2023 1st IEEE International Conference on Smart Technology (ICE-SMARTec)
Bandung, Indonesia, 17-19 July 2023

Comparative Analysis of Machine Learning Models for Relative Humidity Prediction in the Philippines

Pitz Gerald G. Lagrazon
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
pitzgerald.lagrazon@gmail.com

Jennifer Edytha E. Japor
Southern Luzon State University
Lucban, Quezon, Philippines
japorjen@gmail.com

Julie Ann B. Susa
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
jannsusa@gmail.com

Marmelo V. Abante
Graduate School
World Citi Colleges
Quezon City, Philippines
dmva888@gmail.com

Renato R. Maaliw III
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
maaliw@slsu.edu.ph

Arnold B. Platon
Computer Studies Department
Bicol University Polangui
Polangui, Albay, Philippines
abplaton@bicol-u.edu.ph

Ace C. Lagman
Information Technology Dept.
FEU Institute of Technology
Manila, Philippines
aclagman@feutech.edu.ph

Manuel B. Garcia
Educational Innovation and Technology Hub
FEU Institute of Technology
Manila, Philippines
mbgarcia@feutech.edu.ph

2023 1st IEEE International Conference on Smart Technology (ICE-SMARTec) | 979-8-3503-0615-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICE-SMARTECH59237.2023.10461954

***Abstract*—Relative humidity is an important environmental parameter and is widely used in various fields. Prediction of humidity levels is crucial for climate modeling, heat stress, air quality forecasting, and public health. Machine learning techniques have shown potential for predicting humidity due to their nonlinear nature. However, there is a research gap in humidity prediction in the Philippines, specifically the lack of studies utilizing the available parameters provided by PAGASA, presenting an opportunity for further investigation and development of models for predicting humidity levels in the country. In this study, the researchers used a publicly available dataset from PAGASA containing weather measurements from 2000 to 2022 in the Philippines. Various machine learning models were trained and tested, with hyperparameter tuning performed using Bayesian optimization. The Gaussian Process Regression model with optimized hyperparameters achieved the best performance in predicting relative humidity, with the lowest RMSE and highest R-squared values. This study provides a reliable way to predict humidity levels in the Philippines based on weather parameters.**

Keywords—Gaussian Process Regression, machine learning, PAGASA, Philippines, relative humidity

I. Introduction

Relative humidity refers to the level of water saturation in the air [1] and has numerous applications in fields such as metrology, meteorology, climatology, and engineering [2]. Along with temperature and pressure, it is a significant environmental factor and has a connection with particle formation [3]. The magnitude of relative humidity depends not only on the amount of water vapor in the air but also on how temperature and water vapor interact [4].

Humidity prediction holds great importance for various reasons. The presence of water vapor in the atmosphere, acting as humidity, possesses strong greenhouse effects, making it crucial to monitor humidity levels in tropical regions for global climate considerations [5]. Recent studies have underscored the significance of humidity in determining extreme wet-bulb temperatures, which serve as indicators of heat stress [6]. Numerous algorithms have been developed to estimate air humidity by utilizing temperature and precipitation data [7]. When combined with other climatic factors such as temperature and ultraviolet radiation, humidity exerts a notable influence on the skin [8]. Furthermore, it contributes to the correlation between heat and mental health [9].

The use of machine learning-based prediction strategies has gained considerable attention in predicting relative humidity because of their ability to handle nonlinearity. Machine learning techniques are crucial for air quality forecasting and the efficient design of air-dependent energy systems [10]. It is essential to have a proper understanding of the relationships between meteorological and climatological variables before implementing machine learning models in weather prediction and climate modeling [11]. However, there is a current research gap on humidity prediction in the Philippines, specifically the lack of studies utilizing the available parameters provided by PAGASA. This gap presents an opportunity for further investigation into developing a best-fit and reliable model for predicting humidity levels in the country. Such models would be beneficial in various applications such as agriculture, public health, and disaster preparedness. Moreover, filling this void in the research could result in improved comprehension of the variations and patterns in humidity levels within the Philippines, which could provide insights into the impacts of climate change and inform climate adaptation strategies. Therefore, it is important to explore and develop machine learning models that can anticipate humidity levels in the Philippines using the available parameters provided by PAGASA.

II. Related Works

In [12], the importance of controlling temperature and relative humidity for crop cultivation is highlighted. The study uses a multilayer perceptron (MLP) to predict air temperature and relative humidity for a greenhouse cultivating mango. The

MLP model incorporates various inputs, such as information gathered from inside and outside the greenhouse, and the configured and operational values of environmental control equipment. The MLP's performance was tested by comparing three-day data from each of the four seasons in Korea. The MLP was optimized for air temperature and relative humidity, but its accuracy decreased as the prediction time increased. The study emphasizes the need for collecting more data from different greenhouses and modifying the neural network structure for generalization.

In [13], a study was conducted to compare two models for forecasting temperature and humidity in an open office. The models were a linear ARX model and a nonlinear NNARX model. Both models were developed and verified using climate data over a three-month period and aimed to forecast for various timeframes, from 30 minutes to 3 hours ahead. The study found that the nonlinear NNARX model performed better than the linear ARX model and had the potential for integration into HVAC plant controllers, specifically adaptive control systems.

In [14], the study focuses on predicting relative humidity (RH) using XGBoost combined with SVR, RF, and MARS models. The study evaluates the performance of the models at two stations in Iraq, Kut, and Mosul, using numeric and graphic indicators. The results showed that all models were effective in predicting RH, and the XGBoost approach was successful in abstracting essential parameters for RH simulation with fewer input parameters. The most precise prediction outcomes were obtained using the RF model at Kut station, which utilized evaporation and maximum air temperature parameters, whereas the MARS model, which used all climate parameters, produced the best predictions at Mosul station. The study suggests that the proposed coupled machine learning models have the potential to model RH in semi-arid environments.

In [15], researchers explore the prediction of indoor air temperature (IAT) and indoor relative humidity (IRH) in closed barns for livestock animals. The study compares the performance of several machine learning models, such as MLR, SVR, DTR, RFR, and MLP, in forecasting IAT and IRH using external environmental data. Three different input datasets are utilized, and the models are evaluated based on various performance metrics. The study highlights the superior performance of RFR models in predicting both IAT and IRH, especially when using the S3 input dataset. The importance of feature selection from the input data is emphasized for achieving accurate predictions. The results demonstrate the potential of machine learning models in predicting IAT and IRH in barns with natural ventilation.

In [16], the authors conducted a study on precise and dependable temperature and humidity management in industrial manufacturing. They developed an enhanced forecasting model based on BPNN to predict IAT and IRH. The study was conducted in Chongqing, China, known for its humid and hot summers and cold winters. The model was created utilizing an industry-specific cloud database, and the findings revealed that the predictions for IT and IH by the model were strongly correlated with the actual data. The approach proposed in this study proved to be a potent means of temperature prediction and can be implemented for forecasting and regulating indoor temperature and relative humidity in industrial production. This could result in consistent enhancements in productivity.

In [17], the study utilized artificial neural networks (ANNs) and multiple regression analysis to develop models for forecasting daily mean indoor air temperature (IAT) and indoor relative humidity (IRH) in an education building. The models were trained and tested using various parameters, including outdoor climate conditions, day of the year, and indoor thermal comfort. The results indicated that despite limited data, the ANN model demonstrated proficiency in predicting IAT and IRH parameters in educational buildings. Furthermore, the ANN model outperformed the multiple regression analysis model in terms of accuracy. This research highlights the potential of ANNs in predicting indoor thermal comfort conditions, estimating energy needs, and optimizing the sizing of HVAC systems.

In [18], the authors focused on the prediction of relative humidity for precipitation forecasting. They compared the performance of ARIMA and LSTM models using meteorological data from a county in China. The study found that the ARIMA model provided better results in predicting relative humidity than the LSTM model. The study also analyzed the factors that affect the prediction accuracy of both models, including the length of the training data and the hyperparameters of the models. The findings can help improve the prediction accuracy of weather forecasting models, particularly for areas with a high risk of precipitation.

In [19], a study evaluates an LSTM-based model for predicting relative humidity (RH) using observed weather data from a synoptic station location. The LSTM model is trained on two years of climate data and shows strong performance in forecasting complex, non-stationary univariate time series of RH. The results indicate that the LSTM model outperforms traditional forecasting methods and demonstrates its usefulness in predicting RH records. The study emphasizes the potential of using machine learning approaches and computational physics learning theory in weather and climate prediction. The findings may contribute to improving the accuracy of weather forecasting and understanding the underlying mechanisms of climate change.

In [20]. the study investigates the use of MARS and M5T models for relative humidity prediction in Pakistan. The findings indicate that the MARS model demonstrated superior performance compared to the M5T model across all meteorological stations. Among the different input scenarios, scenario S6 yielded the best results. During testing, the MARS model exhibited slightly better performance than the M5T model, but both models performed better during training than in testing. The study emphasizes the need for future research to employ machine learning tools in predicting additional meteorological variables in high-altitude basins..

In [21], the relationship between Industry 4.0 and Machine Learning is discussed, with predictive maintenance being a key application. The paper focuses on the use of the RF method to predict relative humidity in the environment of a smart factory. To ensure data reliability and interoperability, IIoT devices based on the oneM2M standard platform were used to collect data. The implementation of the RF method resulted in an accuracy of 82.49% in predicting relative humidity. This research is expected to benefit the manufacturing industry by reducing maintenance costs and increasing efficiency.

In [22], the study focuses on utilizing feedforward neural networks (FFNN) to predict relative humidity values in

Malaysia based on weather records. The proposed model predicts hourly relative humidity using inputs such as sunshine ratio and cloud cover. The model's accuracy is evaluated using three statistical parameters, namely MAPE, MBE, and RMSE. The findings demonstrate that the FFNN model effectively predicts hourly relative humidity, exhibiting low values for MAPE, RMSE, and MBE. The model also performs well in predicting daily and monthly relative humidity values, as indicated by the reported MAPE values. This research highlights the potential of FFNNs for accurate relative humidity predictions, particularly in tropical climates like Malaysia.

III. Research Methods

A. Dataset

The dataset used in the study consisted of daily measurements of various weather parameters such as maximum, minimum, and mean temperature, wind speed, wind direction, cloud cover, rainfall, and relative humidity. The dataset spanned from the year 2000 to 2022, providing a long-term perspective on the weather conditions in the Philippines. It is publicly available records in the Philippines, specifically from PAGASA. PAGASA is responsible for evaluating and predicting the weather, issuing alerts for floods and typhoons, providing public weather forecasts and advice, and offering specialized information services related to weather conditions[23]. Table I shows the list of attributes used in the study.

TABLE I. List of Attributes

Attributes	Unit	Description
Maximum Temperature	°C	Highest recorded temperature obtained at the main standard time of observations.
Minimum Temperature	°C	Lowest recorded temperature obtained at the main standard time of observations.
Mean Temperature	°C	The sum of Tmax and Tmin divided by two.
Rainfall Amount	mm	Measurement of the vertical depth of water that reaches the ground and is obtained from PAGASA stations using an 8-inch rain gauge.
Wind Speed	m/s	Average speed of the wind observed during a 10-minute interval.
Wind Direction	Degree (°) relative to true North	The nearest 10-degree direction of the wind's origin relative to true north.
Cloud Cover	okta	Amount of cloud present in the sky
Relative Humidity	%	The ratio of actual vapor pressure to the saturation vapor pressure of the air at a specific height (1.25-2.00 m above the ground) corresponding to the prevailing temperature.

B. Data Preprocessing

1) Data Cleaning: This dataset has undergone an extensive data-cleaning process to ensure its quality. Faulty data undermines the reliability of accurate results and algorithms [24]. First, any missing values were identified, and either imputed or removed based on the extent of the missingness. Duplicates were removed to avoid skewing the analysis results. Incorrectly formatted data, such as inconsistent date formats or string representations, were standardized for consistency. Any data inconsistencies were resolved to ensure the data is compatible with the chosen analysis method. The resulting clean dataset was then used for further analysis, ensuring that the findings were accurate and reliable.

2) Validation scheme and data splitting: In addition to the 70:30 data splitting ratio, 5-fold cross-validation was also used to validate the models built on the dataset. This process involved dividing the training data into five equal parts, training the models on four of these parts, and then evaluating their performance on the fifth part. This was repeated five times, with each part serving as the validation set once. The results were then averaged, providing an estimate of the model's performance on unseen data. The use of cross-validation helped to mitigate any potential issues with the data-splitting process, ensuring that the models were robust and generalizable. Additionally, it provided a more comprehensive assessment of model performance, allowing for a more reliable evaluation of their accuracy and predictive power. The combination of data splitting and cross-validation allowed for a thorough and robust validation scheme, ensuring that the models built on the dataset were accurate, reliable, and could generalize well to new, unseen data.

C. Correlation

The correlation coefficients between relative humidity (RH) and different weather variables in Fig. 1. It provides insight into how humidity levels are related to other weather conditions. The negative correlation between RH and temperature variables (TMAX, TMIN, TMEAN) suggests that as the temperature increases, the relative humidity decreases. This is because warm air can hold more moisture than cool air, so as the temperature rises, the air becomes more capable of holding moisture, resulting in lower relative humidity. The weak or insignificant negative correlation between RH and wind speed suggests that wind speed may not have a significant impact on relative humidity. The positive correlation between RH and cloud cover and rainfall indicates that higher humidity levels are associated with cloudier and wetter weather conditions. The positive correlation between RH and wind direction suggests that wind direction may be a factor in determining humidity levels, although the strength of this relationship is weak.

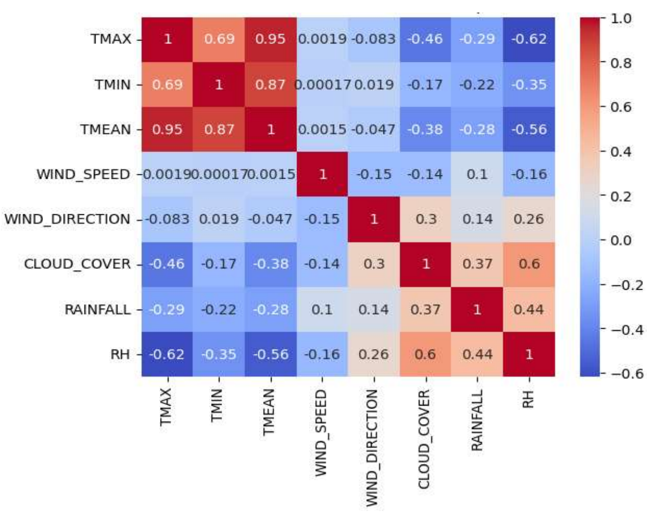

Fig. 1. Correlation Coefficient Heatmap.

While the correlation coefficient between relative humidity and some variables may be weak or insignificant, it is still possible that these variables significantly impact humidity levels in the real world. In practice, machine learning algorithms can help capture complex relationships between variables that may not be immediately apparent from simple statistical analysis. Therefore, it is still possible and even advisable to include all available variables in the training dataset, even if some have weak correlations with the response variable. This can help ensure that the model captures all possible factors that may influence the outcome and improve its predictive power.

D. Training and Testing Dataset

In this section, various machine learning models and their variations were trained and tested to determine which one provided the best accuracy for predicting humidity levels. The models were trained on a dataset containing weather variables. The accuracy of each model was evaluated based on how well it predicted the relative humidity values in a test dataset. This process involved analyzing the accuracy of each model without any hyperparameter adjustments.

RMSE, R-squared, MAE, and MSE are important parameters in regression models for several reasons. Firstly, RMSE quantifies the average magnitude of the residuals, providing a measure of how well the model fits the data and allowing for easy comparison across different models. R-squared, on the other hand, represents the proportion of variance in the dependent variable that is explained by the independent variables, indicating the model's goodness of fit. Lastly, MAE and MSE offer insights into the absolute and squared errors, respectively, enabling a more comprehensive understanding of the model's predictive accuracy and helping to assess the impact of outliers. Together, these metrics aid in evaluating and comparing the performance of regression models. Table II shows the summary of the machine learning model results.

To determine if there are significant differences in the performance across the models, evaluation measures (errors) were subjected to the ANOVA test. An f-value of 0.067, with a p-value of 1.00, was computed. Since the p-value is not less than the desired 0.05 level of significance, it was found that there are no significant differences in the performance of modelling relative humidity across models.

However, the ensemble (bagged trees) was considered the best-fit model since it has the least RMSE, MSE, and MAE and with the highest R squared.

E. Hyperparameter Tuning

The selection of the best values for the hyperparameters that govern the behavior of a machine learning model during the training process is known as hyperparameter tuning. In this study, the process involves selecting the best combination of hyperparameters to optimize the model's performance on a given dataset using MATLAB Regression Learner App. It is essential for building effective machine learning models that can make accurate predictions on new, unseen data.

The hyperparameter tuning using Bayesian optimization was used to improve the performance of five different algorithms – DT, SVM, GPR, Ensemble, and NN- for predicting humidity level. The tuning process involved selecting the best combination of hyperparameters for each algorithm on a 70:30 training and testing dataset split.

TABLE II. SUMMARY OF MACHINE LEARNING REGRESSION MODEL RESULT

Model	**5 cross-fold validation; 70:30 train test split**			
	RMSE	***R-squared***	***MAE***	***MSE***
Linear Regression (Linear)	4.067556	0.603732	3.136209	16.54501
Linear Regression (Interactions Linear)	3.982538	0.620124	3.05281	15.86061
Linear Regression (Robust Linear)	4.078521	0.601592	3.118904	16.63433
Stepwise Linear Regression	3.988475	0.61899	3.060174	15.90794
Tree (Fine Tree)	4.289579	0.559291	3.253593	18.40049
Tree (Medium Tree)	3.945899	0.627081	2.984066	15.57012
Tree (Coarse Tree)	3.879908	0.63945	2.956888	15.05369
SVM (Linear SVM)	4.125789	0.592304	3.145419	17.02213
SVM (Quadratic SVM)	3.860265	0.643092	2.917154	14.90164
SVM (Cubic SVM)	3.860092	0.643124	2.881756	14.90031
SVM (Fine Gaussian SVM)	4.54657	0.504903	3.413419	20.6713
SVM (Medium Gaussian SVM)	3.774332	0.658805	2.871859	14.24559
SVM (Coarse Gaussian SVM)	3.885012	0.638501	2.95457	15.09332
Ensemble (Boosted Trees)	5.009305	0.398996	4.161789	25.09313
Ensemble (Bagged Trees)	*3.634746*	*0.683575*	*2.751787*	*13.21138*
Gaussian Process Regression (Squared Exponential GPR)	3.749861	0.663215	2.867918	14.06146
Gaussian Process Regression (Matern 5/2 GPR)	3.724406	0.667772	2.847139	13.8712
Gaussian Process Regression (Exponential GPR)	3.721019	0.668376	2.824556	13.84598
Gaussian Process Regression (Rational Quadtraic GPR)	3.737151	0.665494	2.857863	13.9663
Neural Network (Narrow Neural Network)	3.774617	0.658754	2.890967	14.24773
Neural Network (Medium Neural Network)	4.036016	0.609853	3.111551	16.28942
Neural Network (Wide Neural Network)	3.902902	0.635164	2.997245	15.23264
Neural Network (Bilayered Neural Network0	3.669717	0.677457	2.82762	13.46683
Neural Network (Trilayered Neural Network)	3.657023	0.679684	2.800738	13.37382
Kernel (SVM Kernel)	4.360765	0.544542	3.281063	19.01627
Kernel (Least Squares Regression Kernel)	4.409173	0.534375	3.254083	19.4408

After training and testing the five algorithms with the tuned hyperparameters, the results were compared in Table III, and it was found that Gaussian Process Regression had the best performance in predicting crop yield. This means that, based on the selected hyperparameters, GPR was able to produce the most accurate predictions on the test dataset compared to the other algorithms.

TABLE III. SUMMARY OF M. L. REGRESSION MODEL RESULTS

Model	5 cross-fold validation; 70:30 train test split; Bayesian Optimization			
	RMSE	*R-squared*	*MAE*	*MSE*
Linear Regression (Linear)	4.067556	0.603732	3.136209	16.54501
Tree	3.825692	0.649456	2.906104	14.63592
SVM	3.740041	0.664977	2.849802	13.98791
Gaussian Process Regression	3.557933	0.696808	2.678641	12.65888
Ensemble	3.670035	0.677401	2.773259	13.46916
Neural Network	3.668477	0.677675	2.815893	13.45773

Similarly, evaluation measures (errors) were subjected to an ANOVA test to determine if there are significant differences in the performance across the models. Results show an f-value of 0.024 and a p-value of 1.00. This indicates the performances of the models are not statistically significant. However, considering RMSE, MSE, MAE, and R squared, Gaussian Process Regression was considered as the best-fit model.

IV. Results and Discussion

A. Model Performance Visualization

The performance of various machine learning models can be effectively visualized using Predicted versus Actual plots. These plots have the actual values of the outcome variable on the x-axis and the predicted values of the outcome variable on the y-axis. Ideally, the plotted points should be near the diagonal line, which represents a close match between the predicted and actual values. The Predicted versus Actual plots of the five models are presented in Fig. 2, Fig. 3, Fig. 4, Fig. 5, and Fig. 6.

Based on the results shown in Table III, the Gaussian Process Regression (GPR) model had the lowest RMSE and highest R-squared values, indicating that it performed the best overall at predicting the outcome variable. This is reflected in Fig. 4, where we can see that the points are more tightly clustered around the diagonal line compared to the other models. This indicates that the GPR model is able to best-fit in predictions.

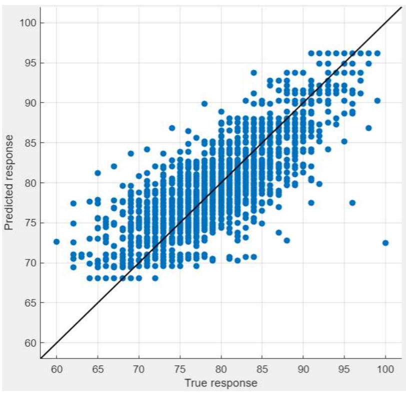

Fig. 2. Predicted versus Actual Plots of Tree

B. Experimental results

The hyperparameters for the Gaussian Process Regression (GPR) model were optimized using 5-fold cross-validation, a 70:30 train-test split, and Bayesian optimization. The optimized hyperparameters include Rational Isotropic Quadratic kernel function, a kernel scale of 0.4803, a sigma value of 0.0067, and a linear basis function. The data was not standardized in this optimization. By optimizing these hyperparameters, the GPR model was able to achieve better performance in predicting the outcome variable, compared to using default hyperparameters. Details of the hyperparameters can be seen in Table IV.

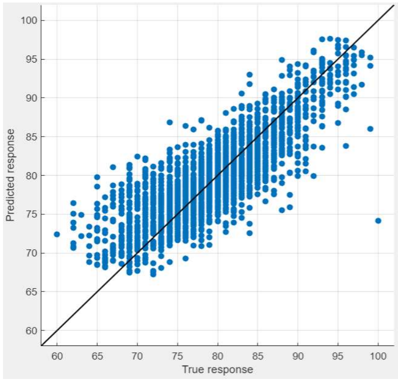

Fig. 3. Predicted versus Actual Plots of SVM

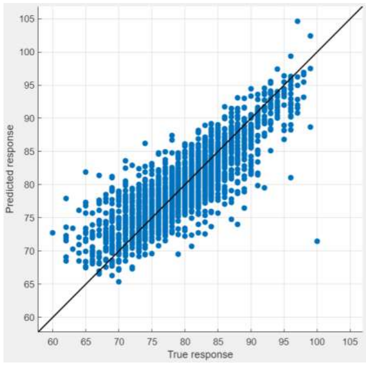

Fig. 4. Predicted versus Actual Plots of GPR

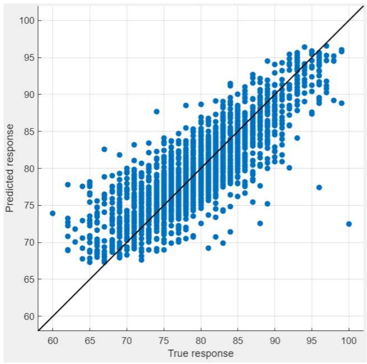

Fig. 5. Predicted versus Actual Plots of Ensemble

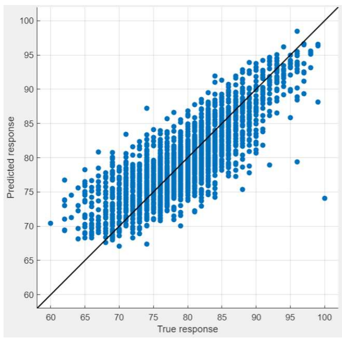

Fig. 6. Predicted versus Actual Plots of Neural Network

V. Conclusion and Future Work

We used machine learning algorithms to predict relative humidity levels based on weather variables such as temperature, wind speed, wind direction, cloud cover, and rainfall. After cleaning the data and validating the models using a 70:30 data splitting ratio and 5-fold cross-validation, we trained various machine learning algorithms to predict humidity levels. To improve their performance, we used Bayesian optimization to tune hyperparameters for five algorithms, including Decision Tree, Gaussian Process Regression (GPR), Support Vector Machines (SVM), Ensemble, and Neural Network. The GPR algorithm with optimized hyperparameters achieved the best performance in predicting humidity levels.

TABLE IV. Hyperparameter Values

Hyperparameter	Value
Sigma	0.0067
Basis function	Linear
Kernel function	Rational Isotropic Quadratic
Kernel scale	0.4803
Standardize data	FALSE

In future work, the focus should be on deploying the best-performing machine learning algorithm for humidity prediction in the Philippines. This would involve building a robust and scalable application that utilizes the selected algorithm to provide users with real-time or near-real-time humidity forecasts. Additionally, integrating data visualization and user-friendly interfaces would enhance the usability and accessibility of the application, making it valuable for various stakeholders such as farmers, healthcare professionals, and disaster management authorities.

Acknowledgment

We would like to extend our sincerest appreciation to PAGASA for providing us with the dataset used in our analysis. Their contribution has been invaluable in achieving the objectives of this research. We are grateful for their generosity and support.

References

[1] L. J. published, “What is relative humidity? And how does it impact you?,” livescience.com, Oct. 11, 2021. https://www.livescience.com/what-is-relative-humidity

[2] R. Feistel and J. W. Lovell-Smith, “Defining relative humidity in terms of water activity. Part 1: definition,” *Metrologia*, vol. 54, no. 4, pp. 566–576, Jul. 2017.

[3] J. A. Tapia-Hernández, F. Rodríguez-Félix, and I. Katouzian, “Nanocapsule formation by electrospraying,” *Nanoencapsulation Technologies for the Food and Nutraceutical Industries*, pp. 320–345, 2017.

[4] Z. Pan et al., “The influence of meteorological factors and total malignant tumor health risk in Wuhu city in the context of climate change,” *BMC Public Health*, vol. 23, no. 1, Feb. 2023.

[5] “NASA - Humidity Relative to Earth’s Temperature,” www.nasa.gov. https://www.nasa.gov/vision/earth/lookingatearth/warmer_humidity.html

[6] S. C. Sherwood, “How Important Is Humidity in Heat Stress?,” *Journal of Geophysical Research: Atmospheres*, vol. 123, no. 21, pp. 11, 808–11, 810, Nov. 2018.

[7] E. Eccel, “Estimating air humidity from temperature and precipitation measures for modelling applications,” *Meteorological Applications*, vol. 19, no. 1, pp. 118–128, May 2011.

[8] N. Goad and D. J. Gawkrodger, “Ambient humidity and the skin: the impact of air humidity in healthy and diseased states,” *Journal of the European Academy of Dermatology and Venereology*, vol. 30, no. 8, pp. 1285–1294, Jun. 2016.

[9] N. Ding, H. L. Berry, and C. M. Bennett, “The Importance of Humidity in the Relationship between Heat and Population Mental Health: Evidence from Australia,” *PLOS ONE*, vol. 11, no. 10, p. e0164190, Oct. 2016.

[10] K. Qadeer, A. Ahmad, M. A. Qyyum, A.-S. Nizami, and M. Lee, “Developing machine learning models for relative humidity prediction in air-based energy systems and environmental management applications,” *Journal of Environmental Management*, vol. 292, p. 112736, Aug. 2021.

[11] B. Bochenek and Z. Ustrnul, “Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives,” *Atmosphere*, vol. 13, no. 2, p. 180, Feb. 2022.

[12] H. Choi, T. Moon, D. H. Jung, and J. E. Son, “Prediction of Air Temperature and Relative Humidity in Greenhouse via a Multilayer Perceptron Using Environmental Factors,” *Protected horticulture and Plant Factory*, vol. 28, no. 2, pp. 95–103, Apr. 2019.

[13] G. Mustafaraj, G. Lowry, and J. Chen, “Prediction of room temperature and relative humidity by autoregressive linear and nonlinear neural network models for an open office,” *Energy and Buildings*, vol. 43, no. 6, pp. 1452–1460, Jun. 2011.

[14] H. Tao, S. M. Awadh, S. Q. Salih, S. S. Shafik, and Z. M. Yaseen, “Integration of extreme gradient boosting feature selection approach with machine learning models: application of weather relative humidity prediction,” *Neural Computing and Applications*, Aug. 2021.

[15] E. Arulmozhi, J. K. Basak, T. Sihalath, J. Park, H. T. Kim, and B. E. Moon, “Machine Learning-Based Microclimate Model for Indoor Air Temperature and Relative Humidity Prediction in a Swine Building,” *Animals*, vol. 11, no. 1, p. 222, Jan. 2021.

[16] X. Shi, W. Lu, Y. Zhao, and P. Qin, “Prediction of Indoor Temperature and Relative Humidity Based on Cloud Database by Using an Improved BP Neural Network in Chongqing,” *IEEE* Access, vol. 6, pp. 30559–30566, 2018.

[17] T. G. Özbalta, A. Sezer, and Y. Yıldız, “Models for Prediction of Daily Mean Indoor Temperature and Relative Humidity: Education Building in Izmir, Turkey,” *Indoor and Built Environment*, vol. 21, no. 6, pp. 772–781, Sep. 2011.

[18] Z. Li, H. Zou, and B. Qi, “Application of ARIMA and LSTM in Relative Humidity Prediction,” *2019 IEEE 19th International Conference on Communication Technology (ICCT)*, Oct. 2019.

[19] M. I. Hutapea, Y. Y. Pratiwi, I. M. Sarkis, I. K. Jaya, and M. Sinambela, “Prediction of relative humidity based on long short-term memory network,” *AIP Conference Proceedings*, 2020.

[20] M. Adnan, R. M. Adnan, S. Liu, M. Saifullah, Y. Latif, and M. Iqbal, Prediction of Relative Humidity in a High Elevated Basin of Western Karakoram by Using Different Machine Learning Models. IntechOpen, 2021. Available: https://www.intechopen.com/chapters/77040

[21] A. T. Prihatno, H. Nurcahyanto, and Y. M. Jang, “Predictive Maintenance of Relative Humidity Using Random Forest Method,” *2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, Apr. 2021.

[22] S. AlSadi and T. Khatib, “Modeling of Relative Humidity Using Artificial Neural Network,” *Journal of Asian Scientific Research*, vol. 2, no. 2, pp. 81–86, Feb. 2012.

[23] M. Villafuerte Ii et al., “Astronomical Services Administration (DOST-PAGASA) 2 Manila Observatory.” Available: https://philjournalsci.dost.gov.ph/images/pdf/pjs_pdf/vol150no1/ClimDatPh_an_online_platform_for_data_acquisition_.pdf

[24] P. G. G. Lagrazon and J. B. Tan, "A Comparative Analysis of the Machine Learning Model for Crop Yield Prediction in Quezon Province, Philippines," *2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT)*, Bhopal, India, 2023, pp. 115-119.