Predictive analytics on crop yield using supervised learning techniques

Julius Olatunji Okesola[1,2], Olaniyi Ifeoluwa[1], Sunday Adeola Ajagbe[3,4], Olubunmi Okesola[5], Adeyinka O. Abiodun[2,6], Francis Bukie Osang[2,6], Olakunle O. Solanke[7]

[1]Department Computer Sciences, Abiola Ajimobi Technical University, Ibadan, Nigeria
[2]Africa Centre of Excellence on Technology Enhanced Learning (ACETEL), Abuja, Nigeria
[3]Department of Computer Engineering, Abiola Ajimobi Technical University, Ibadan, Nigeria
[4]Department of Computer Science, University of Zululand, Kwadlangezwa, South Africa
[5]Library and Information Sciences, University of Ibadan, Ibadan, Nigeria
[6]Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria
[7]Department of Mathematical Sciences, Olabibisi Onabanjo University, Ago-Iwoye, Nigeria

Article Info

Article history:

Received Dec 15, 2023
Revised Aug 13, 2024
Accepted Aug 26, 2024

Keywords:

Crop yield
Crop yield prediction
Machine learning algorithms
Predictive analytics
Supervised learning techniques

ABSTRACT

Agriculture is one of Nigeria's most important economic activities but with climate change is a threat to crop production and a significant impact on the national economy as unforeseen scenarios can cause a drop in crop yield. Machine learning algorithms are now being considered as decision support tools for crop yields prediction and weather forecasting. Maize is the crop selected in this study, and a stochastic gradient model of five popular regression algorithms was evaluated. The prediction system is written in Python programming language and linked to a web-based interface for ease of use and effectiveness. Using performance metrics, the result shows that stochastic gradient descent (SGD) performed best with lower error rates and better R2_score value of 0.98505036. This crop yield prediction system (CYPS) is able to predict the yield of the crop which will help farmers and analysts in decision-making. It will also help industries that make use of the agricultural product in strategizing the logistics of their business.

Corresponding Author:

Sunday Adeola Ajagbe
Department of Computer Engineering, Abiola Ajimobi Technical University
Ibadan, 200255, Nigeria
Email: saajagbe@pgschool.lautech.edu.ng

1. INTRODUCTION

Agriculture is one of Nigeria's most important and largest economic activities with a significant impact on national development. The favorable wide range of climatic variations has made the country a leader in the production of various types of agricultural products such as maize, yam, palm oil, pineapple, cocoa, cassava, millet, and so on. The objective of agricultural production is to achieve maximum crop yield [1], [2]. The amount of agricultural production harvested per unit of land area is referred to as crop yield [3]-[5]. Sometimes, it is also referred to as "agricultural output". The crop yield which is the primary objective of agricultural production is affected by some notable climatic conditions including rainfall, temperature, humidity, and land factors such as soil pH, soil type. Climate-driven crop yield, yield variability and climate change impact studied [6] suggested that weather and climatic factors are the prominent drivers of crop yield. Hence, certain features like rainfall and temperature that have significant effects on crop yields are affected once climatic changes tend towards the unfavorable side.

The study and management of complicated scenarios like crop yield prediction can therefore help to render a larger return yield and promote profitability. Since machine learning is regarded as an effective tool [7], [8] for crop yield prediction [9] as it is used to determine patterns, correlations and knowledge from datasets [10]. This research extends the study of [11] on artificial intelligence (AI) aiming at filling the gaps that unforeseen scenarios like climate change could cause to crop yield. The crop selected for this study is maize, and the supervised machine learning algorithms random forest (RF), stochastic gradient descent (SGD), or AdaBoost regressor are engaged to predict the crop yield, and develop a system for farmers to have an insight into the future and know whether the crops he wants to plant will yield well or not. The research contributions: i) evaluate the performance of different algorithms to determine the best; and ii) develop a prediction system using the best-determined algorithm.

The remainder of the article is structured as follows: section 2 examines the earlier research and analysis. The suggested scheme’s methodology is shown in section 3 for the crop yield prediction system (CYPS). In section 4 presents the implementation, conclusions, and discussion. In section 5 of the study presents its findings.

2. LITERATURE REVIEW

Kaur [12] evaluated the applications areas in the field of agriculture including forecasting, smart irrigation system, crop selection, storage systems, and crop disease prediction to agree with [13] that, crop yield prediction is a critical responsibility of decision-makers, experts, and farmers at the national and regional levels. Hence, maize was as a control to figure out the impact of weather on agricultural produce and admitted that agricultural yields are truly susceptible to extreme weather and that changes in the mean and extreme weather pose a significant danger to governments and organizations [14], [15].

Winter Wheat for instance, is particularly vulnerable to low temperatures (freezing) in the fall, as well as to heat stress during grain filling and stem elongation [16]. This vulnerability to severe temperatures is therefore assumed to be the declining cause of wheat yields throughout Europe [3]. Meanwhile, the factors influencing Fall Armyworm damage on the African maize field and its quantifying impacts have been widely studied by many authors including [3], [16] to conclude that the Fall Armyworm causes substantial damage to maize and some other crops, and that the potential damages may be greatly minimized by regular weeding. Towards improving productivity through crop yield prediction therefore, many researchers including [17] are recently focusing on machine learning algorithms and their uses.

Veenadhari *et al.* [9] and Moraye *et al.* [18] specifically employed most influencing climatic parameters on crop yield to train C4.5 algorithm and demonstrated the use of data mining techniques in predicting crop yields. Their results showed that machine learning techniques have better skills in crop yield predicting compared to the principal regression. Similarly, Adebiyi *et al.* [19] discussed the optimization of farmland and monitoring of crops by developing a prediction system using a machine learning algorithm to analyze and classify dataset containing some parameters related to the yield of crops. This mobile application guarantees farmers instant information and services needed in their farmland. Jeong *et al.* [20] also investigated the influence of the performance of the RF algorithm on the predicting of the yield of wheat, maize, and potato crop to submit that the RF algorithm is an efficient tool to predict crop yield. This agrees with the findings of Sangeeta and Shruthi [21] where the performance of RF, polynomial regression (PR), and decision tree (DT) were evaluated and RF was adjudged the best algorithm for crop prediction based on the accuracy of the testing and training results, lesser processing time, and better performance even when handling large amount of data.

Meanwhile, some authors have been employing performance metrics to evaluate the machine learning algorithms in other to put some speculations about the accuracy of the test results to rest. For instance, Shah *et al.* [22] evaluated multivariate polynomial regression (MPR), support vector machine regression (SVM), and RF base on four performance metrics-root mean square error (RMSE), mean absolute error (MAE), median absolute error (MdAE), and R-squared values. Gonzalez-Sanchez *et al.* [23] also compared the predicting power of various machine learning algorithms making use of performance metrics RMSE, RRSE, and MAE to conclude that M5-Prime was the best with the largest number of crop yields and lowest error rate. Some of these commonly used performance metrics and their mathematical expressions are as depicted in Table 1.

3. METHOD

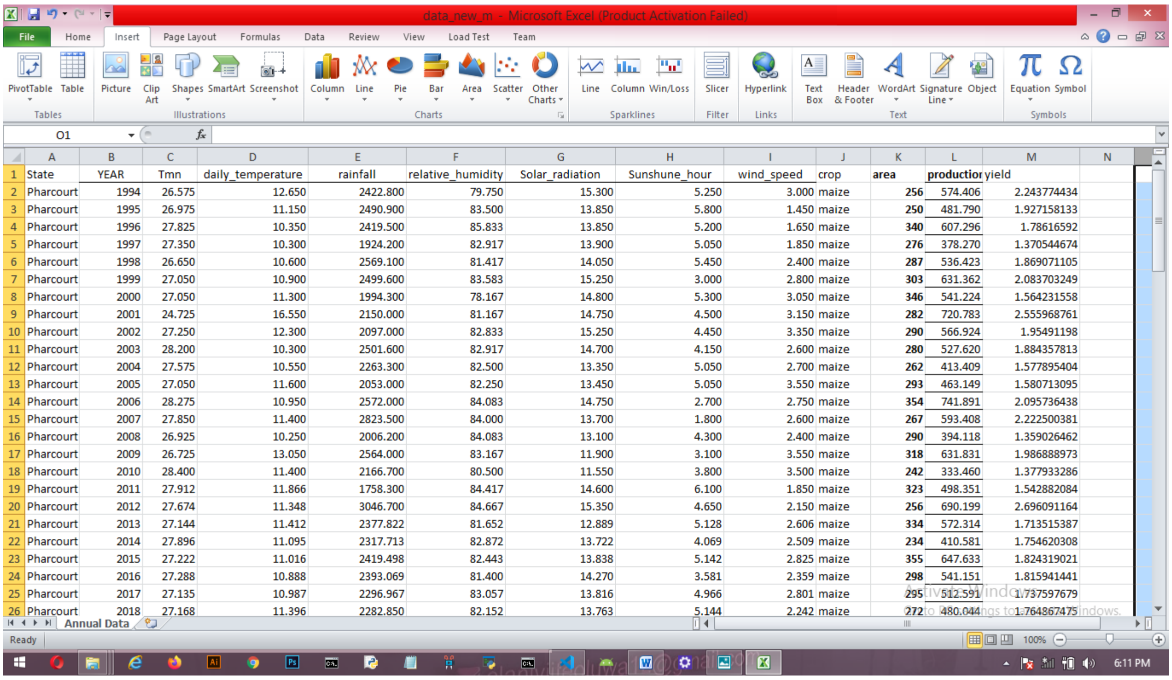
Using labeled data sets to train algorithms that reliably identify data or predict outcomes. The supervised learning approach is used in this work. This is a learning process that converts known input into output.

Table 1. Performance metrics and their mathematical expressions

s/n	Performance metrics	Expression
1	Root-relative square error (RRSE)	$\sqrt{\frac{\sum_{i-1}^{n}(y_i + \hat{y}_i)^2}{\sum_{i-1}^{n}(y_i - \bar{y})^2}}\ .100$
2	RMSE	$\sqrt{\frac{\sum_{i-1}^{n}(y_i + \hat{y}_{\bar{y}})^2}{n}}$
3	MAE	$\left(\frac{\sum_{i-1}^{n}\lvert y_i - \hat{y}_i\rvert}{(n)(\bar{y})}\right).100$
4	MdAE	
5	R-squared values	$\frac{\sum_{i-1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i-1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i-1}^{n}(\hat{y}_i - \bar{\hat{y}}_i)^2}}$

3.1. Dataset

The entire dataset used to implement this prediction system is live on dataworld.org https://data.world/smithcalvin/nigeria-maize-yield/workspace/file?filename=data.xlsx and consists of many features including [24]. Region name, crop year, area (in hectares), yield (in tons), rainfall, relative humidity, and solar radiation as shown in Figure 1. Historical data of these parameters are saved in a file, and divided into two parts a part (80%) of the dataset was used for training the model and the other part (20%) was set aside for testing the model.



State	YEAR	Tmn	daily_temperature	rainfall	relative_humidity	Solar_radiation	Sunshune_hour	wind_speed	crop	area	productior	yield
Pharcourt	1994	26.575	12.650	2422.800	79.750	15.300	5.250	3.000	maize	256	574.406	2.243774434
Pharcourt	1995	26.975	11.150	2490.900	83.500	13.850	5.800	1.450	maize	250	481.790	1.927158133
Pharcourt	1996	27.825	10.350	2419.500	85.833	13.850	5.200	1.650	maize	340	607.296	1.78616592
Pharcourt	1997	27.350	10.300	1924.200	82.917	13.900	5.050	1.850	maize	276	378.270	1.370544674
Pharcourt	1998	26.650	10.600	2569.100	81.417	14.050	5.450	2.400	maize	287	536.423	1.869071105
Pharcourt	1999	27.050	10.900	2499.600	83.583	15.250	3.000	2.800	maize	303	631.362	2.083703249
Pharcourt	2000	27.050	11.300	1994.300	78.167	14.800	5.300	3.050	maize	346	541.224	1.564231558
Pharcourt	2001	24.725	16.550	2150.000	81.167	14.750	4.500	3.150	maize	282	720.783	2.555968761
Pharcourt	2002	27.250	12.300	2097.000	82.833	15.250	4.450	3.350	maize	290	566.924	1.95491198
Pharcourt	2003	28.200	10.300	2501.600	82.917	14.700	4.150	2.600	maize	280	527.620	1.884357813
Pharcourt	2004	27.575	10.550	2263.300	82.500	13.350	5.050	2.700	maize	262	413.409	1.577895404
Pharcourt	2005	27.050	11.600	2053.000	82.250	13.450	5.050	3.550	maize	293	463.149	1.580713095
Pharcourt	2006	28.275	10.950	2572.000	84.083	14.750	2.700	2.750	maize	354	741.891	2.095736438
Pharcourt	2007	27.850	11.400	2823.500	84.000	13.700	1.800	2.600	maize	267	593.408	2.222500381
Pharcourt	2008	26.925	10.250	2006.200	84.083	13.100	4.300	2.400	maize	290	394.118	1.359026462
Pharcourt	2009	26.725	13.050	2564.000	83.167	11.900	3.100	3.550	maize	318	631.831	1.986888973
Pharcourt	2010	28.400	11.400	2166.700	80.500	11.550	3.800	3.500	maize	242	333.460	1.377933286
Pharcourt	2011	27.912	11.866	1758.300	84.417	14.600	6.100	1.850	maize	323	498.351	1.542882084
Pharcourt	2012	27.674	11.348	3046.700	84.667	15.350	4.650	2.150	maize	256	690.199	2.696091164
Pharcourt	2013	27.144	11.412	2377.822	81.652	12.889	5.128	2.606	maize	334	572.314	1.713515387
Pharcourt	2014	27.896	11.095	2317.713	82.872	13.722	4.069	2.509	maize	234	410.581	1.754620308
Pharcourt	2015	27.222	11.016	2419.498	82.443	13.838	5.142	2.825	maize	355	647.633	1.824319021
Pharcourt	2016	27.288	10.888	2393.069	81.400	14.270	3.581	2.359	maize	298	541.151	1.815941441
Pharcourt	2017	27.135	10.987	2296.967	83.057	13.816	4.966	2.801	maize	295	512.591	1.737597679
Pharcourt	2018	27.168	11.396	2282.850	82.152	13.763	5.144	2.242	maize	272	480.044	1.764867475

Figure 1. Dataset

3.2. Machine learning algorithms

The problem to be solved in this study required a regression technique which is a modeling task that involves predicting a numeric value by given input. The following algorithms are therefore considered based on their individual qualities as specified.

- RF: is the most popular and powerful supervised machine learning algorithm and it is capable of solving both classification and regression problems [25]. Overfitting (which occurs when there are so many false positives) of the training set is not an issue [26].
- AdaBoost regressor: this is a boosting approach used in machine learning as an ensemble method and helps to capture various non-linear correlations resulting in improved prediction accuracy on the problem of interest [27].

- Extra tree regressor: this is a trees regression system with extremely randomized trees [27] that involves heavily randomizing both attribute and cut-point selections.
- SGD: SDG is a quick and easy way to fit linear classifiers and regressors (SVM, LR) to convex loss functions.
- Linear regression: LR is a model for determining the connection between input and output numerical variables that was established in the field of statistics [27].

3.3. System architecture

The system architecture is represented in Figure 2 and defines the conceptual model of the system in multiple views and structures. Figure 2 above shows the architectural design of the proposed system for the project. The above architecture clearly explains the processes involve in achieving the crop yield and how all the components of the system communicate with one another, starting from data input to result. Following the processes contained by the architecture, the crop yield is being predicted by the proposed system. This architecture displayed user connection to the system and shows clearly how data is captured, the data is then preprocessed to remove every unwanted data such as NULL, and unwanted features. After preprocessing, the dataset is then divided into two. A part (80%) of the dataset as the training set and the other part (20%) as the testing set. The training set is used to train the model and the testing set to test the model.

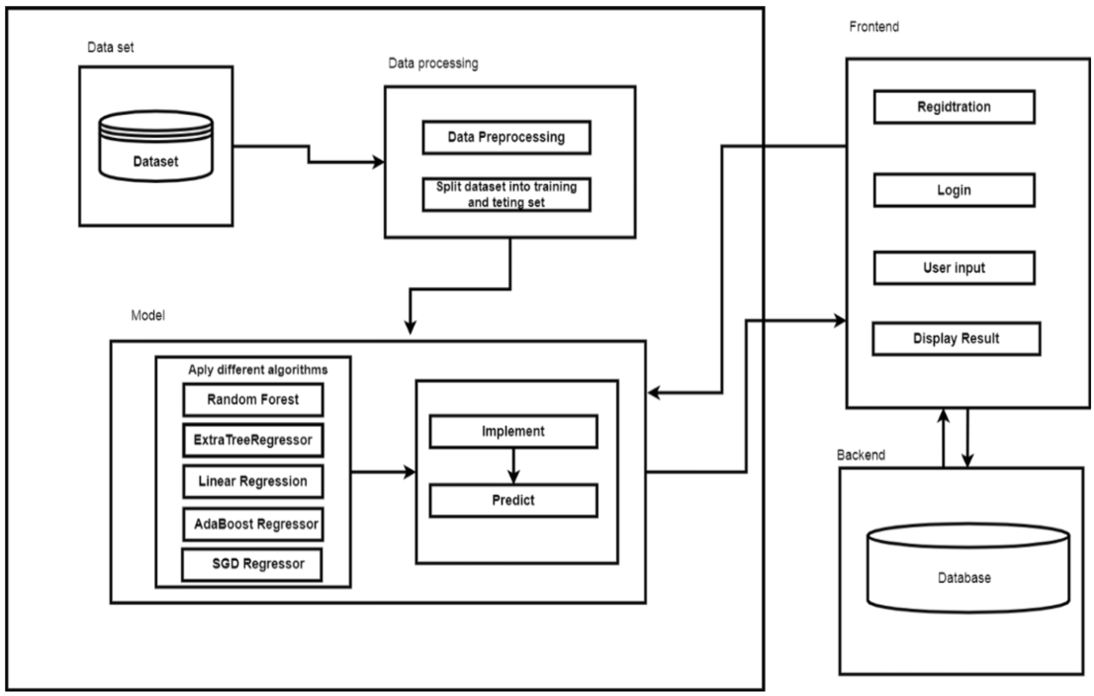


Figure 2. System architecture

Then different machine learning algorithms are applied to build models, the models' performances are then evaluated using different performance metrics. The best performed model is then passed to the implementation stage. The implementation phase contains the model of the best performed model and it takes inputs from the frontend which are processed within the phase. Yield output is being sent back to the frontend. The frontend phase is connected to a database, which allow for user registration and authentication.

4. IMPLEMENTATION, REESULT, AND DISCUSSION

The prediction system was implemented using python programming language. Jupyter platform was used for simulation. The dataset was preprocessed to remove all unwanted parameters, null variables, and also to convert string variables to numbers. The graphs below are to show the distributions, correlations, and relationships between some of the important parameters that mostly contributed to the prediction.

Figure 3 depict the Lineplot of rainfall against production, it is a line plot of rainfall against production. The chart shows that an increase in rainfall leads to an increase in production. This means that rainfall is a crucial parameter in crop yield. Figure 4, represents the bar chart of temperature against region. The chart shows that Kaduna has the highest average temperature, followed by Kano, Portharcourt, and Ibadan respectively.

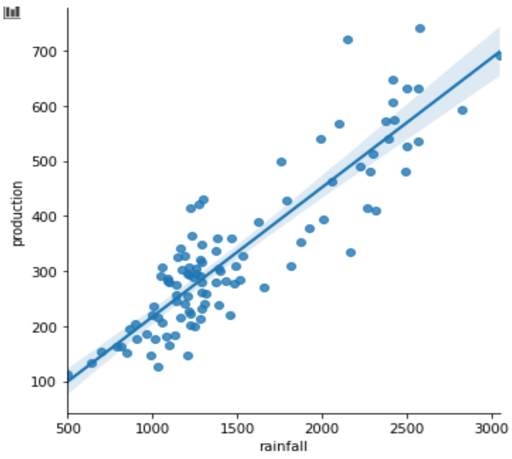


Figure 3. Line plot of rainfall against production

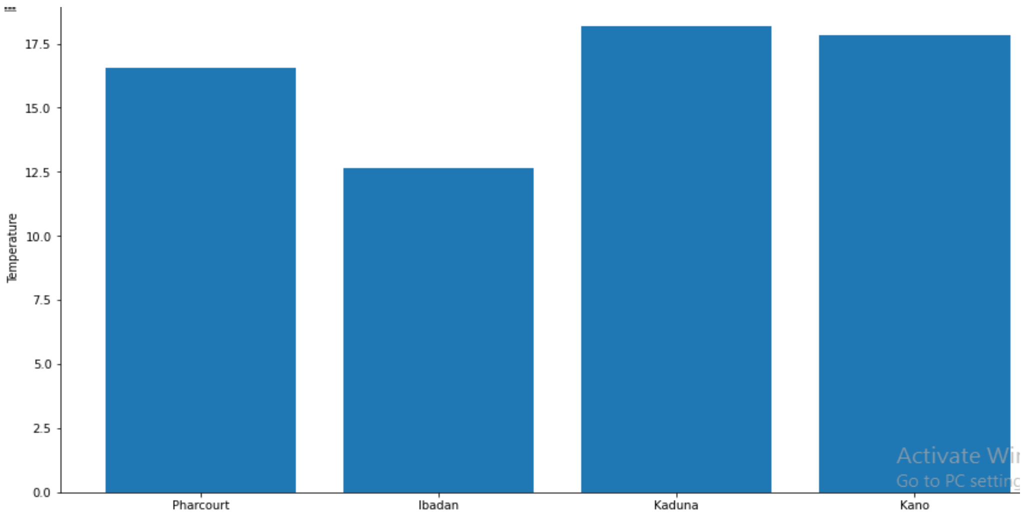


Figure 4. Bar chart of region against temperature

4.1. Training set and testing set

The data set was divided into two 80% of the data set for training data and the remaining 20% for testing. The algorithm for splitting the dataset is depicted in Figure 5. For the x-axis, the training set has a total of 83 rows and 8 columns while the testing set in total has 21 rows and 8 columns. Model development is shown in Figure 6.

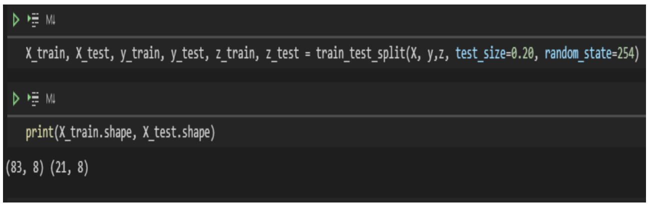


Figure 5. Splitting dataset

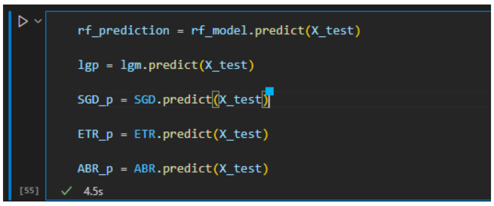


Figure 6. Models' development

4.2. Performance evaluation on regression algorithms

This problem requires regression technique, which is an unsupervised learning method. Various regression algorithms were used to build models. Some of these algorithms were used in existing literature. Five regression algorithms were used for the study. These algorithms are RF, SGD, ExtraTree regressor (ETR), AdaBoost regressor, and linear regression.

4.3. Performance evaluation using performance metrics

Performance metrics used to evaluate models developed with regression algorithms are R^2_score, MAE, MSE, MdAE, RMSE, and mean absolute percentage error (MAPE). Figure 7 shows performance of the various models employed in this study by means of R2_score. The highest attainable R2_score value is 1.0, the closer the value is to 1 the better the model. Figure 8 depict error metrics which are MAE, MSE, MdAE, RMSE, and MAPE of the five models. The lower the value of errors the better the performance of the algorithm.

4.4. Discussion of findings

The result from the performance evaluation of the selected algorithms shows that SGD has performed a lot better compared to others. It has a higher R^2_score and has lower values of errors compared to other algorithms. Therefore, SGD was selected among others to implement the prediction system.

Figure 7 depict R^2_score performance metrics and also shows that the ExtraTree regressor has a better R^2_score compared to linear regression. But the performance evaluation of both algorithms using error metrics as represented in Figure 8 shows that linear regression has lower MAE, MdAE, and MAPE compared to ExtraTree regressor. While ExtraTree regressor has lower MSE and RMSE compared to linear regression. Therefore, if an algorithm has a higher R^2_score does not mean the algorithm is better compared to others. When selecting the algorithm to adapt for a project, the algorithm should not be evaluated based on R^2_score performance alone, different error metrics should also be put into consideration.

[45] ✓ 3.6s

	Models	R2_Score
0	Random Forest	0.947990
1	Stochastic Gradien Descent	0.985050
2	AdBoost Regressor	0.893774
3	Linear Regression	0.970510
4	ExtraTree Regression	0.973098

Figure 7. R^2_score performance metric

[54] ✓ 0.5s

	Models	MAE	MSE	MdAE	RMSE	MAPE
0	Random Forest	0.083481	0.013035	0.056281	0.114169	0.016838
1	Stochastic Gradien Descent	0.040664	0.003747	0.012859	0.061210	0.008030
2	AdBoost Regressor	0.123298	0.026622	0.086082	0.163163	0.024465
3	Linear Regression	0.058291	0.007391	0.044265	0.085969	0.011906
4	ExtraTree Regression	0.064532	0.006742	0.052187	0.082111	0.012888

Figure 8. Error metrics

4.5. Crop yield prediction system

The prediction system was implemented using python programming language as depicted in Figure 9. The prediction system is then connected to a web interface for efficient usage. The algorithm used to build the model for the CYPS is SGD. The prediction system is a python code (predictor.py) that allows the entering of new inputs. These inputs supplied are collected by the model built which is capable of processing additional data to make predictions.

The prediction system processes the supplied data, predicts the value of y (crop yield), and then displays the predicted result of expected production. For easy usage of the prediction system by the users, the python code (the prediction system) is then connected to a web-based frontend through which users will be able to interact with the prediction system effectively. The frontend is developed with PHP, HTML, JavaScript, and CSS. To use the system users are required to have register and login successfully into the system. If this is not satisfied, users are redirected automatically to the login page (Figure 10). The registration and login page are designed such that input errors are handled.

The form on Figure 11 is required to be filled by the user to make a prediction. This form is designed such that the predict button remains disabled until all input boxes are filled. The prediction result is displayed on a modal box (Figure 12) which pops up immediately after the predict button is clicked. It is required the user wait for some seconds for the system to process the data supplied and display the result.

There is a save button on the result page that allows users to save the prediction output to their system. The result is downloaded and saved on the user system as a txt file. The user is expected to enter a name of which the file will be saved, and the name is required to end with ".md" or ".txt".

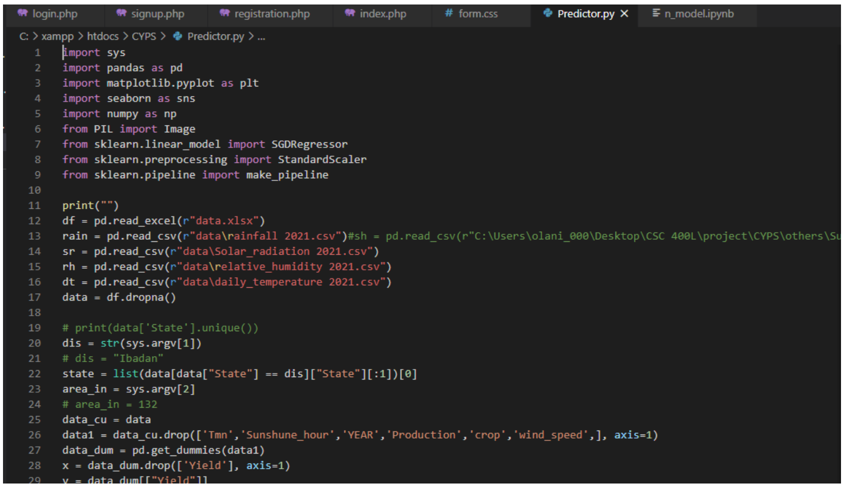


Figure 9. The prediction system

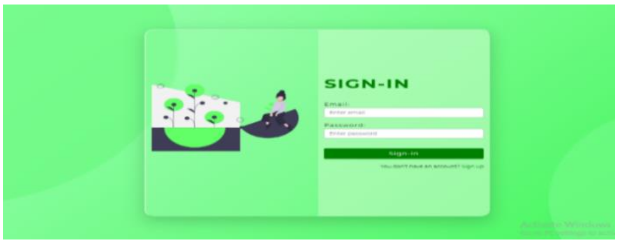


Figure 10. Login

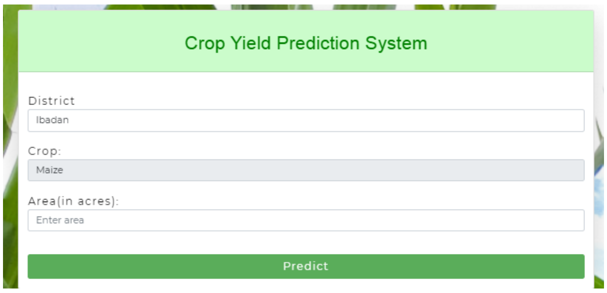


Figure 11. Input form

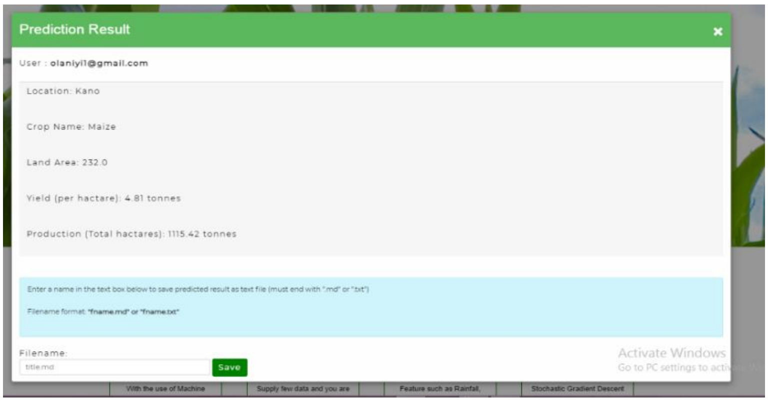


Figure 12. Prediction result

5. CONCLUSION

This project is undertaken using machine learning algorithms and to evaluates the performance of RF, SGD, ExtraTree regressor (ET), AdaBoost regressor, and linear regression. In the developed models, among all the five algorithms, SGD has shown great ability in predicting the yield of crops compared to other models. It has the lowest value of errors and highest value of R^2-score. The implementation of this system (CYPS) will aid in the betterment of this country's agriculture practices. It may also be used to help farmers minimize their losses and increase crop yields in order to increase their capital in agriculture. To aid the country's agricultural progress, the approach might be strengthened by integrating it with other sectors such as horticulture, sericulture, crop disease prediction, smart irrigation system, crop selection, storage system and so on. To summarize, this research has the potential to transform agriculture by offering farmers, policymakers, and other stakeholders' practical insights that will boost productivity, profitability, and resilience to shifting market and environmental conditions as its contributions. In future, farmers can be empowered to make prompt and well-informed decisions by utilizing predictive analytics to develop real-time monitoring systems and decision support tools. In order to support adaptive management practices in agriculture, future work can concentrate on integrating crop yield predictions with practical recommendations, automated alerts, and interactive interfaces.

ACKNOWLEDGEMENT

The APC was funded by the Africa Centre of Excellence on Technology Enhanced Learning (ACETEL), National Open University, Abuja, Nigeria.

REFERENCES

[1] M. O. Arowolo, M. O. Adebiyi, and A. A. Adebiyi, "A genetic algorithm approach for predicting ribonucleic acid sequencing data classification using KNN and decision tree," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 1, p. 310, Feb. 2021, doi: 10.12928/telkomnika.v19i1.16381.

[2] M. O. Arowolo, M. O. Adebiyi, A. A. Adebiyi, and O. J. Okesola, "Predicting RNA-Seq data using genetic algorithm and ensemble classification algorithms," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 2, pp. 1073–1081, 2020, doi: 10.11591/ijeecs.v21.i2.pp1073-1081.

[3] F. Baudron, M. A. Zaman-Allah, I. Chaipa, N. Chari, and P. Chinwada, "Understanding the factors influencing fall armyworm (Spodoptera frugiperda J.E. Smith) damage in African smallholder maize fields and quantifying its impact on yield. A case study in Eastern Zimbabwe," *Crop Protection*, vol. 120, pp. 141–150, Jun. 2019, doi: 10.1016/j.cropro.2019.01.028.

[4] S. A. Ajagbe and M. O. Adigun, "Deep learning techniques for detection and prediction of pandemic diseases: a systematic literature review," *Multimedia Tools and Applications*, vol. 83, no. 2, pp. 5893–5927, 2024, doi: 10.1007/s11042-023-15805-z.

[5] N. Brisson, P. Gate, D. Gouache, G. Charmet, F.-X. Oury, and F. Huard, "Why are wheat yields stagnating in Europe? A comprehensive data analysis for France," *Field Crops Research*, vol. 119, no. 1, pp. 201–212, Oct. 2010, doi: 10.1016/j.fcr.2010.07.012.

[6] M. S. Kukal and S. Irmak, "Climate-driven crop yield and yield variability and climate change impacts on the U.S. great plains agricultural production," *Scientific Reports*, vol. 8, no. 1, p. 3450, Feb. 2018, doi: 10.1038/s41598-018-21848-2.

[7] M. O. Adebiyi, O. O. Adeoye, R. O. Ogundokun, O. J. Olatunji, and A. A. Adebiyi, "Secured loan prediction system using artificial neural network," *Journal of Engineering Science and Technology*, vol. 17, no. 2, pp. 854–873, 2022.

[8] I. Eweoya, A. A. Ayodele, A. Azeta, and O. Olatunji, "Fraud prediction in bank credit administration: a systematic literature review," *Journal of Theoretical and Applied Information Technology*, vol. 97, no. 11, pp. 3135–3157, 2019.

[9] S. Veenadhari, B. Misra, and C. Singh, "Machine learning approach for forecasting crop yield based on climatic parameters," in *2014 International Conference on Computer Communication and Informatics*, Jan. 2014, pp. 1–5, doi: 10.1109/ICCCI.2014.6921718.

[10] T. Van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: a systematic literature review," *Computers and Electronics in Agriculture*, vol. 177, p. 105709, Oct. 2020, doi: 10.1016/j.compag.2020.105709.

[11] T. Ayodele, "Introduction to machine learning," *New Advances in Machine Learning*, vol. 2, no. 9–18, p. 16, 2010.

[12] K. Kaur, "Machine learning: applications in Indian agriculture," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 5, no. 4, pp. 342–344, 2016.

[13] T. Priyanka, P. Soni, and C. Malathy, "Agricultural crop yield prediction using artificial intelligence and satellite imagery," *Eurasian Journal of Analytical Chemistry*, vol. 13, no. SP, pp. 6–12, 2018.

[14] V. S. Konduri, T. J. Vandal, S. Ganguly, and A. R. Ganguly, "Data science for weather impacts on crop yield," *Frontiers in Sustainable Food Systems*, vol. 4, May 2020, doi: 10.3389/fsufs.2020.00052.

[15] B. O. Akinyemi, D. A. Olalere, M. L. Sanni, E. A. Olajubu, G. A. Aderounmu, and I. A. Ibrahim (Pantami), "Performance evaluation of machine learning models for cyber threat detection and prevention in mobile money services," *Informatica*, vol. 47, no. 6, pp. 173–190, Jun. 2023, doi: 10.31449/inf.v47i6.4691.

[16] J. Tack, A. Barkley, and L. L. Nalley, "Effect of warming temperatures on US wheat yields," *Proceedings of the National Academy of Sciences*, vol. 112, no. 22, pp. 6931–6936, Jun. 2015, doi: 10.1073/pnas.1415181112.

[17] C. A. O. Midega, J. O. Pittchar, J. A. Pickett, G. W. Hailu, and Z. R. Khan, "A climate-adapted push-pull system effectively controls fall armyworm, Spodoptera frugiperda (J E Smith), in maize in East Africa," *Crop Protection*, vol. 105, pp. 10–15, Mar. 2018, doi: 10.1016/j.cropro.2017.11.003.

[18] K. Moraye, A. Pavate, S. Nikam, and S. Thakkar, "Crop yield prediction using random forest algorithm for major cities in maharashtra state," *International Journal of Innovative Research in Computer Science & Technology*, vol. 9, no. 2, pp. 40–44, Mar. 2021, doi: 10.21276/ijircst.2021.9.2.7.

[19] M. O. Adebiyi, R. O. Ogundokun, and A. A. Abokhai, "Machine learning–based predictive farmland optimization and crop monitoring system," *Scientifica*, vol. 2020, pp. 1–12, May 2020, doi: 10.1155/2020/9428281.

[20] J. H. Jeong *et al*., "Random forests for global and regional crop yield predictions," *PLOS ONE*, vol. 11, no. 6, p. e0156571, Jun. 2016, doi: 10.1371/journal.pone.0156571.
[21] Sangeeta and G. Shruthi, "Design and implementation of crop yield prediction model in agriculture," *International Journal of Scientific and Technology Research*, vol. 9, no. 1, pp. 544–549, 2020.
[22] A. Shah, A. Dubey, V. Hemnani, D. Gala, and D. R. Kalbande, "Smart farming system: crop yield prediction using regression techniques," in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 19, 2018, pp. 49–56.
[23] A. Gonzalez-Sanchez, J. Frausto-Solis, and W. Ojeda-Bustamante, "Predictive ability of machine learning methods for massive crop yield prediction," *Spanish Journal of Agricultural Research*, vol. 12, no. 2, pp. 313–328, Apr. 2014, doi: 10.5424/sjar/2014122-4439.
[24] "Crop Yield Dataset," Dataworld.com 2020. https://data.world/login?next=%2Fsmithcalvin%2Fnigeria-maize-yield%2Fworkspace%2Ffile%3Ffilename%3Ddata.xlsx.
[25] R. Rawat, O. Oki, R. K. Chakrawarti, T. S. Adekunle, J. L. A. Gonzáles, and S. A. Ajagbe, "Autonomous artificial intelligence systems for fraud detection and forensics in dark web environments," *Informatica*, vol. 47, no. 9, pp. 51–62, Oct. 2023, doi: 10.31449/inf.v46i9.4538.
[26] S. A. Ajagbe, A. A. Adegun, A. B. Olanrewaju, J. B. Oladosu, and M. O. Adigun, "Performance investigation of two-stage detection techniques using traffic light detection dataset," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 4, p. 1909, Dec. 2023, doi: 10.11591/ijai.v12.i4.pp1909-1919.
[27] M. L. Maskey, T. B. Pathak, and S. K. Dara, "Weather based strawberry yield forecasts at field scale using statistical and machine learning models," *Atmosphere*, vol. 10, no. 7, p. 378, Jul. 2019, doi: 10.3390/atmos10070378.

BIOGRAPHIES OF AUTHORS

Julius Olatunji Okesola, Ph.D. is a professor of cybersecurity at the Abiola Ajimobi Technical University, (Formerly First Technical University) Ibadan, Nigeria. He is a Certified Information Security Manager (CISM) and a Certified Information Systems (IS) Auditor (CISA) with a Ph.D. in Computer Sciences. He is a member of IS Audit and Control Association (ISACA), Computer Professionals of Nigeria (CPN), and a fellow of Nigerian Computer Society (NCS). He is a scholar, an Information Security expert and a seasoned banker. Until November 2016, he was the Group Head, for Information Systems Control and Revenue Assurance at Keystone Bank (Nigeria) Ltd, Lagos. An alumnus of both the University of Lagos (UNILAG) and University of South Africa (UNISA), his research interests include cyber security, biometrics, and software engineering. He has several publications in scholarly journals and conference proceedings both local and international. He can be contacted at email: olatunjiokesola@tech-u.edu.ng.

Olaniyi Ifeoluwa earned his B.Sc. degree in Computer Science from Abiola Ajimobi Technical University, (Formerly First Technical University) Ibadan, Nigeria and currently pursuing his M.Sc. degree at the Department of Computer Science, University of Ibadan, Nigeria. His major research focus is in artificial intelligence, machine learning and human centered computing to ensure technology serves real-world needs. He's shaping a future where technology empowers. He can be contacted at email: olaniyiifeoluwa12@gmail.com.

Sunday Adeola Ajagbe is a Ph.D. candidate at the Department of Computer Engineering, Ladoke Akintola University of Technology (LAUTECH), Ogbomoso, Nigeria and a Lecturer, at Abiola Ajimobi Technical University, (Formerly First Technical University) Ibadan, Nigeria. He obtained M.Sc. and B.Sc. in Information Technology and Communication Technology respectively at the National Open University of Nigeria (NOUN), and his Postgraduate Diploma in Electronics and Electrical Engineering at LAUTECH. His specialization includes AI, natural language processing (NLP), information security, data science and internet of things (IoT). He is also licensed by The Council Regulating Engineering in Nigeria (COREN) as a professional Engineer, a student member of the Institute of Electrical and Electronics Engineers (IEEE), and International Association of Engineers (IAENG). He has many publications to his credit in reputable academic databases. He can be contacted at email: saajagbe@pgschool.lautech.edu.ng.

Olubunmi Okesola is a postgraduate student of Library and Information Sciences, University of Ibadan. She is also a graduate Assistant at the same University. A product of Federal Government College Osiele, she holds A B.Sc. (ed) Educational Management degree of Tai Solarin University of Education Ijebu-Ode, Nigeria. She can be contacted at email: sundayadeajagb@gmail.com.

Adeyinka O. Abiodun, Ph.D. obtained her B.Tech. Computer Science from the Federal University of Technology Akure, her M.Sc. and Ph.D. in Computer Science from the University of Ibadan. She is a lecturer and facilitator in the Department of Computer Science, National Open University of Nigeria (NOUN), and also a facilitator at the Africa Centre of Excellence on Technology Enhanced Learning (ACETEL-NOUN). Research interest includes forensic science, data analytics, pattern recognition, AI and data science and she has a number of publications in these fields that have appeared in high impact journals and learned conferences. She is a member of many associations, among which are the Organization of Women in Science for the Developing World, Black-in-AI amongst others. She is the monitoring and evaluation officer for Africa Centre of Excellence on Technology Enhanced Learning, NOUN. She can be contacted at email: aabiodun@noun.edu.ng.

Francis Bukie Osang, Ph.D. obtained his Ph.D. from the ICT University Yaoundé Cameroon in 2016. He is a senior lecturer in the Department of Computer Science, National Open University of Nigeria. He has authored over seventy (70) publications in referred international and national journals and conference proceedings. His research interests include information systems optimization/security, human-computer interactions, smart computing, cyber-security and eLearning. He has won several grants for research. He is a Fellow of the Nigeria Computer Society (FNCS). He is also a member of Computer Professionals Registration Council of Nigeria (CPN) and International Professional Managers Association, UK. He can be contacted at email: fosang@noun.edu.ng.

Olakunle O. Solanke, Ph.D. received his Ph.D. from the Federal University of Technology, Akure, Nigeria in 2016. He is a lecturer in the Department of Mathematical Sciences at Olabisi Onabanjo University, Nigeria. He has authored more than 30 articles. He has been involved in research activities in the area of network intrusion detection, network congestion control, information security, and machine learning. He is a member of the Nigerian Computer Society and the IEEE. He can be contacted at email: solanke.olakunle@oouagoiwoye.edu.ng.