A Comparative Analysis of the Machine Learning Model for Rainfall Prediction in Cavite Province, 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

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

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

Maria Rossana D. De Veluz
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
mrddeveluz@slsu.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

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

2023 IEEE World AI IoT Congress (AIIoT) | 979-8-3503-3761-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/AIIoT58121.2023.10174533

***Abstract*— Rainfall is crucial for flood prevention and comprehending the correlation between rainfall and flooding. Cavite province in the Philippines is vulnerable to flooding caused by heavy rainfall and climate change impacts. Early detection of flooding through early warning systems can prevent excessive damage loss and potentially save lives. It can also provide major savings in terms of monetary benefit and increased interagency coordination for rapid decision-making. Machine learning is an important tool for predicting rainfall which can be used to predict rainfall in the province. The objective of this study is to conduct a comparative analysis of various models for predicting daily rainfall, using relevant atmospheric features such as maximum, minimum, and mean temperature, relative humidity, wind speed, wind direction, cloud cover, pressure, and evaporation. The study seeks to identify the most effective model for accurately predicting rainfall in the Cavite Province to benefit the local community. Among the five machine learning models evaluated, the Gaussian Process Regression model demonstrated the highest accuracy in predicting daily rainfall. The findings of this study can be leveraged to mitigate the damage caused by flooding in the Cavite Province and serve as a useful reference for similar studies in other regions prone to flooding.**

Keywords*— *Cavite province, gaussian process regression, machine learning, rainfall

I. INTRODUCTION

The detection of rainfall plays a vital role in preventing floods, as it enables the identification of suitable thresholds that are liable to cause flood damage [1], understanding of the relationship between rainfall and flood probabilities [2], identify spatiotemporal and fluvial-pluvial sources of flooding [3], and evaluate the impact of climate change on flood and extreme precipitation events [4]. Additionally, rapid onset flooding, commonly known as flash floods, can rise within a brief duration of time, varying from a few minutes to a few hours, triggered by intense rainfall, a sudden release of water, or a failure of a dam or levee [5].

The province of Cavite in the Philippines faces the threat of river flooding due to climate change, including an increase in both the frequency and severity of heavy precipitation days and a rise in the occurrence of extreme rainfall events [6]. The rivers make the lower regions of the province particularly susceptible to flooding [7]. In 2018, Cavite declared a state of calamity due to widespread floods caused by occasional heavy rains [8]. In 2021, thousands were evacuated amid monsoon rain [9], and in 2013, heavy rain brought floods to Cavite with 395 mm falling in parts of the province, leading to a state of calamity being declared [10]. Flooding damages buildings via inundation and other forms of destruction caused by heavy rainfall [11]. Other areas in the Philippines, such as Kalinga province, are prone to flooding and at high risk of experiencing flooding and landslides [12].

Machine learning is an important tool for predicting rainfall and its intensity [13][14][15]. It has the capability to uncover latent patterns in past weather data, recognize pertinent atmospheric characteristics that lead to rainfall, and forecast the magnitude of daily precipitation [13]. It can be used to predict rainfall amounts in Cavite Province. For example, a rainfall forecast model based on the Attentive Interpretable Tubular Net (TabNet) Model was proposed in a study [16], and machine learning algorithm techniques such as linear regression were used to predict daily rainfall amounts using important atmospheric features [17]. Other studies have utilized deep learning mechanisms and historical weather data to build precipitation prediction models [18][19].

The advantages of early detection of flooding include the availability to prepare and warn people of impending danger, preventing excessive damage and loss, and potentially saving lives [20]. Early warning systems can also help reduce the risk of flooding through accurate forecasts and technical expertise [21]. Studies have also shown that continental flood early warning systems can provide major savings in terms of monetary benefit [22] and increased interagency coordination for rapid decision-making [23]. Thus, the primary objective of this study is to assess and compare different models for predicting daily rainfall using relevant atmospheric features. To accomplish this goal, the models will be evaluated based on various metric parameters. However, the study has certain limitations that should be considered, including the fact that the dataset used for analysis and evaluation is limited to the period from 2000 to 2022. Additionally, the research methods rely exclusively on the availability of atmospheric feature data from PAGASA, which may limit the breadth of analysis. Moreover, external factors such as climate change could influence rainfall patterns and cannot be fully accounted for in this study. Nonetheless, this research can still provide useful

979-8-3503-3761-7/23/$31.00 ©2023 IEEE

insights and recommendations for predicting daily rainfall in Cavite Province.

II. Related Works

According to Grace & Suganya [24], the growth rate of agricultural products is heavily dependent on the quantity of rainfall received. In order to assist farmers in crop planning, forecasts are generated to predict the amount of rainfall expected in a particular season. To identify the most effective algorithm for rainfall prediction using Indian data, researchers evaluated various models such as QPF, LR, and MLR. The analysis indicated that MLR achieved better accuracy than two other models, as determined by metrics such as MSE, RMSE, and Correlation.

In the study of Barrera-Animas et al. [25], a comparative analysis of rainfall estimation models is presented, comparing deep learning architectures and traditional machine learning algorithms. The analysis utilizes climatic data collected from five major UK cities between 2000 and 2022. Various models, including LSVR, XGBoost, BLSTM Networks, LSTM, SLSTM, an ensemble of Extra-trees Regressor, and Gradient Boosting, were assessed based on differen metrics such as RMSE, and RMSLE. The results revealed that the SLSTM Network and BLSTM Network with two hidden layers consistently outperformed all other models.

In their study [26], Ridwan et al. employed two distinct methods involved in utilizing the Autocorrelation Function, while the second method employed Projected Error. Data from ten stations located within the Theissen polygon research region were collected and analyzed for various scenarios and time frames. The study employed four machine learning algorithms, including NNR, DFR, BTDR, and BLR, results indicated that BTDR exhibited the highest coefficient determination for method 1, while both BTDR and DRF were found to have acceptable results for method 2.

[27] aimed to identify the most suitable machine learning approach based on data analysis for predicting the mean daily and monthly precipitation. Their evaluation and comparison of various models, including support SVR, KNN, MARS, ANN, and a hybrid multi-model approach, was based on metrics such as Coefficient of Efficiency (CE), Persistence Index (PI), and RMSE. The study found that the hybrid multi-model method produced more accurate predictions for daily rainfall, while SVR was the top-performing model for monthly rainfall compared to the other models.

In their study [28], Appiah-Badu et al. assess the performance of KNN, XGB, MLP, RF, and DT in forecasting the occurrence of rainfall. The climatic attributes were taken from Ghana Meteorological Agency and covered the years 1980-2019. They used accuracy, f1-score, recall, and precision as their evaluation metrics and found that MLP, RF, and XGB performed well.

Aswin et al. [29] performed a research in which they utilized LSTM and ConvNet architecture to model and predict global monthly average rainfall for 10368 geographical locations around the world for a period of 468 months. ConvNet appears to be extremely promising for 100 training epochs when they evaluate the two deep learning architectures using RMSE and MAPE but claim that both are effective and efficient models.

The research of [30] assesses the performance of different models such as SVR, LSTM, BPNN, and LR to predict rainfall. The analysis is conducted using past rainfall data for the years 1901-2015 that were acquired from Narendra Nagar. The study finds that BPPNN outperforms and offers the best inferences compared to other models using the MSE, MAE, and RMSE as performance metrics.

The aim of the investigation in [31] is to explore the potential of a combined model consisting of RNNs and SVMs, known as RSVR, for predicting rainfall depth values. To accomplish this, the authors employ the CPSO to determine the parameters of an SVR model. An example of rainfall data from Northern Taiwan during typhoon periods is used to demonstrate the effectiveness of the proposed RSVRCPSO model. The empirical results demonstrate that the suggested model is highly accurate and effective, indicating that the RSVRCPSO model has the potential to be a valuable alternative for forecasting rainfall values.

In [32], the researchers propose a hybrid deep learning (DL) technique with multi-layer perceptrons to predict daily rainfall over multiple steps. The hybrid model takes input incorporating data with nine variables obtained from GCM. The estimates of rainfall are typically less accurate than those of meteorological variables using GCM, but the suggested scheme leverages the capability of GCM to simulate meteorological variables and thus contributes to enhancing the accuracy of rainfall forecasting. The hybrid Conv1d-MLLP model has been implemented in distinct locations across various meteorological regimes, and the result indicates that it more accurately captures the intricate connection between the predictor variables and the daily variation in rainfall. The benefit of this approach stems from the combination of potentials from several methods for obtaining the hidden characteristics of hydrometeorological associations.

The suggested work in [33] tries to predict rainfall by combining various machine learning and forecasting methodologies. Despite the fact that rainfall depends on a wide range of factors, we may achieve outstanding classification accuracy with a small number of factors. The classification of rainfall into eight separate groups is also shown to provide us with satisfactory accuracy. The RMSE measure is used to validate the anticipated values. Based on empirical evidence, the maximum temperature was found to be most accurately forecasted using ARIMA, while neural networks were found to be the most effective approach for predicting minimum temperature. Additionally, SVR was identified as the optimal method for forecasting relative humidity and wind speed. Accuracy, precision, and recall are used to gauge the validity of the classification. Random forest performs best for classifying rainfall, according to the ROC curve for all classifiers.

III. Research Methods

A. Study Area

Cavite is a province located in the CALABARZON region of the Philippines. It is situated on the southern shore of Manila Bay and is part of the island of Luzon with a total land area of approximately 1,427.06 square kilometers. The province has a total population of 4,831,240 as of the 2020 Consensus. Fig. 1 provides a visual representation of the province's geography.

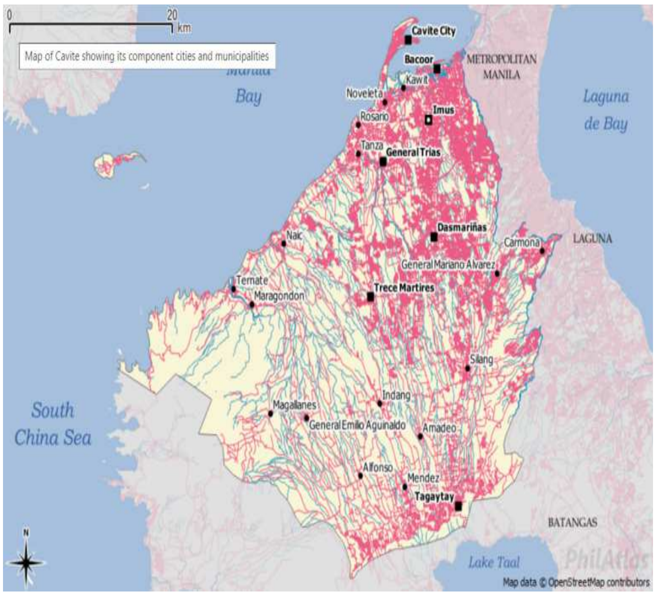

Fig. 1. Cavite Province, Philippines

B. Dataset

The data sources for this study came from publicly available records in the Philippines. The PAGASA is in charge of evaluating and predicting the weather, releasing alerts for floods and typhoons, giving out public weather forecasts and advice, as well as providing specialized information services related to weather conditions. The weather dataset is composed of maximum, minimum, and mean temperature, relative humidity, wind speed, wind direction, cloud cover, pressure, evaporation, and rainfall from the year 2000-2022 on a daily basis.

C. Data Preprocessing

1) Data Cleaning: Data cleaning is a crucial stage in the machine learning pipeline as it ensures the data's accuracy, consistency, and reliability for model training and evaluation. It involves various techniques that depend on the problem and data sources and constitutes a significant initial step in the data analytic process. The dataset was cleaned by fixing incorrect, incomplete, duplicates, filling in missing values, smoothing, and removing outliers.

2) Feature Selection: The dataset contains 9 features-maximum, minimum, and mean temperature, relative humidity, wind speed, wind direction, cloud cover, pressure, evaporation, and 1 target variable which is rainfall.

3) Validation Scheme and Data Splitting: The dataset used 10 cross-fold validation with a 70:30 ratio for the train test split where the large portion was used for training and the smaller portion was used for testing. Cross-validation and train-test split can provide a more robust evaluation of a model's performance. The application of cross-validation assists in ensuring that the model is not overfitting to any particular subset of the data, while the train-test split provides an independent evaluation of the model's performance on unseen data.

D. Training and Testing

The dataset was subjected to training and testing using various models and their respective variations as part of the research study. It is to identify the most optimal model and variation based on their performance and accuracy. Through multiple trials using different models and variations, the collected data was analyzed to determine the most effective approach.

IV. Results and Discussion

Based on the correlation results in fig. 2, it appears that rainfall is negatively correlated with evaporation, maximum, minimum, and mean temperature, with the strongest negative correlation being with pressure. This suggests that as these temperature-related variables decrease, rainfall amount tends to increase. On the other hand, rainfall is positively correlated with relative humidity and cloud cover, indicating that higher levels of these variables are associated with the increased amount of rainfall. Wind speed and wind direction have relatively weak correlations with rainfall.

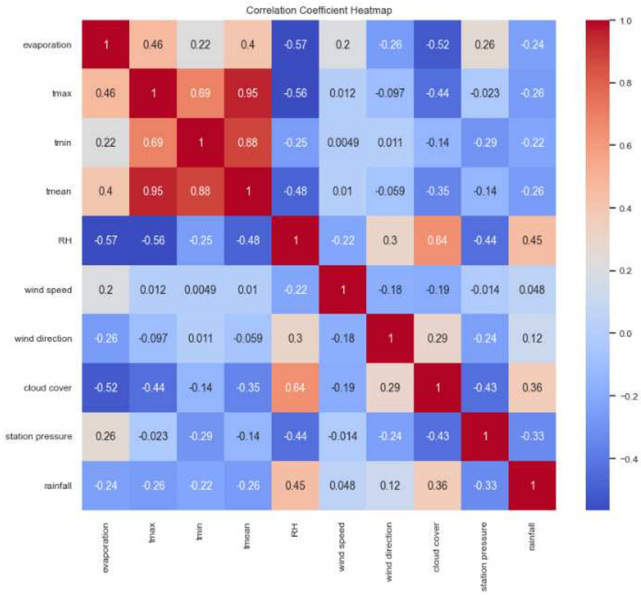

Fig. 2. Correlation Coefficient Heatmap

The table presents a summary of various machine learning models, analyzed using four parameter metrics, without any hyperparameter tuning. The results show the root mean squared error (RMSE) and R-squared values for each model, obtained through 10 cross-fold validation and a 70:30 train-test ratio. The models include linear regression, tree-based models, SVMs, ensemble models, Gaussian process regression, neural networks, and kernel regression. The top-performing models with the lowest RMSE values and highest R-squared values are Gaussian process regression (with the squared exponential and Matern 5/2 kernels), tree-based models (Fine Tree and Bagged Trees), and neural networks (Trilayered Neural Network and Wide Neural Network).

TABLE I. Summary of Machine Learning Regression Model Results

Model	**10 cross-fold validation, 70:30 train-test ratio**			
	RMSE	***Rsquared***	***MSE***	***MAE***
Linear Regression (Linear)	0.00415	0.981113	1.72E-05	0.003256
Linear Regression (Interactions Linear)	0.001618	0.99713	2.62E-06	0.001227
Linear Regression (Robust Linear)	0.00422	0.980461	1.78E-05	0.003198
Stepwise Linear Regression	0.001618	0.997128	2.62E-06	0.001236

Tree (Fine Tree)	0.000849	0.99921	7.20E-07	0.000463
Tree (Medium Tree)	0.001027	0.998843	1.06E-06	0.00064
Tree (Coarse Tree)	0.001701	0.996825	2.89E-06	0.001226
SVM (Linear SVM)	0.004212	0.980542	1.77E-05	0.003281
SVM (Quadratic SVM)	0.002295	0.994224	5.27E-06	0.001912
SVM (Cubic SVM)	0.002263	0.994382	5.12E-06	0.001922
SVM (Fine Gaussian SVM)	0.003075	0.989625	9.46E-06	0.002614
SVM (Medium Gaussian SVM)	0.002286	0.994266	5.23E-06	0.001901
SVM (Coarse Gaussian SVM)	0.003525	0.986372	1.24E-05	0.00285
Ensemble (Boosted Trees)	0.002008	0.995575	4.03E-06	0.001556
Ensemble (Bagged Trees)	0.000775	0.999341	6.01E-07	0.000458
Gaussian Process Regression (Squared Exponential GPR)	0.000526	0.999696	2.77E-07	0.000339
Gaussian Process Regression (Matern 5/2 GPR)	0.000476	0.999751	2.27E-07	0.000288
Gaussian Process Regression (Exponential GPR)	0.000459	0.999769	2.10E-07	0.00027
Gaussian Process Regression (Rational Quadtraic GPR)	0.000472	0.999755	2.23E-07	0.000277
Neural Network (Narrow Neural Network)	0.001774	0.996548	3.15E-06	0.001347
Neural Network (Medium Neural Network)	0.001547	0.997375	2.39E-06	0.00117
Neural Network (Wide Neural Network)	0.001447	0.997702	2.10E-06	0.001123
Neural Network (Bilayered Neural Network0	0.00312	0.98932	9.74E-06	0.002498
Neural Network (Trilayered Neural Network)	0.001256	0.998269	1.58E-06	0.000934
Kernel (SVM Kernel)	0.002208	0.994654	4.87E-06	0.001852
Kernel (Least Squares Regression Kernel)	0.003169	0.988982	1.00E-05	0.002446

Hyperparameter Tuning

After tuning the hyperparameters, the performance of various machine learning models was evaluated using 10 cross-fold validation and a 70:30 train-test ratio. The Tree model achieved a low RMSE of 0.000812 and a high R-squared value of 0.999277, indicating a good fit to the data. The SVM model had a slightly higher RMSE of 0.00089 but still performed well with an R-squared value of 0.999132. The Gaussian Process Regression model had the lowest RMSE of 0.000362 and the highest R-squared value of 0.999856, indicating excellent performance. The Ensemble model and Neural Network model achieved comparable results, with RMSE values of 0.00061 and 0.001173, and R-squared values of 0.999592 and 0.998492, respectively. Overall, the results suggest that the Tree, SVM, Gaussian Process Regression, Ensemble, and Neural Network models are all capable of achieving good performance in this dataset, with the Gaussian Process Regression model being the most accurate.

TABLE II. SUMMARY OF OPTIMIZED VALUES

Model	**10 cross-fold validation, 70:30 train-test ratio**			
	RMSE	***Rsquared***	***MSE***	***MAE***
Tree	0.000812	0.999277	6.59E-07	0.000446
Support Vector Machine	0.00089	0.999132	7.91E-07	0.000735
Gaussian Process Regression	0.000362	0.999856	1.31E-07	0.000208
Ensemble	0.00061	0.999592	3.72E-07	0.000317
Neural Network	0.001173	0.998492	1.38E-06	0.000818

Experimental Results

Table III shows the optimized hyperparameter values for the Gaussian Process Regression model, which resulted in its optimal performance. By tuning the hyperparameters, the model was able to achieve better accuracy in its predictions.

TABLE III. HYPERPARAMETER VALUES

Hyperparameter	**Value**
Sigma	0.0059
Basis function	Constant
Kernel function	Nonisotropic Exponential
Standardize data	true

Model Performance Visualization

Predicted versus actual plots are shown in fig. 3 are a common way to visually evaluate the performance of regression models. The plots show how well the predicted values from the model match the actual values. The Tree, SVM, Ensemble, and Gaussian Process Regression models all showed excellent performance with low RMSE values ranging from 0.000362 to 0.000812 and high R-squared values ranging from 0.999132 to 0.999856. These models are expected to have a high correlation between their predicted and actual values, resulting in a plot that follows a diagonal line. However, the Neural Network model showed lower performance, with a higher RMSE of 0.001173 and a lower R-squared of 0.998492. As a result, the predicted versus actual plot for this model is expected to have a looser cluster of data points that deviate more from the diagonal line, indicating a higher variance between the predicted and actual values.

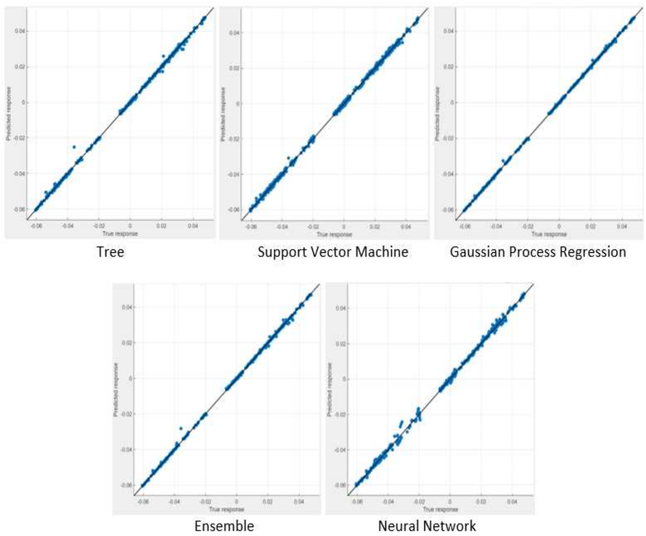

Fig. 3. Predicted versus Actual Plots

V. Conclusion

This study provides a comparative analysis of different machine learning models to predict daily rainfall in the Cavite Province of the Philippines. The results indicate that the Tree, SVM, Gaussian Process Regression, Ensemble, and Neural Network models are all capable of achieving good performance in this dataset. Among these models, the Gaussian Process Regression model showed the best performance, with the lowest RMSE and highest R-squared values. The findings of this study can be utilized to improve early warning systems and flood preparedness in the Cavite Province, potentially preventing excessive damage and loss, and saving lives. Moreover, the approach used in this study can serve as a template for similar studies in other regions prone to flooding.

VI. Future Work

In the future, the trained and tested machine learning algorithm can be deployed to create an application that predicts the amount of rainfall in Cavite. This application will serve as a valuable tool for the people of Cavite, helping them to prepare and take necessary precautions in the event of heavy rainfall.

Acknowledgment

We wish to express our gratitude to the PAGASA for their generous support and assistance in providing the data needed for this study. Their expertise and dedication to meteorology and climatology have been invaluable to the success of this research. The study would not have been possible without their unwavering commitment to advancing scientific knowledge in the Philippines.

References

[1] K. Papagiannaki *et al.*, "Identification of Rainfall Thresholds Likely to Trigger Flood Damages across a Mediterranean Region, Based on Insurance Data and Rainfall Observations," *Water*, vol. 14, no. 6, p. 994, Mar. 2022, doi: https://doi.org/10.3390/w14060994.

[2] K. Breinl, D. Lun, H. Müller-Thomy, and G. Blöschl, "Understanding the relationship between rainfall and flood probabilities through combined intensity-duration-frequency analysis," *Journal of Hydrology*, vol. 602, p. 126759, Nov. 2021, doi: https://doi.org/10.1016/j.jhydrol.2021.126759.

[3] S. Try, T. Sayama, C. Oeurng, T. Sok, S. Ly, and S. Uk, "Identification of the spatio-temporal and fluvial-pluvial sources of flood inundation in the Lower Mekong Basin," Geoscience Letters, vol. 9, no. 1, Jan. 2022, doi: https://doi.org/10.1186/s40562-022-00215-0.

[4] H. Tabari, "Climate change impact on flood and extreme precipitation increases with water availability," *Scientific Reports*, vol. 10, no. 1, p. 13768, Aug. 2020, doi: https://doi.org/10.1038/s41598-020-70816-2.

[5] US, "Floods," *Weather.gov*, 2019. https://www.weather.gov/pbz/floods

[6] "Think Hazard - Cavite - River flood," *www.thinkhazard.org*. https://www.thinkhazard.org/en/report/24228-philippines-region-iv-a-calabarzon-cavite/FL

[7] "THE STUDY ON COMPREHENSIVE FLOOD MITIGATION FOR CAVITE LOWLAND AREA IN THE REPUBLIC OF THE PHILIPPINES FINAL REPORT Volume 1: Master Plan Study," 2009. Available: https://openjicareport.jica.go.jp/pdf/11925864_01.pdf

[8] E. Amoroso, "Bacoor City in Cavite under state of calamity," *Philstar.com*. https://www.philstar.com/nation/2018/07/20/1834995/bacoor-city-cavite-under-state-calamity

[9] "Thousands evacuate amid Philippines monsoon rain," *UPI*. https://www.upi.com/Top_News/World-News/2021/07/24/Phillippines-monsoon-flood/8571627145460/

[10] "Heavy rain brings floods to Philippines; markets, offices shut," Reuters, Aug. 19, 2013. Available: https://www.reuters.com/article/us-philippines-floods/heavy-rain-brings-floods-to-philippines-markets-offices-shut-idUKBRE97I04720130819

[11] "PROVINCIAL HAZARD PROFILE." Available: https://wvphilippineshea.files.wordpress.com/2015/06/cavite-pdrrm-plan.pdf

[12] "List Of Flood Prone Areas In The Philippines: 10 Heavily Flooded Zones," *philtoyota.com*. https://philtoyota.com/stories/list-of-flood-prone-areas-in-the-philippines-10-heavily-flooded-zones-str93

[13] A. Rahman *et al.*, "Rainfall Prediction System Using Machine Learning Fusion for Smart Cities," *Sensors*, vol. 22, no. 9, p. 3504, May 2022, doi: https://doi.org/10.3390/s22093504.

[14] C. M. Liyew and H. A. Melese, "Machine learning techniques to predict daily rainfall amount," *Journal of Big Data*, vol. 8, no. 1, Dec. 2021, doi: https://doi.org/10.1186/s40537-021-00545-4.

[15] B. Neo, "Predicting Rain with Machine Learning," *Medium*, Mar. 29, 2022. https://towardsdatascience.com/predicting-rain-with-machine-learning-2acf80017c44

[16] J. Yan, T. Xu, Y. Yu, and H. Xu, "Rainfall Forecast Model Based on the TabNet Model," *Water*, vol. 13, no. 9, p. 1272, Apr. 2021, doi: https://doi.org/10.3390/w13091272.

[17] M. S. Hanoon *et al.*, "Developing machine learning algorithms for meteorological temperature and humidity forecasting at Terengganu state in Malaysia," *Scientific Reports*, vol. 11, no. 1, Sep. 2021, doi: https://doi.org/10.1038/s41598-021-96872-w.

[18] A. Alnawas, N. Al-Khafaji, and H. Azeez, "Precipitation Forecast for Thi-Qar Province of Iraq Utilizing Machine Learning Approaches," *Proceedings of 2nd International Multi-Disciplinary Conference Theme: Integrated Sciences and Technologies, IMDC-IST 2021, 7-9 September 2021, Sakarya, Turkey*, 2022, doi: https://doi.org/10.4108/eai.7-9-2021.2314897.

[19] "Build software better, together," *GitHub*. https://github.com/topics/rainfall-prediction (accessed Feb. 21, 2023).

[20] O. Hudson, "The Methods and Benefits of Flood Monitoring," *AZoCleantech.com*, Mar. 10, 2022. https://www.azocleantech.com/article.aspx?ArticleID=1476

[21] "Flood early warning systems: A review of benefits, challenges and prospects," www.preventionweb.net. https://www.preventionweb.net/publication/flood-early-warning-systems-review-benefits-challenges-and-prospects

[22] J. Thielen Del Pozo *et al.*, "The benefit of continental flood early warning systems to reduce the impact of flood disasters," *JRC Publications Repository*, Jan. 20, 2016. https://publications.jrc.ec.europa.eu/repository/handle/JRC97266 (accessed Apr. 12, 2022).

[23] S. Potter, S. Harrison, and P. Kreft, "The benefits and challenges of implementing impact-based severe weather warning systems: Perspectives of weather, flood, and emergency management personnel," *Weather, Climate, and Society*, Jan. 2021, doi: https://doi.org/10.1175/wcas-d-20-0110.1.

[24] R. K. Grace and B. Suganya, "Machine Learning based Rainfall Prediction," *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Mar. 2020, doi: https://doi.org/10.1109/icaccs48705.2020.9074233.

[25] A. Y. Barrera-Animas, L. O. Oyedele, M. Bilal, T. D. Akinosho, J. M. D. Delgado, and L. A. Akanbi, "Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting," *Machine Learning with Applications*, p. 100204, Nov. 2021, doi: https://doi.org/10.1016/j.mlwa.2021.100204.

[26] W. M. Ridwan, M. Sapitang, A. Aziz, K. F. Kushiar, A. N. Ahmed, and A. El-Shafie, "Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia," *Ain Shams Engineering Journal*, Nov. 2020, doi: https://doi.org/10.1016/j.asej.2020.09.011.

[27] S. M. Sumi, M. F. Zaman, and H. Hirose, "A rainfall forecasting method using machine learning models and its application to the Fukuoka city case," *International Journal of Applied Mathematics and Computer Science*, vol. 22, no. 4, pp. 841–854, Dec. 2012, doi: https://doi.org/10.2478/v10006-012-0062-1.

[28] N. K. A. Appiah-Badu, Y. M. Missah, L. K. Amekudzi, N. Ussiph, T. Frimpong, and E. Ahene, "Rainfall Prediction Using Machine Learning Algorithms for the Various Ecological Zones of Ghana," *IEEE Access*, vol. 10, pp. 5069–5082, 2022, doi: https://doi.org/10.1109/access.2021.3139312.

[29] S. Aswin, P. Geetha, and R. Vinayakumar, "Deep Learning Models for the Prediction of Rainfall," *IEEE Xplore*, Apr. 01, 2018. https://ieeexplore.ieee.org/abstract/document/8523829 (accessed Nov. 10, 2022).

[30] S. Srivastava, N. Anand, S. Sharma, S. Dhar, and L. K. Sinha, "Monthly Rainfall Prediction Using Various Machine Learning Algorithms for Early Warning of Landslide Occurrence," *IEEE Xplore*, Jun. 01, 2020. https://ieeexplore.ieee.org/document/9154184

[31] W.-C. Hong, "Rainfall forecasting by technological machine learning models," *Applied Mathematics and Computation*, vol. 200, no. 1, pp. 41–57, Jun. 2008, doi: https://doi.org/10.1016/j.amc.2007.10.046.

[32] M. I. Khan and R. Maity, "Hybrid Deep Learning Approach for Multi-Step-Ahead Daily Rainfall Prediction Using GCM Simulations," *IEEE Access*, vol. 8, pp. 52774–52784, 2020, doi: https://doi.org/10.1109/access.2020.2980977.

[33] U. Shah, S. Garg, N. Sisodiya, N. Dube, and S. Sharma, "Rainfall Prediction: Accuracy Enhancement Using Machine Learning and Forecasting Techniques," *IEEE Xplore*, Dec. 01, 2018. https://ieeexplore.ieee.org/document/8745763