PRINT ISSN 1119-8362
Electronic ISSN 2659-1499

Full-text Available Online at
https://www.ajol.info/index.php/jasem
https://www.bioline.org.br/ja

J. Appl. Sci. Environ. Manage.
Vol. 29 (3) 825-829 Mar. 2025

Design and Implementation of Stock Market Prediction System for Used Cars in Nigeria

[1]*IBRAHIM, AA; [2]ALABI, AA; [3]AYILARA-ADEWALE, OA; [4]OLOKUN-OLUKOTUN, FR

*[1]*Department of Mathematical Sciences, Oduduwa University, Ipetumodu, Nigeria*
[2]Department of Information Systems, [3]Department of Information Technology, [4]Department of Software Engineering, Osun State University, Osogbo, Nigeria

**Corresponding Author Email: adebisiibrahim97@gmail.com*
**ORCID: https://orcid.org/0000-0002-8797-1369*
**Tel: +2347034845265*

Co-Authors Email: ayilara_adewale@uniosun.edu.ng; akeem.alabi@uniosun.edu.ng; funmilayo.olukotun@uniosun.edu.ng

ABSTRACT: Stock market prediction of commodities have undergone changes from the traditional to modern methods of using machine learning. Hence, the objective of this study was to design and implement a stock market price prediction system for used cars in Nigeria using machine learning techniques, the extra tree algorithm and support vector machine (SVM). The dataset used included such attributes as fuel type, number of doors, number of cylinders, drive wheel and price amongst others. The model was designed by training and testing using pre-processed data. Python programming language was used in the implementation. The results obtained for the mean square error and the R-squared showed high accuracy and therefore made the model ideal for car price prediction in the automobile Nigerian market.

DOI: https://dx.doi.org/10.4314/jasem.v29i3.18

Cite this Article as: IBRAHIM, A. A; ALABI, A. A; AYILARA-ADEWALE, O. A; OLOKUN-OLUKOTUN, F. R (2025 Design and Implementation of Stock Market Prediction System for Used Cars in Nigeria. *J. Appl. Sci. Environ. Manage.* 29 (3) 825-829

Dates: Received: 24 February 2025; Revised: 04 March 2025; Accepted: 18 March 2025; Published: 31 March 2025

Keywords: Tokunbo; Used cars; Extra tree; Machine learning;, Stock market

Trade in used automobiles form a significant part of foreign trade between advanced countries and their third world counterparts. While it is relatively easier to own brand new automobiles in advanced countries, third world countries depend mostly on foreign used automobiles, popularly called “Tokunbo” in Nigeria. The used car market history dates back to the early 20th century. The influx of automobiles and the need for affordable transportation led to the emergence of second-hand car sales (Smith, 2005). After World War II, there was a significant growth in the used car market. Returning soldiers sought vehicles and surplus military jeeps and trucks flooded the market (Brown, 1999), thus expanding used car dealerships and offering a wide range of models. Auto auctions became popular, providing a platform for buyers and sellers to trade pre-owned vehicles (Johnson, 2010). In the mid-20th century, quality concerns arose due to unscrupulous practices. Consumer protection laws were enacted to regulate the industry and ensure fair transactions (Garcia, 1987). With the advent of digital revolution of late 20th century, online platforms like Auto-Trader and Craigslist transformed the used car market, allowing buyers to search and compare listings easily (Lee, 2002). In recent times, environmental awareness has influenced the used car market. Consumers nowadays consider fuel efficiency, emissions as well as hybrid/electric

**Corresponding Author Email: adebisiibrahim97@gmail.com*
**ORCID: https://orcid.org/0000-0002-8797-1369*
**Tel: +2347034845265*

options when purchasing pre-owned vehicles (Green, 2015). Modern used cars benefit from advanced safety features, navigation systems, and improved reliability. Certified pre-owned programs by manufacturers assure buyers of quality (White, 2018).An understanding of market trends and demand patterns is crucial. Factors like economic conditions and consumer preferences also impact used car prices. Market trends significantly impact the demand and supply dynamics of used cars. High demand for specific models or features can drive up prices, while oversupply may lead to price reductions. Machine learning models consider historical demand patterns to estimate future prices based on market trends (Adhikary *et al*., 2022). Features such as make, model, year, mileage, engine type, and transmission directly affect pricing. Desirable features (e.g., leather seats, sunroof) increase value, while outdated features may decrease it. Using historical data, algorithms learn to understand feature importance and predict prices accordingly. Urban areas with higher population density often have greater demand for used cars. Regional preferences and local economic conditions influence prices. Geospatial features (e.g., city, state) are considered by machine learning algorithms. Similarly, economic stability, inflation rates, and interest rates affect consumer spending. During economic downturns, used car prices may decrease. Models may incorporate macroeconomic indicators to adjust price predictions (Asghar, *et al* 2021). Buyers seek modern technology (e.g., infotainment systems, safety features). Cars with advanced tech tend to fetch better prices. Algorithms learn from historical sales to assess the impact of technology on prices Scarcity or abundance of specific models affects prices. Limited availability raises demand, leading to higher prices. Models learn from historical data to anticipate supply-demand imbalances (Das Adhikary *et al*., 2022). Used car price prediction methods encompass a range of techniques used to develop models that estimate the value or future prices of used cars. These methods can be broadly categorized into traditional approaches and machine learning techniques. Traditional approaches to used car price prediction include statistical and econometric models that rely on regression analysis. These models use historical data on used car prices and a set of predictor variables to estimate the relationship between the predictors and the target variable (used car price). Common traditional approaches include hedonic pricing models and multiple regression analysis (Sopranzetti, 2015; John, 2007). Statistical and econometric models of used car prediction belong to traditional approaches that rely on regression analysis. These models use historical data of used car prices and a set of predictor variables to estimate the relationship between the predictors and the target variable, the used car price (John,2007, Cox, 2006). With the advent of technology, machine learning techniques have gained popularity in used car price prediction due to their ability to handle complex relationships and nonlinearities in the data. These techniques utilize algorithms that learn patterns and make predictions based on training data (Li, and Lin, 2021). Decision trees, including variants like gradient boosting and XGBoost, are powerful machine learning algorithms used for used car price prediction. Decision trees recursively split the data based on different features, creating a hierarchical structure that predicts the target variable. Gradient boosting and XGBoost enhance decision trees by iteratively optimizing the model's performance. Machine learning techniques offer advantages such as the ability to handle large and complex datasets, automatic feature selection, and adaptability to nonlinear relationships. However, they may require more computational resources and extensive data pre-processing. The choice of used car price prediction method depends on the characteristics of the dataset, the complexity of the relationships, and the desired level of accuracy. Hybrid models that combine traditional approaches with machine learning techniques are also common, leveraging the strengths of both methodologies. Hence, the objective of this study is to design and implement a stock market price prediction system for used cars in Nigeria.

MATERIALS AND METHODS

The dataset used was that of used cars dataset from Kaggle.com which consists of such attributes as car name, brand and model as well as specific version, fuel type, aspiration, number of doors, body type, drive wheel, engine type and mileage. The model was designed by developing two algorithms; one for data processing to prepare (clean and organize) the raw data to make it suitable for building and training the model and the other for the actual training of the model which is an ensemble learning method that builds multiple decision trees and aggregates their results for improved predictive accuracy and robustness. It was implemented using python programming language. Libraries like NumPy, pandas, matplotlib, which enables data pre-processing and training of the model were imported. Also, the pandas library was used to import the dataset which contains the variables to be used for the price prediction. The system was evaluated using mean squared error (MSE) (equation 1).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(i - \hat{Y}_i\right)^2 \qquad (1)$$

The average squared difference between actual and predicted values and the R square (equation 2)

$$R^2 = 1 - \frac{SS_R}{SS_M} \qquad (2)$$

Where: SS_R is the sum of squared error by regression line and SS_M represents the sum of squared error by mean line.

The data was preprocessed using python programming language such as NumPy, pandas and matplotlib, and presented as reading data set (Table 2) and further statistically evaluated and presented (Table 3).

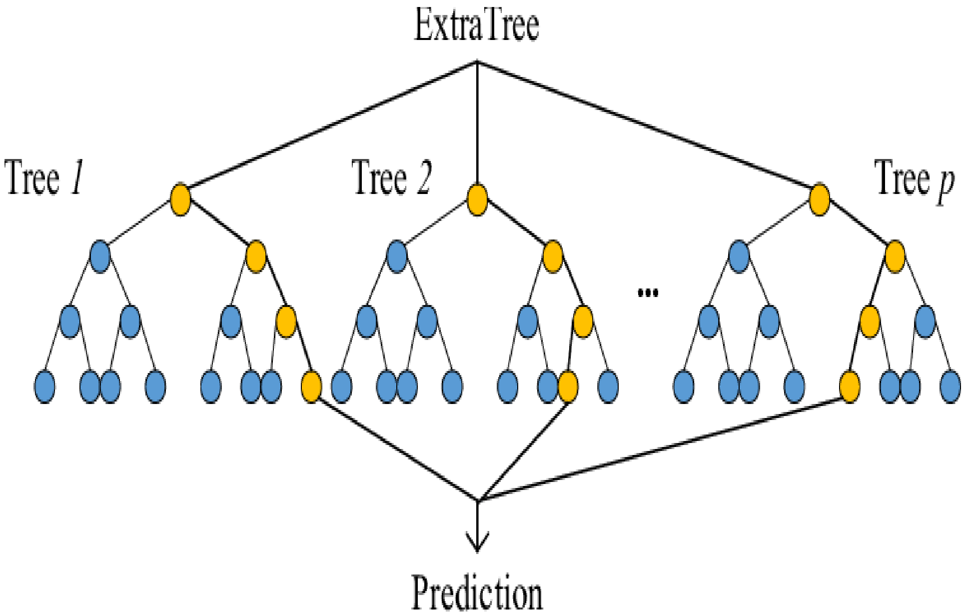

Fig.1: Extra Tree Model.

Table 1: Used Car Dataset

Car ID	Car Name	Fuel type	Door number	Car Body	Drive wheel	Engine location	Car width	Engine type	Cylinder number	Engine Size	Fuel system	Price
1	alfa-romero giulia	gas	Two	convertible	rwd	front	64.1	dohc	Four	130	mpfi	13495
2	alfa-romero stelvio	gas	Two	convertible	rwd	front	64.1	dohc	Four	130	mpfi	16500
3	alfa-romero Quadrifoglio	gas	Two	hatchback	rwd	front	65.5	ohcv	Six	152	mpfi	16500
4	audi 100 ls	gas	Four	sedan	fwd	front	66.2	ohc	Four	109	mpfi	13950
5	audi 100ls	gas	Four	sedan	4wd	front	66.4	ohc	Five	136	mpfi	17450
6	audi fox	gas	Two	sedan	fwd	front	66.3	ohc	Five	136	mpfi	15250
7	audi 100ls	gas	Four	sedan	fwd	front	71.4	ohc	Five	136	mpfi	17710
8	audi 5000	gas	Four	wagon	fwd	front	71.4	ohc	Five	136	mpfi	18920
9	audi 4000	gas	Four	sedan	fwd	front	71.4	Ohc	Five	131	mpfi	23875
10	audi 5000s (diesel)	gas	Two	hatchback	4wd	front	67.9	Ohc	Five	131	mpfi	17859.17
11	bmw 320i	gas	Two	sedan	rwd	front	64.8	Ohc	Four	108	mpfi	16430
12	bmw 320i	gas	four	sedan	rwd	front	64.8	Ohc	Four	108	mpfi	16925
13	bmw x1	gas	two	sedan	rwd	front	64.8	Ohc	Six	164	mpfi	20970
14	bmw x3	gas	four	sedan	rwd	front	64.8	Ohc	Six	164	mpfi	21105
15	bmw z4	gas	four	sedan	rwd	front	66.9	Ohc	Six	164	mpfi	24565
16	bmw x4	gas	four	sedan	rwd	front	66.9	Ohc	Six	209	mpfi	30760
17	bmw x5	gas	two	sedan	rwd	front	67.9	Ohc	Six	209	mpfi	41315
18	bmw x3	gas	four	sedan	rwd	front	70.9	Ohc	Six	209	mpfi	36880
19	chevrolet impala	gas	two	hatchback	fwd	front	60.3	L	Three	61	2bbl	5151
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
201	volvo 145e (sw)	gas	four	sedan	rwd	front	68.9	Ohc	Four	141	mpfi	16845
202	volvo 144ea	gas	four	sedan	rwd	front	68.8	Ohc	Four	141	mpfi	19045
203	volvo 244dl	gas	four	sedan	rwd	front	68.9	Ohcv	Six	173	mpfi	21485
204	volvo 246	diesel	four	sedan	rwd	front	68.9	Ohc	Six	145	Idi	22470
205	volvo 264gl	gas	four	sedan	rwd	front	68.9	Ohc	Four	141	mpfi	22625

Table 2: Reding Dataset.

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PRATIO	B	LSTAT	MEDV
0	0.000632	18	2.31	0	0.538	6.575	65.2	4.09	1	296	15.3	396.9	4.98	24
1	0.02731	0	7.07	0	0.469	6.421	78.9	4.9671	2	242	17.8	396.9	9.14	21.6
2	0.02727	0	7.07	0	0.469	7.185	61.1	4.9671	2	242	17.8	392.83	4.03	34.7
204	0.03237	0	2.18	0	0.458	6.998	45.8	6.0622	3	222	18.7	394.63	2.94	33.4
205	0.06905	0	2.18	0	0.458	7.147	54.2	6.0622	3	222	18.7	396.9	5.33	36.2

Table 3: Summary Statistics

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PRATIO	B	LSTAT
0	0.000632	18	2.31	0	0.538	6.575	65.2	4.09	1	296	15.3	396.9	4.98
1	0.02731	0	7.07	0	0.469	6.421	78.9	4.9671	1	242	17.8	396.9	9.14
2	0.02727	0	7.07	0	0.469	7.185	61.1	4.9671	1	242	17.8	392.83	4.03
204	0.03237	0	2.18	0	0.458	6.998	45.8	6.0622	1	222	18.7	394.63	2.94
205	0.06905	0	2.18	0	0.458	7.147	54.2	6.0622	1	222	18.7	396.9	5.33

RESULTS AND DISCUSSION

The model was tested using user input for the features required of the car and prediction made from that preference. The model was evaluated using mean squared error (MSE) and R-squared score. Mean Squared Error measured the average squared difference between actual and predicted values, penalizing larger errors more. The predicted and actual used car prices were computed as presented in Table 4.

Table 4: Actual versus Predicted Car

S/N	Actual Value	Predicted Value	Difference
1	30760	35854.09	-5094.09
2	17859.2	18938.93	-1079.76
3	9549	9021.80	527.20
4	11850	12894.80	-1044.80
5	28248	26908.96	1339.04
6	7799	6506.24	1292.76
7	7788	7762.08	25.92
8	9258	7990.34	1267.66
9	10198	9875.34	322.66
10	7775	8301.86	-526.86
11	13295	14157.18	-862.18
12⋮	8238⋮	7858.89⋮	379.11
203	18280	13555.17	4724.83
204	9988	10722.95	-734.95
205	409060	39655.12	1304.90

The performance evaluation metrics for the model on the test set were performed, giving MSE and R^2 results. This study shows that it is possible to predict used cars price by training an extra tree regression model. The model outperformed traditional linear regression, capturing complex relationships and providing reliable predictions, along with the evaluation of the model's performance with MSE of 8.73 and R^2 of 0.90

Conclusion: A predictive model for stock market of used cars in Nigeria was proposed as against the traditional methods. The developed model used dataset from Kaggle.com with attributes that are generally desired by car owners. The results obtained showed that it has a higher accuracy than traditional linear regression models in terms of accuracy and robustness. The model achieved a mean square error of 8.73 and R^2 value of 0.90 respectively with a dataset of 206 entries. This indicates that the model is more accurate for used cars price prediction and will be of immense use to both sellers and buyers of cars in the Nigerian automobile market.

Declaration of Conflict of Interest: The authors declare that there is no conflict of interest among them.

Data Availability: Data are available upon request from first author.

REFERENCES

Adhikary, DR; Sahu, R., and Panda, SP. (2022). Machine Learning Models for Used Car Price Prediction. *J. Mach. Learn. Res.* 40(4): 300-315.

Soprazentti, BJ, (2015). Hedonic Regression Models for Used Car Appraisal. *App. J.* 18(2): 95-108.

Asghar, N; Mehmood, Z; Yasin, A; Khan, MA. (2021). Forecasting pre-owned car prices in Pakistan using machine learning algorithms. *Int. J Adv. Comp. Sc. App.* 12(5): 221-229.

Cox, S. (2006). Regression Analysis Techniques. *Stat. Meth. J* 15(4): 180-193.

Brown, R. (1999). Post-World War II Expansion of the Used Car Market. *Econ. Rev.* 15(2): 112-125.

Garcia, L. (1987). Consumer Protection Laws and the Used Car Industry. *Leg. St J* 12(1): 30-42.

Green, P. (2015). Environmental Influence on the Used Car Market. *Gr Tech. Rev* 25(2): 78-91.

Smith, J. (2005). History of Used Cars. *J Auto Hist* 20(3): 45-58.

John, R. (2007). Statistical Models in Car Valuation. *Econ. Rev*. 22(3): 210-225.

Johnson, M. (2010). Evolution of Auto Auctions in the Used Car Market. *Auc Ins*, 5(4): 225-238.

Lee, S. (2002). Digital Transformation of the Used Car Market. *On Mark*, 8(3): 135-148.

Li, S; Lin, Z. (2021). Deep learning models for predicting used car prices: a case study. *Neu Comp App*, 33(14): 10067-10076.

White, A. (2018). Quality Assurance through Certified Pre-Owned Programs. *Auto Qtly*, 30(1): 15-28.