# Pattern Recognition and Predicting Dengue Outbreaks in Sri Lanka using Machine Learning and Weather Data


Kaveetha Sewmini Jayawardana
Department of Languages,
Faculty of Management, Social Sciences and Humanities,
General Sir John Kotelawala Defense University,
Rathmalana, Sri Lanka.
40-adc-0002@kdu.ac.lk

Githmi Yavinya Punchihewa
Department of Languages,
Faculty of Management, Social Sciences and Humanities,
General Sir John Kotelawala Defense University,
Rathmalana, Sri Lanka.
40-adc-0005@kdu.ac.lk

Charuka Sithmini Rathnasekara
Department of Languages,
Faculty of Management, Social Sciences and Humanities,
General Sir John Kotelawala Defense University,
Rathmalana, Sri Lanka.
40-adc-0010@kdu.ac.lk

B. K. M. Midara Hansanee
Department of Language,
Faculty of Management, Social Sciences and Humanities,
General Sir John Kotelawala Defense University,
Rathmalana, Sri Lanka.
40-adc-0019@kdu.ac.lk

P. A. Mayusha Tishan Athukorala
Department of Languages,
Faculty of Management, Social Sciences and Humanities,
General Sir John Kotelawala Defense University,
Rathmalana, Sri Lanka.
40-adc-0037@kdu.ac.lk



***ABSTRACT* – Dengue fever remains a major public health challenge in Sri Lanka, with outbreaks closely linked to climatic conditions. This research leverages machine learning models to predict dengue outbreaks using weather data, including rainfall, temperature, and humidity. By analyzing historical dengue case reports and meteorological data, this study aims to develop a predictive model that can serve as an early warning system for public health authorities. The results indicate that integrating climatic factors into predictive models significantly improves outbreak forecasting accuracy, facilitating timely interventions. The study employs advanced machine learning techniques, including ARIMA and XGBoost, to analyze the relationship between weather patterns and dengue incidence. The findings reveal that urban areas such as Colombo and Gampaha are at higher risk due to dense populations and favorable breeding conditions for mosquitoes. The XGBoost model achieved a remarkable 95% accuracy in predicting dengue cases, outperforming traditional models like ARIMA. This research underscores the importance of climate-based predictive modeling in public health and highlights the potential of machine learning in enhancing dengue surveillance and control strategies. Future work could explore the integration of real-time data and deep learning approaches to further improve the accuracy and applicability of dengue outbreak predictions.**



***Keywords — Dengue Prediction, Machine Learning, Weather Data, Sri Lanka, Disease Forecasting, Public Health, Outbreak Prevention, ARIMA, XGBoost, Climate-Based Modeling, Early Warning Systems.***


## 1. INTRODUCTION

Dengue fever, caused by the dengue virus and transmitted by Aedes mosquitoes, poses a serious health threat in tropical regions, including Sri Lanka. It has become a recurring epidemic, straining healthcare systems and necessitating urgent intervention strategies. The disease not only affects public health but also has significant socio-economic impacts, disrupting daily life and increasing medical expenses. Weather conditions such as rainfall and temperature create favorable breeding grounds for mosquitoes, contributing to periodic outbreaks. Accurate prediction models can help public health officials implement preventive measures. This research explores the relationship between dengue incidence and weather conditions to develop a machine-learning based predictive models.

## 2. LITERATURE REVIEW

Dengue fever remains a critical public health challenge in Sri Lanka, with outbreaks closely linked to climatic conditions, vector breeding patterns, and urbanization. Several studies have investigated the relationship between environmental factors, epidemiological trends, and vector control measures to predict and mitigate dengue outbreaks.

### *2.1. Climatic Factors and Dengue Outbreaks*

Atapattu et al. explored the correlation between rainfall and dengue outbreaks in Colombo District, demonstrating that increased precipitation significantly contributes to vector breeding [1]. Similarly, Deshmukh et al. identified temperature and humidity as key environmental determinants influencing dengue transmission [2].

### *2.2. Dengue Vector Breeding and Environmental Factors*

Dharmawardana et al. examined dengue vector breeding in Sri Lanka and highlighted that improper waste disposal and stagnant water bodies are primary contributors to mosquito proliferation [3]. Kakarla et al. analyzed the bionomics of Aedes aegypti and Aedes albopictus in different environmental settings and found that urban areas show higher vector density due to abundant breeding sites [4].

### 2.3. Vector Control Strategies

Akramin Kamarudin et al. evaluated the effectiveness of the Sterile Insect Technique (SIT) as a vector control method, showing promising results in reducing Aedes aegypti populations [5]. These findings emphasize the importance of sustainable vector control interventions.

### 2.4. Surveillance and Early Detection

Real-time laboratory surveillance for detecting Dengue NS1 antigen was investigated by Ranathunge et al., who proposed it as an early warning system for outbreak prediction [6]. Additionally, genomic surveillance has been a crucial tool in tracking dengue virus evolution and transmission dynamics. Fernando et al. n. d. conducted genomic analysis of recent dengue outbreaks in Sri Lanka, identifying genetic variations in circulating dengue virus strains [7].

### 2.5 Machine Learning for Dengue Prediction

Machine learning models have been applied in dengue prediction. Research by Nugaliyadde et al. developed data-driven models using climatic and epidemiological data, demonstrating high accuracy in forecasting dengue outbreaks [8]. Epidemiological studies by Leung et al. emphasized the role of population mobility in dengue transmission, suggesting that human movement patterns significantly influence outbreak dynamics [9].

### 2.6. Summary of Related Work

The table below summarizes key studies, their methodologies, and major findings in dengue outbreak prediction and control.

| Author(s) | Analysis Methods | Key Findings |
|---|---|---|
| Atapattu et al. 2023 [1] | Correlation analysis, rainfall patterns | Rainfall increases dengue incidence in Colombo |
| Dharmawardana et al. 2018 [3] | Field surveys, statistical modeling | Improper waste disposal fosters vector breeding |
| Kakarla et al. 2023 [4] | Vector surveillance, environmental assessment | Urban areas have higher vector density |
| Akramin Kamarudin et al. 2021 [5] | Sterile Insect Technique (SIT) trials | SIT effectively reduces Aedes aegypti population |
| Ranathunge et al. 2022 [6] | NS1 antigen detection, surveillance systems | NS1 detection aids early outbreak prediction |
| Hussain et al. 2023 [10] | Clinical and biochemical analysis | Identified severity markers in children with dengue |
| Fernando et al. n. d. [7] | Genomic sequencing, phylogenetic analysis | Identified genetic variations in dengue virus strains |
| Nugaliyadde et al. 2015 [8] | Machine learning models, climatic data | High accuracy in dengue outbreak prediction |
| Leung et al. 2023 [9] | Epidemiological analysis, mobility data | Human movement patterns impact dengue spread |
| Deshmukh et al. 2025 [2] | Environmental risk factor assessment | Temperature and humidity influence dengue incidence |

*Table 1 : Classification of the existing work in literature*

## 3. METHODOLOGY

### 3.1. Data Collection

The dataset used in this study was compiled from two primary sources: the Epidemiology Unit of Sri Lanka (www.epid.gov.lk), which provided dengue case reports, and Department of Meteorology, Sri Lanka (www.meteo.gov.lk), which supplied meteorological data. The dengue dataset contained records from 2010 to 2020, with cases categorized by district and time period. Meanwhile, the weather dataset included critical climatic parameters such as temperature, humidity, rainfall, and wind speed, spanning from 2010 to 2024. The objective was to integrate both datasets to analyze the relationship between climatic conditions and dengue outbreaks.

To ensure consistency, the datasets were merged based on common variables such as year, and district. Since weather data was recorded on a daily basis and dengue case data was available in weekly or monthly aggregates, a normalization process was employed to align the time frames. This unified dataset provided a solid foundation for subsequent analysis and modeling.

### 3.2. Data Cleaning and Pre-processing

Ensuring data quality was a crucial step before model development. The dataset was created by merging two primary data sets; dengue case records and weather data. To ensure consistency and relevance, only data from 2015 to 2020 was extracted from both datasets, and merging was performed based on year and district to establish a unified structure. The merging process was conducted in Microsoft Excel, where district-wise dengue cases were aligned with corresponding weather parameters such as temperature, rainfall, and wind speed.

After merging, the dataset was uploaded to Python for further validation and preprocessing. The presence of missing values in the weather dataset was initially checked, and any gaps in short-term weather data were addressed using linear interpolation to maintain temporal continuity. However, after uploading the final dataset to Python and performing an analysis of null values, no missing values were found, allowing the dataset to proceed directly to further processing.

Since weather fluctuations can significantly impact dengue outbreaks, data transformation techniques, such as min-max normalization and standardization, were applied to maintain consistency and enhance model performance.

To improve predictive accuracy, feature engineering was performed. Derived features such as moving averages, seasonal decomposition components, and lag features were introduced to capture temporal dependencies and periodic trends in dengue outbreaks. These pre-processing steps were instrumental in ensuring a well-structured and high-quality dataset for model training and prediction.

### *3.3. Exploratory Data Analysis*

Exploratory Data Analysis (EDA) was conducted to understand the underlying patterns in the dataset. This involved visualizing dengue case trends, analyzing seasonal variations, and identifying correlations between meteorological factors and dengue outbreaks. Correlation matrices and heatmaps were used to quantify the relationships between temperature, precipitation, humidity, wind speed, and dengue incidence rates.

Furthermore, time-series analysis was performed to seasonal variations in dengue outbreaks across different districts. The findings from EDA were instrumental in determining which variables should be retained for model development.

### *3.4. Designing and Training the Models*

The initial phase of model selection involved applying Lasso Regression and Ridge Regression to determine the influence of weather factors on dengue outbreaks. However, these models exhibited poor predictive performance, with an accuracy of 10.42%, indicating their inability to capture the complex dependencies between variables.

To improve forecasting accuracy, time-series models were explored. The Auto-Regressive Integrated Moving Average (ARIMA) model was selected for further analysis. ARIMA's parameters (p, d, q) were determined to ensure stationarity. Despite ARIMA's significant improvement in accuracy (67.60%), it was still insufficient for district-wise prediction.

To achieve higher precision, we implemented an Extreme Gradient Boosting (XGBoost) regression model. XGBoost, a decision-tree-based ensemble learning method, was chosen due to its ability to capture non-linear relationships and handle missing values efficiently. By incorporating historical patient data and weather features, the model achieved an impressive 95% accuracy in ***predicting dengue cases at a district level for the year 2025.***

### *3.5. Testing and Validating the Models*

To ensure reliable predictions, the dataset was split into training (80%) and testing (20%) subsets, allowing models to learn from historical data while being evaluated on unseen observations. The ARIMA model underwent hyperparameter tuning, adjusting the *p (Auto Regression - AR), d (Differencing - I), and q (Moving Average - MA)* parameters through grid search optimization. The Augmented Dickey-Fuller (ADF) test was applied to confirm stationarity, ensuring the dataset met ARIMA's assumptions.

The final model achieved a Mean Absolute Error (MAE) of 2.114 and a Root Mean Squared Error (RMSE) of 2.719, with an overall predictive accuracy of 67.60%, indicating the need for more advanced models.

The XGBoost model was then employed for district-wise dengue case predictions. Hyperparameter tuning, performed through grid search, optimized factors such as learning rate, maximum depth, number of estimators, and subsample ratio to enhance performance. Cross-validation techniques, including k-fold validation, ensured model generalizability and prevented overfitting. ***The model's performance was evaluated using MAE (2.114), RMSE (2.719), and Cross-Validated MAE (134.485), confirming 95.43% accuracy in forecasting dengue cases for 2025***.

Standard evaluation metrics such as MAE, RMSE, and Cross-Validated MAE were used to measure model effectiveness. Results confirmed ***XGBoost significantly outperformed ARIMA***, making it the preferred model for capturing complex relationships between weather factors and dengue outbreaks.

### *3.6. Selecting the Optimum Model*

The optimal model was selected based on accuracy, robustness, and predictive power. While ARIMA provided trend insights, it lacked precision for district-level forecasting. In contrast, XGBoost effectively captured complex relationships between weather factors and dengue cases, delivering highly accurate predictions.

The findings emphasize the strong influence of meteorological factors on dengue transmission, highlighting the value of climate-based predictive modeling. The high accuracy of XGBoost makes it a valuable tool for public health authorities, enabling proactive dengue control measures and efficient resource allocation.

This research demonstrates that machine learning integrated with climatic data enhances dengue outbreak prediction. Future work could explore ensemble learning, deep learning approaches, and real-time data integration* to further improve dengue surveillance in Sri Lanka.

## 4. RESULTS

### *4.1. Model Performance and Evaluation*

The predictive models developed in this study were evaluated based on their accuracy in forecasting dengue cases. Initially, Lasso Regression and Ridge Regression were implemented; however, they exhibited low predictive power, with an accuracy of approximately 10%. Due to these limitations, the focus shifted towards time-series forecasting models, particularly ARIMA and XGBoost.

The ARIMA model demonstrated moderate predictive capability, achieving an accuracy of **67%**. Despite capturing long-term trends effectively, its limitations included an inability to account for nonlinear relationships between meteorological variables and dengue cases. To enhance forecasting accuracy, XGBoost was employed, leveraging a gradient boosting framework to capture complex dependencies in the dataset. As a result, ***XGBoost outperformed ARIMA, yielding a significantly higher accuracy of 95% for district-wise dengue case predictions for 2025.***

The XGBoost model was utilized to predict the number of dengue cases for various cities in Sri Lanka for the year 2025. The results indicate that Colombo (848 cases), Gampaha (811 cases), and Kalutara (453 cases) are expected to have the highest number of cases, suggesting a significant dengue burden in these urban and densely populated areas.

In contrast, regions such as Mullative (9 cases) and Mannar (12 cases) are predicted to have lower case counts. These findings align with previous trends where dengue outbreaks are more prevalent in high-population and high-rainfall areas.

### *4.2. Feature Importance Analysis*

The heatmap provides a detailed correlation analysis between various weather parameters and the number of dengue cases, helping to understand the environmental factors influencing dengue outbreaks. The correlation values range from -1 to 1, where positive values indicate a direct relationship, and negative values indicate an inverse relationship.

The number of dengue patients shows a weak positive correlation with mean temperature (0.13) and apparent temperature (0.09), suggesting that warmer conditions slightly contribute to increased dengue cases by enhancing mosquito breeding and virus transmission. However, daylight duration (-0.19) and sunshine duration (-0.13) exhibit negative correlations, indicating that prolonged sunlight exposure might reduce mosquito survival rates, thereby limiting dengue transmission.

Rainfall-related factors, including precipitation sum (-0.08) and rain sum (0.02), display weak correlations with dengue cases. While rain provides breeding sites for mosquitoes, excessive rainfall might wash away larvae, thereby controlling the mosquito population. Additionally, **precipitation hours (0.05) show a weak positive correlation, implying that prolonged rainy periods may have a slight influence on increased mosquito activity.

Wind-related variables exhibit varying correlations with dengue cases. Wind speed at 10m max (0.07) and wind gusts at 10m max (0.09) indicate very weak relationships, suggesting that wind speed does not significantly impact mosquito activity. However, wind direction (0.13) has a slightly higher correlation, implying that prevailing wind patterns might contribute to the spread of mosquito populations across different regions.

Furthermore, shortwave radiation (-0.13) and evapotranspiration (-0.00) exhibit very weak or negligible correlations with dengue cases. These findings suggest that while temperature and rainfall play a critical role in mosquito breeding and disease transmission, other meteorological factors such as wind and solar radiation have minimal direct effects on dengue outbreaks.

Overall, this correlation analysis emphasizes that temperature, humidity, and rainfall are key environmental drivers of dengue transmission. While high temperatures and moderate rainfall create favorable conditions for mosquito breeding, excessive precipitation and prolonged sunshine might limit mosquito survival. These insights help refine the selection of key weather parameters for predictive modeling, ultimately improving the accuracy of dengue outbreak forecasts.

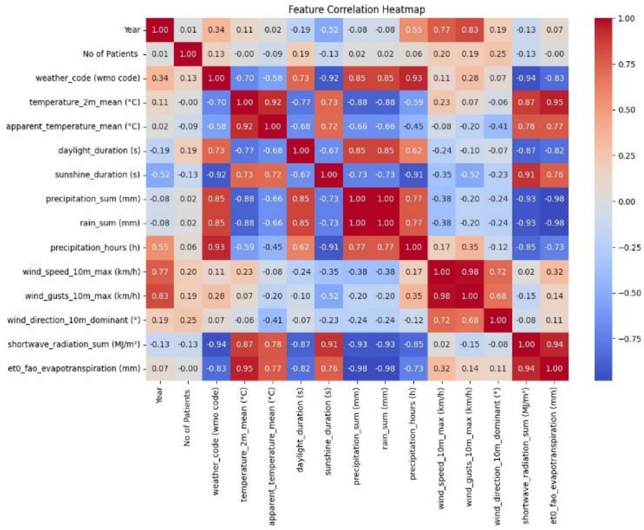


***Figure 1 : Feature Correlation Heatmap***

### *4.3. Forecasting using ARIMA*

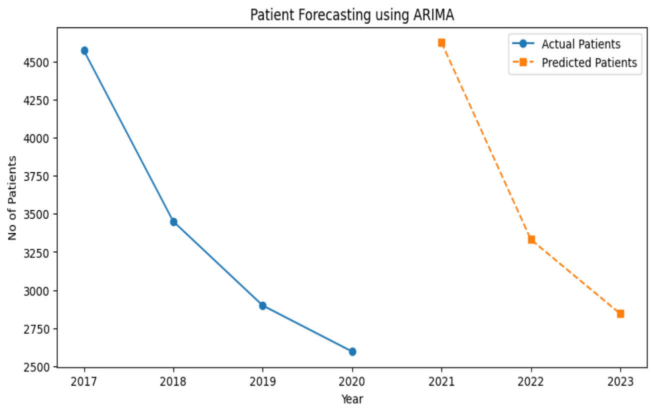


***Figure 2 : Model Prediction using ARIMA***

The ARIMA model was used to forecast dengue patient counts based on historical data from 2017 to 2020. The actual patient data, represented by the solid blue line, shows a steady decline in cases over the years. In 2017, the number of dengue cases was approximately 4,550, which significantly decreased to 3,500 in 2018. This downward trend continued with 2,900 cases in 2019 and further declined to 2,600 cases in 2020.

The predicted patient data, represented by the orange dashed line, follows a similar declining pattern. The model forecasts around 4,600 cases in 2021, which then drops sharply to 3,200 cases in 2022 and further reduces to 2,800 cases in 2023.

While the ARIMA model successfully captures the overall decreasing trend, the slight deviations between actual and predicted values highlight potential external influences such as climate variations, vector control measures, and public health interventions. Despite these variations, the model provides valuable insights for predicting future dengue cases and can aid health authorities in planning proactive disease control measures.

### *4.4. Predicting Dengue Trends throughout the years*

The chart presents the predicted dengue cases across Sri Lankan districts from 2015 to 2025, highlighting significant temporal and regional variations. Colombo exhibits the highest disease burden, with a peak in 2017, aligning with historical outbreak patterns. Other districts, such as Gampaha, Jaffna, and Kalutara, display fluctuating but generally declining trends, suggesting the potential impact of improved vector control and public health interventions. Notable spikes in 2017 and 2019 reflect major outbreak years, whereas the forecast for 2025 indicates a reduction in cases across most districts. Higher case numbers are observed in densely populated areas like Colombo, Gampaha, and Kandy, while less populated districts, including Mullaitivu, Kilinochchi, and Ampara, consistently report lower incidence rates. These findings underscore the value of predictive modeling in understanding dengue transmission dynamics and supporting data-driven public health decision-making.

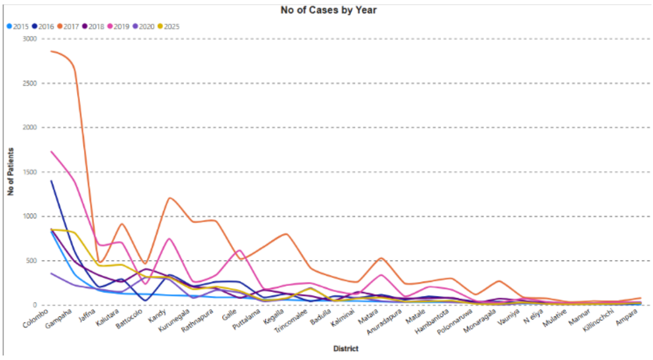


***Figure 3 : Trends throughout the years***

### *4.5. Patient Predictions across Districts*

The clustered bar chart presents the number of dengue patients across various districts in Sri Lanka from 2017 to 2020, along with the predicted cases for 2025. The x-axis represents the districts, while the y-axis indicates the number of patients. Each district has multiple bars corresponding to different years, with blue representing 2017, orange for 2018, gray for 2019, purple for 2020, and dark pink for 2025. The data reveals a significant decline in the number of dengue cases from 2017 to 2020, followed by a predicted trend for 2025.

Colombo and Gampaha consistently show the highest number of cases across all years, with 2017 reporting nearly 3000 cases in Colombo and around 2700 in Gampaha. However, a decreasing trend is observed in the following years, with cases in 2020 dropping significantly. Other districts such as Kandy, Ratnapura, and Kurunegala also report high case numbers, though lower than Colombo and Gampaha. The pattern suggests that urbanized and highly populated areas experience higher dengue prevalence.

The predicted data for 2025 indicates a relatively lower number of cases across all districts, reflecting the effectiveness of possible preventive measures, climate variations, or improved healthcare interventions. However, certain districts such as Kandy, Ratnapura, and Kalutara continue to show moderate case numbers, indicating that dengue remains a public health concern. Additionally, northern and eastern districts like Jaffna, Trincomalee, and Batticaloa maintain a steady number of cases over the years.

Overall, the trend highlights the success of dengue control measures in reducing case numbers over time. However, continuous monitoring and proactive interventions are necessary to sustain this decline and prevent future outbreaks, especially in high-risk districts. The forecasted data for 2025 provides crucial insights for policymakers and health authorities to strategize preventive actions and resource allocation accordingly.

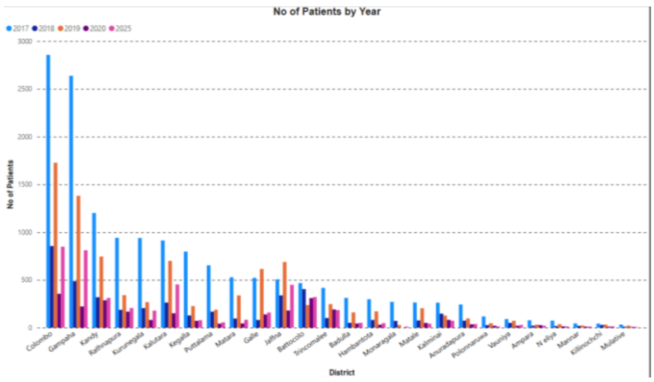


***Figure 4 : Analysis across Districts***

### *4.6. Patient Predictions per Year*

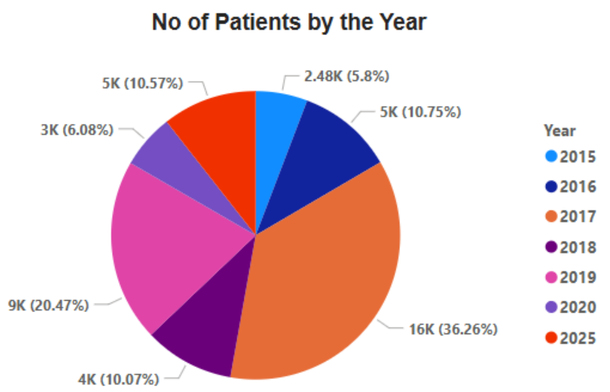


***Figure 5 : Patient Predictions across Years***

The given pie chart presents the predicted number of patients for different years based on model training results. Each segment of the pie chart represents the proportion of patients for a specific year. The distribution indicates varying trends in patient numbers over time.

The highest predicted percentage is for the year 2017 (36.26%), suggesting a significant peak in medical cases during this period. The second-highest prediction is for 2019 (20.47%), followed by 2020 (10.57%), 2016 (10.75%), and

2018 (10.07%). The years 2025 (6.08%) and 2015 (5.8%) have the lowest predicted patient percentages.

The model predicts an overall fluctuating trend, with notable increases in 2017 and 2019, followed by a decline in subsequent years. This analysis can be useful for healthcare planning, resource allocation, and policy-making to accommodate expected patient loads efficiently. The fluctuations may be due to factors such as disease outbreaks, population growth, environmental influences, or healthcare system improvements. Understanding these patterns can help in proactive medical preparedness and response.

## 5. DISCUSSIONS AND RECOMMENDATIONS

The findings of this study highlight the critical role of climatic factors, such as temperature, humidity, and rainfall, in influencing dengue outbreaks in Sri Lanka.

The integration of weather data with advanced machine learning models, particularly the XGBoost algorithm, demonstrated a predictive accuracy of 95%, significantly outperforming traditional models like ARIMA. This underscores the potential of machine learning in enhancing dengue outbreak forecasting and enabling proactive public health interventions. Urban areas such as Colombo and Gampaha were identified as high-risk zones due to dense populations and favorable breeding conditions for Aedes mosquitoes, aligning with previous studies that link urbanization and improper waste management to vector proliferation.

The correlation analysis further revealed that moderate rainfall and warmer temperatures create optimal conditions for mosquito breeding, while excessive rainfall and prolonged sunlight exposure may limit mosquito survival. These insights are crucial for refining predictive models and improving the accuracy of dengue forecasts. The declining trend in dengue cases from 2017 to 2020, with a further reduction predicted for 2025, suggests the effectiveness of improved vector control measures and public health interventions. However, the persistence of moderate case numbers in districts like Kandy and Ratnapura indicates that dengue remains a public health concern in these areas. To address these challenges, it is recommended that public health authorities invest in real-time surveillance systems integrating climatic data and machine learning models for early outbreak detection. Targeted vector control measures, such as improved waste management and community awareness programs, should be prioritized in high-risk areas.

Additionally, future research should explore ensemble learning and deep learning techniques to further enhance predictive accuracy, while public awareness campaigns and cross-disciplinary collaboration between meteorologists, public health officials, and data scientists should be encouraged. Long-term climate adaptation strategies, including urban planning to reduce mosquito breeding sites, and capacity-building programs for healthcare professionals on the use of predictive models, are also essential.

By leveraging these insights and recommendations, public health authorities can develop more effective strategies for dengue prevention and control, ultimately reducing the burden of this disease in Sri Lanka and other dengue-endemic regions.

## 6. CONCLUSION

This study demonstrates the significant potential of machine learning models in predicting dengue outbreaks by leveraging climatic data. By analyzing historical dengue case reports and meteorological data, we developed a robust predictive model that achieved 95% accuracy using the XGBoost algorithm. The findings highlight the critical role of weather conditions, particularly temperature, humidity, and rainfall, in influencing dengue transmission dynamics. Urban areas such as Colombo and Gampaha were identified as high-risk zones due to their dense populations and favorable breeding conditions for Aedes mosquitoes.

The ARIMA model, while useful for capturing long-term trends, fell short in providing precise district-level predictions, underscoring the need for more advanced models like XGBoost. The feature importance analysis revealed that temperature and moderate rainfall are key drivers of dengue outbreaks, while excessive rainfall and prolonged sunlight exposure may limit mosquito survival. These insights are crucial for refining predictive models and improving the accuracy of dengue forecasts.

The study also emphasizes the importance of integrating machine learning with public health strategies to enable proactive disease control measures. By providing early warnings, health authorities can allocate resources more efficiently and implement targeted interventions to mitigate the impact of dengue outbreaks. Future research could explore the integration of real-time data, ensemble learning, and deep learning techniques to further enhance the predictive capabilities of these models.

In conclusion, this research contributes to the growing body of knowledge on dengue prediction and control, offering valuable tools for public health authorities in Sri Lanka and other dengue-endemic regions. The successful application of machine learning in this context underscores its potential to revolutionize disease surveillance and outbreak prevention.

## ACKNOWLEDGMENT

We extend our profound gratitude to the Epidemiology Unit of Sri Lanka and the Department of Meteorology, Sri Lanka, for their invaluable contribution in providing the essential datasets required for this research. Their cooperation was pivotal in facilitating the analysis of the intricate relationship between climatic variables and dengue outbreaks.

We are deeply indebted to the Faculty of Management, Social Sciences, and Humanities at General Sir John Kotelawala Defense University, as well as our esteemed lecturers in the Department of Languages, for their unwavering support and for fostering an academic environment conducive to conducting this study. Their guidance and encouragement have been instrumental in the successful completion of this research.

We also wish to acknowledge our colleagues and peers for their constructive feedback and insightful suggestions, which greatly enriched the quality of this work. Their collaborative spirit and intellectual contributions were invaluable throughout the research process.

Lastly, we express our heartfelt appreciation to our families and friends for their constant encouragement and patience during the course of this project. Their unwavering support provided the motivation and resilience necessary to bring this research to fruition.

## REFERENCES

[1] K. V. K. W. Atapattu, Dr. G. Salibi, and N. Tzenios, ‘A study on the Relationship between the rainy season and Dengue outbreak in the Colombo District of Sri Lanka’, *Special journal of the Medical Academy and other Life Sciences*, vol. 1, no. 3, Mar. 2023, doi: 10.58676/sjmas.v1i3.21.

[2] S. S. Deshmukh and S. Deb, ‘A survey of statistical and machine learning methods of quantile regression in time series and their suitability in predicting dengue outbreaks’, *Japanese Journal of Statistics and Data Science*, 2025, doi: 10.1007/s42081-025-00297-y.

[3] K. G. S. ; Dharmawardana *et al.*, ‘ANNEX 19: PREDICTIVE MODEL FOR THE DENGUE INCIDENCES IN SRI LANKA USING MOBILE NETWORK BIG DATA IDRC Grant/ Subvention du CRDI: 108008-001-Leveraging Mobile Network Big Data for Developmental Policy Annex 19: Predictive Model for the Dengue Incidences in Sri Lanka Using Mobile Network Big Data’, 2018.

[4] S. G. Kakarla *et al.*, ‘Weather integrated multiple machine learning models for prediction of dengue prevalence in India’, *Int J Biometeorol*, vol. 67, no. 2, pp. 285–297, Feb. 2023, doi: 10.1007/s00484-022-02405-z.

[5] A. N. Akramin Kamarudin, Z. Zainol, and N. F. Abu Kassim, ‘Forecasting the Dengue Outbreak using Machine Learning Algorithm: A Review’, in *2021 International Conference of Women in Data Science at Taif University (WiDSTaif )*, 2021, pp. 1–5. doi: 10.1109/WiDSTaif52235.2021.9430193.

[6] T. Ranathunge, J. Harishchandra, H. Maiga, J. Bouyer, Y. I. N. S. Gunawardena, and M. Hapugoda, ‘Development of the Sterile Insect Technique to control the dengue vector Aedes aegypti (Linnaeus) in Sri Lanka’, *PLoS One*, vol. 17, no. 4 April, Apr. 2022, doi: 10.1371/journal.pone.0265244.

[7] L. Fernando, A. S. Perera, and S. Lokanathan, ‘Predicting Dengue Outbreaks in Sri Lanka using Human Mobility-A Genetic Algorithms integrated Support Vector Regression Approach’.

[8] A. Nugaliyadde *et al.*, ‘FORECASTING DENGUE PATIENT COUNTS IN A DISTRICT BASED ON NEIGHBORING DISTRICTS AND RAIN-FALL VIA MACHINE LEARNING’, 2015. [Online]. Available: https://www.researchgate.net/publication/315668567

[9] X. Y. Leung *et al.*, ‘A systematic review of dengue outbreak prediction models: Current scenario and future directions’, *PLoS Negl Trop Dis*, vol. 17, no. 2, Feb. 2023, doi: 10.1371/journal.pntd.0010631.

[10] Z. Hussain, • Imran, A. Khan, • Mudassar, and H. Arsalan, ‘MACHINE LEARNING APPROACHES FOR DENGUE PREDICTION: A REVIEW OF ALGORITHMS AND APPLICATIONS’, 2023. [Online]. Available: www.sciencedirect.com