iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 Issue: V Month of publication: May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5290

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

The Price Prediction for used Cars using Multiple Linear Regression Model

Chadraprakash Trivendra[1], Shashank Girepunje[2], Rahul Chawda[3]
[1, 2, 3]Computer Science Department, Kalinga University

Abstract: ***In this paper, we research the use of regulated AI method to foresee the cost of trade-in vehicles or vehicles. The expectations depend on authentic information gathered from a legitimate source. Distinctive method numerous straight relapse examination have been utilized to make the Prediction. The expectation depends on certain highlights that aren't mulled over at the hour of estimation. In this work, the estimation depends on a blend of highlights that will help to decide the cost for utilized vehicles. The investigation has demonstrated great proficiency over datasets to predict the model.***
Keywords**:** ***AI, Prediction model, Prediction.***

I. INTRODUCTION

Predicting the cost of trade-in vehicles is both a significant and fascinating issue in vehicle selling market. As per information acquired from the National Transport Authority [3], the quantity of vehicles enlisted somewhere in the range of 2003 and 2013 has seen a staggering increment of 234%. From 68, 524 vehicles enrolled in 2003, this number has now arrived at 160, 701. With troublesome financial conditions, almost certainly, deals of second-hand imported (reconditioned) vehicles and trade-in vehicles will increment. It is accounted for in [2] that the deals of new vehicles have enrolled an abatement of 8% in 2013.

In many created nations, it is entirely expected to rent a vehicle as opposed to getting it altogether. A rent is a coupling contract between a purchaser and a dealer (or an outsider – typically a bank, protection firm or other budgetary foundations) in which the purchaser must compensation fixed portions for a pre-characterized number of months/years to the merchant/financer. After the rent time frame is finished, the purchaser has the likelihood to purchase the vehicle at its lingering esteem, for example its normal resale esteem. In this manner, it is of business enthusiasm to vender/financers to have the option to anticipate the rescue esteem (remaining estimation) of vehicles with precision. On the off chance that the remaining worth is under-evaluated by the merchant/financer toward the start, the portions will be higher for the customers who will absolutely then decide on another dealer/financer. In the event that the remaining worth is over-assessed, the portions will be lower for the customers however then the merchant/financer may have a lot of trouble at selling these expensive trade-in vehicles at this over-evaluated lingering esteem. Subsequently, we can see that assessing the cost of trade-in vehicles is of high business significance too.

Anticipating the resale estimation of a vehicle is definitely not a straightforward undertaking. It is trite information that the estimation of trade-in vehicles relies upon various variables. The most significant ones are generally the age of the vehicle, its make (and model), the beginning of the vehicle (the first nation of the producer), its mileage (the quantity of kilometers it has run) and its pull. Because of rising fuel costs, efficiency is likewise of prime significance. Sadly, practically speaking, a great many people don't know precisely how much fuel their vehicle expends for every km driven. Different factors, for example, the kind of fuel it utilizes, the inside style, the slowing mechanism, quickening, the volume of its chambers (estimated in cc), wellbeing file, its size, number of entryways, paint shading, weight of the vehicle, shopper surveys, renowned honors won by the vehicle producer, its physical state, regardless of whether it is a games vehicle, whether it has journey control, whether it is programmed or manual transmission, whether it had a place with an individual or an organization and different alternatives, for example, forced air system, sound framework, power guiding, astronomical wheels, GPS pilot all may impact the cost too. Some uncommon elements which purchasers append significance are the neighborhood of past proprietors, regardless of whether the vehicle had been associated with genuine mishaps and whether it is a woman driven vehicle. The look and feel of the vehicle absolutely contributes a great deal to the cost. As should be obvious, the cost relies upon an enormous number of variables. Sadly, data pretty much every one of these variables are not constantly accessible and the purchaser must settle on the choice to buy at a specific cost dependent on hardly any components as it were.

In this work, we have considered only a small subset of the factors mentioned above. More details are provided in Section III. This paper is organized as follows. In the next section, a review of related work is provided. Section III describes the methodology while in section IV, we describe the result of machine learning technique to predict the price of used cars. Finally, we end the paper with a conclusion with some pointers towards future work.

©IJRASET: All Rights are Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

II. RELATED WORK

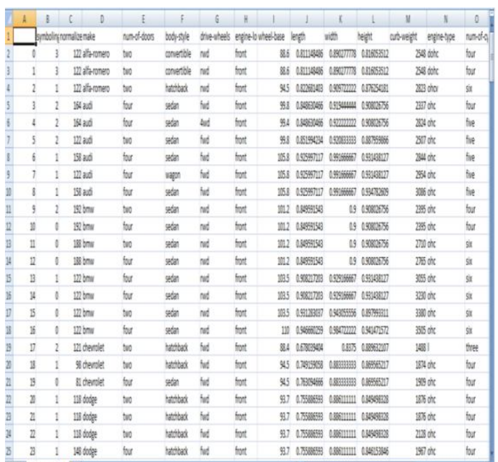
The author Sameerchand Pudaruth [1] investigate the research the utilization of regulated AI methods to foresee the cost of trade-in vehicles in Mauritius. The expectations depend on recorded information gathered from day by day papers. Various systems like numerous direct relapse investigation, k-closest neighbors, gullible byes and choice trees have been utilized to make the expectations. The expectations are then assessed and contrasted all together with discover those which give the best exhibitions. An apparently simple issue ended up being in fact hard to determine with high exactness. All the four techniques gave similar execution. Sun et al. [2] proposed the utilization of online trade-in vehicle value assessment model utilizing the improved BP neural system calculation. They presented another improvement technique called Like Block-Monte Carlo Method (LB-MCM) to advance shrouded neurons. The outcome indicated that the advanced model yielded higher exactness when it contrasted with the non-upgraded model Based on the past related works, we understood that none of them had executed inclination boosting strategy in the expectation of trade-in vehicle cost at this point. In this way, we chose to construct a trade-in vehicle value assessment model utilizing slope helped relapse trees.

Peerun et al. [5] did an examination to assess the exhibition of the neural system in utilized vehicle value expectation. The anticipated worth, in any case, are not extremely near the real cost, particularly on vehicles with a more significant expense. They inferred that help vector machine relapse marginally beat neural system and straight relapse in anticipating utilized vehicle cost.

Gonggi [6] proposed another model dependent on counterfeit neural systems to estimate the leftover estimation of private trade-in vehicles. The primary highlights utilized in this investigation were: mileage, maker and gauge valuable life. The model was improved to deal with nonlinear connections which is impossible with straightforward direct relapse techniques. It was discovered that this model was sensibly precise in anticipating the lingering estimation of trade-in vehicles.

III. METHODOLOGY

Data understanding and Data preparation The used car data used in this research were collected from www.kaggle.com under the public domain license. This dataset consists of 6000 car observations and the attributes of used car, a German e-commerce site as shown in Table I.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		symboling	normalize	make	num-of-doors	body-style	drive-wheels	engine-lo	wheel-base	length	width	height	curb-weight	engine-type	num-of-c
2	0	3	122	alfa-romero	two	convertible	rwd	front	88.6	0.811148486	0.890277778	0.816053512	2548	dohc	four
3	1	3	122	alfa-romero	two	convertible	rwd	front	88.6	0.811148486	0.890277778	0.816053512	2548	dohc	four
4	2	1	122	alfa-romero	two	hatchback	rwd	front	94.5	0.822681403	0.909722222	0.876254181	2823	ohcv	six
5	3	2	164	audi	four	sedan	fwd	front	99.8	0.848630466	0.919444444	0.908026756	2337	ohc	four
6	4	2	164	audi	four	sedan	4wd	front	99.4	0.848630466	0.922222222	0.908026756	2824	ohc	five
7	5	2	122	audi	two	sedan	fwd	front	99.8	0.851994234	0.920833333	0.887959866	2507	ohc	five
8	6	1	158	audi	four	sedan	fwd	front	105.8	0.925997117	0.991666667	0.931438127	2844	ohc	five
9	7	1	122	audi	four	wagon	fwd	front	105.8	0.925997117	0.991666667	0.931438127	2954	ohc	five
10	8	1	158	audi	four	sedan	fwd	front	105.8	0.925997117	0.991666667	0.934782609	3086	ohc	five
11	9	2	192	bmw	two	sedan	rwd	front	101.2	0.849591543	0.9	0.908026756	2395	ohc	four
12	10	0	192	bmw	four	sedan	rwd	front	101.2	0.849591543	0.9	0.908026756	2395	ohc	four
13	11	0	188	bmw	two	sedan	rwd	front	101.2	0.849591543	0.9	0.908026756	2710	ohc	six
14	12	0	188	bmw	four	sedan	rwd	front	101.2	0.849591543	0.9	0.908026756	2765	ohc	six
15	13	1	122	bmw	four	sedan	rwd	front	103.5	0.908217203	0.929166667	0.931438127	3055	ohc	six
16	14	0	122	bmw	four	sedan	rwd	front	103.5	0.908217203	0.929166667	0.931438127	3230	ohc	six
17	15	0	122	bmw	two	sedan	rwd	front	103.5	0.931283037	0.948055556	0.897993311	3380	ohc	six
18	16	0	122	bmw	four	sedan	rwd	front	110	0.946660259	0.984722222	0.941471572	3505	ohc	six
19	17	2	121	chevrolet	two	hatchback	fwd	front	88.4	0.678039404	0.8375	0.889632107	1488	l	three
20	18	1	98	chevrolet	two	hatchback	fwd	front	94.5	0.749159058	0.883333333	0.869565217	1874	ohc	four
21	19	0	81	chevrolet	four	sedan	fwd	front	94.5	0.763094666	0.883333333	0.869565217	1909	ohc	four
22	20	1	118	dodge	two	hatchback	fwd	front	93.7	0.755886593	0.886111111	0.849498328	1876	ohc	four
23	21	1	118	dodge	two	hatchback	fwd	front	93.7	0.755886593	0.886111111	0.849498328	1876	ohc	four
24	22	1	118	dodge	two	hatchback	fwd	front	93.7	0.755886593	0.886111111	0.849498328	2128	ohc	four
25	23	1	148	dodge	four	hatchback	fwd	front	93.7	0.755886593	0.886111111	0.846153846	1967	ohc	four

Fig. 1: sample dataset for used cars

In prescient measurement and AI, a characteristics with high relationship coefficient frequently, however not generally, have more effect on forecast variable. The connection coefficient, as its name infers, is a factual measure that portrays the connection between factors. The connection coefficient of two traits is consistently territory between 1 (Positive relationship) to í1 (Negative relationship).

Similar examination on value expectation This exploration executes a few AI calculations accessible in Scikit-learn AI library. The model is prepared utilizing same preparing information and assesses utilizing same testing information. The outcome at that point thought about and depicted in the following area. In regulated AI, the relapse based strategy has been demonstrated to be solid in foreseeing constant variable. For essential prescient demonstrating, single direct relapse model as communicated in (2) is sufficient to foresee Y where Y is needy variable and X is the autonomous variable. By finding the Y-capture and slant of relapse line in addition to commotion, the model can gauge the future estimation of Y.

©IJRASET: All Rights Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

In this exploration paper, we directed a similar report on straight relapse trees to discover which model are the best when it will be utilized to take care of relapse issue. For this situation, our relapse issue is a model for utilized vehicle value forecast.

These datasets may contain countless trade-in vehicle data with a few probably; they require some tweaking and designing. For instance, copied perceptions may impact on model execution and they should be expelled in advance. The investigation utilized python programming language for this activity. The header segment of the information is appeared here.

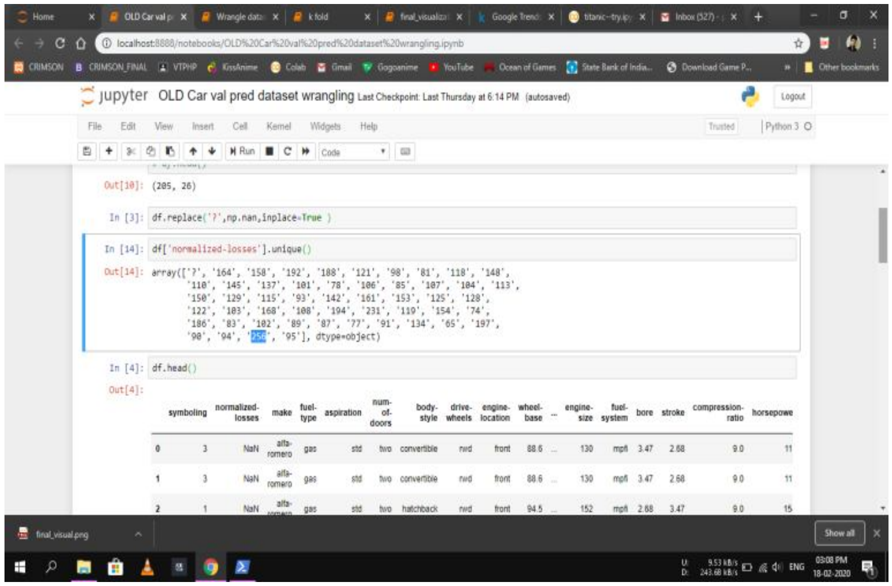


Fig. 2: set of attributes for vehicles for sample dataset

IV. EXPERIMENT

There is a prerequisite at an exchange vehicle value expectation framework to satisfactorily choose the estimation of the vehicle using a grouping of features. In spite of the way that there are destinations that offers this organization, their desire strategy may not be the best. What's more, different models and systems may contribute on predicting power for an exchange vehicle's veritable market regard. It is basic to consider the absolute information about the vehicle like particular or non specific credits to measure their authentic market regard while both buying and selling. The Experiment is performed on dataset using Multiple Linear Regression Model.

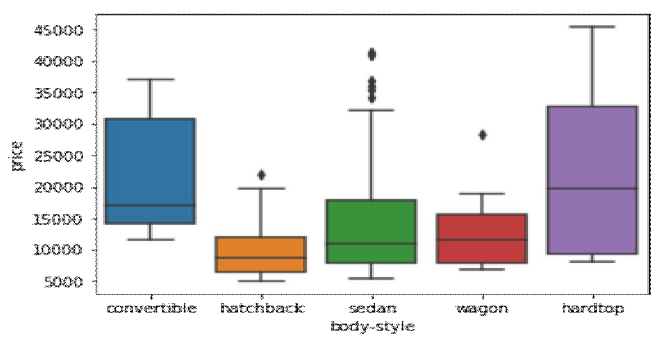


Fig. 3: linear plot for price and body type cars

Data cleaning is one of the critical bits of AI. It has a significant effect in building a model. Data Cleaning is some random thing that everyone excepts no one genuinely talks about. It unmistakably isn't the fanciest bit of AI and at the same time, there aren't any covered tricks or insider realities to uncover. In any case, real data cleaning can speak to the critical point in time your assignment. Capable data analysts normally spend a very tremendous piece of their time on this movement. If we have an inside and out cleaned dataset, we can get needed results even with an astoundingly direct estimation, which can show very profitable once in a while. Obviously, different sorts of data will require different sorts of cleaning. Nevertheless, this effective philosophy can for the most part fill in as a not too bad starting stage. One of the plot are also given to understand the data in better way. & also the overall plot for multiple linear regressions is shown here:

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

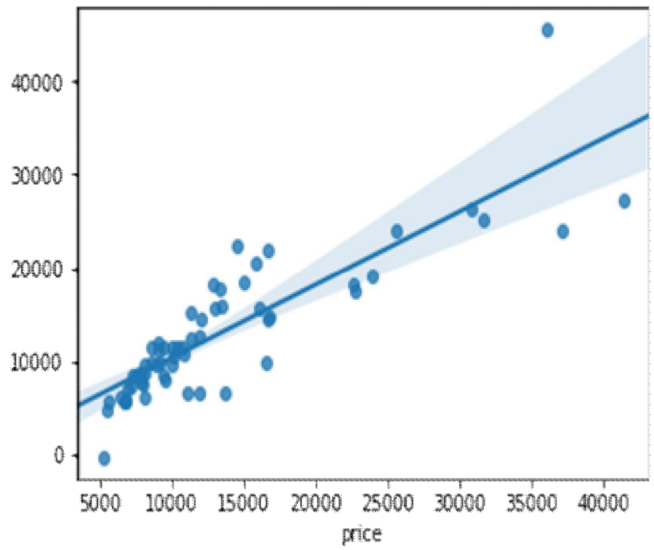


Fig. 4: overall plot for regression model for vehicle prediction

V. CONCLUSIONS

In this paper, distinctive AI procedure has been utilized to estimate the cost of trade-in vehicles. The outcome gives some minor advancement in term terms of estimation of vehicle cost. Henceforth, the value credit must be arranged into classes which comprise specialized and non specialized traits of vehicles. The fundamental confinement of this examination is the low number of records that have been utilized. As future work, we expect to gather more information and to utilize further developed methods like fake neural systems, fluffy rationale and hereditary calculations to predict car prices.

REFERENCES

[1] Pudaruth,"Predicting the Price of Used Cars using Machine Learning Techniques", ISSN 0974-2239 Volume 4, Number 7 (2014), pp. 753-764

[2] N.Sun, H. Bai, Y. Geng, and H. Shi, "Price evaluation model in second-hand car system based on BP neural network theory," in 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), jun 2017, pp. 431–436.

[3] NATIONAL TRANSPORT AUTHORITY. 2014. Available from: http://nta.gov.mu/English/Statistics/Pages/Archive.aspx [Accessed 15 January 2014].

[4] N. Kanwal and J. Sadaqat, "Vehicle Price Prediction System using Machine Learning Techniques," International Jounal of Computer Ap-plications, vol. 167, no. 9, pp. 27–31, 2017.

[5] S. Peerun, N. H. Chummun, and S. Pudaruth, "Predicting the Price of Second-hand Cars using Artificial Neural Networks," The Second International Conference on Data Mining, Internet Computing, and Big Data, no. August, pp. 17–21, 2015.

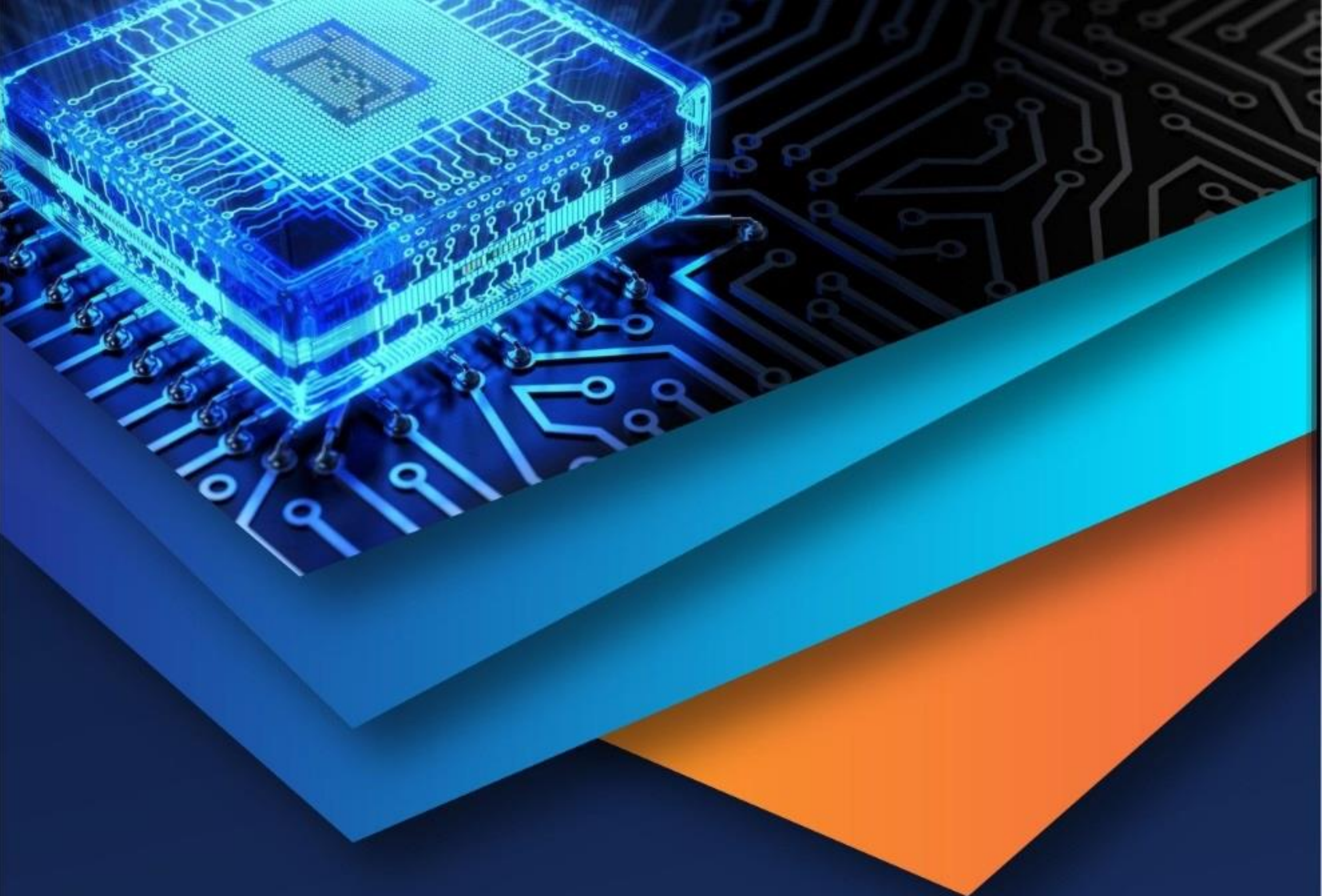
[6] G.Rossum, "Python Reference Manual," Amsterdam, The Netherlands, The Netherlands, Tech. Rep., 1995.

[7] GONGGI, S., 2011. New model for residual value prediction of used cars based on BP neural network and non-linear curve fit. In: Proceedings of the 3 rd IEEE International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Vol 2. pp. 682-685, IEEE Computer Society, Washington DC, USA.

[8] S. Peerun, N. H. Chummun, and S. Pudaruth, "Predicting the Price of Second-hand Cars using Artificial Neural Networks," The Second International Conference on Data Mining, Internet Computing, and Big Data, no. August, pp. 17–21, 2015.

[9] N.Sun, H. Bai, Y. Geng, and H. Shi, "Price evaluation model in second-hand car system based on BP neural network theory," in 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), jun 2017, pp. 431–436.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)