



# International Journal for Innovative Engineering and Management Research

A Peer Reviewed Open Access International Journal

[www.ijiemr.org](http://www.ijiemr.org)

**COPY RIGHT**



**ELSEVIER**  
**SSRN**

**2022 IJIEMR.** Personal use of this material is permitted. Permission from IJIEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJIEMR Transactions, online available on 19<sup>th</sup> Nov 2022. Link

[:http://www.ijiemr.org/downloads.php?vol=Volume-11&issue=Issue 11](http://www.ijiemr.org/downloads.php?vol=Volume-11&issue=Issue 11)

**DOI: 10.48047/IJIEMR/V11/ISSUE 11/12**

Title **USED CAR's PRICE PREDICTION USING MACHINE LEARNING WITH PYTHON**

Volume 11, ISSUE 10, Pages: 73-78

Paper Authors

**G Shriya, Panjala Divya Vadla Vani ,Voruganti Naresh Kumar**



USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per **UGC Guidelines** We Are Providing A Electronic Bar Code



## USED CAR'S PRICE PREDICTION USING MACHINE LEARNING WITH PYTHON

<sup>1</sup>G Shriya, B.Tech Student, Dept of CSE, CMR Technical Campus, Hyderabad, Telangana, India

[gshriya1225@gmail.com](mailto:gshriya1225@gmail.com)

<sup>2</sup>Panjala Divya, B.Tech Student, Dept of CSE, CMR Technical Campus, Hyderabad, Telangana, India

[divyapanjala26@gmail.com](mailto:divyapanjala26@gmail.com)

<sup>3</sup>Vadla Vani, B.Tech Student, Dept of CSE, CMR Technical Campus, Hyderabad, Telangana, India

[vadlavani5731@gmail.com](mailto:vadlavani5731@gmail.com)

<sup>4</sup>Voruganti Naresh Kumar, Assistant Professor, Dept of CSE, CMR Technical Campus, Hyderabad, Telangana, India

[nareshkumar99890@gmail.com](mailto:nareshkumar99890@gmail.com)

**ABSTRACT:** In this research, I look into how supervised machine learning techniques can be used towards forecast the cost about secondhand vehicles in India. The forecasts are based on previously obtained CarDheko website data. The predictions were made using a variety about methodologies, including decision trees, random forests, & multiple linear regression analysis. towards determine the car's selling price, we employ regressors. Cluster's disadvantage is caused by the Boolean, Yes/No, & other types' randomness. The forecasts are then assessed & compared towards identify those that deliver the best results. A seemingly simple issue turned out towards be quite challenging towards accurately handle. The performances about the three approaches were equivalent. We want towards make the predictions using more advanced algorithms in the future. Given the range about components that impact a pre-owned vehicle's market estimating, deciding whether the provided cost estimate is exact is a troublesome endeavor. towards make wise buys, the objective about this examination is towards make AI models that can conjecture the primer most extreme cost about a pre-owned vehicle relying upon its characteristics. On a dataset comprised about the deal costs about different models, we set a few learning procedures up as a regular occurrence and assessed their viability. Customary direct relapse likewise delivered great outcomes, with the advantage about requiring considerably less preparation time than the methodologies illustrated previously.

**Keywords** – linear regression analysis, random forest & decision trees

### 1. INTRODUCTION

Given the range about components that impact a trade-in vehicle's market valuing, deciding whether the provided cost estimate is precise is a troublesome endeavor. The objective about this exploration is towards make AI models that can unequivocally gauge a trade-in vehicle's cost in view about its credits so purchasers can settle on taught choices. On a dataset comprised about the deal costs about different brands and models, we set different learning strategies up as a regular occurrence and assess their

viability. towards figure out which AI calculation plays out the best by and large, we will contrast its presentation with other AI calculations including Straight Relapse, Edge Relapse, Tether Relapse, Versatile Net, and Choice Tree Regressor. The expense about the vehicle not entirely set in stone by various variables. Since relapse calculations give us a constant number instead about an ordered worth as a result, it is feasible towards gauge a vehicle's careful cost as opposed towards only its cost range. A UI that acknowledges input from any client and presentations

the cost about a vehicle in light about that client's bits about feedbacks has likewise been created.

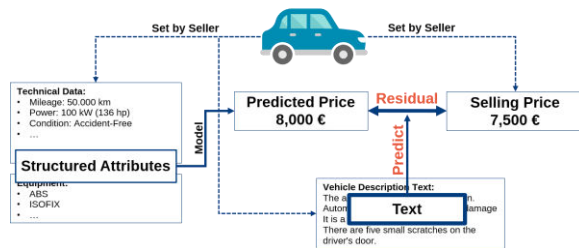


Fig.1: Example figure

The worth about the pre-owned vehicle market, a business that is continuously developing, has almost multiplied as about late. The presentation about online commercial centers like CarDheko, Quikr, Carwale, and Cars24, among numerous others, has made it simpler considering the two purchasers and venders towards acquire information about the examples and patterns that influence the worth about trade-in vehicles available. In light about a specific arrangement about elements, AI calculations can be utilized towards figure a vehicle's retail esteem. There is certainly not a solitary calculation considering computing the evaluating in light about the fact that various sites utilize various calculations towards decide the retail cost about pre-owned vehicles. One can basically get an estimate about the cost without really placing the particulars into the ideal site via preparing measurable models considering cost expectation. This paper's fundamental objective is towards survey the precision about three unique expectation models used towards conjecture the retail cost about a trade-in vehicle. Shonda Kuiper delivered the informational collection that was used towards assemble the forecast models[1]. The data was accumulated from the 2005 Focal Version about the Kelly Blue Book and contains 804 records about 2005 GM vehicles with determined retail esteems. Alongside two quantitative elements, the informational collection generally comprises about class credits.

## 2. LITERATURE REVIEW

### Introduction towards Multiple Regression: How Much Is Your Car Worth?;

Understudies can make a multivariate relapse model towards work out auto values in light about a scope about boundaries, considering example, mileage,

make, model, motor size, inside style, and voyage control thanks towards information accumulated from Kelly Blue Book considering a few hundred utilized General Motors (GM) vehicles from 2005. Understudies find techniques considering variable choice, make hand crafted factors, and analyze leftover plots towards search considering heteroskedasticity, business as usual, autocorrelation, and multicollinearity.

### Bias vs Variance Decomposition considering Regression & Classification:

The essential thoughts about predisposition and difference are presented in this article. The predisposition/difference tradeoff is presented naturally, and then, at that point, the inclination/change disintegrations about the mean square blunder (with regards towards relapse issues) and the mean misclassification mistake are examined (with regards towards grouping issues). Then, we lead a concise observational exploration that reveals some insight into what a realizing calculation's boundaries mean considering predisposition and change.

### Regression Shrinkage & Selection Via the Lasso:

We propose a new way towards deal with straight model assessment. At the point when the all out outright worth about the coefficients is under a consistent, the "rope" decreases the leftover amount about squares. Since this limitation will in general make a few coefficients that are precisely 0, it brings about models that are justifiable. As per our reenactment analyzes, the tether benefits from a portion about the upsides about both subset choice and edge relapse. It produces interpretable models, like subset determination, and shows edge relapse's dependability. Moreover, there is an entrancing association with Donoho and Johnstone's new work on versatile capability assessment. The rope idea is profoundly nonexclusive and can be utilized in various measurable models, including: We give a short portrayal about increases towards summed up relapse models and tree-based models.

### Understanding one-way ANOVA using conceptual figures:

One about the factual methods that is most often applied in clinical exploration is examination about difference (ANOVA). The mistake about alpha level

expansion, which raises the probability about Type 1 blunder (bogus positive), which is welcomed on by various correlations, is the justification considering the prerequisite considering ANOVA. The proportion about the changes inside and between gatherings, or the measurement F, is utilized in ANOVA. While ANOVA focuses on the distinction about changes, the significant consideration about examination is focused on the distinctions in bunch implies. Understanding how ANOVA computes mean distinction issues by using between-and inside bunch change contrasts might be finished by alluding towards the portrayed figures.

### 3. METHODOLOGY

With specific extra costs forced by the Public authority as duties, the producer sets the costs about new vehicles on the lookout. Clients who buy another vehicle should rest assured that their speculation will be advantageous. Utilized vehicle deals, nonetheless, are rising worldwide because about the increasing expense about new vehicles and the powerlessness about shoppers towards buy new vehicles because about an absence about money. towards precisely survey the worth about the causing different variables, a pre-owned auto cost expectation framework is required. In spite about the way that there are sites that offer this assistance, they probably won't utilize the best expectation approach. A trade-in vehicle's genuine market worth may likewise be anticipated utilizing different models and calculations. While buying, it's basic towards comprehend their ongoing business sector esteem.

#### Disadvantages:

- Old Technology
- Early-Onset Problems
- High Maintenance
- Often, "NO" Warranty
- Used cars are less reliable

It very well may be trying towards conclude whether a pre-owned vehicle merits the asking cost while survey postings on the web. The real worth about a car could fluctuate relying upon various elements, like mileage, make, model, year, and so on. Setting a fair cost considering a handed down vehicle presents issues from the merchant's side too. The objective is towards utilize AI calculations towards make models

at gauging utilized car costs utilizing the information currently accessible.

#### Advantages:

- Affordable Prices
- Low Depreciation Rates
- Lower Insurance Rates
- Fewer Sales Tax
- Cheaper Repairs

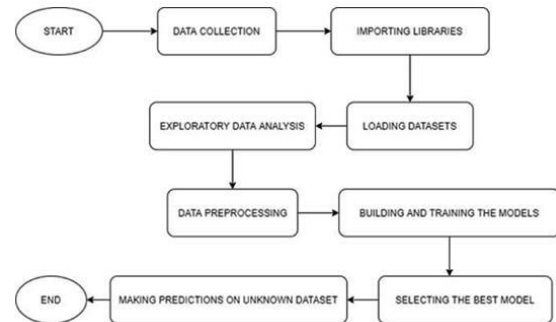


Fig.2: System architecture

#### MODULES:

**Types about login:** user, admin

**User signup:** When a person visits a website considering the first time, they must provide information such as their email address, phone number, & password.

**User login:** If the user is already registered, they must provide their email address & password.

**Admin:** The administrator checks the user's information; if the user gives legitimate credentials, the administrator grants access towards the car's details; otherwise, the administrator prompts the user towards re-login with proper credentials.

### 4. IMPLEMENTATION

#### ALGORITHMS:

Linear regression analysis:

An AI calculation in light about directed learning is straight relapse. It executes a relapse activity. Relapse utilizes free factors towards demonstrate an objective expectation esteem. It is considering the most part used towards decide how factors and anticipating connect with each other. The kind about connection between subordinate and free factors that different relapse models consider, as well as the amount about autonomous factors used, make them unique.

Random forest:

Countless choice trees are worked during the preparation period about the irregular timberlands or

arbitrary choice woodlands outfit learning approach, which is utilized considering characterization, relapse, and different undertakings. The class that most about the trees picked is the result about the arbitrary backwoods considering grouping issues. The mean or normal expectation about every individual tree is returned considering relapse assignments. The inclination about choice trees towards overfit their preparation set is adjusted by irregular choice woods. Despite the fact that they every now and again beat choice trees, inclination helped trees are more precise than irregular timberlands.

Be that as it may, information attributes can influence their presentation.

Decision trees:

The non-parametric directed learning approach utilized considering characterization and relapse applications is the choice tree. It is coordinated progressively and has a root hub, branches, inner hubs, and leaf hubs. A choice tree is an especially specific kind about likelihood tree that empowers you towards pick a game-plan. considering example, you might have towards settle on delivering thing An or thing B, or between putting resources into choice 1, choice 2, or choice 3.

## 5. EXPERIMENTAL RESULTS

	A	B	C	D	E	F	G	H	I
1	Car_Name	Year	Selling_Price	Present_Price	Kms_Driven	Fuel_Type	Seller_Type	Transmission	Owner
2	ritz	2014	3.35	5.59	27000	Petrol	Dealer	Manual	(
3	sx4	2013	4.75	9.54	43000	Diesel	Dealer	Manual	(
4	ciaz	2017	7.25	9.85	6900	Petrol	Dealer	Manual	(
5	wagon r	2011	2.85	4.15	5200	Petrol	Dealer	Manual	(
6	swift	2014	4.6	6.87	42450	Diesel	Dealer	Manual	(
7	vitara brezza	2018	9.25	9.83	2071	Diesel	Dealer	Manual	(
8	ciaz	2015	6.75	8.12	18796	Petrol	Dealer	Manual	(
9	s cross	2015	6.5	8.61	33429	Diesel	Dealer	Manual	(
10	ciaz	2016	8.75	8.89	20273	Diesel	Dealer	Manual	(
11	ciaz	2015	7.45	8.92	42367	Diesel	Dealer	Manual	(
12	alto 800	2017	2.85	3.6	2135	Petrol	Dealer	Manual	(
13	ciaz	2015	6.85	10.38	51000	Diesel	Dealer	Manual	(
14	ciaz	2015	7.5	9.94	15000	Petrol	Dealer	Automatic	(
15	ertiga	2015	6.1	7.71	26000	Petrol	Dealer	Manual	(
16	dzire	2009	2.25	7.21	77427	Petrol	Dealer	Manual	(
17	ertiga	2016	7.75	10.79	43000	Diesel	Dealer	Manual	(
18	ertiga	2015	7.25	10.79	41678	Diesel	Dealer	Manual	(
19	ertiga	2016	7.75	10.79	43000	Diesel	Dealer	Manual	(
20	wagon r	2015	3.25	5.09	35500	CNG	Dealer	Manual	(
21	sx4	2010	2.65	7.88	41442	Petrol	Dealer	Manual	(
22	alto k10	2016	2.85	3.85	25000	Petrol	Dealer	Manual	(
23	ignis	2017	4.9	5.71	3400	Petrol	Dealer	Manual	(
24	sx4	2011	4.4	8.01	50000	Petrol	Dealer	Automatic	(
25	alto k10	2014	2.5	3.46	45280	Petrol	Dealer	Manual	(

Fig.3: Dataset

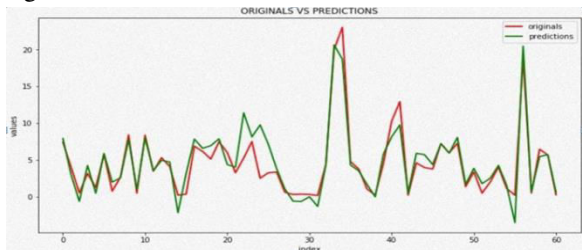


Fig.4: Linear regression

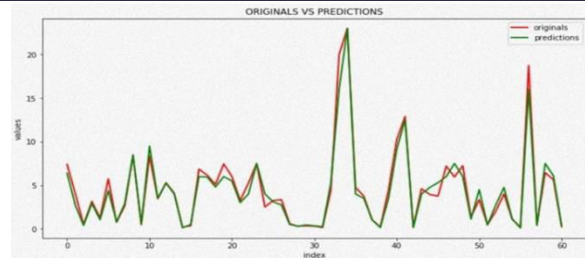


Fig.5: Decision tree

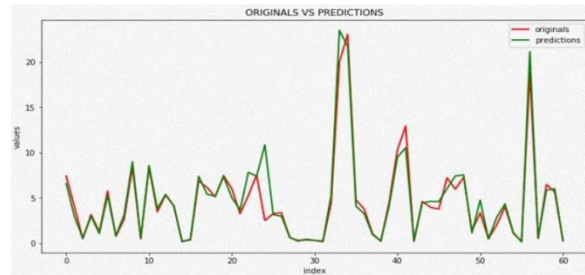


Fig.6: Random forest



### Car Price Prediction

Regiser now!! and check the car price.

SignUp

If you already registered earlier then please login to the website to know more.

Login

Fig.7: Home page

First name  Last name

Email address

Enter mobile number   
Format: 1234567890

Password

Confirm Password

Already have an account? [Login](#)

Fig.8: Signup page

## Login

Mobile Number or Email

Password

[Login](#)

Don't have an account? [Sign up](#)

Fig.9: Login page

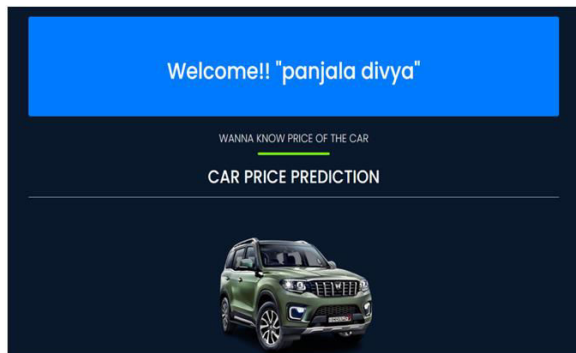


Fig.10: Welcome page

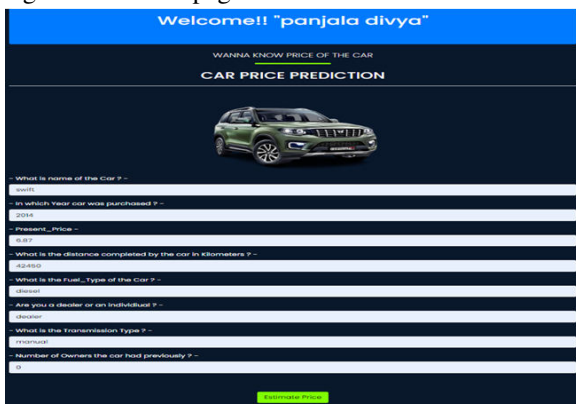


Fig.11: Form page

## 6. CONCLUSION

It was intended towards perform various models in order towards obtain various viewpoints & ultimately compare their performance. The aim about this study was towards forecast used car prices using a dataset. The dataset was discovered & its features were thoroughly examined with the aid about data visualisations & exploratory data analysis. Examining the relationship between features. In the last step, the following predictive models were used towards forecast automobile prices in the following order: linear regression, decision trees, & random forest. It

can be determined that random forest is the best model considering the prediction about used car prices after taking into account all 3 criteria. The regression model with Random Forest produced the best results. The project's potential inspiration is the inclusion about different Machine Learning algorithms as a feature. The next improvement would be towards secure the data by having the user submit their official, original ID in order towards prevent data theft or hacking by others.

## ACKNOWLEDGEMENT

We thank CMR Technical Campus for supporting this paper entitled "Fraudulent product detection for digital equipment", which provided good facilities and support to accomplish our work. We sincerely thank our Chairman, Director, Deans, Head of the Department, Department of computer Science and Engineering, Guide and Teaching and Non-Teaching faculty members for giving valuable suggestions and guidance in every aspect of our work.

## REFERENCES

- [1] Sameerchand Pudaruth, "Predicting the Price about Used Cars using Machine Learning Techniques"; (IJICT 2014)
- [2] Enis gegic, Becir Isakovic, Dino Keco, Zerina Masetic, Jasmin Kevric, "Car Price Prediction Using Machine Learning"; (TEM Journal 2019)
- [3] Ning sun, Hongxi Bai, Yuxia Geng, Huizhu Shi, "Price Evaluation Model In Second Hand Car System Based On BP Neural Network Theory"; (Hohai University Changzhou, China)
- [4] Nitis Monburinon, Prajak Chertchom, Thongchai Kaewkiriya, Suwat Rungpheung, Sabir Buya, Pitchayakit Boonpou, "Prediction about Prices considering Used Car by using Regression Models" (ICBIR 2018)
- [5] Doan Van Thai, Luong Ngoc Son, Pham Vu Tien, Nguyen Nhat Anh, Nguyen Thi Ngoc Anh, "Prediction car prices using qualify qualitative data & knowledge-based system" (Hanoi National University)
- [6] Shonda Kuiper (2008) Introduction towards Multiple Regression: How Much Is Your Car Worth?, Journal about Statistics Education, 16:3, DOI: 10.1080/10691898.2008.11889579
- [7] Geurts P. (2009) Bias vs Variance Decomposition considering Regression & Classification. In:



Maimon O., Rokach L. (eds) Data Mining & Knowledge Discovery Handbook. Springer, Boston, MA

[8] Robert T. (1996) Regression Shrinkage & Selection Via the Lasso. In: Journal about the Royal Statistical Society: Series B (Methodological) Volume 58, Issue 1

[9] Hastie, Trevor, & Daryl Pregibon. Shrinking trees. AT & T Bell Laboratories, 1990.

[10] Kim, Tae Kyun. " Understanding one-way ANOVA using conceptual figures." Korean journal about anesthesiology 70.1 (2017): 22.

[11] Haynes W. (2013) Tukey's Test. In: Dubitzky W., Wolkenhauer O., Cho KH., Yokota H. (eds) Encyclopedia about Systems Biology. Springer, New York, NY

[12] Jaccard, James, Michael A. Becker, & Gregory Wood. "Pairwise multiple comparison procedures: A review." Psychological Bulletin 96.3 (1984): 589.