruth, S. (2014). Predicting the price of used cars using machine learning techniques. Int. J. Inf. Comput. Technol, 4(7), 753-764.

[8] Noor, K., & Jan, S. (2017). Vehicle Price Prediction System using Machine Learning Techniques. International Journal of Computer Applications, 167(9), 27-31.

[9] Auto pijaca BiH. (n.d.), Retrieved from: https://www.autopijaca.ba. [accessed August 10, 2018].

[10] Weka 3 - Data Mining with Open Source Machine Learning Software in Java. (n.d.), Retrieved from: https://www.cs.waikato.ac.nz/ml/weka/. [August 04, 2018].

[11] Ho, T. K. (1995, August). Random decision forests. In Document analysis and recognition, 1995., proceedings of the third international conference on (Vol. 1, pp. 278-282). IEEE.

[12] Russell, S. (2015). Artificial Intelligence: A Modern Approach (3rd edition). PE.

[13] Ben-Hur, A., Horn, D., Siegelmann, H. T., & Vapnik, V. (2001). Support vector clustering. Journal of machine learning research, 2(Dec), 125-137. [14] Aizerman, M. A. (1964).

[14] Noor, K., & Jan, S. (2017). Vehicle Price Prediction System using Machine Learning Techniques. International Journal of Computer Applications, 167(9), 27-31.

[15] Auto pijaca BiH. (n.d.), Retrieved from: https://www.autopijaca.ba. [accessed August 10, 2018].

[16] Weka 3 - Data Mining with Open Source Machine Learning Software in Java. (n.d.), Retrieved from: https://www.cs.waikato.ac.nz/ml/weka/. [August 04, 2018].

Table 2.1 Literature Review

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sr**  **No.** | **Author** | **Year** | **Title** | **Paper Outcomes** | **Limitation** |
| 1 | Agencija za statistiku BiH. (n.d.) | December,2020 | AI Based Car Monitoring System | The car Monitoring system using artificial Intelligence concept mainly works using the concept recognisation system. | System has to recognize the future growing price and according to demand of model. |
| 2 | Listiani, M. | November,2009 | Automatic Dataset recognition and prediction of price. | Support vector regression analysis for price prediction in a car leasing application. | Classified all type of car model according to the quality to sell. |
| 3 | Richardson, M. S. | January,2009 | The car Monitoring system using artificial Intelligence concept mainly works using the concept recognisation system. | Determinants of used car resale value. | This process will be repeated for each & every model price and deal accordingly. Maintaining a set of previous data. |
| 4 | Wu, J. D., Hsu, C. C., & Chen, H. C. | June, 2019 | Predict the most labeled and supervised class of car according to price. | Automatic Dataset recognition and prediction of price. | An expert system of price forecasting for used cars using adaptive neuro-fuzzy inference. |
| 5 | J., Xie, L., & Schroeder, S. | Sept, 2018 | According to prediction of previous data it predict. | Automatic modulation classification of vehicles. | Practice Prize Paper—PIN Optimal Distribution of Auction Vehicles System: Applying Price Forecasting, Elasticity Estimation, and Genetic Algorithms to Used-Vehicle Distribution. |
| 6 | Gongqi, S., Yansong, W., & Qiang, Z | January,2011 | The objectiveof the study entited. | New Model for Residual Value Prediction of the Used Car Based on BP Neural Network and Nonlinear Curve Fit. | This process will be repeated for each & every model price and deal accordingly. Maintaining a set of previous data. |
| 7 | Pudaruth, S. | May ,2014 | Over the year the manual car prediction checking. | Automatic modulation classification of vehicles. | Predicting the price of used cars using machine learning techniques. Int. J. Inf. Comput. Technol. |
| 8 | Noor, K., & Jan, S. | August, 2017 | Vehicles classification system. | Vehicle Price Prediction System using Machine Learning Techniques. International Journal of Computer Applications. | The car Monitoring system using artificial Intelligence concept mainly works using the concept recognisation system. |

CHAPTER 3

PROPOSED SYSTEM

**3.1 System Architecture**

Data is collected from a local web portal for selling and buying cars autopijaca.ba [9], during winter season, as time interval itself has high impact on the price of the cars in Bosnia and Herzegovina. The following attributes were captured for each car: brand, model, car condition, fuel, year of manufacturing, power in kilowatts, transmission type, millage, color, city, state, number of doors, four wheel drive (yes/no), damaged (yes/no), navigation (yes/no), leather seats (yes/no), alarm (yes/no), aluminum rims (yes/no), digital air condition (yes/no), parking sensors (yes/no), xenon lights (yes/no), remote unlock (yes/no), electric rear mirrors (yes/no), seat heat (yes/no), panorama roof (yes/no), cruise control (yes/no), abs (yes/no), esp (yes/no), asr (yes/no) and price expressed in BAM (Bosnian Mark). Since manual data collection is time consuming task, especially when there are numerous records to process, a “web scraper” as a part of this research is created to get this job done automatically and reduce the time for data gathering. Web scraping is well known technique to extract information from websites and save data into local file or database. Manual data extraction is time consuming and therefore web scrapers are used to do this job in a fraction of time. Web scrapers are programed for specific websites and can mimic regular users from website’s point of view. After raw data has been collected and stored to local database, data preprocessing step was applied. Many of the attributes were sparse and they do not contain useful information for prediction. Hence, it is decided to remove them from the dataset. The attributes “state”, “city”, and “damaged” were completely removed. Accurate car price prediction involves expert knowledge, because price usually depends on many distinctive features and factors. Typically, most significant ones are brand and model, age, horsepower and mileage. 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 of transmission, 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. This paper is organized in the following manner: Section II contains related work in the field of price prediction of used cars. In section III, the research methodology of our study is explain. Section IV elaborates various machine learning algorithms and examine their respective performances to predict the price of the used cars. Finally, in section V, a conclusion of our work are given, together with the future works plan. Car price prediction is somehow interesting and popular problem. As per information that was gotten from the Agency for Statistics of BiH, 921.456 vehicles were registered in 2014 from which 84%.

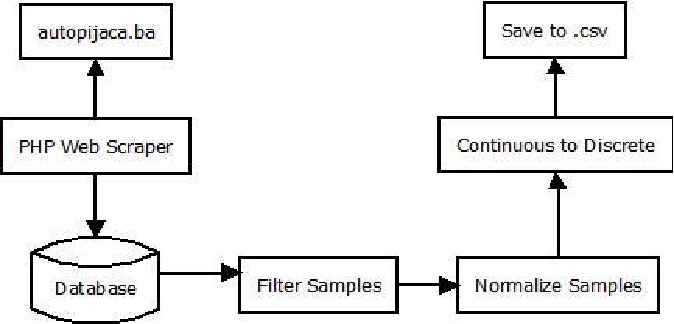


Fig. 3.1 Proposed System

CHAPTER 4

SYSTEM REQUIREMENT

4.1 Hardware Requirement

1. Very less requirement of hardware in a car price prediction because it is just used for predicting the current and future value. Hence this topic of Hardware is eliminated.

**4.2 Software Requirement**

Software which are mostly used in Monitoring the face using many IDE some of them is below given:-

* Jyputer Notebook
* OpenCV libraries
* Pandas packages
* Numpy Packages
* Anaconda IDE
* Pycharm IDE
* Dataset from Kaggle

CHAPTER 5

SYSTEM DESIGN

5.1 Flow Chart

The system flow diagram is a visual representation of all processed in sequential order.  
The System flow chart diagram is a graphical representation of the relation between all the major parts or step of the system. Flow chart diagram can not include minor parts of the system. Below given is flow chart of AI based monitoring admin system:-

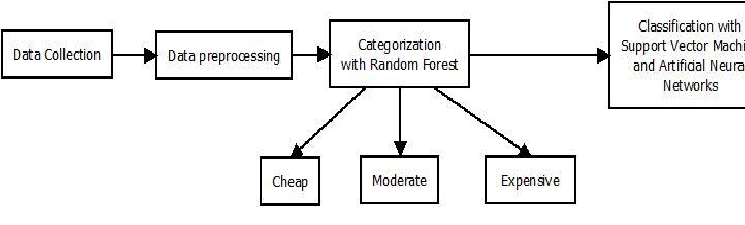


Fig 5.1 Flow Chart of Admin System for Car price prediction monitoring

**5.2 Use Case Diagram**

The use case diagram are usually referred to as behavior diagram used to describe the actions of all user in a system. All user describe in use case are actors and the functionality as action of system.  
The Use case diagram is a collection of diagram and text together that make action on goal of a process.

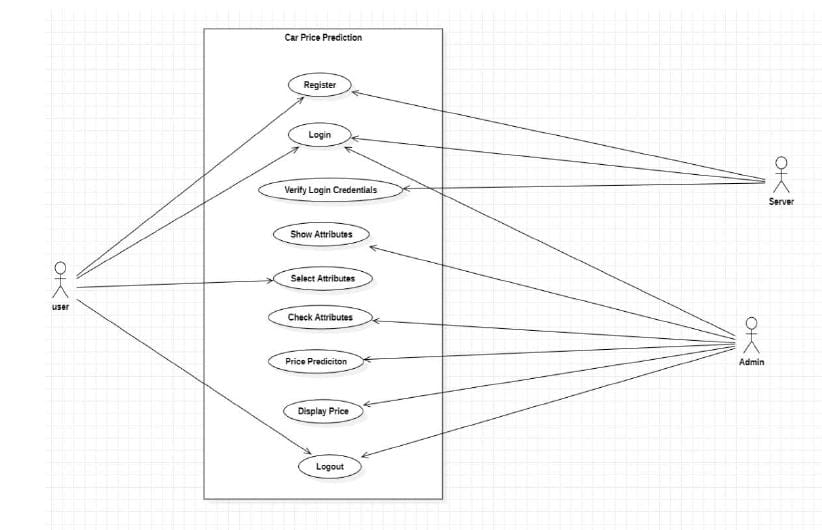


Fig 5.2 Use Case Diagram of AI Based Car Price system

**5.3 Methodology**

This research aims to develop a good regression model to offer accurate prediction of car price. In order to do this, we need some previous data of used cars for which we use price and some other standard attributes. Car price is considered as the dependent variable while other attributes as the independent variables. Let X be the input and Y be the output, the linear regression correlation can be expressed as: (1) In the above equation, β0, β1shows the regression coefficients, Y is the output or required variable and X shows the input. The above equation represents relation in case of single input. In case of multiple inputs, equation will be: (2) In the above equation, X1 , X2,…………Xn represent multiple inputs. For n inputs there are n+1 regression coefficients. Used cars have very high dimension input space. There are various attributes and features that have impact on car price which naturally generates large set of data leading to complexity in analyzing it [7]. The focus of this research is to build such a model which has the capabilities of dealing with high complexity and gives accurate results irrespective of the magnitude of data set. The input data is gathered from pakwheels in a month or two. In the beginning, 2000 records of used cars were recorded. The collected data included variable values for price, engine capacity, color, advertisement date, number of views, mileage in kilometer, power steering, alloy rims, transmission, type of engine, registered city, city, version, model, make and model year. Once the data collection was over, we processed data using multiple linear regression technique for price prediction. In this research, statistical software Minitab was used in which we input the data and analyze the results via linear regression application. Initially, all attributes were considered, but later we applied the variable selection techniques on our input data and found the most significant variables and skipped all other insignificant variables.

CHAPTER 6

PROJECT IMPLEMENTATION

6.1 Outputs

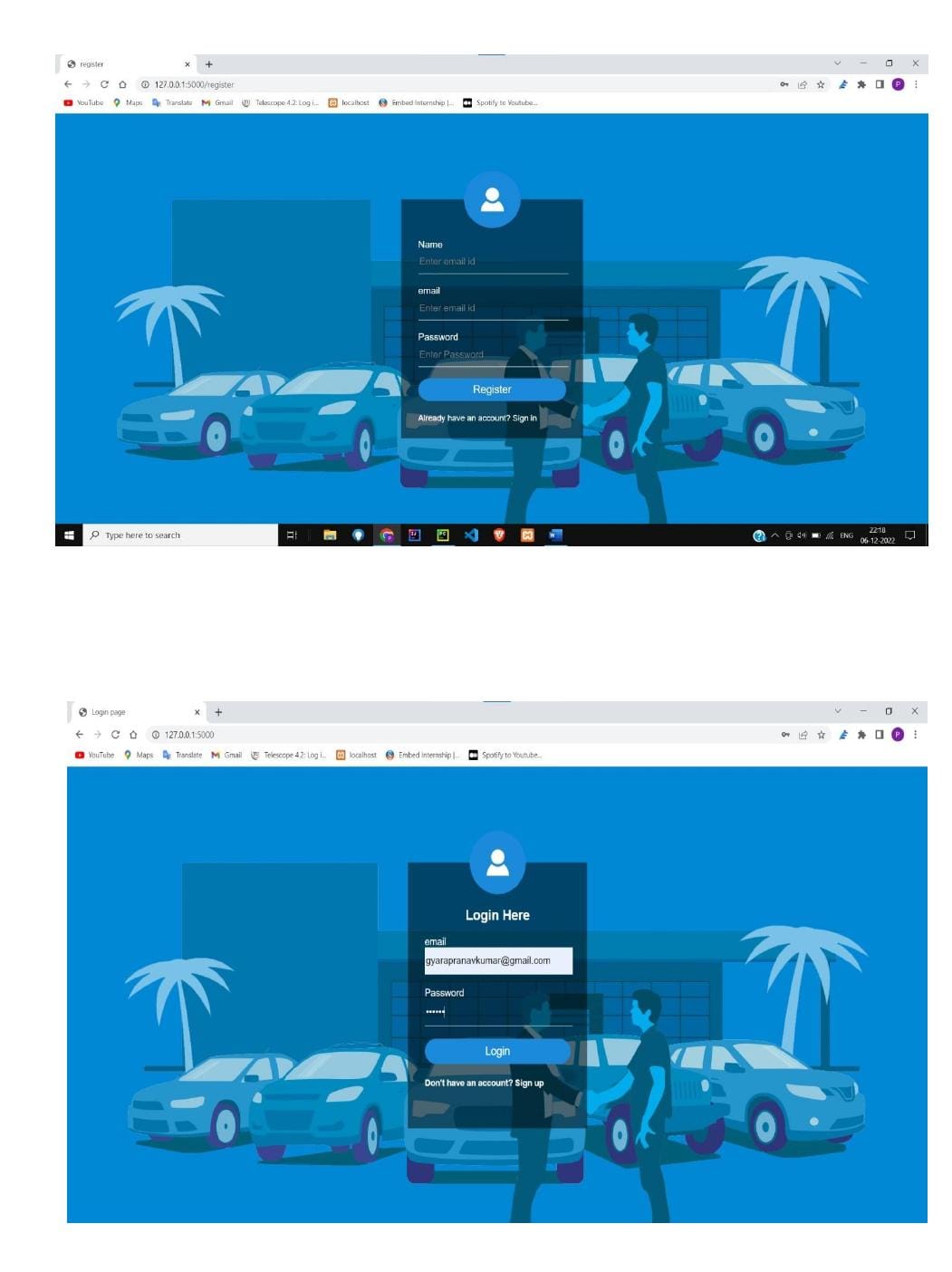


Fig 6.1 Output of Car Price Monitoring Interface

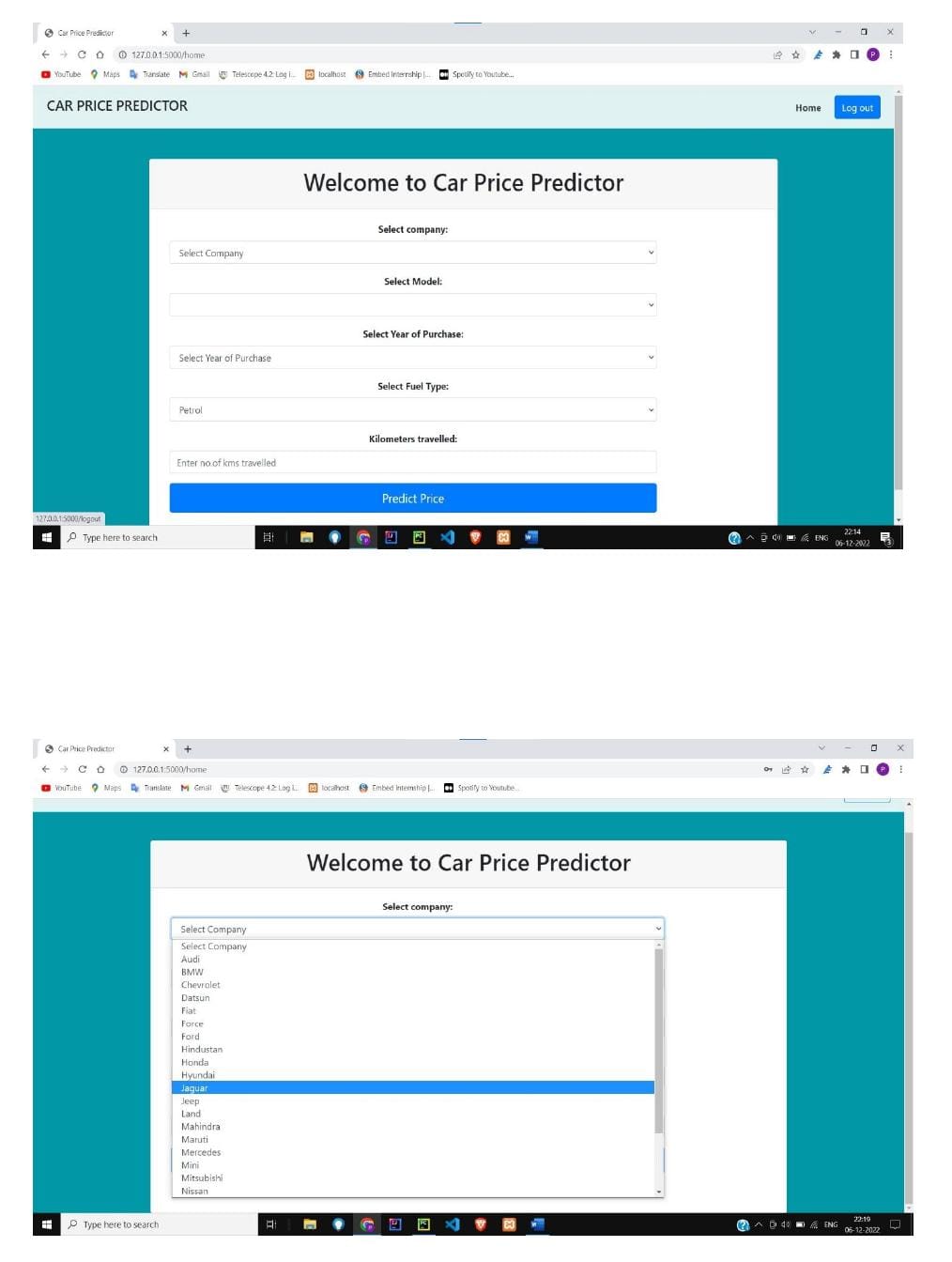


Fig 6.2 Car Price cost Interface

Figure 6.1 Show that car price prediction predict the price the cost of future scope. fter test planning phase is over test analysis phase starts, in this phase we need to dig deeper into project and figure out what testing needs to be carried out in each SDLC phase. Automation activities are also decided in this phase, information needs to be done for software product.

Figure 6.2 Car price cost show the optimum solution of any predict by using regression analysis ca ategy to be followed, what will be the test environment, what test methodologies will be followed, hardware and software availability, resources, risks etc. A high level test plan document is created which includes all the planning inputs mentioned above and circulated to the stakeholders..

CHAPTER 7

CONCLUSION

**Conclusion**

The prediction error rate of all the models was well under the accepted 5% of error. But, on further analysis, the mean error of the regression tree model was found to be more than the mean error rate of the linear regression model. Even though for some seeds the regression tree has better accuracy, its error rates are higher for the rest. This has been confirmed by performing an ANOVA. Also, the post-hoc test revealed that the error rates in multiple regression models and lasso regression models aren’t significantly different from each other. To get even more accurate models, we can also choose more advanced machine learning algorithms such as random forests, an ensemble learning algorithm which creates multiple decision/regression trees, which brings down overfitting massively or Boosting, which tries to bias the overall model by weighing in the favor of good performers. More data from newer websites and different countries can also be scraped and this data can be used to retrain these models to check for reproducibility.

**REFERENCE**

REFERENCE

[1.] Patel UA, Swaminarayan Priya R. Development of a Car and preception management system using Linear Regression: a review. Int J Adv Res Comput Sci Manag Stud. 2014;2(8):109–19.

[2.] Jacksi K. Design and Implementation of Online Submission And Peer Review System: A Case Study Of E-Journal Of University Of Zakho. Int J Sci Technol Res. 2015;4(8):83–5.

[3.] Jacksi K, Badiozamany S. General method for data indexing using clustering methods. Int J Sci Eng. 2015 Mar;6(3):641–4.

[4.] Jacksi K, Dimililer N, and Zeebaree SR. State of the Art Exploration Systems for Linked Data: A Review. Int J Adv Comput Sci Appl IJACSA. 2016;7(11):155–64.

[5.] Gangagowri G, Muthuselvi J, Sujitha S. vehicle and car Management System.

[6.] Anitha V Pai, Krishna A, Kshama PM, Correa M. Web service for Car speed checking management system. www.ijarse.com. 2016 Mar;5(3).

[7.] Arbain N, Nordin NF, Isa NM, Saaidin S. LAS: Web-based laboratory car system by integrating ID-ARDUINO technology. In IEEE; 2014. p. 89–94.

[8.] Srinidhi M, Roy R. A web enabled secured system for car and help assistance system and monitoring and real time location tracking using Biometric and Radio Frequency Identification (RFID) technology. In IEEE; 2015. p. 1–5.

[9.] Arulogun O, Olatunbosun A, Fakolujo O, Olaniyi O. RFID-based Car cost predicting management system. Int J Sci Eng Res. 2013;4(2):1–9.

[10.] Kassim M, Mazlan H, Zaini N, Salleh MK. Webbased Track insurance detection system using RFID technology. In IEEE; 2012. p. 213–8.

[11.] Ahmad BI. TouchIn: an NFC supported Vechicles system in a university environment. Int J Inf Educ Technol. 2014;4(5):448.

[12.] Benyo B, Sodor B, Doktor T, Fördős G. Car Monitoring system and monitoring at the university using NFC. In IEEE; 2012. p. 1–5.

[13.] Srinidhi M, Roy R. A web enabled secured system for car monitoring and real time location tracking using Owner and ages of car used and Radio Frequency Identification (RFID) technology. In IEEE; 2015. p. 1–5.

[14.] Arulogun O, Olatunbosun A, Fakolujo O, Olaniyi O. Car Price and owner management system. Int J Sci Eng Res. 2013;4(2):1–9.

[15.] Kassim M, Mazlan H, Zaini N, Salleh MK. Car detection using dataset kaggles 1-5.

PUBLICATION DETAILS