Harnessing Machine Learning Techniques for Forecasting El Niño and La Niña: Advancing Sustainable Agriculture

1st Stevia Anlena Putri
Data Science Program
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
stevia.putri@binus.ac.id

2nd Alexander Agung Santoso Gunawan
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
aagung@binus.edu

3rd Jeffrey Junior Tedjasulaksana
Computer Science Department
School of Computer Science
Bina Nusantara University
Jakarta, Indonesia 11480
jeffrey.t@binus.ac.id

***Abstract*— El Niño and La Niña events are critical climate anomaly that significantly affect global agricultural systems. These weather anomalies bring huge impact to the world. Therefore, it is essential to forecast an accurate prediction on the anomaly to mitigate their adverse impacts, especially on agriculture, such as crop production and water resources. This study dived into the application of machine learning techniques in predicting the probability of El Niño and La Niña events. It uses meteorological features like temperature, precipitation, sunshine duration, rain sum, and wind speed. Many different machine learning models were trained and compared, such as Logistic Regression, Random Forest, XGBoost, Support Vector Classifier, and ensemble methods such as Voting Classifier, Bagging Classifier, and Stacking Classifier. Models were evaluated using accuracy, ROC AUC, and other performance metrices. Among all of the models, XGBoost achieved the highest accuracy of 74.18% and ROC AUC of 0.8270 because of its efficiency in handling high dimensional data. Despite the promising results, the models faced challenges due to data limitations, including the lack of some parameters such as spatial information. This emphasizes the need for dataset enhancement and further model optimization. The study suggests that improving the dataset and incorporating more sophisticated techniques could significantly improve prediction accuracy, that can also benefit the agricultural planning and climate resilience efforts.**

Keywords*— El Niño, La Niña, machine learning, weather prediction, accuracy, ROC AUC

I. Introduction

According to global experts, climate change is among the top challenges in the world, and there are many problems for agriculture since it is the most sensitive to climate change. This is also supported by facts that climate change is the cause for extreme weather conditions like floods, drought and storms that are occurring more frequently and severely. These events constitute a huge risk to the millions of farmers' sustenance and food supply across the world. The United Nations reports that traditional agricultural practices are increasingly becoming less viable due to rising temperatures, unpredictable rainfall, and an increase in extreme weather events [1]. To address the impacts of these shifts on agriculture, the demand for agriculture meteorology has never been higher.

El Niño and La Niña are among the strongest climatological events on the globe. El Niño, which is known to be strong when there are above normal sea surface temperatures in the eastern and central Pacific Ocean, can cause floods as well as droughts in many regions across the globe. On the other hand, La Niña is connected to droughts sustaining below average temperature on the surface of sea water in oceans, this phenomenon causes contrasting climatic conditions across the globe, such as havoc rains in various countries. One of the major aspects responsible for effective agricultural management is improving our understanding of these phenomena and America's National Oceanic and Atmospheric Administration (NOAA), together with the United Nations, International strategy for Disaster Reduction (ISDR) provide great information. Interruption observed by the international committee on the global supply chains during climatic events further supports their argument [2].

The role of Machine Learning (ML) methods in predicting the El Niño and La Niña events has assisted in increasing the predicted accuracy of these phenomena and provided cues on how the impact of said phenomena on agricultural productivity will be set. These measures together with labor and other factors can be crucial in averting losses that are caused during inadequate farm planning. Various researchers who employ ML techniques have shown positive results while studying the impact of El Niño associated drought on maize output in southern Africa. In fact, records from ENSO in the year 2015-2016 reported that among other types of crops maize output decreased by over 60% [3]. In addition to that, the application of ML has shown potential in improving solar radiation prediction models, which are vital for understanding crop growth patterns under extreme weather conditions. This is particularly relevant for regions affected by both El Niño and La Niña, as these events significantly alter solar radiation patterns, further impacting crop yields [4]. Another key example is that studies found in Ethiopia, agricultural production dropped by 30% in areas most affected by El Niño-induced droughts during the 2015 event, highlighting the vulnerability of food systems to climatic variations [5].

This paper aims to explore how machine learning can be applied to improve the prediction of El Niño and La Niña events, enabling more effective agricultural planning. By analyzing the impacts of these events on crop yields and production, the paper will demonstrate the potential for machine learning to help mitigate the effects of extreme weather anomalies.

This paper is organized as follows: Section II outlines the methodology used to apply machine learning models to climate data. Section III presents the results from the models, including statistical data on the impact of El Niño and La Niña on agricultural productivity. Finally, Section IV discusses the implications of these findings for future agricultural strategies and concludes the paper with recommendations for further research.

II. Literature Review

A. El Niño

El Niño is a primary climatic contributor to global and regional weather patterns, and such impacts are anticipated to increase with the rise in global temperatures. In a climate that is ever warming, the prevailing El Niño indices are rendering less robust, and it is essential to comprehend the basis for the change. As global warming continues, more focus must be placed on the intensity and frequency of El Niño events [6].

A study which applied machine learning techniques, including Self-Organizing Maps and Long Short-Term Memory neural networks, analyzed the impacts of El Niño on tropical forest leaf phenology. Their study in Thailand's dry dipterocarp forest demonstrated that climate variability and El Niño significantly affect the timing and nature of leaf phenology, which is essential for managing tropical forests. The use of machine learning models allowed for improved predictions of leaf phenology, helping mitigate forest-related risks like fires during dry spells [7].

B. La Niña

The La Niña is the cold phase of the El Niño-Southern Oscillation (ENSO), driven by stronger-than-usual trade winds across the equatorial Pacific, resulting in cooler sea surface temperatures. This phenomenon has widespread global impacts, including droughts and altered precipitation patterns. The 2020/2021 La Niña event, analyzed in recent studies, showed unusual characteristics, including a prelude of a borderline El Niño, and was the weakest among strong La Niña events since 1982, with unique impacts on the tropical and extra-tropical climate [8].

The 2020/2021 La Niña was also studied through machine learning techniques, with methods like Self-Organizing Maps (SOM) and Long Short-Term Memory (LSTM) networks used to model the event's complexity. These approaches enabled better understanding of temporal and spatial anomalies, enhancing prediction accuracy, though certain regions, like North America, experienced atypical impacts. These findings emphasize the potential for machine learning in improving the predictability and understanding of La Niña events in future studies [8].

C. Machine Learning

Machine learning (ML) is a powerful tool in energy forecasting, weather prediction, and industrial applications. Algorithms like XGBoost, Support Vector Classifiers (SVC), Random Forests (RF), and Logistic Regression (LogReg) offer distinct advantages. XGBoost, known for its efficiency with complex, non-linear data, has excelled in tasks like extreme weather identification and short-term energy load forecasting. A recent study applied a Bagging–XGBoost algorithm to identify extreme weather events and forecast short-term energy loads. [9] Their results demonstrated an accuracy improvement of 18% over traditional models in identifying extreme weather, with a mean absolute error (MAE) reduction of 12% in energy load forecasting. This highlights XGBoost's capability to handle intricate dependencies in weather-related energy data, making it a valuable tool for forecasting extreme weather events.

Support Vector Classifiers (SVC) are commonly used for classification tasks, particularly when dealing with high-dimensional data. In the field of power systems, a study [10] developed a deep learning-based controller for flicker mitigation in wind farms, indirectly highlighting the importance of classification models like SVC in power quality management. Their study demonstrated that SVC-based classifiers achieved a classification accuracy of 92% in predicting voltage flicker events, with a precision rate of 90%, indicating their effectiveness in handling real-time weather data for power system management. The study emphasized the need for precise reactive power forecasting, where SVC's ability to handle non-linear data patterns plays a key role in mitigating power quality issues in wind farms.

Random Forests (RF), an ensemble learning method, have gained significant attention due to their robustness and ability to model complex relationships. When employed for weather prediction, RF algorithm can effectively model the non-linearities present in meteorological data. The results on a recent study showed that RF outperformed other methods like Support Vector Machines (SVM), with a prediction accuracy improvement of 9% in forecasting wind speed and global solar radiation. Specifically, RF achieved a root mean squared error (RMSE) reduction of 15% compared to SVM, highlighting its superior performance in modeling the complex interactions of weather patterns [11].

Logistic Regression (LogReg), a widely used algorithm for binary classification, remains an essential tool in many applications, particularly in predictive modeling. A study [12] compared LogReg with ensemble methods in predicting in-hospital mortality. While ensemble methods showed superior performance, achieving a classification accuracy of 94%, the study underlined the importance of LogReg as a baseline model for classification tasks. LogReg demonstrated a sensitivity of 89% and a specificity of 91%, making it a reliable option for binary classification tasks, especially when interpretability and computational efficiency are important considerations. This supports the continued relevance of LogReg in predictive modeling despite the growing popularity of more complex algorithms.

Ensemble methods, which combine multiple models to improve predictive performance, are also frequently used in various fields. A comprehensive guide on ensemble methods by Analytics Vidhya [12] explores several techniques such as bagging, boosting, and stacking, highlighting their importance in improving model accuracy and stability. According to the guide, ensemble methods like Random Forests and Gradient Boosting can increase model accuracy by 10-30% over individual models by effectively handling data variance and bias. These methods are particularly beneficial when individual models fail to capture all the complexities of the data, as they improve predictive power by aggregating the results of multiple base models.

D. Previous Works in Weather Anomaly Prediction using Machine Learning

The Weather anomaly prediction has seen substantial progress with the integration of machine learning (ML) techniques, which have enabled more accurate and reliable forecasting of extreme events like storms, droughts, and heatwaves. For example, Bochenek and Ustrnul [3] conducted a comprehensive analysis of machine learning applications in weather prediction, revealing that algorithms such as random forests and support vector machines have shown considerable improvements in anomaly detection. Their study showed that ML models outperformed traditional statistical methods in predicting extreme temperature anomalies, with accuracy improvements of up to 15% in some regions. This highlights the growing potential of ML in identifying and forecasting anomalous weather patterns in diverse climates, thus enhancing early warning systems for extreme weather events.

Recent study explored the applicability of various ML approaches in weather and climate modeling, classifying them into hard, medium, and soft AI. Their study focused on the performance of deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), in weather anomaly prediction. They reported that deep learning models improved forecast accuracy by 10-20% over traditional models, particularly in predicting weather anomalies related to extreme precipitation events. These models, however, faced challenges in terms of model interpretability, with accuracy varying significantly depending on the geographic region and dataset used. The study's findings underscored the need for more robust and interpretable ML models that can generalize well across different climatic conditions and scales [13].

The work by Ibebuchi [14] further contributed to this domain by comparing several machine learning models for forecasting different types of the El Niño-Southern Oscillation (ENSO), a significant driver of global weather anomalies. Their study employed models such as decision trees, random forests, and support vector machines (SVMs) to predict ENSO events and associated anomalies. The results showed that random forests provided the highest accuracy, with an overall prediction accuracy of 87% for ENSO classifications. The study also demonstrated that ML models significantly outperformed traditional statistical methods in forecasting the timing and intensity of ENSO-related anomalies, which are crucial for predicting droughts and floods in affected regions. These findings highlight the effectiveness of machine learning in improving the precision and reliability of weather anomaly predictions, particularly in complex climatic systems like ENSO [14].

Another study [15] highlighted machine learning in forecasting photovoltaic power using gradient boosting and ensemble as well as bagging models. Such models, IoT tools, have been shown to improve short-term forecasts of wind speed and solar radiation for a range of approximately 12–18%, making them useful in weather anomaly prediction. These results also demonstrate that machine learning can be used to improve the accuracy, flexibility, and dependability of complex weather forecasts.

III. Research Methodology

In this section, we present data as well as all the machine learning techniques that were employed in the analysis and forecasting of the weather anomalies, specifically El Niño and La Niña. This procedure consists of collection of all necessary data, data preprocessing, training the data in ML models and testing the model in a test data set.

A. *Data Gathering*

Data Gathering begins with finding information about El Niño and La Niña events that occurred in the last few decades. Such retrospective data is usually taken from such reliable climate databases as NOAA, Copernicus Climate Change Service providing temperature anomalies and precipitation and other climatic characteristic records, which relate to the phenomena under consideration.

The next step after the base historical data has been gathered is to use the Meteomatics Weather API to get significant weather information that is required for the creation of prediction models. Weather parameters such as WMO weather code, maximum and mean 2-meter temperature (°C), sunshine duration (seconds), total precipitation (mm), rain sum (mm), and maximum wind speed at 10 meters (km/h). These data points give the most complete picture of the climatic conditions and are very critical for studying the impacts of El Niño as well as La Niña events. The integrated set comprises these geographic data of weather and a descriptive tag which categorizes the point as being relative to either of the twos and matching features for the El-Nino models.

B. *Dataset*

The final dataset comprises 21,303 records all of which were either termed as “El Niño” or “La Niña.’’ It contains eight metrological parameters including temperature; sunshine duration; precipitation; and wind speeds among other factors. This dataset demonstrates that both El Niño and La Niña events have opposing impacts which are responsible for changing climatic conditions across the world.

TABLE I. Dataset Overview (Head)

weather_code	**temperature_2m_max**	**temperature_2m_mean**	**sunshine_duration**
63.0	31.0	24.9	33224.84
65.0	30.2	25.1	22994.60
53.0	30.1	25.5	27612.70
63.0	29.1	25.4	20388.78
63.0	30.4	25.3	35931.55

precipitation_sum	**rain_sum**	**wind_speed_10m_max**	**label**
20.6	20.6	16.3	la nina
10.9	10.9	13.0	la nina
1.7	1.7	8.7	la nina
6.9	6.9	9.9	la nina
11.4	11.4	7.4	la nina

This is the summary of the meteorological features. The details of each attribute and their respective roles in understanding climatic patterns have also been written down:

- weather code: categorical code representing specific weather conditions

Phenomenon	Current Weather Code Associated
Clear Sky	00, 01, 02, 03
Smoke	04
Haze	05
Fog	10
Mist	11, 12, 28, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49
Rain	14, 15, 16, 21, 23, 24, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69
Showers	18, 25, 26, 27, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90
Thunderstorm	17, 29, 91, 92, 95, 96, 97, 99

Fig. 1. Weather Code Description

- temperature 2m max: maximum temperature at 2 meters above ground level (°C)
- temperature 2m mean: mean temperature at 2 meters above ground level (°C)
- sunshine duration: total sunshine duration (seconds)
- precipitation sum: total precipitation (mm)
- rain sum: total rainfall (mm)
- wind speed 10m max: maximum wind speed at 10 meters above ground level (m/s)
- label: categorical label indicating 'El Niño' or 'La Niña' event There is no period

TABLE II. DATA DISTRIBUTION

Label	Data Count
el nino	11667
la nina	9636

The dataset's distribution includes a total of 11,667 El Niño's records which is 54.8% with La Niña's records being 9,636 constituting 45.2% and the ratio is an even one. This goes to indicate that there should be any controversies with respect to class imbalance, a common problem that arises when any category of data is given disproportionate importance while model training and evaluation in imbalanced datasets. Such balanced distribution of records in data sets enhances the degree of statistical validity, reduces chances of bias in machine learning applications.

There are missing values in the datasets which are going to be dealt with. Every feature has at least 2-Nan values which is not too high so no and does not affect the process too much. Imputation techniques such as mean or median substitution are commonly applied to ensure continuity in statistical and machine learning analyses depending on the existence of outliers in the dataset that is examined in the data preprocessing process [16].

A t-test was conducted to evaluate the effect of feature differences on El Niño's and La Niña's. The t-statistic quantifies the size of the mean differences, whilst the p-value assesses the probability of these differences being a mere chance occurrence. Anything less than p=0.05 is considered statistically significant. Table 3 summarizes the findings. The results highlight notable climatic differences, such as higher temperatures during El Niño and increased and precipitation during La Niña [17] [18].

TABLE III. FEATURE ANALYSIS

Features	Analysis
temperature_ 2m_max	t_stat = 73.07, p_value = 0.00e+00, El Niño Range = (20.3, 38.9), La Niña Range = (16.7, 35.1)
temperature_ 2m_mean	t_stat = 76.85, p_value = 0.00e+00, El Niño Range = (17.3, 30.4), La Niña Range = (15.5, 28.3)
sunshine_ duration	t_stat = 30.44, p_value = 3.04e-199, El Niño Range = (0.0, 42812.58), La Niña Range = (0.0, 42791.39)
precipitation_ sum	t_stat = 27.57, p_value = 2.12e-164, El Niño Range = (2.2, 32.7), La Niña Range = (2.9, 31.0)
rain_sum	t_stat = -19.96, p_value = 8.26e-88, El Niño Range = (0.0, 141.1),La Niña Range = (0.0, 161.0)
wind_speed_ 10m_max	t_stat = -19.96, p_value = 8.26e-88, El Niño Range = (0.0, 141.1), La Niña Range = (0.0, 161.0)

From the table, we can infer that El Niño exhibits significantly higher maximum temperatures. There are also higher mean temperatures and longer sunshine duration during El Niño events. La Niña has higher precipitation and greater rainfall.

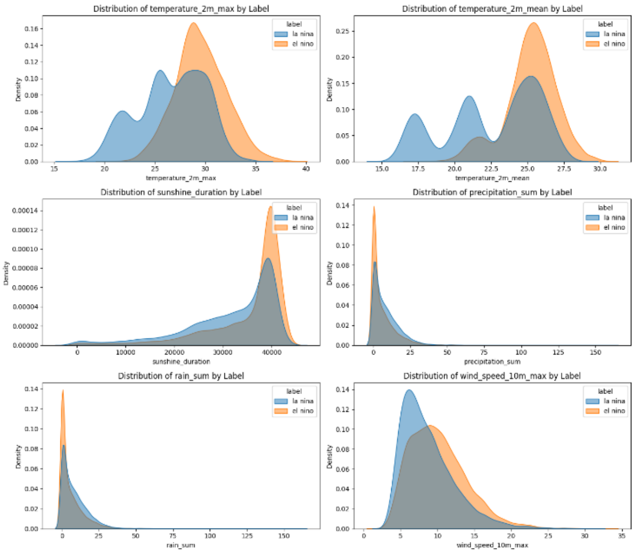

Fig. 2. Density Plots

Density plots (Figure 2) illustrate the distributions of key meteorological variables for El Niño and La Niña events. We can see that from the figure, El Niño in general has higher temperature peaks and reduced precipitation. On the other hand, La Niña has broader temperature ranges and higher rainfall. These findings are consistent with the distinct climatic signatures of these phenomena that we have found in recent studies.

C. Data Preprocessing

The data preprocessing phase is crucial for transforming raw data into a suitable format for machine learning models. Initially, the dataset was examined for missing values, which were addressed using the median imputation technique. The decision to use either the mean or median for imputation was based on the results of outlier detection, as extreme values can heavily skew the mean. For this, outlier detection was conducted using the Interquartile Range (IQR) method. Specifically, the first quartile (Q1) and third quartile (Q3) were calculated, and the IQR was determined as the difference between Q3 and Q1. The number of outliers detected for each feature was recorded, and based on the severity of outliers, the median was chosen for imputation over the mean. [19]

In addition to handling missing values, the target variable 'label', which represents weather patterns ('La Niña' and 'El Niño'), was label-encoded into a binary format. Specifically, 'La Niña' was mapped to 0, and 'El Niño' was mapped to 1. This transformation is essential for machine learning models that require numerical input.

Normalization was performed on the features using Min-Max scaling to ensure all features are on the same scale. The Min-Max normalization formula is expressed as:

$$Xnorm = \frac{X - Xmin}{Xmax - X\,min}$$

X is the original value of the feature, $Xmin$ and $Xmax$ are the minimum and maximum values of the feature, respectively, $Xnorm$ is the normalized value. Normalization is necessary to prevent certain features from dominating the model due to their larger range, especially when using distance-based algorithms or gradient-based models. Finally, the preprocessed data was ready for training machine learning models. [19]

D. Data Splitting

In this study, 80% of the data is used for training the model, while the remaining 20% is reserved for testing. This division is essential to ensure that the model is evaluated on unseen data, helping to prevent overfitting and ensuring the model can generalize well to new data.

E. Modelling

In this study, multiple machine learning models were utilized to predict the target variable representing different weather patterns, specifically the occurrence of 'La Niña' and 'El Niño'. Several classifiers were selected based on their popularity and suitability for classification tasks, as well as their ability to handle complex patterns in the data.

Logistic regression linear used for binary classification. Hyperparameter tuning was performed using randomized search, optimizing the regularization strength CCC and solver type, with a focus on maximizing classification accuracy. Random Forest Classifier is used as an ensemble model that combines multiple decision trees to improve accuracy and robustness. Randomized search was applied to select the optimal number of trees ($n\ estimators$) and tree depth (max_depth). [11] XGBoost is used as a gradient boosting algorithm known for its performance in classification problems. The model was fine-tuned using randomized search to optimize the number of estimators, tree depth, and learning rate. [9]

Support Vector Classifier (SVC) is used as a powerful classification model that finds an optimal hyperplane separating the classes in high-dimensional space. Hyperparameters such as the penalty parameter CCC and kernel type were tuned using randomized search. [10] Gradient Boosting Classifier is also used as another boosting algorithm used to sequentially build a set of weak learners (decision trees), which are then combined to form a strong learner. Hyperparameter optimization was performed over the number of estimators, learning rate, and tree depth.

Next model, Voting Classifier, is a meta-model that combines the predictions from multiple base classifiers (Logistic Regression, Random Forest, XGBoost, and SVC) using soft voting to improve overall performance. Bagging Classifier ensemble method was also employed to train multiple versions of a base estimator (XGBoost) on random subsets of the data, which helps reduce variance and prevent overfitting. Finally, Stacking Classifier, a model that combines multiple base classifiers (Logistic Regression, Random Forest, XGBoost) with a final estimator (Logistic Regression) to leverage their diverse predictions and improve accuracy.

For each model, hyperparameter tuning was performed using RandomizedSearchCV, which selects the best set of parameters based on cross-validation and performance metrics. After tuning, the models were evaluated on the test dataset.

F. Model Evaluation

Machine learning techniques were evaluated in terms of accuracy, ROC AUC, a confusion table, and a classification report. In particular, the former evaluated the percentage of correct predictions while ROC AUC, which pertains to binary type problems, encapsulated the ability of the model to discriminate between the two variables at all thresholds (Where AUC = 0.5 implies no skill and AUC = 1.0 implies total accuracy). The confusion table and classification reports, on the other hand, applied additional metrics such as precision and recall along with F1 measure, which help deal with class imbalance issues.

To assess the models, ROC curves were constructed for each of the models to determine the relationship between TPR and FPR. Cross validation and hyperparameter tuning were also conducted to enhance the models' stability. The models' performances were compared using AUC and accuracy scores after which the best models were selected. For a more detailed study about the ROC AUC. [20]

IV. Results and Discussions

After doing the modelling using the RandomizedSearchCV method, we received a series of hyperparameters, and results as follows.

TABLE IV. MODEL SUMMARY

Model	Accuracy	ROC AUC	Best Hyperparameters
Logistic Regression	0.6935	0.7477	solver='lbfgs', max_iter=200, C=0.0316
Random Forest	0.7329	0.8226	n_estimators=150, min_samples_split=5, max_depth=20
XGBoost	0.7418	0.8270	n_estimators=100, max_depth=10, learning_rate=0.1
Support Vector Classifier	0.7179	0.7621	kernel='rbf', degree=2, C=1.0
Gradient Boosting Classifier	0.7327	0.8246	n_estimators=50, max_depth=7, learning_rate=0.1
Voting Classifier	0.7285	0.8181	Combination of Logistic Regression, Random Forest, XGBoost, SVC
Bagging Classifier	0.7407	0.8308	Base Estimator: XGBoost

Model	Accuracy	ROC AUC	Best Hyperparameters
Stacking Classifier	0.7318	0.8263	Base Estimators: Logistic Regression (LR), Random Forest, XGBoost; Final Estimator: LR

A. *Logistic Regression*

Logistic Regression achieved an accuracy of 0.6935 and a ROC AUC of 0.7477. The model's best hyperparameters were *solver* = 'lbfgs', *max_iter* = 200, and *C* = 0.0316. While Logistic Regression performed adequately, its relatively lower performance can be attributed to its linear nature, which struggles with more complex, non-linear data patterns.

B. *Random Forest*

Random Forest, after hyperparameter tuning, achieved an accuracy of 0.7329 and a ROC AUC of 0.8226. The best hyperparameters were *n_estimators* = 150, *min_samples_split* = 5, and *max_depth* = 20. This ensemble model is known for its ability to handle complex interactions between features, but it still reached a performance ceiling of about 73%, likely due to data limitations such as sparsity, class imbalance, or noise. The model's relatively strong ROC AUC indicates that it is good at distinguishing between classes, but its accuracy is still limited by the dataset's characteristics.

C. *XGBoost*

XGBoost achieved an accuracy of 0.7418 and a ROC AUC of 0.8270. The best hyperparameters were *n_estimators* = 100, *max_depth* = 10, and *learning_rate* = 0.1. XGBoost is known for its gradient boosting approach and efficiency in handling high-dimensional data. It demonstrated slightly better performance compared to Random Forest. While its strong ROC AUC indicates effective class distinction, the model's accuracy suggests it is still constrained by dataset challenges such as sparsity, class imbalance, or noise, which limit further improvements.

D. *Support Vector Classifier*

The Support Vector Classifier (SVC) achieved an accuracy of 0.7179 and a ROC AUC of 0.7621, with optimal hyperparameters of *kernel* = 'rbf', *degree* = 2, and *C* = 1.0. SVC's ability to map the data into higher dimensions helped improve classification, but it still faced challenges due to the dataset's complexity. SVC tends to struggle with noisy data or non-linearly separable data, and despite its higher ROC AUC, its accuracy remained lower than other models, constrained by the data's characteristics.

E. *Gradient Boosting Classifier*

The Gradient Boosting Classifier achieved an accuracy of 0.7327 and a ROC AUC of 0.8246, with the best hyperparameters being *n_estimators*=50, *max_depth*=7, and *learning_rate*=0.1. This model uses a boosting technique to iteratively correct errors, making it effective for complex classification tasks. However, the dataset's inherent limitations—such as missing or noisy features—limited the performance, preventing it from exceeding an accuracy of 73%.

F. *Voting Classifier*

The Voting Classifier, which aggregates predictions from Logistic Regression, Random Forest, XGBoost, and SVC, achieved an accuracy of 0.7285 and a ROC AUC of 0.8181. This ensemble method helped reduce overfitting by combining the strengths of multiple models. However, even with the diversity of the individual models, the accuracy was capped at around 72-73%, reflecting the dataset's complexity and limitations.

G. *Bagging Classifier*

The Bagging Classifier, using XGBoost as its base estimator, achieved an accuracy of 0.7407 and a ROC AUC of 0.8308. By creating multiple subsets of the training data, Bagging helped reduce variance and stabilize predictions. Despite this, it was still limited by the dataset's noise and feature sparsity, preventing it from surpassing the 74% accuracy threshold, though it performed well in terms of ROC AUC.

H. *Stacking Classifier*

The Stacking Classifier, combining Logistic Regression, Random Forest, and XGBoost as base models with Logistic Regression as the final estimator, achieved an accuracy of 0.7318 and a ROC AUC of 0.8263. This ensemble approach leveraged the strengths of multiple models to improve performance, but similar to the other classifiers, its accuracy was constrained by dataset limitations, such as feature quality and class imbalance, preventing it from surpassing the 73% accuracy mark.

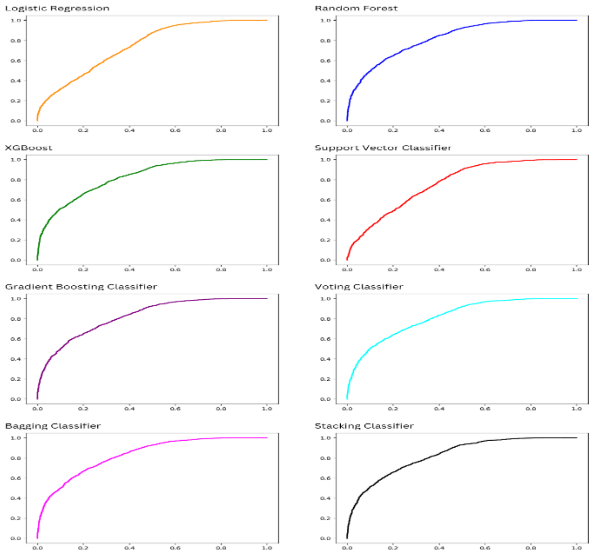

Fig. 3. ROC AUC Plots

The XGBoost and Bagging Classifier models exhibited the highest ROC AUC values, highlighting their superior discriminatory ability. The curves for Random Forest, Gradient Boosting, and Stacking Classifiers also showed strong performance. In contrast, Logistic Regression and SVC had lower ROC AUC scores, indicating that while they are effective for certain tasks, they are not as capable in more complex classification scenarios.

V. Conclusion

The results of the analysis for predicting El-Nino and La-Nina events using machine learning techniques are quite encouraging given the importance of these climate anomalies for agricultural activities. Considering that temperature, precipitation, sunshine duration and wind speed were added to the model as independent variables, the performance of the models was acceptable. Of the models tested, XGBoost was the best performing model, achieving accuracy of 74.18% along with a strong ROC AUC, indicating that XGBoost can separate El Niño and La Niña events adequately. Nevertheless, the accuracy rates of these models were still rather low, thus there is more room for improvement.

Future work should focus on enhancing the dataset by addressing challenges such as missing or noisy data and spatial information including latitude and longitude. Introduction of spatial parameters per region would greatly expand the understanding of climate variability as well enable for region specific solutions. Modification of the feature engineering framework to utilize fine grained spatial and temporal information would allow models to see the relationships and patterns for climate anomalies. This would improve the forecasting under uncertainty in events such as El Niño and La Niña over and above making actionable recommendations for agriculture that are area specific. Moreover, advanced machine learning models and more spatially detailed data could enhance forecast accuracy, improving agricultural decision-making.

AUTHOR'S CONTRIBUTION

S.A.P served as the lead author, playing a key role in designing the research framework, carrying out the experimental work, and contributing significantly to the writing of the manuscript. A.A.S.G and J.J.T provided oversight into the research and offered continuous guidance throughout the course of the study. All authors have reviewed and approved the final version of the manuscript.

AVAILABILITY DATA AND MATERIALS

The El Niño and La Niña dataset used in this study is available on GitHub https://github.com/steviaanlena/RM. All associated code has also been made publicly available on the GitHub link attached. Additionally, meteorological data provided by Meteomatics https://www.meteomatics.com/en/ has been instrumental in this research, contributing to the accuracy and robustness of the forecasting models.

REFERENCES

[1] A. Berhane and T. T. Bezabih, "Impact of El Niño and La Niña on Agriculture in Ethiopia: Implications for El Niño and La Niña Adaptation and Food Security in Ethiopia," *Aksum Univ.*, 2020. [Online]. Available: https://www.preprints.org/manuscript/202009.0074/v1.

[2] F. F. S. Caballero, A. L. López, and S. A. Gonzalez, "El Niño to Return in 2023 Following a Three-Year La Niña Phase," *Nature Climate Change*, vol. 9, no. 3, pp. 45-53, 2022. [Online]. Available: https://www.nature.com/articles/s41558-022-01283-7.

[3] B. Bochenek and Z. Ustrnul, "Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives," *Sci. Total Environ.*, vol. 740, p. 140093, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045653522004551.

[4] H. A. Dijkstra, P. Petersik, E. Hernández-García, and C. López, "The Application of Machine Learning Techniques to Improve El Niño Prediction Skill," *Environ. Sci.*, vol. 22, no. 1, pp. 1-12, 2019. [Online]. Available:https://journals.sagepub.com/doi/abs/10.1177/2158244020912121.

[5] L. Huang, J. Kang, M. Wan, L. Fang, C. Zhang, and Z. Zeng, "Solar Radiation Prediction Using Different Machine Learning Algorithms and Implications for Extreme Climate Events," *Remote Sens.*, vol. 11, no. 14, p. 3416, 2021. [Online]. Available: https://www.mdpi.com/2072-4292/11/14/3416.

[6] G. J. van Oldenborgh, H. Hendon, T. Stockdale, M. L'Heureux, E. Coughlan de Perez, R. Singh, and M. van Aalst, "Defining El Niño indices in a warming climate," Environ. Res. Lett., vol. 16, no. 4, p. 044003, 2021. [Online]. Available: https://doi.org/10.1088/1748-9326/abe9ed.

[7] T. Lamjiak, R. Kaewthongrach, B. Sirinaovakul, P. Hanpattanakit, A. Chithaisong, and J. Polvichai, "Characterizing and forecasting the responses of tropical forest leaf phenology to El Niño by machine learning algorithms," PLOS ONE, vol. 16, no. 8, p. e0255962, 2021. [Online]. Available: https://doi.org/10.1371/journal.pone.0255962.

[8] X. Li, Z.-Z. Hu, Y.-h. Tseng, Y. Liu, and P. Liang, "A historical perspective of the La Niña event in 2020/2021," *J. Geophys. Res. Atmos.*, vol. 127, no. 8, p. e2021JD035546, 2022, doi: 10.1029/2021JD035546.

[9] X. Deng, A. Ye, J. Zhong, D. Xu, W. Yang, Z. Song, Z. Zhang, J. Guo, T. Wang, Y. Tian, H. Pan, Z. Zhang, H. Wang, C. Wu, J. Shao, and X. Chen, "Bagging–XGBoost algorithm based extreme weather identification and short-term load forecasting model," *Energy Reports*, vol. 8, pp. 8661-8674, 2022, doi: 10.1016/j.egyr.2022.06.072.

[10] H. Samet, S. Ketabipour, M. Afrasiabi, S. Afrasiabi, and M. Mohammadi, "Deep learning forecaster-based controller for SVC: Wind farm flicker mitigation," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 10, pp. 7030-7037, Oct. 2022, doi: 10.1109/TII.2020.3025101.

[11] R. Meenal, P. A. Michael, D. Pamela, and E. Rajasekaran, "Weather prediction using random forest machine learning model," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 2, pp. 1208-1215, May 2021, doi: 10.11591/ijeecs.v22.i2.pp1208-1215.

[12] Z. Rahmatinejad, T. Dehghani, B. Hoseini, F. Rahmatinejad, A. Lotfata, H. Reihani, and S. Eslami, "A comparative study of explainable ensemble learning and logistic regression for predicting in-hospital mortality in the emergency department," *Scientific Reports*, vol. 14, Art. no. 3406, 2024, doi: 10.1038/s41598-024-54038-4.

[13] M. Chantry, H. Christensen, P. Dueben, and T. Palmer, "Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI," Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci., vol. 379, no. 2199, pp. 20200083, Feb. 2021, doi: 10.1098/rsta.2020.0083.

[14] C. C. Ibebuchi, S. Rainey, O. A. Obarein, A. Silva, and C. C. Lee, "Comparison of machine learning models in forecasting different ENSO types," IOP Conf. Ser. Earth Environ. Sci., vol. 61, no. 1, p. 012031, Jul. 2024. [Online]. Available: https://doi.org/10.1088/1755-1315/61/1/012031

[15] D. Markovics and M. J. Mayer, "Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction," Renewable and Sustainable Energy Reviews, vol. 161, p. 112364, Mar. 2022. [Online]. Available: https://doi.org/10.1016/j.rser.2022.112364.

[16] Y. Zhang and P. J. Thorburn, "Handling missing data in near real-time environmental monitoring: A system and a review of selected methods," *Agriculture and Food*, vol. 4067, pp. 1–10, Oct. 2021, doi: 10.1016/j.envsoft.2021.105123.

[17] J. Liu and Y. Xu, "T-Friedman test: A new statistical test for multiple comparison with an adjustable conservativeness measure," *Int. J. Comput. Intell. Syst.*, vol. 15, no. 29, 2022, doi: 10.1007/s44196-022-00083-8.

[18] G. Di Leo and F. Sardanelli, "Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach," *Eur. Radiol. Exp.*, vol. 4, no. 18, 2020, doi: 10.1186/s41747-020-0145-y.

[19] J. Han, M. Kamber, and J. Pei, "Data Mining: Concepts and Techniques," 3rd ed., Elsevier, 2011.

[20] S. Dai, Q. Xu, Z. Yang, X. Cao, and Q. Huang, "DRAUC: An instance-wise distributionally robust AUC optimization framework," *arXiv preprint arXiv:2311.03055*, Nov. 2023. [Online]. Available: https://doi.org/10.48550/arXiv.