IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 11, No. 4, December 2022, pp. 1535~1544
ISSN: 2252-8938, DOI: 10.11591/ijai.v11.i4.pp1535-1544

Weather prediction performance evaluation on selected machine learning algorithms

Muyideen AbdulRaheem[1], Joseph Bamidele Awotunde[1,3,4], Abidemi Emmanuel Adeniyi[2,3,4,5], Idowu Dauda Oladipo[1], Sekinat Olaide Adekola[1]

[1]Department of Computer Science, University of Ilorin, Ilorin, Nigeria
[2]Department of Computer Science, Landmark University, Omu-Aran, Nigeria
[3] Industry, Innovation and Infrastructure Research Group, Landmark University, Omu-Aran, Nigeria, SDG 9
[4] Climate Action Research Group, Landmark University, Omu-Aran, Nigeria, SDG 13
[5] Life on Land Research Group, Landmark University, Omu-Aran, Nigeria, SDG 15

Article Info

Article history:

Received Jul 27, 2021
Revised May 31, 2022
Accepted Jun 29, 2022

Keywords:

Data mining
Decision trees
K-nearest neighbors
Logistic regression
Normalization
Weather forecasting

ABSTRACT

Prediction of weather has been proved useful in the early warning on the impacts of weather on several areas of human livelihood. For example, the provision of decisions for autonomous transportation to reduce traffic congestion and accidents during the rainy season. However, providing the most accurate and effective forecasting model for weather forecasts has been a challenge. Hence, machine learning (ML) techniques and factors influencing weather prediction need to be investigated. Data scientists are yet to discover the best models for weather prediction. Therefore, this study compares three ML classification techniques for weather prediction. A web-based software application was developed using Flask App to demonstrate weather modeling using three ML models, and the data used for the study was obtained from Kaggle. For the weather prediction; a decision tree (DT), k-nearest neighbor (k-NN), and logistic regression (LR) classifier method were suggested, and comparisons were made between the three classifications techniques. The accuracy results show that with a 100% accuracy rate, the DL classifier outperforms the k-NN with a 78% accuracy rate and LR with a 93% accuracy rate. The results show that the application of ML models gives accurate results on weather prediction.

Corresponding Author:

Abidemi Emmanuel Adeniyi
Department of Computer Science, Landmark University
Omu-Aran, Kwara State, Nigeria
Email: adeniyi.emmanuel@lmu.edu.ng

1. INTRODUCTION

Nowadays, data mining is playing a vital role in weather prediction and climate change studies and it is one of the most important fields of research to extract useful knowledge from large datasets [1], [2]. Weather forecasting and climate change studies are significant research areas that aid in the prediction of useful information from weather data to enhance weather usability, better understanding, and more reliable prediction of weather-related activities. Data mining, also known as knowledge discovery, is becoming increasingly important because it aids in the analysis of data from various perspectives and the summarization of that data into useful information [3], [4]. Every data scientist is faced with the following fundamental questions: which predictive model is more appropriate for the problem at hand? What type of programming language to use? And what type of tools can be used for efficient outcomes. Python has several built-in machine learning libraries that provide a strong computational capability in data mining.

Weather forecasting is the use of science and technology to predict atmospheric conditions for a specific location and at a specific time. For centuries, people have attempted to predict the weather informally, and formally since the 19th century [5], [6]. Weather predictions are used by a wide range of users such as agriculture [7], [8]. Weather warnings are important predictions because they protect people and property. Forecasts dependent on temperature and precipitation, for example, are important for agriculture as well as commodity traders. People use weather forecasts daily to decide what to wear. Since heavy rain, high temperatures, snow, and wind chill greatly limit outdoor activities, forecasts can be used to schedule activities around these events, as well as to prepare for and withstand them. The reliance on weather forecasts by farmers, traders, transportation industries, and individual call for an accurate weather forecast. Data mining is a technique for training a computer on how to make data-driven decisions. This decision could be as simple as predicting the weather for tomorrow, blocking a spam email from accessing your inbox, detecting the language of a website, or discovering a new relationship on a dating website. Data mining has a wide range of applications, with new ones appearing regularly. Moreover, prediction is used to foretell the next event based on the current state of events. In intelligence environments, prediction is important because it detects the repetitive trend and predicts what will happen in the future. Modern farmers and business people need accurate real-time prediction tools to forecast the weather and climate change.

Data mining is widely used in weather and climate change studies to accurately forecast the weather. Weather forecasting applications tend to be complex to apply, inaccurate and unreliable as a result of the methods used in the application. Accurate weather prediction requires correct weather parameters, an efficient data mining approach, and apparatus. Weather prediction has become one of the most challenging scientific and technological problems in the last century all over the world [9]. The inaccuracy always calls for an alternative approach to address the issue because accurate weather prediction requires correct weather parameters, an efficient data mining approach, and apparatus. However, the proposed system employs three supervised learning algorithms, including logistic regression (LR), k-nearest neighbors (k-NN), and decision tree (DT) classification, with Spyder as the IDE adopting Python as the programming language to determine the most suitable technique for weather prediction. Data mining techniques may be used for weather forecasting and climate change studies if there is enough case data. Weather forecasting has historically been conducted traditionally with disturbances and uncertainty when calculating the initial atmospheric conditions. Table 1 provided the summary of the related works in weather and rainfall prediction.

Table 1. Related works

Authors	Research Topics	Techniques/Algorithms	Tools
Anton *et al.* [7]	Collaborative data mining in agriculture for prediction of soil moisture and temperature	k-nearest neighbor model (k-NN) local polynomial regression (LPR) neural net model (NN) and support vector machine (SVM)	The data analysis was conducted with the aid of a SQL command and a Microsoft Access database.
Shah *et al.* [10]	Rainfall prediction: accuracy enhancement using machine learning and forecasting techniques	Neural network Decision tree ARIMA	A text tool was used (MS Word as the formatting tool)
Zaman [11]	Machine learning model on rainfall - a predicted approach for Bangladesh	Classification algorithms (Naive Bayes, random forest classifier, and decision tree algorithm) Regression algorithm (linear regression, random forest regression)	Apache Spark was used for the machine learning library and scala as a development platform.
Shivang and Sridhar [12]	Weather Prediction for Indian location using machine learning	Linear regression Functional regression Neural network	Python (manual programming)
PosPieszny [13]	Application of data mining techniques in project management-an overview	C4.5 tree Random tree (RT) Classification and regression tree (CART)	Not indicated
Talib *et al.* [14]	Application of data mining techniques in weather data analysis	K-means clustering algorithm Decision tree algorithm	Waikato Environment for Knowledge Analysis (WEKA) data mining tools were used
Gouda and Chandrika [15]	Data mining for weather and climate studies	Classification Clustering	Not indicated
Joshi *et al.* [9]	Weather forecasting and climate changing using data mining application	Decision tree classifiers	Both decision trees and decision tree rules were created using the See5 program.
Olaiya and Adeyemo [16]	Application of data mining techniques in weather prediction and climate change Studies	Artificial neural network Decision tree algorithms	The C5 Decision Tree classifier algorithm has been implemented in See5.

Table 1. Related works *(continue)*

Authors	Research Topics	Techniques/ Algorithms	Tools
Cortez and Morais [17]	A Data mining approach to predict forest fires using meteorological data	Multiple regression (MR) Decision trees (DT) and Random forests (RF) Neural networks (NN) Support vector machines (SVM)	The open-source library Rainer (for the R statistical environment) was used.
Oladipo *et al.* [18]	Prediction and analysis of student performance by data mining in WEKA	Classification Association	WEKA tool was used as the software for data mining

Many other studies have contributed to the production of weather prediction systems that have yielded significant results and achieved the goals for which they were created; each study focuses on a specific challenge and investigates solutions from a unique perspective [19]–[21]. However, these techniques exhibit several defects such as; they may not completely experience from the existing weather data, and may not have the most accurate results. This study aims to compare the performance of three different classification machine learning algorithms in predicting weather and climate change for a specific location using Python because of its speed and libraries

Predictive modeling has gone through a revolution in recent years, thanks to advances in computer computing power [22]–[24]. Thousands of models can now be run on multiple cores at high GHz speeds, making predictive modeling more effective and affordable than ever before [25]. However, every data scientist is faced with the following fundamental questions: which predictive model is more appropriate for the problem at hand? What type of programming language to use? And what type of tools can be used for efficient outcomes. Since Python has several built-in machine learning libraries that provide a strong computational capability in data mining, this research contribution will be an adequate, efficient, and reliable resource for weather prediction applications.

2. RESEARCH METHOD

In this paper, the methodology used various classifiers to forecast weather. The daily meteorological data consisting of DATE, PRCP (which is the precipitation for that day), TMAX (which is the maximum temperature for that day), TMIN (which is the minimum temperature for that day), and RAIN of Seattle-Tacoma International Airport will be trained using Python inbuilt library, Scikit-learn library which consists of DT, LR and K-NN machine learning algorithms. These three machine learning algorithms will be used in this study to conduct classification rules and provide a model for weather forecasting. The target variable in this study is the Rainfall which is 1 (TRUE) if rain was observed on that day and 0 (FALSE) if it was not. Inconsistency and unreliability of meteorological factors make this research focus on the use of Python as the tool for analyzing the weather data to develop approaches that can recognize the best technique that can handle such an inconsistent pattern effectively. The main problem of this research work is to test the ability of some classification machine learning algorithms to predict weather data and to analyze, based on the data used, the accuracy of the machine learning algorithms.

2.1. Architecture of research methodology

Machine learning, also known as predictive modeling, is mainly concerned with reducing a model's error, and also, the most precise predictions are expected. Algorithm tuning is one of the most popular approaches to improving the accuracy of an algorithm. Machine learning algorithms are known to be influenced by parameters. These variables have a major impact on the learning process' outcome. Parameter tuning aims to find the best value for each parameter to improve the model's accuracy. To tune these parameters, you'll need to know what they mean and how they affect the model individually. To understand the main problem of the research work, it can be considered in the following two steps as shown in Figure 1.

2.1.1. Step 1

Weather data consists of complete records of daily rainfall pattern from January 1st, 1978 to December 12, 2017 (69 years) and constitutes parameters (25,551 rows x 5 columns) namely, DAY, MONTH, PRCP (inch), TMAX °F, TMIN °F, RAIN (TRUE or FALSE). The main problem of this part is to do data wrangling on the weather data. The variables have to be chosen in such a way that it includes the effect of rainfall. Then the problem comes to decide the weather variables which have a significant impact on the output variable. We need to decide various inputs and output parameters and check for anomalies using various Pandas library tools to check for missing values, performing data wrangling, and then process the data.

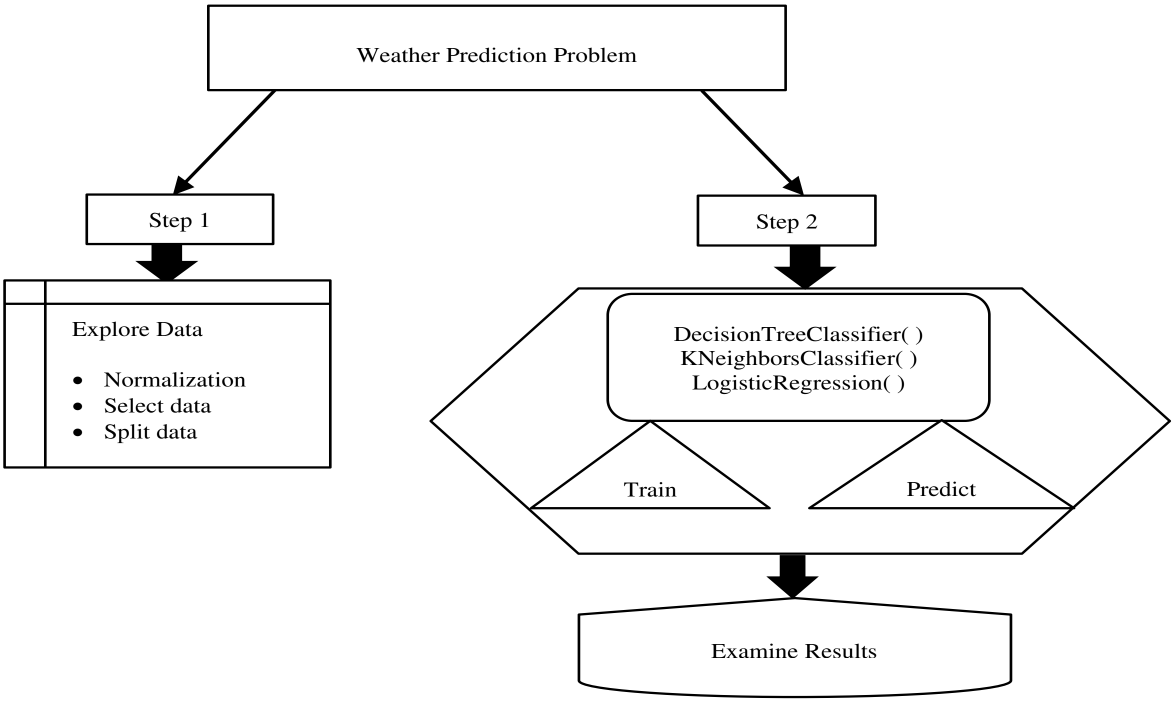


Figure 1. Methodology process

These steps are as: i) Import data: Using Pandas to import the data set, which is a two-dimensional structure in an arrangement of rows and columns; ii) Data normalization: This involves tasks such as removing and/or modifying duplicate data; iii) Data splitting: To split the data into training sets and testing sets using a higher percentage for the training sets; and iv)Data selection: This involves the process of dividing and selecting the dataset into two parts X and y, which are variables data and feature data respectively.

2.1.2. Step 2

The authors need to develop the model using the machine learning algorithms (DecisionTreeClassifier(), KneighborClassifier(), and LogisticRegression()) to train data and predict the output value of rain. The same data is used for accuracy check and error is calculated between the predicted and observed value of rain. Lastly, error evaluation and analysis visualization are done on each model to examine the results given. These steps are as: i) Data modelling: applying the machine learning algorithms which are DecisionTreeClassifier(), KneighborsClassifier() and LogisticRegression() that are provided in Scikit-learn library; ii) Model evaluation: This is the stage whereby the models are tested to measure their accuracy using error evaluation metrics such as accuracy score, and confusion matrix.

2.2 Data collection

The data collected contains the 49 years of weather data, from Seattle-Tacoma International Airport. It contains for each day the minimum temperature, maximum temperature, precipitation, and rainfall. The temperature is measured in Fahrenheit and precipitation is measured in inches. Rain contains Boolean values. The data is already in the .csv file. Table 2 shows the sample from the dataset.

In this study, to perform prediction, the weather dataset is split into two sets, input and output variables. Input data is loaded as X vector whose columns are DAY, MONTH, PRCP, TMAX, and TMIN, these are the independent variables and the target is loaded as y, which is the dependent variable containing the data of Rainfall (Boolean) as well-defined in Table 3. To predict the weather, we need to characterize the parameters which indirectly affect the rainfall. The variables of weather prediction are maximum temperature (TMAX ^{o}F) and minimum temperature (TMIM ^{o}F).

Figure 2 displayed the proposed model framework used for weather prediction. The proposed model was use for the prediction of weather using selected machine learning algorithms, the framework consists of the following: i) Pre-processing, ii) normalization, iii) prediction, and iv) evaluation of the proposed system using various metrics.

Table 2. Data sample

DATE	DAY	MONTH	PRCP	TMAX	TMIN	RAIN
01/01/2017	1	1	0.43	37	28	TRUE
02/01/2017	2	1	0	34	26	FALSE
03/01/2017	3	1	0	33	21	FALSE
04/01/2017	4	1	0	36	22	FALSE
05/01/2017	5	1	0	35	21	FALSE
06/01/2017	6	1	0	40	20	FALSE
07/01/2017	7	1	0	37	29	FALSE
08/01/2017	8	1	0.45	45	35	TRUE
09/01/2017	9	1	0.05	42	34	TRUE
10/01/2017	10	1	0.07	40	32	TRUE

Table 3. List of input and output variables

Variable	Weather Parameter	Type
X1	DAY (int)	Input
X2	MONTH (int)	Input
X3	PRCP (float)	Input
X4	TMAX (int)	Input
X5	TMIN (int)	Input
y1	Rainfall (int)	Target

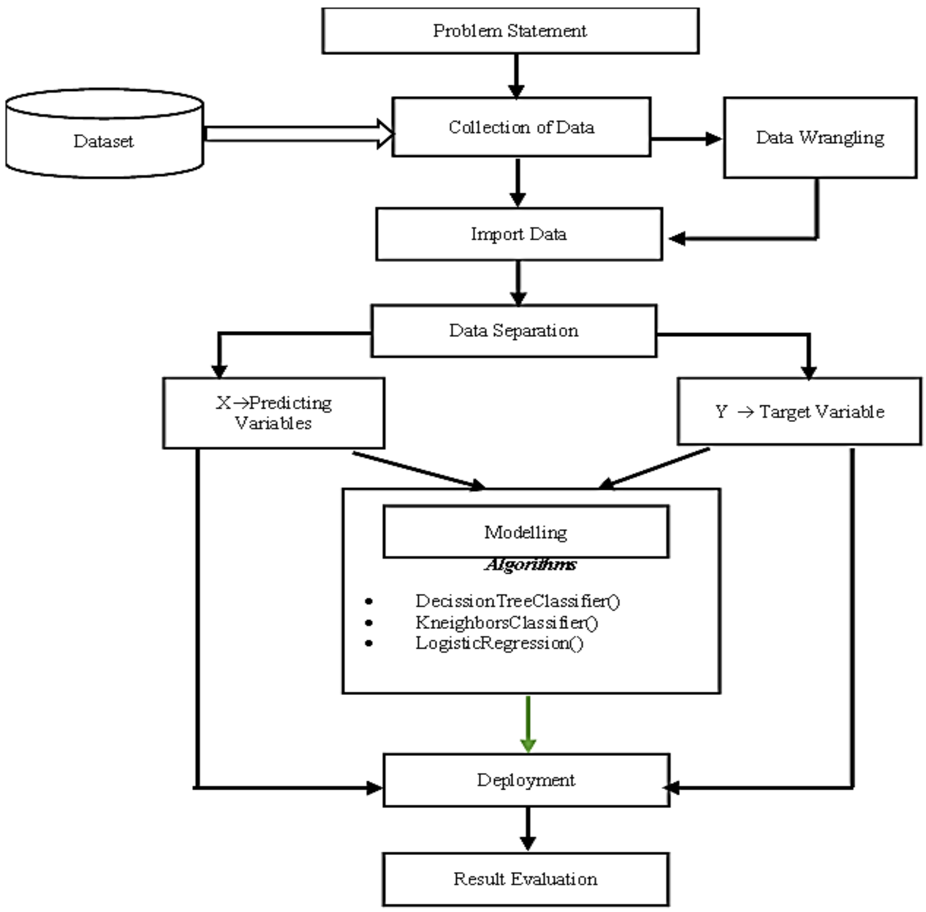


Figure 2. Framework model for weather forecast

Many researchers globally have been working on weather prediction using historical data and machine learning models in recent years to prevent weather prediction unreliability [26]–[28]. In this article, the study used classification algorithms for the weather prediction model to classify data for the target column used to predict rain in Seattle weather data. If the qualified model is for predicting any of two target classes, the main aim of the classification algorithm is to predict the target class (Yes/No). When used correctly, it will assist in predicting how a variable will appear in the future based on other variables. Spyder

IDE was used for data collection, model building, preparation, testing, and prediction. I had introduced various types of functionalities for common tasks during this process. The data are trained, tested, and saved the model to disk for live data testing in the future. Testing for those models is going to be difficult and inconvenient. As a result, the study decided to create a web application that would allow anyone to use the developed model to predict rainfall using a web browser.

3. RESULTS AND DISCUSSION

Building an efficient machine learning program necessitates evaluating a model. Model output is explained using evaluation criteria, which often aid in distinguishing between model outcomes. The study used a confusion matrix as well as four evaluation metrics: accuracy score, precision score, recall score, and F1-score, which are the most commonly used evaluation metrics for machine learning models. The following are the model evaluation and outcomes. This experiment's decision tree classifier resultant tree is shown in Figure 3. The top of the tree (known as the root node) has arrows pointing to and away from it, while intermediate nodes have arrows pointing to and away from them. Leaves are nodes at the bottom of a tree that do not have any edges pointing away from them.

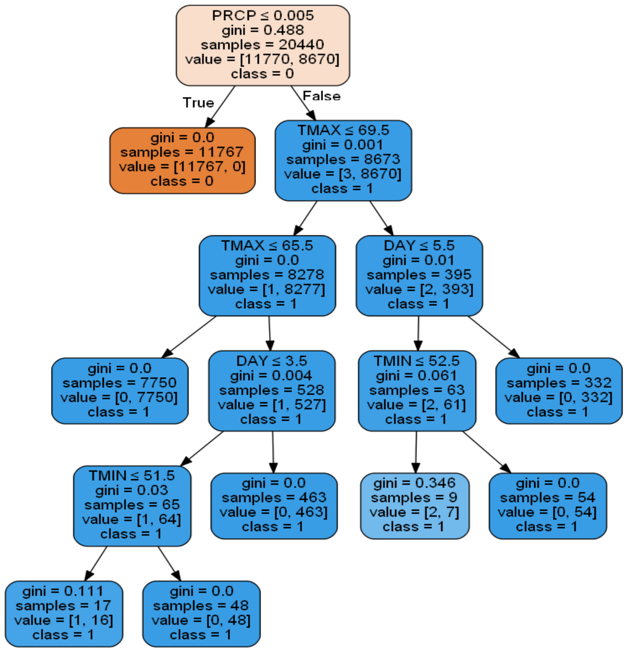
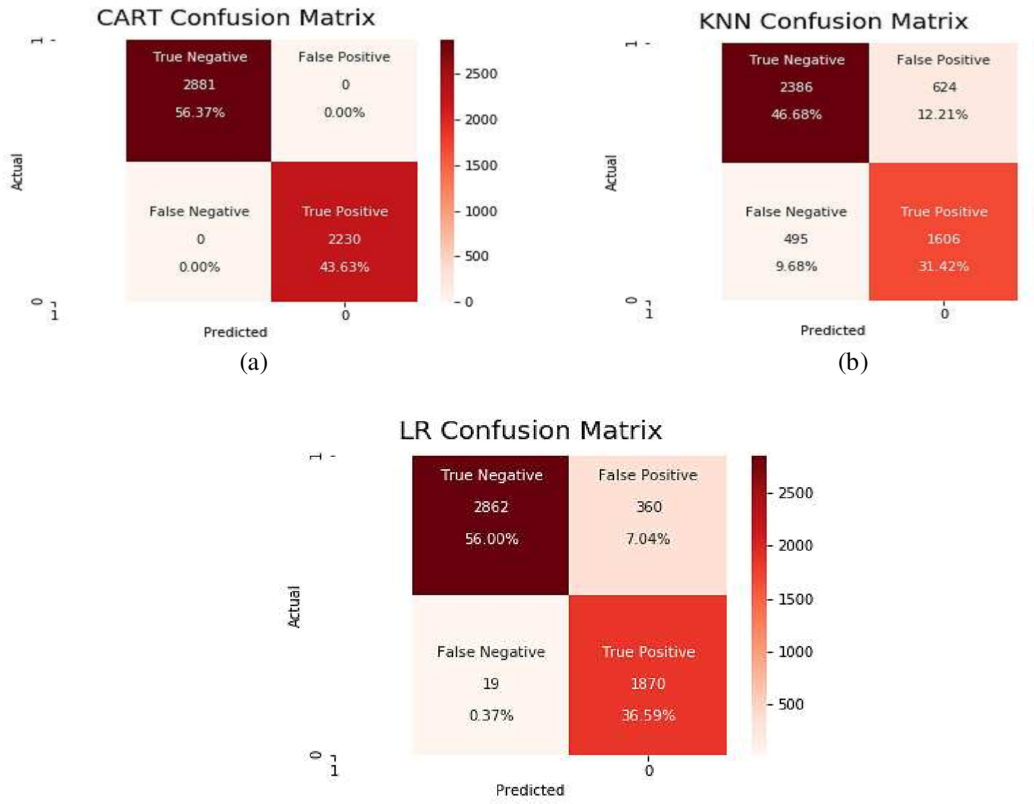


Figure 3. Decision tree classification

The experiment's decision tree classifier confusion matrix results are shown in Figure 4(a). The true positive (TP) of the entire test set in this decision tree classifier confusion matrix was 95.37% of the dataset belongs to the positive class, whereas 4.63% of data points belong to the negative category. The false positive (FP) and false negative (FN) rates are both 0%, indicating that the decision tree model did not classify any data incorrectly. The experiment's k-nearest neighbor confusion matrix results shown in Figure 4(b) reveals that the TP of the entire test set in this k-Neighbor classifier confusion matrix was 46.68%, indicating that the model correctly identified 46.68% of positive class data points. In addition, the true negative (TN) was 31.42%, indicating that the model correctly identified 31.42% of negative class data points. The FP was 12.21%, which means the model incorrectly classified 12.21% of negative class data points as positive, and the FN was 9.68%, which means the model incorrectly classified 9.68% of positive class points as negative.

The experiment's logistic regression confusion matrix results shown in Figure 4(c). The TP of the entire test set in this logistic regression classifier confusion matrix was 56.00%, indicating that the model correctly identified 56.00% of positive class data points. Furthermore, the TN was 36.59%, indicating that the model correctly identified 36.59% of negative class data points. The FP rate was 7.04%, indicating that the model incorrectly classified 7.04% of negative class data points as positive, and the FN rate was 19%, indicating that the model incorrectly classified 19% of positive class points as negative.



(a) (b)

(c)

Figure 4. Comparing the confusion matrix results for weather prediction in (a) CART, (b) K-NN, and (c) LR machine learning

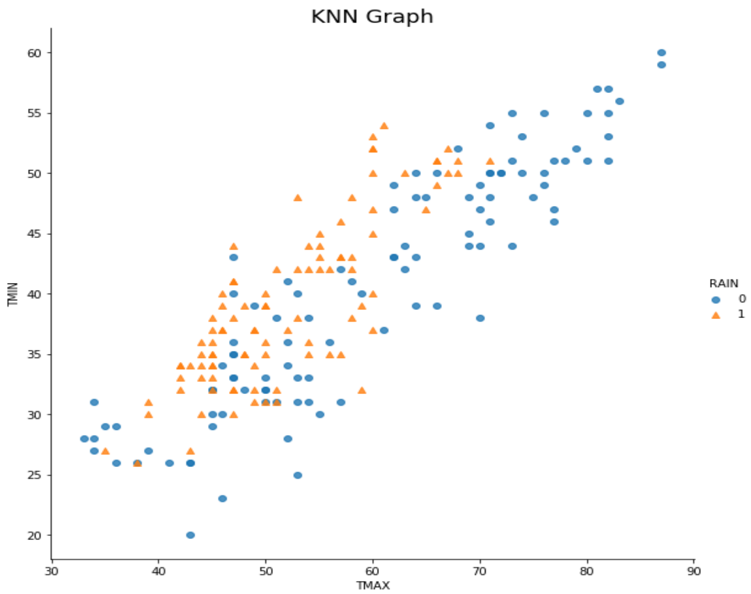
The K-NN algorithm predicts the values of new data points based on ‘feature similarity.' This implies that a value is given to the new point based on how similar it is to the points in the training set. It effectively measures the difference between a new data point and the training set's previous data points. It places the data point in the class that contains the majority of the K data points. The resultant classification model for this experiment k-nearest neighbor classification illustration is shown in Figure 5(a), and the logistic regression is displayed in Figure 5(b).

The dependent variable in logistic regression is a binary variable that contains data coded as 1 (yes, performance, etc.) or 0 (no, failure, etc.). To put it another way, the logistic regression model predicts P(Y=1) as a function of X and outputs a constant value. The TMAX represents the x-axis in this logistic regression graph, while the TMIN represents the y-axis. The blue circles represent observations that are categorized as zeros, indicating that rain is not occurring, while the orange circles represent observations that are classified as ones, indicating that rain is occurring. Table 4 displayed the model evaluation results using various metrics like accuracy, precision, recall, and F1-score.

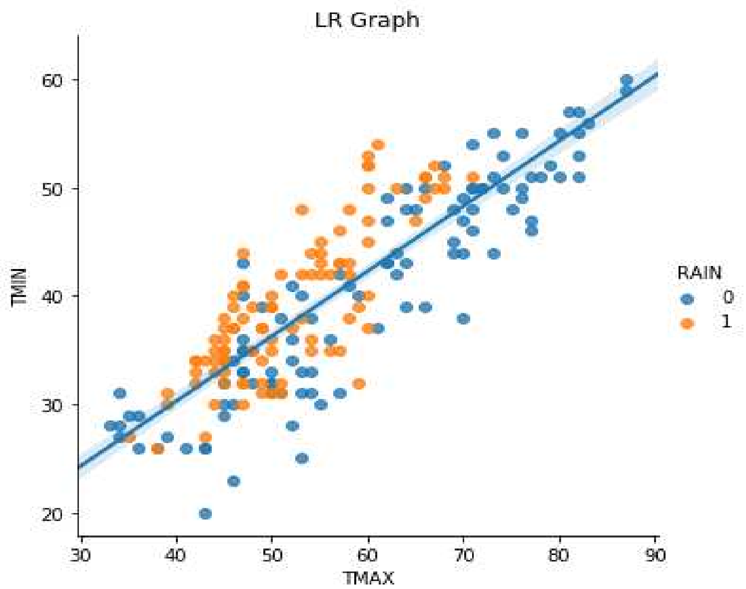
Table 4. Model evaluation (evaluation metrics)

MODELS	Accuracy	Precision	Recall	F1-Score
Decision Tree	100	1.0	1.0	1.0
k-Nearest Neighbor (k-NN)	78.10	0.76	0.72	0.74
Logistic Regression (LR)	93	0.98	0.83	0.90

The models were measured using four different evaluation metrics. The evaluation metrics allow for a quick assessment of the model's performance. To demonstrate how the models worked, the accuracy score, precision score, recall score, and F1-score evaluation metrics were used in this experiment. The results showed that the DT classifier performs better with an accuracy of 100%, and 100% all through other metrics used. According to the evaluation metrics, the DT classifier has the highest evaluation scores, followed by LR, and k-NN classifier.



(a)



(b)

Figure 5. Comparing scatter graph results for weather prediction in (a) k-nearest neighbor and (b) logistic regression

4. CONCLUSION

Improving a model's productivity can be challenging at times, and a predictive model can be built in various ways since there are no set rules to follow. The production of huge data from weather generated datasets have helped the prediction of weather forecast using machine learning (ML)-models. This has helped in the "proof to speak for itself" rather than relying on assumptions and bad associations. Since variables like humidity and wind speed influence the weather, providing more data leads to better and more accurate models. This intuition develops over time as a result of practice and experience. This is compounded by the fact that certain algorithms are better suited to some types of datasets than others. therefore, this paper predicts weather forecasts using three classifiers using the Seattle weather dataset to test the effectiveness of these three models. These models are validated using Kaggle's meteorological data, which includes, the date, maximum temperature, minimum temperature, precipitation, and rain. The decision tree algorithm, which has a 100% accuracy rate, outperforms the logistic regression (which has a 93% accuracy rate), and the k-nearest neighbor algorithm (which has a 78% accuracy rate). The study achieved an appropriate label of accuracy for the decision tree algorithm in terms of rain prediction. Future work can make use of the open works up by training the weather datasets using intuitive optimization on the datasets parameter this will result in better and more accurate models.

REFERENCES

[1] H. A. Issad, R. Aoudjit, and J. J. P. C. Rodrigues, "A comprehensive review of Data Mining techniques in smart agriculture," *Eng. Agric. Environ. Food*, vol. 12, no. 4, pp. 511–525, Oct. 2019, doi: 10.1016/j.eaef.2019.11.003.

[2] K. Soomro, M. N. M. Bhutta, Z. Khan, and M. A. Tahir, "Smart city big data analytics: An advanced review," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 9, no. 5, p. e1319, Jun. 2019, doi: 10.1002/widm.1319.

[3] Z. Ge, Z. Song, S. X. Ding, and B. Huang, "Data mining and analytics in the process industry: The role of machine learning," *IEEE Access*, vol. 5, pp. 20590–20616, 2017, doi: 10.1109/access.2017.2756872.

[4] A. M. Abdu, M. M. M. Mokji, and U. U. U. Sheikh, "Machine learning for plant disease detection: an investigative comparison between support vector machine and deep learning," *IAES Int. J. Artif. Intell.*, vol. 9, no. 4, p. 670, Dec. 2020, doi: 10.11591/ijai.v9.i4.pp670-683.

[5] P. Steer, "The climates of the victorian novel: seasonality, weather, and regional fiction in Britain and Australia," *PMLA/Publications Mod. Lang. Assoc. Am.*, vol. 136, no. 3, pp. 370–385, May 2021, doi: 10.1632/s0030812921000286.

[6] I. Loor and J. Evans, "Understanding the value and vulnerability of informal infrastructures: Footpaths in Quito," *J. Transp. Geogr.*, vol. 94, p. 103112, Jun. 2021, doi: 10.1016/j.jtrangeo.2021.103112.

[7] C. A. Anton, O. Matei, and A. Avram, "Collaborative data mining in agriculture for prediction of soil moisture and temperature," in *Computer Science On-Line Conference*, Springer International Publishing, 2019, pp. 141–151.

[8] B. Das, B. Nair, V. K. Reddy, and P. Venkatesh, "Evaluation of multiple linear, neural network and penalised regression models for prediction of rice yield based on weather parameters for west coast of India," *Int. J. Biometeorol.*, vol. 62, no. 10, pp. 1809–1822, Jul. 2018, doi: 10.1007/s00484-018-1583-6.

[9] A. Joshi, B. Kamble, V. Joshi, K. Kajale, and N. Dhange, "Weather forecasting and climate changing using data mining application," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 4, no. 3, pp. 19–21, Mar. 2015, doi: 10.17148/ijarcce.2015.4305.

[10] U. Shah, S. Garg, N. Sisodiya, N. Dube, and S. Sharma, "Rainfall prediction: accuracy enhancement using machine learning and forecasting techniques," in *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Dec. 2018, pp. 776–782, doi: 10.1109/pdgc.2018.8745763.

[11] Y. Zaman, "Machine learning model on rainfall-a predicted approach for Bangladesh," United International University, 2018.

[12] J. Shivang and S. S. Sridhar, "Weather prediction for indian location using Machine learning," *Int. J. Pure Appl. Math.*, vol. 118, no. 22, pp. 1945–1949, 2018.

[13] P. PosPieszny, "Application of data mining techniques in project management–an overview," *CEA Ann.*, vol. 43, pp. 199–220, 2017, [Online]. Available: http://rocznikikae.sgh.waw.pl/p/roczniki_kae_z43_12.pdf.

[14] M. R. Talib, T. Ullah, M. U. Sarwar, M. K. Hanif, and N. Ayub, "Application of data mining techniques in weather data analysis," *Int. J. Comput. Sci. Netw. Secur.*, vol. 17, no. 6, pp. 22–28, 2017, [Online]. Available: http://paper.ijcsns.org/07_book/201706/20170604.pdf.

[15] K. C. Gouda and and Chandrika M, "Data mining for weather and climate studies," *Int. J. Eng. Trends Technol.*, vol. 32, no. 1, pp. 29–32, Feb. 2016, doi: 10.14445/22315381/ijett-v32p206.

[16] F. Olaiya and A. B. Adeyemo, "Application of data mining techniques in weather prediction and climate change studies," *Int. J. Inf. Eng. Electron. Bus.*, vol. 4, no. 1, pp. 51–59, Feb. 2012, doi: 10.5815/ijieeb.2012.01.07.

[17] P. Cortez and A. de J. R. Morais, "A data mining approach to predict forest fires using meteorological data," *Environmental Science, Computer Science*. Sep. 2007, [Online]. Available: http://www3.dsi.uminho.pt/pcortez/fires.pdf.

[18] I. D. Oladipo *et al.*, "An improved course recommendation system based on historical grade data using logistic regression," in *Communications in Computer and Information Science*, Springer International Publishing, 2021, pp. 207–221.

[19] D. F. Sengkey, A. Jacobus, and F. J. Manoppo, "Effects of kernels and the proportion of training data on the accuracy of SVM sentiment analysis in lecturer evaluation," *IAES Int. J. Artif. Intell.*, vol. 9, no. 4, p. 734, Dec. 2020, doi: 10.11591/ijai.v9.i4.pp734-743.

[20] H. Hertina *et al.*, "Data mining applied about polygamy using sentiment analysis on twitters in Indonesian perception," *Bull. Electr. Eng. Informatics*, vol. 10, no. 4, pp. 2231–2236, Aug. 2021, doi: 10.11591/eei.v10i4.2325.

[21] N. Razali, S. Ismail, and A. Mustapha, "Machine learning approach for flood risks prediction," *IAES Int. J. Artif. Intell.*, vol. 9, no. 1, p. 73, Mar. 2020, doi: 10.11591/ijai.v9.i1.pp73-80.

[22] M. A. Febriantono, S. H. Pramono, R. Rahmadwati, and G. Naghdy, "Classification of multiclass imbalanced data using cost-sensitive decision tree C5.0," *IAES Int. J. Artif. Intell.*, vol. 9, no. 1, p. 65, Mar. 2020, doi: 10.11591/ijai.v9.i1.pp65-72.

[23] M. Y. I. Basheer, S. Mutalib, N. H. A. Hamid, S. Abdul-Rahman, and A. M. A. Malik, "Predictive analytics of university student intake using supervised methods," *IAES Int. J. Artif. Intell.*, vol. 8, no. 4, p. 367, Dec. 2019, doi: 10.11591/ijai.v8.i4.pp367-374.

[24] J. B. Awotunde, C. Chakraborty, and A. E. Adeniyi, "Intrusion detection in industrial internet of things network-based on deep learning model with rule-based feature selection," *Wirel. Commun. Mob. Comput.*, vol. 2021, Sep. 2021, doi: 10.1155/2021/7154587.

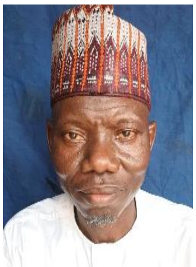
[25] L. Tiwari *et al.*, "Detection of lung nodule and cancer using novel Mask-3 FCM and TWEDLNN algorithms," *Measurement*, vol. 172, p. 108882, Feb. 2021, doi: 10.1016/j.measurement.2020.108882.

[26] J. B. Awotunde, S. O. Folorunso, A. K. Bhoi, P. O. Adebayo, and M. F. Ijaz, "Disease diagnosis system for IoT-based wearable body sensors with machine learning algorithm," in *Intelligent Systems Reference Library*, vol. 209, Springer Singapore, 2021, pp. 201–222.

[27] S. Susanto, D. D. Budiarjo, A. Hendrawan, and P. T. Pungkasanti, "The implementation of intelligent systems in automating vehicle detection on the road," *IAES Int. J. Artif. Intell.*, vol. 10, no. 3, p. 571, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp571-575.

[28] X. Ren *et al.*, "Deep learning-based weather prediction: A survey," *Big Data Res.*, vol. 23, p. 100178, Feb. 2021, doi: 10.1016/j.bdr.2020.100178.

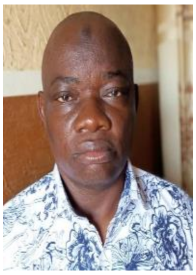
BIOGRAPHIES OF AUTHORS



Dr. Muyideen AbdulRaheem is a Lecturer in the Department of Computer Science, University of Ilorin, Ilorin, Nigeria. He obtained a Bachelor of Technology Computer Science at the Prestigious Abubakar Tafawa Balewa University Bauchi, in 1997, Master of Science in Mathematics (Computer Option) and Ph.D. Computer Science at the University of Ilorin, Ilorin, Nigeria. His research area is on computer security, grid computing security and data and information security. He can be contacted at: email: muyideen@unilorin.edu.ng.

Joseph Bamidele Awotunde received his B.Sc. degree in Mathematics/ Computer Science from Federal University of Technology, Minna, Nigeria, in 2007. M.Sc. and Ph.D. degrees in Computer Science from the University of Ilorin, Ilorin, Nigeria, in 2014 and 2019 respectively. From 2012 to 2015, and 2018, he was a Computer Science Instructor with the University School, University of Ilorin, Ilorin, Nigeria. From 2017 to 2018, he was a Lecturer II with the McPherson University, Ijebo, Seriki-Sotayo, Nigeria. Since 2019, he has been a Lecturer II with the Computer Science Department, University, of Ilorin, Ilorin, Nigeria. He is the author of more than 40 articles, 30 chapter in book, and more than 25 Conference Proceedings. His research interests include Information Security, Cybersecurity, Bioinformatics Artificial Intelligence, Internet of Medical Things, Wireless Body Sensor Networks, Wireless Networks, Telemedicine, m-Health/e-health, and Medical Imaging, Information Security, Software Engineering, and Biometrics. He is a member of International Association of Engineers and Computer Scientist (MIAENG), Computer Professional Registration Council of Nigeria (MCPN), Nigeria Computer Society (MNCS), and Internet Society. He can be contacted at email: awotunde.jb@unilorin.edu.ng, jabonnetbylinks@gmail.com.

Abidemi Emmanuel Adeniyi is a Staff in the Department of Computer Science, College of Pure and Applied Sciences, Landmark University, Omu-Aran, Nigeria. He received Bachelor of Science in Computer Science from Ladoke Akintola University of Technology, Ogbomosho. He also begged Masters of Science in Computer Science from Landmark University, Omu-Aran. His area of research includes computer security, Computational Complexity Analysis, Cryptographic Algorithms and Information Security. He can be contacted at email: adeniyi.emmanuel@lmu.edu.ng.



Idowu Dauda Oladipo received his B.Sc. (Edu.) degree in Computer Science from Ekiti State University, Ado-Ekiti, Ekiti State, Nigeria in 2005. He earned his M.Sc. and Ph.D. degree in Computer Science from University of Ilorin, Ilorin, Nigeria, in 2010 and 2018 respectively. Since 2019, he has been a Lecturer II with the Department of Computer Science, University of Ilorin, Ilorin, Nigeria. He is the author of more than 20 articles, and more than 10 conference Proceedings. His research interests include Software Engineering, Bioinformatics, Information Security, Artificial Intelligence, Cyber Security and Computer Education. He is a member of International Computer Professional Registration Council of Nigeria (MCPN) and Nigeria Computer Society (MNCS). He can be contacted at email: odidowu@unilorin.edu.ng, idoladipo@gmail.com.

Sekinat Olaide Adekola received her B.Sc in Computer Science from University of Ilorin, Ilorin. She received her ordinary National Diploma, OND, in computer science from Osun State Polytechnic, Iree, Nigeria. She enrolled from the National Institute of Information Technology, NIIT, in Ajah, Lagos State, Nigeria, with a Microsoft certification. She can be contacted at email: sekinat0648@gmail.com.