IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol. 9, Issue.7, July 2021 E-ISSN: 2347-2693

# KNN and Decision Tree Model to Predict Values in Amount of One Pound Table

**Stanley Ziweritin[1*], Iduma Aka Ibiam[2], Taiwo Adisa Oyeniran[3], Godwin Epiahe Oko[4]**

[1,2]Department of Estate Management, Akanu Ibiam Federal Polytechnic, Unwana-Afikpo, Ebonyi State, Nigeria
[3,4]Department of Computer Science, Akanu Ibiam Federal Polytechnic, Unwana-Afikpo, Ebonyi State, Nigeria

**Corresponding Author: stanlo4godsluv@yahoo.com*

DOI: https://doi.org/10.26438/ijcse/v9i7.1621 | Available online at: www.ijcseonline.org

Received: 20/Jul/2021, Accepted: 20/Jul/2021, Published: 31/Jul/2021

***Abstract*—** Machine learning is one of the fast growing areas of interest in artificial intelligence adopted by professional in every spheres of life that uses algorithms with data to sytematically learn patterns and improve from experience. The increasing competitive and robust predicting methods of machine learning are becoming more interesting and popular. This is valuable to investors, surveyors and valuers against manually computed payment table values that depends on emperical results. There are tedious and rigorous processes in valuation practice that involves some aspects of financial analysis in computations for the one pound table values. The aim is to build K-nearest neigbr and decision tree model to predict the nemeric values in amount of one pound table at a give rate of interest and period of years.This model is useful to investors, accountants, data professionals, surveyors and valuers interested in financial analysis and its applications. A cross validation test was carried out with predicted R-squared test to detect overfitting and generalize model performance on testing dataset. We introduced noisy data with smoothing curve expeoneintial function to overcome the risk of overfitting in predicting target varaible. The K-nearest neighbor and decision tree techniques were trained, tested and resulted into 95.76% and 99.86% respectively.

***Keywords*—**Artificial intelligence, Decision tree, K-nearest neighbor, Machine learning

## I. INTRODUCTION

Machine learning techniques such as decision tree(DT) and K-nearest neighbors(KNN) are widely used by professionals to systematically learn from events of past experience without human assisitance[1],[2],[3],[4],[5] . These Machine learning(ML) techniques can learn from training data without being traditionally programmed through the concept of supervised/unsupervised learning[6]. The supervised is ML and are employed to solve classification and regression problems[7],[8]. The unspuervised ML algorithms are used to solve clustering and association rule-based mining problems[9],[10],[11]. The DT is described to be an inverted form of a normal tree-like structure where each sub-node is used to represent a test on attribute value, tree leaf-nodes denotes classes and branches, depicts results of the test[12]. It is very useful in solving classification and regression problems and there are several types of popularly known DT techniques namely: CHAID, ID3, C4.5 etc. In constructing DTs, sub-nodes belonging to same parent nodes are divided and grouped with respect to the best attribute value that discriminates these sub-nodes based on class. The method of attribute selection is done using gain ratio (Gini index) and information gain.

The increasing competitive and robust predicting methods of machine learning(ML) under artificial intelligence(AI) are becoming more popular, effective and relaible. This is valuable to investors, surveyors and valuers against manually computed payment values in the financial market that depends on limited emperical results. In valuation practice and other studies under land use are some aspects of financial analysis involving tedious and rigorous processes in carrying out property valuation (present value of one pound) and investment values, but little or no work has been done in the area of predicting payment values on real estate investment. The aim is to build and train KNN and DT regressor model capable of predicting numeric values in amount of one pound table at a given rate of interest based on assumptions made on valuation theories in real estate practice. This model will be useful to researchers, accountants, surveyors and valuers, data and ML professionals interested in dealing with investment and financial computations relating to real estate management and valuation practice and its applications.

This paper is divided into sections as follows: section one contains the introduction; section two presents a brief review of some of the previous techniques or methods to the study area and the gap in exploring the proposed KNN and DT model; Section three, introduces the materials and methods employed for building the model; section four, focuses on the results and detailed discussion of results; Section five presents the conclusion.

## II. RELATED WORKS

Mongia and Singh[13]; adopted the C4.5 type of decision tree using the concept of future reduction technique to determine the best investment options such as bank deposits, mutual funds, equity, shares, post office investment options and etc. The ranked algorithm was employed as a feature technique to organize attributes in the order of priority and according to their gain ratio applied to 500 tuples of dataset. The training time complexity of DT was measured to be faster and produced 89.47% accuracy. The general performance was very poor and was less than 95%. A combination of linear regression(LR), K-nearest neighnor (KNN) and random forest(RF) was adopted to estimate the trend of real estate housing prices with some features[14]. A quarterly dataset of 5000 items ranging from the year 2015 to 2016 was collected through web scrapping and divided into 4:1 ratio of training and testing dataset respectively. The performance was not encouraging in terms of accuracy and required more training dataset to perform well as required. Jadhav and Channe[15] employed different data mining classification techniques such as K-NN, Naive Bayes(NB) and DT to learn training data patterns in solving classification problems. The DT technique was developed with some rule-based concept easier for human to understand. The KNN produced 89.84%, NB and DT produced 63.71% with 4627 instances. The proposed model could not work well with large volume of structured and unstructured data. Alkhatib *et al.*[16] adopted non-linear regression and K-NN technique to predict stock market prices. The dataset was obtained from five registered companies within the Jordanian Stock market. The KNN algorithm performed better with small error margin but required more training data and advanced techniques to produce useful stock market prediction model. Singh *et al*,[17] developed KNN, random forest(RF) and Naive Bayes classifiers to identify text, numeric and alphanumeric data types. The model was created and fitted in storing known classes of training dataset to learn patterns to carry out predictions. The changing number of features in the dataset had no significant effect on the performance of RF while the success metrics of KNN and NB fluctuates (increases and decreases) that resulted into low accuracy rate.

## III. METHODOLOGY

DTs are the most popularly used supervised ML techniques by professionals because of its ability to perform well with missing values or noisy data without failure. It can work well under several conditions in forming more robost predictors[18],[19].

**Dataset:**The dataset was obtaned from Parry's investment and valuation table[20]. computed using a randomized function in Python. The dataset was divided into (80% or 12,000) training and (30% or 300) testing set to form 15,000 items. The attributes of the dataset contained interest rate and number of years in estimating target variable as the present value(PV) in amount of one pound table.

**Table I**: The Proposed system dataset

| | YEAR | RATE | TARGET |
|---|---|---|---|
| 0 | 6 | 7.25 | 0.849785 |
| 1 | 20 | 3.75 | 0.478892 |
| 2 | 18 | 2.50 | 0.641166 |
| 3 | 11 | 6.25 | 0.513312 |
| … | … | … | … |
| 14995 | 66 | 2.00 | 0.2706379 |
| 14996 | 2 | 2.50 | 0.9518140 |
| 14997 | 99 | 2.50 | 0.1104920 |
| 14998 | 48 | 3.00 | 0.2419990 |
| 14999 | 83 | 3.25 | 0.0143080 |

The PV of one pound table values are predicted by the NN and DT techniques with the features given bellow in equation 1 and 2 after training stage:

$$PV=\frac{FV}{(1+r)^{n}} \tag{1}$$

$$r=\frac{i}{100} \tag{2}$$

Where n is the number of periods until payment represented with the variable "YEAR", FV as "TARGET" which is the future value and r the internal rate represented with the attribute "RATE" and i is the interest.

### The Decision Tree regressor

The DT regressor is employed to estimate numeric values for valuation purpose. and the problem of black-box in ML models can be simplified with its visualization[21],[22]. The concept of building DT regressor is the same as classification trees but the attribute selection criteria varies or changes[23],[24]. The equation(1) given bellow is used as a measure to overcome the problem of variance in DT.

$$D=\frac{1}{\ell}\sum_{\ell=1}^{1}\left(y_i-\frac{1}{\ell}\sum_{j=1}^{\ell}y_j\right)^2 \tag{3}$$

Where $\ell$ = no. of items in DT leaf nodes and $y_i$ is the target variable representing valuation and investmenet values. We are searching for features that divide training sets into target features approximately equal or same to solve the problem of variance in DT with some distributions of dataset containing noise using an exponential smoothing.

A DT regressor class was created with some instances containing maximum depth set to 15 and random states to 0. The model was trained with (80% or 12,000) of the total dataset, tested and predicted with the remaining 20% or 300 items for validation purpose. We are generating some recursive distribution of data using piecewise function [f(x)] with noisy content given as:

$$f(x) = e^{-x^2} + 1.5e^{-(x-2)^2} \tag{4}$$

Where f(x) is the piecewise function and x is the training and testing data samples. We defined and created a function that returns noise and n_samples values within the interval -4 to 4.

**K-nearest Neighbor**

The KNN requires no assumptions about data distribution in predicting target variables[25],[26],[27],[28],[29],[30]. It works based on observations about data features similar to the points in one particular class known as nearest neighbors[31],[32],[33],[34]. The k parameter specifies the number of neighboring points used to classify similar data points into one class by the concept of voting[35],[36]. We created KNN classifier object and passed the k-neighbors argument values to the function in the K-neighbors Class. The model was fitted and trained with training set with fit command. The predict() command is employed in Python to perform prediction using the test dataset.

**Algorithm 1:** Decision Tree Regressor

| Step | Processes involved |
|---|---|
| 1 | Start |
| 2 | Choose the best attribute as the root node from dataset |
| 3 | Search for subgroups with same value of attributes |
| 4 | Repeat step 1&2 until a leaf node is found for all DT branches |
| 5 | Compare root node with values and move to next node |
| 6 | Return node |

**Algorithm 2:** K-Nearest Neighbors(KNN)

| Step | Processes involved |
|---|---|
| 1 | Start |
| 2 | Gather the unsorted dataset with known categories |
| 3 | Carryout clustering on sample data points |
| 4 | Sort data points and select the first as training samples |
| 5 | Find K by using K= sqrt(n), where n = No. of samples |
| 6 | Locate K-NN values |
| 7 | Categorize new points using the concept of majority vote |
| 8 | Return |

**Training and testing process**: The KNN and DT classifiers are trained and tested using a cross validation test. The total dataset was divided into (80% or 12000) training and (20% or 3000) of testing dataset to form a total of 15000 items. The both model was trained and tested with some distribution of data generated by a piecewise function to learn and estimate (predict) data patterns. The noisy data variable set to 0.1 and developed in a modular fashion [37].

**Parameter tuning**: We defined what can make the model perform better with some twicking hyper-parameter values(learning rate and intial conditions) after training stage. The fine-tuning of these variables still remains a bit more of heart than science to overcome the identified problem as an experiemental process that heavily depends on the specifics of dataset, model and training process[38]. We implicitly adjusted the learning rate to improve the performance of the model which potentialy leads to higher accuracy in performance metrics. The learning rate defines how far we can shift the line accuracy in training process based on the information from the previous step. The initial conditions played a significant role for determining the outcome of training but depends on whether a model starts training with values iniatialized zeros versus some distribution of values.

**Prediction**: The power of machine learning is that; it can detect, estimate, make prediction , forecast and identify hidden data patterns from existing training dataset. ML uses data to answer questions rather than using human judgement. This granted computer systems new abilities to learn from training data patterns in making predictions with the testing(unknown) dataset

**Model Evaluation**: is the final stage where we can test the performance of our model against data that has never been used before(unseen or new born data). The model evaluation is carried out with the testing dataset of smaller percentage, but depends on the source of the original dataset[39]. The accuracy, RMSE, SSE, f1-score, precision and recall metrics are used as diagnositic tools in measuring the performance of KNN and DT techniques.

**The root mean square error(RMSE)**: is used to measure the real difference between the acutal or observed($X_{obs}$) and the estimated/predicted($X_{model}$) values(Holmes, 2000)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{obs,i} - X_{model,i}\right)^2}{n}} \quad (5)$$

Where n is the number of values in one dataset

## IV. RESULTS AND DISCUSSION

We discussed about results of the cross validation test obtained from the KNN and DT regressor class with some evaluation tools. The accuracy metrics, RMSE, R-squared error, Mean square error values and time complexity were fully discussed to ascertain its efficiency.

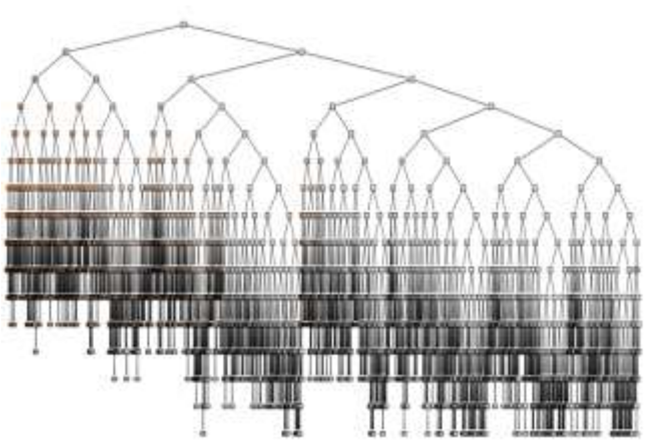


**Figure 1**: DT graph obtained from the training dataset

Figure 1 is the structure of DT generated from the proposed system dataset. The model randomly select patterns from original set in growing the tree and variables to represent random subset at each step. This resulted into a DT of height fifteen as shown in figure 1.

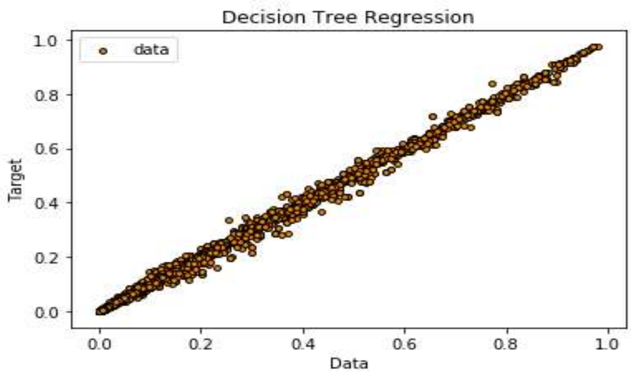



**Figure 2**: The DT regression model

The scattered plot of DT regression class shows that; the data patterns have a linear or strong relationship because the data points follow a linear pattern as shown in figure 2. The presence of noise made few points to fall outside the line of regression but not too far. The DT was able to learn from the noisy data in forming a linear pattern.

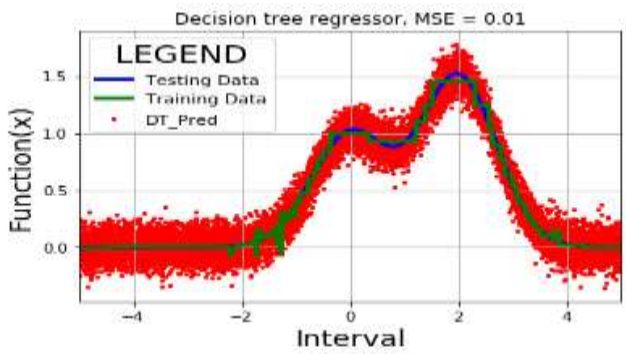



**Figure 3**: The mean square error(MSE) graph of DT regressor

The DT predicted values follows the pattern of training and testing data as represented with green and blue colors shown in figure 2. The simulation was done to visualize the learning data trend or pattern using a function. For each value f(x) within the interval of -4 to 4, the model produced the best possible predicted values and was able to approximate data with a constant piecewise function.

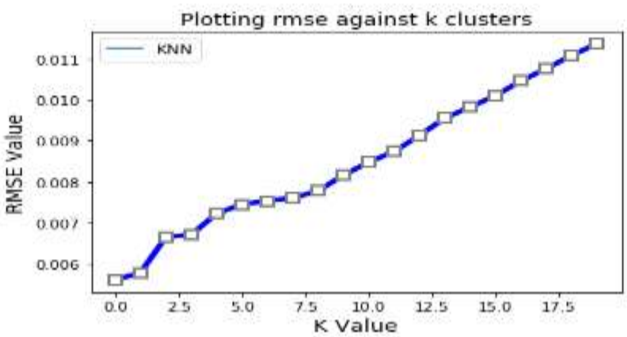



**Figure 4**: The KNN plot of RMSE against k values

Figure 4 shows the graph of root mean square(RMSE) against the K-nearest neighbor value(K). The RMSE value of KNN shoots up along the plane as the k neighbor values increases.

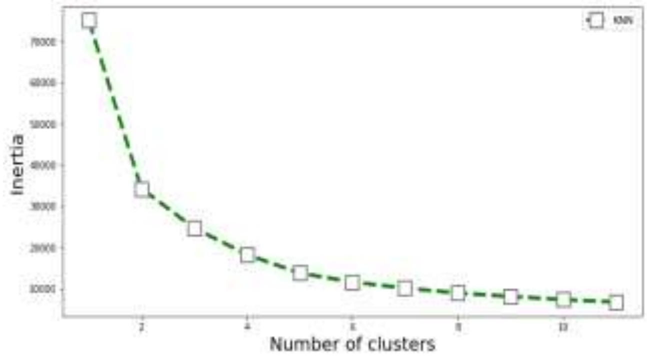



**Figure 5**: The KNN Sum of square error(SSE) graph

Figure 5 depicts the k-NN plot of SSE against the k values ranging from 1 to 12. The best value k is the elbow on the arm at 2. The SSE value decreases gradually to 0 when k is set to loop through the number of data points in the dataset. The SSE value gets closer to 0 point when we increase the number of clusters.

**Table II:** The RMSE value of KNN

| K | RMSE of KNN | K | RMSE of KNN |
|---|---|---|---|
| 1 | 0.005608549031937772 | 11 | 0.008477838288919032 |
| 2 | 0.005792398037887912 | 12 | 0.008739102802204532 |
| 3 | 0.006658118998030837 | 13 | 0.009130543885583226 |
| 4 | 0.006712592916260361 | 14 | 0.00955111052869568 |
| 5 | 0.007229844885707784 | 15 | 0.009823547125459324 |
| 6 | 0.007451746978952787 | 16 | 0.010100174295226135 |
| 7 | 0.007537060529943211 | 17 | 0.010457316836198617 |
| 8 | 0.007609030974402126 | 18 | 0.010755561093739572 |
| 9 | 0.007778680470888901 | 19 | 0.011077310013169254 |
| 10 | 0.00816418015009959 | 20 | 0.011362635116874565 |

The RMSE value of KNN increases down the table with the k ranging from 1 to 20 at training stage as shown in table II.

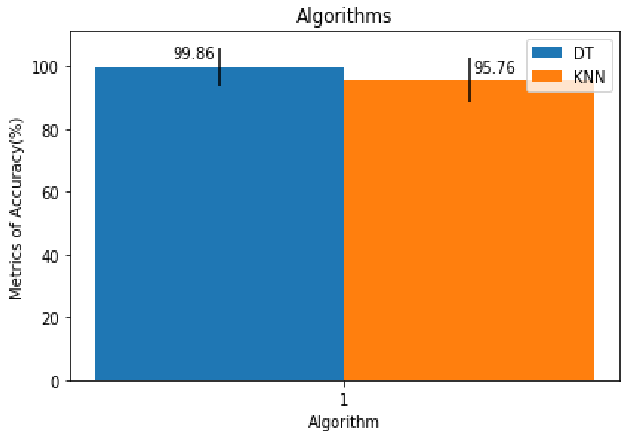



**Figure 6**: Percentage accuracy of KNN and DT

Figure 6 shows the performance of DT and KNN techniques measured in percentage at testing stage. The DT regressor class produced 99.86% metrics of accuracy which was higher compared to the KNN with 95.76% success rate.

**Table I:** Variation in time complexity and R-squared value

| | **Training Time** | **R-Squared** |
|---|---|---|
| Decision tree | 0.0190 | 0.9998 |
| KNN | 0.0120 | 0.9984 |

Table 1 shows the variation of training time complexity and R-squared values for both models. The time complexity of KNN was measured to be 0.0120 and DT as 0.0190 seconds. The KNN was faster than the DT regressor class in terms of time complexity.

## V. CONCLUSION

The existing techniques are inefficient in terms of accuracy, running time complexity and error rates in estimating investment and payment values. The performance metrics of DT model was higher in terms of accuracy but its training time was longer compared to the KNN technique. The experimental results proved to be highly efficient and accurate for predicting the target variable. We therefore; evidently conclude that the DT regressor class performed better as required than the KNN.

## REFERENCES


[1] Y. LeCun, Y. Bengio, and G. Hinto, "Deep Learning," *Nature*, vol.**521**, no.**7553**, pp.**436-444**, **2015**.

[2] L. Breiman, "Random Forests, "Machine Learning", vol.**45**, no.**1**, pp.**5-12**, **2001**.

[3] A. Gupa, A. Mohammad, A. Syed, and M. N. Halgamuge, "A Comparative Study of Classification Algorithms Using Data Mining: Crime and Accidents in Denver City the USA," International Journal of Advanced Computer Science and Applications(IJACSA), vol.**7**, no.**7**, pp.**374-381**, **2016**.

[4] K. Wisaeng, "A Comparison of Different Classification Techniques for Bank Direct Marketing," International Journal Of Soft Computing and Engineering(IJSCE), vol.**3**, no.**4**, **116–119**, **201**3.

[5] M. Fatima, M. Pasha, "Survey of Machine Learning Algorithms for Disease Diagnostic," Journal of Intelligent Learning Systems and Applications,vol.**9**, no.**1**, pp.**1-16**, **2017**.

[6] M. Abedini, N. C. F. Codella, R. Connell, R. Garnavi, M. Merler, S. Pankanti, J. R. Smith, and T. Yeda-Mahmood, "A Generalized Framework for Medical Image Classification and Recognition," IBM J Res Dev. vol.**59**, no.(**2**/**3**), pp.**1-18**, **2015**.

[7] G. Kashyap, and E. Chauhan, "Parametic Comparisons of Classification Techniques in Data Mining Applications," International Journal of Engineering Development and Research(IJEDR), vol.**4**, no.**2**, pp.**1116-1121**, **2016**.

[8] A. Upadhayay, S. Shukla, and S. Kumar, "Empirical Comparison by data mining Classification algorithms (C 4.5 A & C 5.0) For thyroid cancer data set," International Journal of Computer Science and Communication Networks(IJCSCN), vol.**3**, no.**1**, pp.**64–68**, **2013**.

[9] K. H. Rao, G. Srinivas, A. Damodhar, and M. V. Krishna, (2011). Implementation of anomaly detection technique using machine learning algorithms, International Journal of Computer Science and Telecommunication(IJCST), vol.**2**, no.**3**, pp.**25-31**, **2011**.

[10] S. Ruggieri, "Efficient C4.5 Classification Algorithm," Knowledge and Data Engineering, IEEE Transactions, vol.**14**, no.**2**, pp.**438-444**, **2002**.

[11] J. G. Moreno-Torres, J. A. Saez, and F. Herrera, "Study on the Impact of Partition-Induced Dataset Shift on K-Fold Cross-Validation," *IEEE* Transactions on Neural Networks and Learning Systems, vol.**23**, no.**8**, pp.**1304-1312**, **2012**.

[12] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, "Practical Machine Learning Tools and Techniques," Morgan Kaufmann, pp.**34-89**, **2016**.

[13] V. Mongia and G. Singh, "Automatic Evaluation of Best Investment Options for Investors Using Optimal Decision Tree Algorithm," International Journal of Computer Applications(IJCA), vol.**96**, no.**20**, pp.**39-44**, **2014**.

[14] S. Borde, A. Rane, G. Shende, and S. Shetty, "Real Estate Investment Advising Using Machine Learning," International Research Journal of Engineering and Technology(IRJET), *vol.***4**, no.**3**, pp.**1821-1825**, **2017**.

[15] D. S. Jadhav, and H. P. Channe, "Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques," International Journal of Science and Research (IJSR), vol.**5**, no.**1**, pp.**1842-1845**, **2014**.

[16] K. Alkhatib, H. Najadat, I. Hmeidi, and M. K. Ali-Shatnawi,. "Stock Price Prediction Using K-Nearest Neighbor (kNN) Algorithm," International Journal of Business, Humanities and Technology(IJBHT), vol.**3**, no.**3**, pp.**32-44**, **2013**.

[17] A. Singh, M. N. Halgamuge, and R. Lakshmiganthan, "Impact of Different Data Types on Classifier Performance of Random Forest, Naive Bayes, and K-Nearest Neighbors Algorithm," International Journal of Advanced Computer Science and Applications(IJACSA), vol.**8**, no.**12**, pp.**1-11**, **2017**.

[18] A. Liaw, and M. Wiener, "Classification and Regression by Random Forest," R News, vol. **2**, no.**3**, pp.**18-22**, **2003**.

[19] T. Mary-Huard, and S. Robin, "Tailored aggregation for classification," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.**31**, no.**11**, pp.**2098–2105**, **2009**.

[20] A. W. Davidson, "Parry's Valuation and Investment Tables," Estates Gazette, 13th Edition, pp.**1-550**, **2013**.

[21] V. Podgorelec, P. Kokol, B. Stiglic, I. Rozman, "Decision Trees: An Overview andTheir Use in Medicine," Journal of Medical Systems Kluwer Academic/Plenum Press, vol.**25**, no.**5**, pp.**445-463**, **2002**.

[22] M. Anyanwu, and S. Shiva, "Comparative Analysis of Serial Detection Tree Classification Algorithms," International Journal of Computer Science and Security, vol.**3**, no.**3**, pp.**229-240**, **2019**.

[23] L. Wei-yin, "Fifty Years of Classification and Regression Trees," International Statistical Review, vol.**82**, no.**3**, pp.**329-348**, **2014**.

[24] M. Xu, J. Wang, and T. Chen, "Improved Decision Tree Algorithm ID3+, Intelligent Computing in Signal Processing and Pattern Recognition, vol.**345**, pp.**141-149**, **2006**.

[25] S. B. Imandoust, and M. Bolandraftar, "Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background", International Journal of Engineering Research and Applications, vol.**3**, no.**5**, pp.**605-610**, **2013**.

[26] P. Hall, B. U. Park, and R. J. Samworth, "Choice of Neighbor Order in Nearest-Neighbor Classification," *The Annals of Statistics*, vol.**36**, no.**5**, pp.**2135-2152**, **2008**.

[27] A. Solichin, "Comparison of Decision Tree, Naïve Bayes and KNearest Neighbors for Predicting Thesis Graduation," Proceedings of the Electrical Engineering Computer Science and Informatics, pp.**18-20**, **2019**.

[28] M. T. Van, N. V. Tuan, T. T. Son, H. Le-Minh, and A. Burton, "Weighted K-nearest Neighbor Model for Indoor VLC Positioning," *IET Communications*, vol.**11**, no.**6**, pp.**864-871**, **2017**.

[29] R. Benetis, C. Jensen, G. Karciauskas, and S. Saltenis, "Nearest and Reverse Nearest Neighbor Queries for Moving Objects," The International on Very Large Data Bases, vol.**15**, no.**3**, pp.**229-250**, **2006**.

[30] Y. S. Chen, Y. P. Hung, T. F. Yen, and C. S. Fuh, "Fast and Versatile Algorithm for Nearest Neighbor Search Based on a Lower Bound Tree," *Pattern Recognition*, vol.**40**, no.**2**, pp.**360–375**, **2007**.

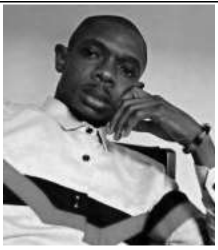
[31] J. Geu, L. Du, Y. Zhang, and T. Xiog, "A New Distance-Weighted K-nearest Neighbor Classifier," *Journal of Information and Computational Science*, vol.**9**, no.**6**, pp.**1429-1436**, **2012**.

[32] V. Athitsos, J. Alon, S. Sclaroff, and G. Kollios, "Boostmap: An Embedding Method for Efficient Nearest Neighbor Retrieval,"

IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.**30**, no.**1**, pp.**89–104**, **2007**.
[33] C. Domeniconi, J. Peng, and D. Gunopulos, "Local Adaptive Metric Nearest-Neighbor Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol.**24**, no.**9**, pp.**1281-1285**, **2002**.
[34] H. Wang, "Nearest Neighbors by Neighborhood Counting," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.**28**, no.**6**, pp.**942–953**, **2009**.
[35] E. Bax, "Validation of Nearest Neighbor Classifiers," IEEE Transaction Information Theory, vol.**46**, pp.**2746-2752**, **2007**.
[36] S. Zhao, C. Rui, and Y. Zhang, "MICKNN Multi-instance Covering KNN Algorithm," Tsinghua Science and Technology, vol.**18**, no.**4**, **360-368**, **2013**.
[37] X. Zhang, and S. A. Jiang, "A Splitting Criteria Based on Similarity in Decision Tree Learning," *JSW*, vol. 7, no. 8, pp.82-1775, 2012.
[38] Koli, S. "Sentiment Analysis with Machine Learning," International Journal of Computer Science and Engineering(IJCSE), Vol.**9**, no.**6**, pp.**77-82**, **2012**.
[39] Ignatious, J., Kulkarni, Y., Bari, S. and Naglot, D. "Comparative Analysis of Machine Learning Algorithms For Credit Card Fraud Detection," International Journal of Computer Science and Engineering(IJCSE), vol.**8**, no.**6**, pp.**6-9**, **2020.**

## AUTHORS PROFILE



*Mr. S. Ziweritin* holds HND(Computer Science) from Rivers State(Now Kenule Benson Saro-Wiwa) Polytechnic, Bori. PGD(Computer Science) and M.Sc(Computer Science) from University of Port Harcourt(UPH) respectively. He is presently a lecturer in the Department of Estate Management, School of Environmental Design and Technology, Akanu Ibiam Federal Polytechnic, Unwana, Ebonyi State. His research interest revolves around: Artificial Intelligence(AI), Data Mining, Fuzzy Logic, Software Engineering, Machine Learning(ML), Algorithms, Database System Design and Programming. He has published in several international journals relating to AI, Data mining and Machine learning.

*Mr. I. A. Ibiam* Holds ND (Building Technology) from College of Technology(Now Federal Polytechnic) Nekede Owerri, BSc(Estate Management) from Imo(Now Abia) State University Uturu, MBA and MSc(Environmental Resource Management) from the University of Calabar(Unical). He is presently a Chief lecturer in the department of Estate Management and Dean, School of Environmental Design and Technology, Akanu Ibiam Federal Polytechnic, Unwana, Ebonyi State. He is an Associate member (ANIVS) of the Nigerian Institution of Estate Surveyors and Valuers (NIESV), and a registered Estate Surveyor and Valuer (RSV). He has published in several national and international journals.

*Mr. T. A. Oyeniran* Completed HND(Computer Science) from Federal Polytechnic, Offa, Kwara State. B.Sc(Computer Science) from Imo State University Owerri. Presently pursuing M.Sc(Computer Science) at Imo State University, Owerri. He is an academic Staff(Asst. Chief Technologist) in the Department of Computer Science, Akanu Ibiam Federal Polytechnic, Unwana with over 16 years of teaching and research experience. His research interest include: Information Technology(IT), AI and Expert System. He has published several reputable journals, books and conference papers.

*Mr. G. E. Oko* holds BSc(Computer Science) from University of Calabar (Unical). He is presently a Lecturer in Computer Science Department, Akanu Ibiam Federal Polytechnic, Unwana(AIFPU), Ebonyi State. He has several years of teaching and research experience.