

Discover the New Factor for Dengue Fever Outbreaks and Predicted using Bayes Network-PSO (BN-PSO)

RAVI KUMAR SUGGALA^{1,+}, DR. M. VAMSI KRISHNA² AND DR. SANGRAM KESHARI SWAIN³

^{1,+}*Research Scholar, Department of Computer Science and Engineering
Centurion University of Technology and Management
Paralakhemundi, Odisha, 761200 India*

E-mail: ravi.suggala@gmail.com

²*Department of Computer Science and Engineering
Aditya Engineering College
East Godavari District, Andhra Pradesh, 533437 India*

³*Department of Computer Science and Engineering
Centurion University of Technology and Management
Khurda, Odisha, 752050 India*

Dengue fever is a mosquito-borne pathological infection that is the nation's most dangerous widespread human illness disorder, posing a critical threat to humankind. Moreover, accuracy is a major challenge during dengue epidemic prediction that must be addressed. A few research studies have looked into the factors influencing dengue outbreak prediction. Furthermore, only a tiny fraction of the infected population can be properly predicted using a forecasting approach for dengue infection disorders based solely on meteorological variables. This limitation is caused by a low mosquito population below infection transmission thresholds. Therefore, an Improved Deep Learning Model for Predicting Dengue Outbreaks is proposed, in which novel climatic parameters such as the TempWind factor are evaluated. Then to estimate the dengue risk incidence level, the Bayes network model combined with Particle Swarm Optimization (PSO) is introduced. As a result, the proposed model has proven that using the correct and relevant factor of putting aspects for epidemic forecasting yields improved and accurate findings.

Keywords: dengue outbreak prediction, deep learning, particle swarm optimization (PSO), Bayes network mode, TempWindFactor (TWF).

1. INTRODUCTION

By 2050, the current global population of 7.8 billion people (as of 2020) is expected to increase to 9.7 billion [1]. Population growth, unfortunately, speeds up the development of infectious diseases [2]. Several elements lead to disease outbreaks. Among these variables are climate change, globalization, and urbanization, the bulk of which are caused by humans somehow. Diseases may be prone to emergence, and emerging infections are more likely to have rapidly evolving viruses. When a parasite transmits to another living creature from one body, it is called infectious disease. It has the potential to do significant harm on a large scale, making it a serious societal issue. It affects not only people moreover society as a whole and is thus classified as a joint problem [3]. As a result, identifying high-risk locations for fatal pathogenic and non-pathogenic infection epidemics is critical. As a result, the fatal illness eruption may remain predicted and detected. Fatal illness epidemics

Received May 25, 2022; revised August 16, 2022; accepted September 26, 2022.

Communicated by Tzung-Pei Hong.

⁺Corresponding author

could be responded to much better, and the unique Coronavirus disease, which is currently causing a worldwide health pandemic, can be predicted and detected (COVID-19).

Even though the COVID-19 virus has gained international attention as an emerging global epidemic, global dengue fever cases have continued to rise due to urbanization, climate variability, and increased travel [4]. Dengue Infection is an arboviral infection affected by one of four virus pathotypes that are directly connected. It is now the most common human-infecting mosquito-borne virus, posing a serious worldwide health threat [5]. Over 3 billion individuals are at risk of infection worldwide, with around 40 million symptom episodes and 20,000 deaths reported each year. Dengue fever has now been detected in more than 100 countries worldwide, with Southeast Asia accounting for over 70% of dengue morbidity and mortality worldwide. However, the true prevalence of the condition is estimated to be many times higher because several medically evident instances escape unreported during the public through care facility monitoring techniques, which are commonly employed to report illness [6]. Furthermore, the examination into why dengue has altered the worldwide risk posed through human illness infection would assist influencers around the globe in effectively anticipating and adapting to possible developments in dengue threat in 2020, 2050, and 2080 [7].

One strategy for controlling and preventing the spread of this infection is to have a good prognosis on the outbreak's existence, which allows higher experts and users to develop tactics and deal with this problem efficiently and successfully [8, 9]. The fundamental and most important concern for managing dengue fever is the accuracy of a forecast system for epidemics. As a result, identifying relevant risk variables for prediction systems is crucial [10]. Given the importance of climatic factors in this illness, connecting climate data and dengue eruption prevalence seems to be crucial for constructing a reliable forecasting system for future epidemics [11-13]. Warmth, moisture content, and rainstorm quantity were the important weather risk variables examined. Climate-based forecast systems' recent accuracy ranges from 82.39 to 90.5 percent [14-19].

In various methods, health organizations can apply Artificial Intelligence (AI) technologies to prevent the spread of serious infectious disease epidemics [20, 21], and it might be accomplished by using artificial intelligence to forecast and identify severe infectious diseases, as well as to respond to lethal infectious diseases. This research aims to assess the accuracy of dengue forecasting methods and the number of threat variables that go into dengue pandemic forecast methods and discover novel risk factor correlations. The precise parameters are then utilized to forecast dengue outbreaks. The existing challenge in dengue epidemic prediction that has to be addressed is accuracy. A few researches have looked in-depth at the elements that influence dengue outbreak prediction. As a result, Improved Deep Learning Model For Predicting Dengue Outbreaks was developed as an improved architecture based on artificial intelligence methodologies. The following is the research work's contribution:

The current study solely considers meteorological characteristics of dengue cases that affect prediction outcomes. However, our proposed method incorporates the TempWind component to predict dengue outbreaks accurately.

In addition, the Bayes network model is used in conjunction with Particle Swarm Optimization (PSO) to anticipate dengue disease epidemics.

The following is the format of the review paper: The survey of existing methodologies is described in Section 2, and then the suggested framework is completed in Section 3.

Section 4 presents the execution outcomes of the proposed method. At last, Section 5 concludes the paper.

2. RELATED WORKS

Dengue outbreaks are currently being tracked using various early warning and monitoring systems worldwide. Dengue estimation methods had already remained examined; nevertheless, few of these designs continue to have high accuracy restrictions in predicting dengue outbreaks. Several models have been created using a variety of styles and procedures for forecasting dengue epidemics. Many research works have also used artificial neural networks to develop prediction models for dengue epidemics. Fusion designs were used in epidemic forecasting studies, and there are several ways based on dynamic techniques to calculate the weight of an interconnected model. The majority of categorization problems, on the other hand, revolve around anticipating disease outbreaks.

K. Duangchaemkarn *et al.* [22] proposed another method for detecting disease outbreaks that employ symptoms. A location-based algorithm uses k-NN clustering to locate a new probable epidemic cluster. Moreover, to assess the accuracy and efficiency of outbreak prediction, additional testing using empirical epidemiology datasets is required.

Chen *et al.* [23] proposed independent Least Absolute Shrinkage and Selection Operator concepts remain to be built for various syndrome/state/estimate windows with varying design difficulty by adding multiple sets of classifiers to analyze the role of various classifiers under different situations using LASSO regression models to predict occurrences for four viral diseases. In the future, the authors could evaluate the effects of disaggregation for nations like Japan where data at the prefectural level is provided.

Scavuzzo *et al.* [24] proposed six models for correlation, such as two linear regression (Simple and Ridge) and four non-linear concepts. In the modeling, the Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), and Meteorological Parameters (LST) at dark, Meteorological Parameters (LST) during the day, and the Moderate-resolution imaging rain gauge were all used (floral life, humidity, heat, and storm). The Artificial Neural Network (ANN), Multilayer Perceptron (MLP) produces the most presentable results compared to other models. With a larger dataset, these algorithms' performance could be significantly enhanced. Although, the used dataset is still somewhat tiny from a machine learning perspective.

Y. Wu *et al.* [25] developed a hybrid technique that combines Convolutional Neural Networks-Recurrent Neural Networks (CNN-RNN), and residual connections to create a group prototype that is sensitive and capable of more robust epidemiological data prediction. The results revealed that this combined approach exceeded AR, VAR, and GPS methodologies, with a mean square error of 0.259. It's important to observe that CNN-RNN does not routinely outperform because all of the datasets are quite tiny (only containing a few hundred training samples), thus adding additional parameters (the result of including the CNN modules) might reduce performance due to overfitting.

V. Vijayakumar *et al.* [26] created a Fog computing-based intelligent healthcare method to detect and control vector-borne infections. According to the empirical evaluation, the Fuzzy k-Nearest Neighbour identifier gives the best performance, with 95.9 percent prediction accuracy. However, several concerns need to be rectified in the future. The proposed approach is mainly developed for diabetes, and it has to be enhanced to effectively assess several diseases at once based on important aspects and risk factors.

Tapak *et al.* [27] used time-series adaptation to examine and compare the efficacy of multiple ml algorithms in predicting the regular amount of influenza-like illness (ILI) cases: Support Vector Machine, Auto-Regressive Integrated Moving Average, Random Forest, and Artificial Neural Network. According to the findings, the ANN's sensitivity for the validation set (86.2 percent) remained higher than the other three approaches. The used data, though, was general national data. However, because Iran's environment varies substantially from region to region, it was not possible to obtain weekly ILI statistics segmented by climate. Therefore, the authors are incapable of researching the effect of these factors.

Somboonsak *et al.* [28] identified the model with the best restriction virtues to forecast the incidence of dengue sickness using the SARIMA model constructed on dengue patient data. The Box-Jenkins technique is combined with the temporary autoregressive integrated moving average concept to project the dengue sick people. The results show that the SARIMA method is useful for monitoring dengue outbreaks.

Xu *et al.* [29] developed a timely, accurate dengue forecasting model using recurrent long short-term memory (LSTM) neural networks. Compared to other candidate models, the LSTM model decreased the overall RMSE of forecasts by 12.99 percent to 24.91 percent. It reduced the average RMSE of predictions during the pandemic by 15.09 percent to 26.82 percent. But there are several gaps in the study. First, compared to other machine learning models, the LSTM model requires more time for training; however, this difference is not significant because the data used in this study originated from a small dataset.

A. Baldominos *et al.* [30] presented the creation of a medical result support system aimed at detecting diseases based on patient diagnostic indicators. This system is designed to manage a minor quantity of records, ideal for applications that adhere to strict guidelines. In addition, the authors supplemented the models with data from other sources to improve their quality. The social data derived from online searches, as gathered from Google Trends and meteorological data, have been considered. Nevertheless, in situations where clinical data may be limited due to a lack of technical and human resources, other widely accessible and easy-to-access sources of data might be employed to improve diagnosis performance.

Sood *et al.* [31] proposed an intelligent healthcare system that employs cloud services, IoT devices, and fog computing frameworks to identify, monitor, and inform dengue-affected persons and other decision-makers in live time and handle the Dengue virus disease eruption. The suggested structure practices a Naive Bayesian Network (NBN) to detect people who may be infected with the dengue virus. It generates real-time notifications to advise and inform key stakeholders to take suitable measures at the fog subsystem as soon as possible. In addition, the proposed system employs Public System Investigation in the web domain to deliver a GoogleMaps-created international hazard analysis of Dengue infectious agent and prevent dengue virus pathogen epidemics. The proposed approach has potential future applications in the field of technology-based healthcare research, where it can be used to establish a system for monitoring diseases based on common symptoms.

Gangula *et al.* [32] focused on using ensemble-based learning to improve the accuracy and outcomes of forecasting dengue fever. The most important meteorological conditions that contribute to dengue epidemics have been discovered by Nejad *et al.* [33]. These features were discovered through relationship research and used as process variables for neural network techniques. Bayes network designs stood picked among the top five categorization models for machine learning: Support Vector Machine, Radial Basis Function tree,

visual representation, and Naive Bayes. This epidemic prediction model is anticipated to help government agencies, health organizations, and other concerned parties become aware and create better prevention measures in the future.

Salim *et al.* [34] used an SVM linear model classifier to forecast dengue epidemics using meteorological data, including heat, cyclones, moisture, and storm. Moreover, the interactions between variables are outside the control of SVM. The study aims to look into SVM as a tool for outbreak prediction as well as the week of the year as the main predictor of dengue outbreaks at different spatial scales and in various types of models.

Kamarudin *et al.* [35] assessed the existing Mosquito-borne disease eruption forecasting framework. The authors proposed an updated framework that included the Entomological Index feature to improve future Mosquito-borne disease epidemic predictability.

However, no research has been done to determine the essential factors in regulating the reproduction of *Aedes* mosquitoes, which cause dengue fever. This proposal is directed to find the essential element, and the emergence of dengue disease was predicted using artificial intelligence algorithms based on that factor.

3. DISCOVER THE NEW FACTOR FOR DENGUE FEVER OUTBREAKS AND PREDICTED USING BAYES NETWORK-PSO (BN-PSO)

Dengue fever is an arboviral infection caused by dengue viruses spread by *Aedes* mosquitoes. Mosquito-borne illness control is not one-size-fits-all since various conditions can promote vector multiplication and raise infection rates. As a result, the predictive ability of AI techniques will be directly influenced. In the production of prediction risk maps and the modeling of dengue cases, climatic and meteorological data were regularly used. According to researchers, inter-annual and seasonal climatic fluctuations substantially impacted dengue virus transmission. As a result of our research, an Improved Deep Learning Model For Predicting Dengue Outbreaks was developed. Initially, the input is preprocessed, and then the processed input is analyzed using PCC analysis.

Furthermore, the study identifies a new meteorological element, the TempWind factor, which combines the lagged lowest temperature with the current month temperature and lags cumulative wind speed. Furthermore, a Bayes network model linked with the stochastic method based on population is developed to assess dengue fever’s risk incidence level. As a result, dengue fever epidemic predictors may provide health officials with timely information to implement preventative measures. Fig. 1 depicts the proposed architecture.

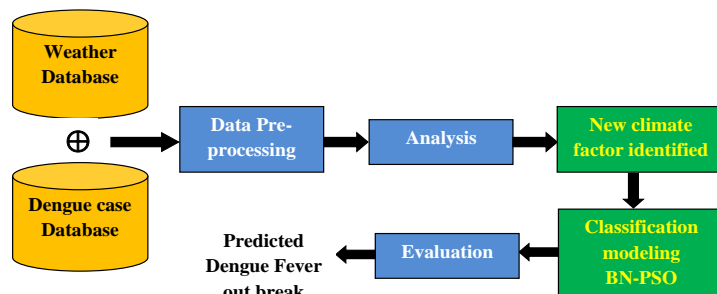


Fig. 1. Proposed architecture.

3.1 Novel Climatic Factor Identification

Dengue fever is caused by a combination of weather and meteorological factors. Initially, the data is collected from two different databases: the weather database and the dengue case database. The incidence of dengue cases is presented monthly, while weather data gives daily weather information. The data was preprocessed at first, and every month, it was standardized and divided into two categories: 'low risk' and 'high risk.' After that, the data is consolidated and cleansed as needed. The Bivariate Correlation analysis is then applied to the preprocessed data. As a result, the data was evaluated, and the Bivariate Correlation was used to determine the association in the middle of the monthly prevalence of dengue instances x_{z_i} and weather information every month y_{z_i} .

$$R_z = \frac{\sum x_{z_i} y_{z_i} - \frac{\sum x_{z_i} \sum y_{z_i}}{N_z}}{\sqrt{\left(\sum x_{z_i}^2 - \frac{(\sum x_{z_i})^2}{N_z} \right) \left(\sum y_{z_i}^2 - \frac{(\sum y_{z_i})^2}{N_z} \right)}} \quad (1)$$

The most important climatic factors were identified using Pearson correlation analysis, in which new climatic parameters were used as direct determinants of dengue cases to predict better future Mosquito-borne disease outbreaks, in which the lagged minimum temperature through the current month temperature and covered growing wind speed was integrated to find new climatic parameters called the TempWind Factor (TWF). According to the findings, the lowest temperature has the strongest link to dengue frequencies, guided by accumulated wind speed and the number of dengue incidences recorded in various months. The lowest temperature and routine wind speed remain two major climatic threat issues for dengue illness. Eq. (2) is used to calculate the average minimum temperature.

$$\text{Average Minimum temperature month}(i_z) = \frac{\text{Minimum temperature(current month)} + \sum_{n_z=0}^5 \text{Minimum temperature[month}(i_z-1)]}{6} \quad (2)$$

Where i_z is the number of months used to calculate the normal least temperature, and month $(i_z - 1)$ is the lowest temperature from the previous month that is added to the current month's minimum temperature [$n_z = 0$]. Divide the result by 6 [the previous five months + the current month] to get the average. The cumulative wind for the month i_z is calculated using Eq. (3) as follows,

$$\text{Cumulative wind speed per month}(i_z) = \sum_{n_z=0}^{i_z-1} [\text{Total wind month}(i_z-1)] \quad (3)$$

where i_z is the cumulative month for total wind speed analysis, cumulative wind speed per month i_z is the final calculation, and month $(i_z - 1)$ is the month preceding the month (n_z). The PCCs between meteorological factors and dengue case incidence. The emphasized and underscored great optimistic figures displayed the maximum connection and constants in the middle of climatic conditions and the occurrence of dengue disease. Based on the data from the previous months, the best value for the normal lowest temperature is determined. Months before the current month, the highest cumulative wind speed value was recorded.

According to the correlation analysis, the average lowest temperature for the 5th month (plus the current month) and collective wind speed for the 2nd month (before the current month) are strongly linked with dengue cases.

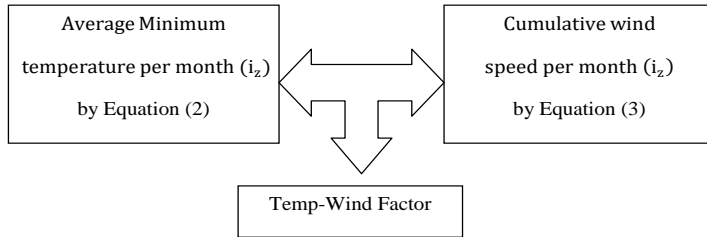


Fig. 2. Evaluation of Temperature Wind Factor (TWF).

TWF will be the name given to the two components, which will be applied as initial parameters to estimate the probability of a dengue eruption. Fig. 2 displays the interplay of variables then evaluated and fed into deep learning models as integrated inputs.

3.2 Classification using Bayes Network Model Combined with Particle Swarm Optimization (BN-PSO)

Following the discovery of temperature wind factors, the study forecasted the likelihood of dengue fever. According to our findings, there is a strong possibility of dengue epidemics and small threats of an eruption not occurring. Furthermore, the deep learning models predict this risk level by inserting parameters presence and absence of TWF. Table 1 summarizes the exact data variables and their characterizations. Furthermore, the Bayes network model paired with Particle Swarm Optimization is used to achieve high accuracy.

The BN-PSO algorithm is accurately explained using the following equation: $BN-PSO = (F_z, X_z, V_z, S_{z \times v_z}, P_{z \times v_z}, M_{z \times v_z}, P_{z \times v_z} \lambda, Ginit, U_z)$, F_z is a weight factor, X_z is particle position space, V_z is particle velocity group, $S_{z \times v_z}$ is a subtraction process (point, point), $P_{z \times v_z}$ represents a continuing procedure (position plus speed), $M_{z \times v_z}$ represents a calculation process (correlation times velocity), $P_{z \times v_z}$ represents arithmetic operation (velocity plus velocity), λ is the length of the swarm, Ginit is the initial clusters, and U_z is the terminating condition.

Table 1. The forecasting model relevant variables in the presence and absence of TWF.

Relevant variables in the absence of TWF	
Weather Factors	Lowest thermal reading (°C)
	Average thermal reading (°C)
	Highest thermal reading (°C)
	Cumulative storm speed (km/h)
Relevant variables in the presence of TWF	
Weather Factors	Average thermal reading(°C)
	Highest thermal reading (°C)
	Cumulative storm speed (km/h)
TWF Factors	Average Minimum temperature (°C) Cumulative wind speed (km/h)

Fitness Function: The most probable structure is $P_z(D_z|B_{zS'})$. These structures are ranked depending on how likely the data came from them. In addition, because the network’s parameters are known in advance, the probability $P(D|BS)$ may be calculated in closed form.

$$F_z = P_z(D_z | B_{zS'}) = \prod_{i_z=1}^{n_z} \prod_{j_z=1}^{q_{zj_z}} \frac{\Gamma(N'_{zj_z})}{\Gamma(N'_{zj_z} + N_{zj_z})} \prod_{k_z=1}^{r_{zj_z}} \frac{\Gamma(N'_{zj_zk_z} + N_{zj_zk_z})}{\Gamma(N'_{zj_zk_z})} \tag{4}$$

where n_z is the quantity of factors in the file, r_{zj_z} indicates the majority of states that could exist for variable X_{zj_z} , and q_{zj_z} represents the number of positions for the parameters (X_{zj_z}), $N_{zj_zk_z}$ indicates sufficient database statistics, and $N'_{zj_zk_z}$ indicates the hyper defined for the variable previously (considering an u_z n_z revealing previous as in the earlier architecture, we fixed the input variables to 1) and the number of alternative states for variables.

$$N_{zj_z} = \sum_{k_z=1}^{r_{zj_z}} N_{zj_zk_z} \tag{5}$$

$$N'_{zj_z} = \sum_{k_z=1}^{r_{zj_z}} N'_{zj_zk_z} \tag{6}$$

Due to computational constraints, the sum of individuals assigned to a parameter is restricted. This counting measure stands as the Bayesian network model (4). In practice, the logarithm of (4) is frequently employed in rating channels. When information is reliable, the Bayesian routing algorithm, a fitness value for BN, occurs in the secured method. As a result, we can consume the score decomposition principles in our research, which makes calculating the ranking method easier. Moreover, using the Bayesian network model, search many architectures for the one that gets optimal connections from entire collections, which indicates certain instances in the information have entries with all factors (4).

The numeric value for the overall material can be calculated by summing the scores for each local position, with the exponential of the evaluating tool serving as the total of points for every element. As a result, the scope of the search is vast. The number of parents at any given device point can vary from $0 - n_z - 1$, in which n_z refers to the dataset’s number of variables.

A local position can have $\sum_{i_z=1}^{k_z} i_z^{n_z-1}$ values, where k_z the maximum number of individuals is a value and n_z is the total amount of items in the dataset. As a result, the number of nodes is as follows,

$$\sum_{j_z=i_z}^{n_z} \sum_{i_z=j_z}^{k_z} (i_z^{n_z-1})Z. \tag{7}$$

An adjacency list can be used to express a particle’s position $P_z = ((x_{zi}))$, $i_z = 1, \dots, n_z$, if a provided BN consists n_z variables. According to our findings, while used to a point in each stage, the switch operator yields another position. As a result, there are three types of SO: $+x_{zi}$, $-x_{zi}$ and φ . $+x_{zi}$ indicates the addition of a variable x_{zi} to its original place $-x_{zi}$ represents the reduction of a variable x_{zi} and φ indicates null.

Addition (P_{zxv_z}) point desirable speed: Let P_z be a point and v_z be a speed. By implementing the first switch of v_z to P_z , then the next to the output, and so on, the position $P'_z = P_z \pm v_z$ is determined.

Subtraction (S_{zxv_z}) point minus point: Let P_{z1} and P_{z2} be two positions. The velocity v_z , as determined by a given algorithm, is defined as the difference $P_{z2} B_z P_{z1}$, so that applying

v_z to P_{z1} gives P_{z2} . The state “discovered using a specified algorithm” is required since, as we have observed, two velocities of the same size can be equivalent. The algorithm is specifically selected so that $P_{z1} = P_{z2} \rightarrow v_z = P_{z2} B_z P_{z1} = \varphi$.

Addition (P_{zx,v_z}) speed desirable speed: Allow v_{z1} as well as v_{z2} are the different speeds. The shift list holds. The first switch unit of v_{z1} , the first switch unit of v_{z2} . The second switch part of v_{z1} . The second switch part of v_{z2} and so on is used to calculate $v_{z1} \oplus v_{z2}$.

Multiplication (M_{zv_z}) coefficient: In the multiplication (M_{zv_z}) correlation times speed, let be an actual correlation and v_z be speed. There are numerous options based on the value. The original cluster Ginit, and the speeds, could be produced randomly or using a Sobol control action, which uniformly distributes the D-dimensional variables across the solution area.

The cluster length must not be reserved too large due to the computing period needed to score the suitability process; however, it cannot be maintained too tiny to improve the variety of swarm elements and prevent early coherence. The halting condition for the algorithm is defined when $g1$ generations have been completed, or the amount of the strongest material’s optimization process corresponds to the mean price of the optimization process in subsequent generations of $g2$.

BN-PSO Algorithm:

$$V_{zi,d}^{k_z+1} = w_z * V_{zi,d}^{k_z} + c_{z1} * rand() * (P_{zi,d} - X_{zi,d}^{k_z})$$

$$c_{z1} * Rand() * (P_{zgd} - K_{zi,d}^{k_z}) \tag{8}$$

$$X_{zi,d}^{k_z+1} = X_{zi,d}^{k_z} + V_{zi,d}^{k_z+1} \tag{9}$$

Where $i_z = 1, 2, \dots, N_z$; N_z denotes the swarm size; the d -dimensional search space is denoted as d ; and w_z is the inertia weight factor. $rand()$ and $Rand()$ are two rectangular distribution arbitrary integers in the region $[0,1]$; c_{z1} and c_{z2} are the positive constants termed the social and cognitive variables, respectively. The particle speed i_z at iteration k_z is the $V_{zi,d}^{k_z} \cdot X_{zi,d}^{k_z}$ the particle’s current position i_z at iteration k_z ; $x_{zi,d}$ is the greatest earlier situation of an element i_z at iteration k_z ; P_{zgd} is the earlier location of the best neighbor at repetition k_z .

Step 1: To begin, give each particle a random first result/point and a shift rows/speed.

Step 2: If the ending requirement has been fulfilled, go to Step 5.

Step 3: Using the current point $X_{zi,d}$ of the particle, find the next position $X'_{zi,d}$ of the particle.

- Find the difference α by $\alpha = P_{zi,d} - X_{zi,d}$, where α is a simple switch list that is used to $X_{zi,d}$ to get $P_{zi,d}$.
- Analyze the variance β by $\beta = P_{zgd} - X_{zi,d}$, where β is also a basic switch list.
- Estimate the speed $V'_{zi,d}$, in terms of Eq. (8) and convert $V'_{zi,d}$ into a simple switch list.
- Determine the recent result $X'_{zi,d}$ using Eq. (9).
- If a better solution is discovered, $P_{zi,d}$ should be modified.

Step 4: Update $P_{zi,d}$ and go to Step 2 if an improved result for the particles in the group is discovered.

Step 5: Display the ideal output.

Furthermore, the classified output is examined, with various accuracy measures and parameters available to assess classifier performance. In addition, the gap between the expected and actual known values is quantified using several exactness and deviation assessments. The accuracy measure was utilized to assess the learners’ effectiveness in this investigation. Eq. (10) depicts the process of achieving precision.

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} * 100 \tag{10}$$

4. RESULT AND DISCUSSIONS

This section addresses the results of the implementation and the performance of our proposed system. Following that, comparative analyses with existing research are presented.

- Tool** : PYTHON 3
- OS** : Windows 7 (64-bit)
- Processor** : Intel Premium
- RAM** : 8GB RAM

4.1 Dataset Description

The research consumes the information from two independent databases: a weather forecasting database and a Kaggle database of confirmed dengue cases. The dengue confirmed cases database comprises the year (1990 to 2010), the week of the year, and the total number of cases. In contrast, the weather forecasting database provides metrological data such as warmth, air direction, moisture, and thunderstorm.

4.2 Data Preprocessing

Figs. 3 (a) and (b) displays the preprocessed data, which were divided into two features, City 1 and City 2, based on the normalized temperature, month, and wind speed data. Then, the data is combined and cleaned as required.

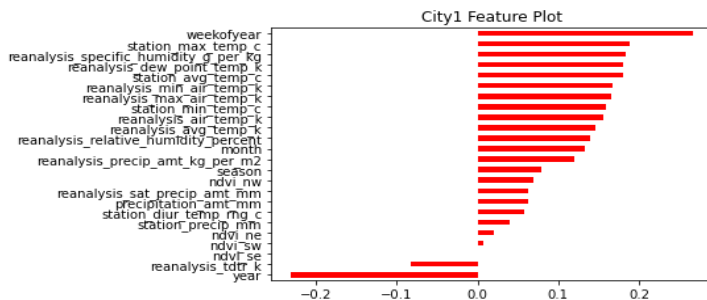


Fig. 3. (a) Preprocessed data.

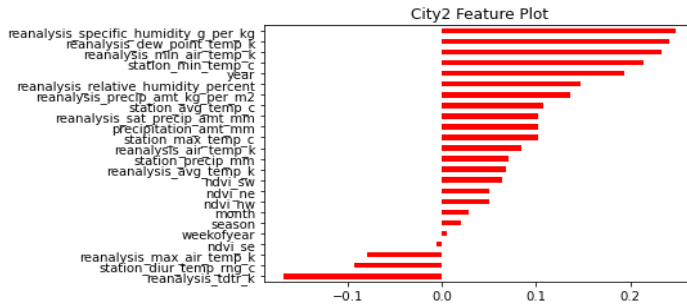


Fig. 3. (b) Preprocessed data.

4.3 Monthly Average Temperature Over the Years

The average temperature is depicted in Fig. 4, which is calculated using Eq. (2). The average temperature is calculated, by adding the minimum temperature from the previous month to the present month’s minimum temperature.

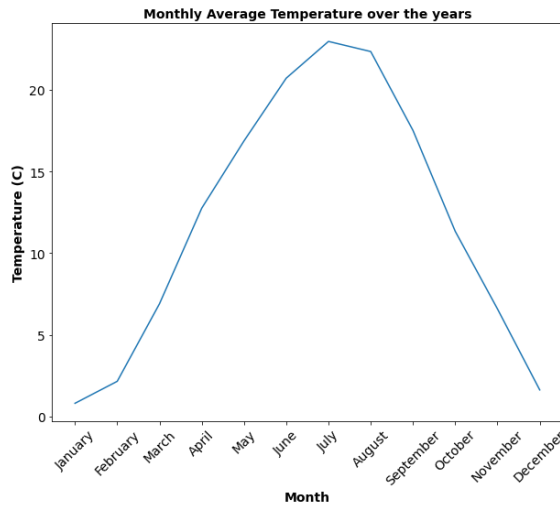


Fig. 4. Average monthly temperature over the years.

From Fig. 4, the average temperature in January month is 0.814°C, February month is 2.16°C, March month is 6.9°C, April month is 12.76°C, May month is 16.87°C, June month is 20.72°C, July month is 22.96°C, August month is 22.35°C, September month is 17.52°C, October month is 11.34°C, November month is 6.59°C, and December month is 1.63°C.

4.4 Monthly Average Wind Speed Over the Years

Fig. 5 shows the cumulative wind speed per month, evaluated by the above-illustrated Eq. (3). Cumulative wind speed analysis is calculated by the cumulative wind speed per month, such as January, February, and March, to December and the preceding month.

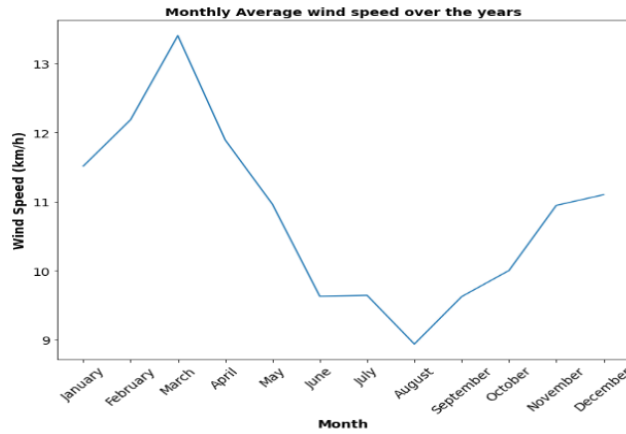


Fig. 5. Monthly wind speed over the years.

From Fig. 5, the cumulative wind speed in January month is 11.512 km/h, February month is 12.19 km/h, March month is 13.41 km/h, April month is 11.9 km/h, May month is 10.96 km/h, June month is 9.63 km/h, July month is 9.64 km/h, August month is 8.94 km/h, September month is 9.62 km/h, October month is 10 km/h, November month is 10.94 km/h, and December month is 11.09 km/h.

4.5 Temperature Aggregation

The evaluation and aggregation of temperature data, such as the minimum, maximum, and mean temperatures, are shown in Fig. 6. The minimum temperature is -21.822°C , the maximum temperature is 39.91°C , and the mean temperature is 11.93°C , according to Fig. 6.

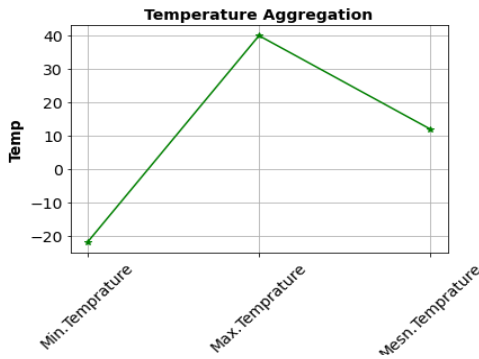


Fig. 6. Aggregation of temperature.

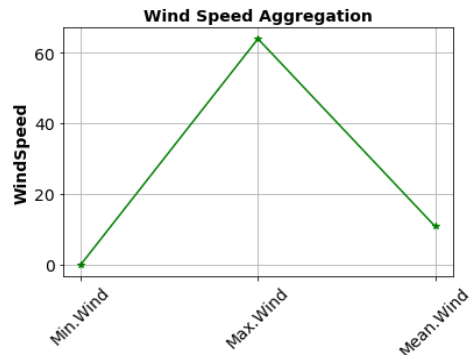


Fig. 7. Aggregation of wind speed.

4.6 Wind Speed Aggregation

Fig. 7 depicts the aggregation of wind speed, including the least, supreme, and mean wind speeds, which are evaluated and aggregated. Minimum, maximum, and mean wind speeds are optimal at 0 km/h, 63.85 km/h, and 10.81 km/h, respectively.

4.7 TempWind Parameter Over the Years

The orange line indicates the temperature factor in Fig. 8, while the blue indicates the wind factor. The temperature aggregation (such as minimum, maximum, and mean) and wind speed (such as minimum, maximum, and mean speed) demonstrated a strong relationship with dengue cases. The Temperature Wind Factor (TWF) through time is depicted in Fig. 8.

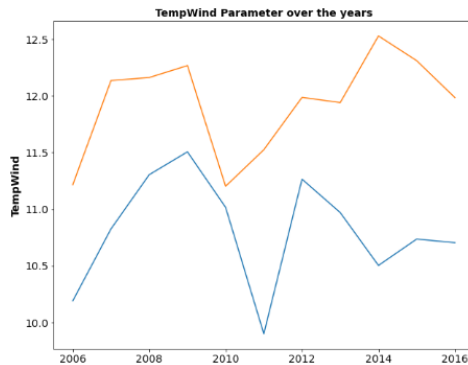


Fig. 8. TempWind parameter over the years.

4.8 Dengue Predicted Cases

Fig. 9 illustrates the forecasting results of the dengue outbreak. The Bayes Network in the proposed method predicts the dengue outbreak with Particle Swarm Optimization, which shows the effectiveness of the results. This proposed model is used to create an early warning system that could assist in monitoring the outbreak. Through this, it will be possible to establish in place effective and efficient reactive approaches to predict dengue outbreaks.

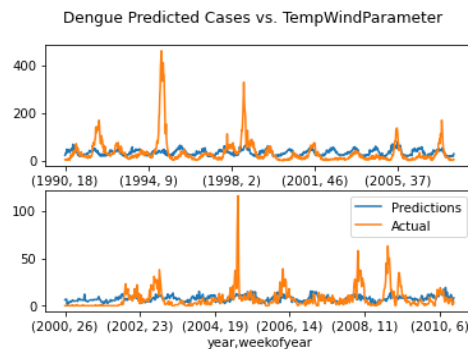


Fig. 9. Dengue predicted cases.

4.9 Comparative Analysis

Our novel technique is compared to baseline approaches such as Bayes net with and without TRF [36], SVM with and without TRF [36], RBF tree with and without TRF [36],

Decision table with and without TRF [36], Correlation and autoregressive model [36], C-SVC kernel and RBF [36], Poisson multivariate regression models [36], and Bayes network model using TRF [36].

Fig. 10 illustrates the total dengue cases predicted by existing techniques compared to the proposed TempWind Factor. Thus, our proposed factor outperforms compare to the existing approach.

Table 2. Comparison of accuracy report with TWF.

Methods	Accuracy (%)
Bayes net per TRF	92.35
SVM with TRF	88.04
RBF tree per TRF	89.47
Decision table per TRF	90.41
Proposed Method	97.5

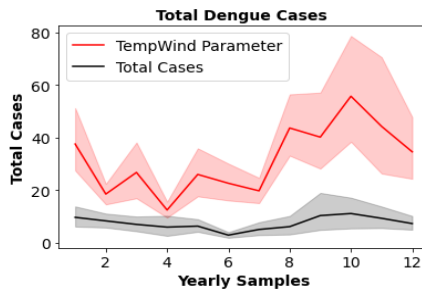


Fig. 10. Total dengue cases.

Fig. 11 depicts the overall accuracy compared with the Temperature Wind Factor (TWF). The proposed technique achieves improved accuracy by integrating the Temperature Wind component (TWF). With 92.35 percent, 88.04 percent, 89.47 percent, and 90.41 percent, our proposed approach compared the baseline Bayes net with TRF [36], SVM with TRF [36], RBF tree with TRF [36], and Decision table with TRF [36]. As a result, our unique technique has a 97.5 percent accuracy, higher than previous procedures.

Fig. 12 shows the overall accuracy comparison without the Temperature Wind Factor (TWF). The proposed technique achieves improved accuracy by integrating the Temperature Wind Factor (TWF). Our proposed approach scored 91.39 percent, 88.00 percent,

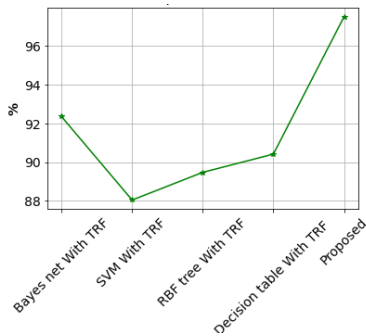


Fig. 11. Comparison report with TWF.

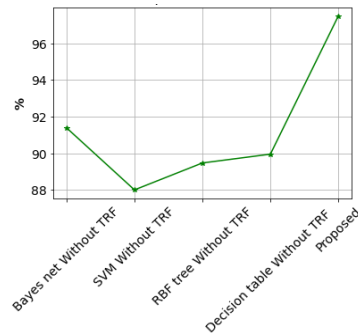


Fig. 12. Comparison report without TWF.

89.47 percent, and 89.95 percent when compared to the baseline Bayes net without TRF [36], SVM without TRF [36], RBF tree without TRF [36], and Decision table without TRF [36]. As a result, our unique novel technique has a greater accuracy of 96.72 percent than existing procedures.

Table 3. Comparison of accuracy report without TWF.

Methods	Accuracy (%)
Bayes net without TRF	92.35
SVM absence TRF	88.04
RBF tree without TRF	89.47
Decision table absence of TRF	90.41
Proposed Method	97.5

Table 4. Comparison of accuracy of dengue outbreak prediction model.

Methods	Accuracy (%)
Correlation and the autoregressive method	84.90
C-Support Vector Classification kernel and Radial Basis Functions	90.50
Multivariate Poisson regression models	90.00
Bayes network model using TRF	92.35
Proposed	97.5

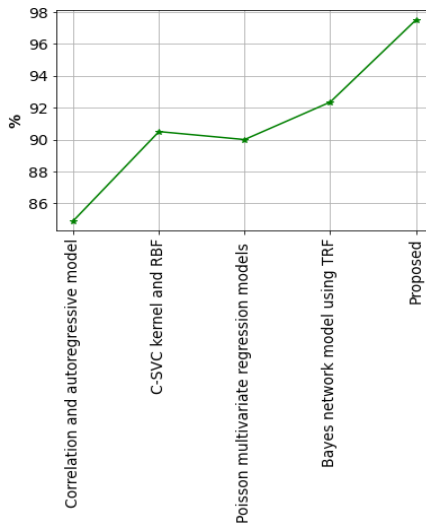


Fig. 13. Comparison of accuracy of dengue outbreak prediction model.

Fig. 13 shows the overall accuracy of the proposed Temperature Wind Factor (TWF). The proposed technique achieves greater accuracy by combining the Bayes Network with Particle Swarm Optimization (BN-PSO). Our proposed method compared the baseline Correlation and autoregressive model [36], C-SVC kernel and RBF [36], Poisson multivariate regression models [36], and Bayes network model using TRF [36] such as 84.90 percent, 90.50 percent, 90.00 percent, and 92.35 percent. Thus, our novel technique has obtained an accuracy of 97.5 %, higher than the existing techniques.

5. CONCLUSION

A variety of circumstances can cause dengue fever epidemics. Our study focused solely on a detailed investigation of heat and wind speed hazard issues for dengue eruptions, identified as the maximum essential determinants due to their importance and accessibility. By combining the 5-month lagged average lowest temperature by the current month and 2-month delayed collective wind speed, a new substantial risk factor TWF was found. TWF has improved in predicting dengue outbreaks, and these lagged meteorological variables are useful in more precisely forecasting the outbreak. This study shows that using exact and relevant input criteria when predicting outbreaks yields more precise and accurate results. The TWF model was merged into the BN-PSO model, yielding a 97.5 percent accuracy. According to the findings, TWF in the Baysien Network-Particle Swarm Optimization paradigm beat all other epidemic forecasting methods. Further studies should investigate and validate the TWF factors using more datasets from various regions, and countries. Moreover, the accuracy of dengue outbreak predictions should improve with the use of deep learning approaches.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to the Department of Science and Technology (DST) for their generous support provided for this research project (Ref. No. DST/SEED/SCSP/STI/2019/140). The valuable resources and infrastructure made available by Shri Vishnu Engineering College For women, Bhimavaram were instrumental in the successful completion of this study. We would also like to express our sincere gratefulness to Primary Health Center, Palakoderu for providing the dataset used in this research project. Their valuable contribution has been instrumental in enabling us to conduct our study and draw meaningful conclusions. We are grateful for their willingness to share the dataset and for their efforts in collecting and organizing the data. This research would not have been possible without their generous support.

REFERENCES

1. Growing at a slower pace, world population is expected to reach 9.7 billion in 2050 and could peak at nearly 11 billion around 21, <https://www.un.org/development/desa/en/news/population/world-population-prospects-2019.html>, 2020.
2. Population growth drives infectious disease rate upward, <https://www.healio.com/infectious-disease/emerging-diseases/news/print/infectious-disease-news/%7B90650258-654e-481e-ab79-176ff7de0676%7D/population-growth-drives-infectious-disease-rate-upward>.
3. D. E. Bloom and D. Cadarette, "Infectious disease threats in the twenty-first century: strengthening the global response," *Frontiers in Immunology*, Vol. 10, 2019, p. 549.
4. C. M. Barker and W. K. Reisen, "Epidemiology of vector-borne diseases" in *Medical and Veterinary Entomology*, 2019, pp. 33-49.
5. R. V. Gibbons and D. W. Vaughn, "Dengue: an escalating problem," *BMJ*, Vol. 324, 2002, pp. 1563-6.

6. M. Fatima and M. Pasha, "Survey of machine learning algorithms for disease diagnostic." *Journal of Intelligent Learning Systems and Applications*, Vol. 9, 2017, p. 1.
7. J. P. Messina, O. J. Brady, N. Golding, *et al.*, "The current and future global distribution and population at risk of dengue," *Nature Microbiology*, Vol. 4, 2019, pp. 1508-1515.
8. N. A. Husin, N. Mustapha, M. N. Sulaiman, R. Yaakob, "A hybrid model using genetic algorithm and neural network for predicting dengue outbreak," in *Proceedings of the 4th Conference on Data Mining and Optimization*, 2012, pp. 23-27.
9. B. Andrick, B. Clark, K. Nygaard, *et al.*, "Infectious disease and climate change: detecting contributing factors and predicting future outbreaks," in *Proceedings of IEEE International Geoscience and Remote Sensing Symposium: Remote Sensing-A Scientific Vision for Sustainable Development*, Vol. 4, 1977, pp. 1947-1949.
10. H. M. Aburas, B. G. Cetiner, and M. Sari, "Dengue confirmed-cases prediction: A neural network model," *Expert Systems with Applications*, Vol. 37, 2010, pp. 4256-4260.
11. S. S. Mathulamuthu, V. S. Asirvadam, S. C. Dass, *et al.*, "Predicting dengue incidences using cluster based regression on climate data," in *Proceedings of the 6th IEEE International Conference on Control System, Computing and Engineering*, 2016, pp. 245-250.
12. T. Soemsap, S. Wongthanavas, and W. Satimai, "Forecasting number of dengue patients using cellular automata model," in *Proceedings of International Electrical Engineering Congress*, 2014, pp. 1-4.
13. Y. L. Hii, R. A. Zaki, N. Aghamohammadi, and J. Rocklöv, "Research on climate and dengue in Malaysia: a systematic review," *Current Environmental Health Reports*, Vol. 3, 2016, pp. 81-90.
14. R. A. K. Tazkia, V. Narita, and A. S. Nugroho, "Dengue outbreak prediction for GIS based early warning system," in *Proceedings of International Conference on Science in Information Technology*, 2015, pp. 121-125.
15. D. Rahmawati and Y. P. Huang, "Using C-support vector classification to forecast dengue fever epidemics in Taiwan," in *Proceedings of International Conference on System Science and Engineering*, 2016, pp. 1-4.
16. Y. L. Hii, H. Zhu, N. Ng, L. C. Ng, and J. Rocklöv, "Forecast of dengue incidence using temperature and rainfall," in *PLoS Neglected Tropical Diseases*, 2012, Vol. 6, p. e1908.
17. D. Thitiprayoonwongse, P. Suriyaphol, and N. Soonthornphisaj, "Data mining of dengue infection using decision tree," *Entropy*, Vol. 2, 2012, p. 2.
18. L. Tanner, M. Schreiber, J. G. Low, *et al.*, "Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness," *PLoS Neglected Tropical Diseases*, Vol. 2, 2008, p. e196.
19. F. Ibrahim, T. Faisal, M. I. Mohamad Salim, and M. N. Taib, "Non-invasive diagnosis of risk in dengue patients using bioelectrical impedance analysis and artificial neural network," *Medical and Biological Engineering and Computing*, Vol. 48, 2010, pp. 1141-1148.
20. Z. S. Wong, J. Zhou, and Q. Zhang, "Artificial intelligence for infectious disease big data analytics," *Infection, Disease and Health*, Vol. 24, 2019, pp. 44-48.
21. S. Agrebi and A. Larbi, "Use of artificial intelligence in infectious diseases," *Artificial Intelligence in Precision Health*, Vol. --, 2020, pp. 415-438.

22. K. Duangchaemkarn, V. Chaovatut, P. Wiwatanadate, and E. Boonchieng, "Symptom-based data preprocessing for the detection of disease outbreak," in *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2017, pp. 2614-2617.
23. Y. Chen, C. W. Chu, M. I. Chen, and A. R. Cook, "The utility of LASSO-based models for real time forecasts of endemic infectious diseases: A cross country comparison," *Journal of Biomedical Informatics*, Vol. 81, 2018, pp. 16-30.
24. J. M. Scavuzzo, F. Trucco, M. Espinosa, *et al.*, "Modeling Dengue vector population using remotely sensed data and machine learning," *Acta Tropica*, Vol. 185, 2018, pp. 167-175.
25. Y. Wu, Y. Yang, H. Nishiura, and M. Saitoh, "Deep learning for epidemiological predictions," in *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2018, pp. 1085-1088.
26. V. Vijayakumar, D. Malathi, V. Subramaniaswamy, *et al.*, "Fog computing-based intelligent healthcare system for the detection and prevention of mosquito-borne diseases," *Computers in Human Behavior Reports*, Vol. 100, 2019, pp. 275-285.
27. L. Tapak, O. Hamidi, M. Fathian, and M. Karami, "Comparative evaluation of time series models for predicting influenza outbreaks: application of influenza-like illness data from sentinel sites of healthcare centres in Iran, BMC Res," *Notes*, Vol. 12, 2019, p. 353.
28. P. Somboonsak, "Time series analysis of Dengue fever cases in Thailand utilizing the sarima model," in *Proceedings of the 7th International Conference on Information Technology: IoT and Smart City*, 2019, pp. 439-444.
29. J. Xu, K. Xu, Z. Li, *et al.*, "Forecast of dengue cases in 20 Chinese cities based on the deep learning method," *International Journal of Environmental Research and Public Health*, Vol. 17, 2020, p. 453.
30. A. Baldominos, H. Ogul, R. Colomo-Palacios, *et al.*, "Infection prediction using physiological and social data in social environments," *Information Processing & Management*, Vol. 57, 2020, p. 102213.
31. S. K. Sood, V. Sood, and I. Mahajan, "An intelligent healthcare system for predicting and preventing dengue virus infection," *Computing*, 2020, pp. 1-39.
32. R. Gangula, L. Thirupathi, R. Parupati, *et al.*, "Ensemble machine learning based prediction of dengue disease with performance and accuracy elevation patterns," *Materials Today: Proceedings*, Vol. 80, 2023, pp. 3458-3463.
33. F. Y. Nejad and K. D. Varathan, "Identification of significant climatic risk factors and machine learning models in dengue outbreak prediction," *BMC Medical Informatics and Decision Making*, Vol. 21, 2021, pp. 1-12.
34. N. A. M. Salim, Y. B. Wah, C. Reeves, *et al.*, "Prediction of dengue outbreak in Selangor Malaysia using machine learning techniques," *Scientific Reports*, Vol. 11, 2021, pp. 1-9.
35. N. A. Kamarudin, Z. Zainol, and N. F. A. Kassim, "Forecasting the Dengue outbreak using machine learning algorithm: A review," in *Proceedings of International Conference of Women in Data Science at Taif University*, 2021, pp. 1-5.
36. F. Y. Nejad, and K. D. Varathan, "Identification of significant climatic risk factors and machine learning models in dengue outbreak prediction," *BMC Medical Informatics and Decision Making*, Vol. 21, 2021, pp. 1-12.



Ravi Kumar Suggala is currently pursuing his Ph.D. in Computer Science and Engineering at Centurion University of Technology and Management (CUTM), Odissa, India. He has fifteen years of teaching experience and four years of research experience. His research interests include artificial intelligence and machine learning, data science applications in health care domain. He is strong in developing real time applications in the web domain.



Dr. M. Vamsi Krishna is a leading Researcher and Academician in Computer Science and Engineering and holds Ph.D. degree from Centurion University of Technology and Management. He is currently working as a Professor in the Department of Computer Science and Engineering at Aditya Engineering College (A), Surampalem, Andhra Pradesh, India. He has 79 research papers of which are Scopus-indexed and Web of Science-indexed. He has twenty years of teaching experience and ten years of research experience. He had received International Best Research Award in the area of Cloud Computing. His Research interests include cloud computing, data mining, big data analytics, data science, artificial intelligence and machine learning.



Dr. Sangram Keshari Swain is an Associate Professor in the Department of Computer Science and Engineering along with holding the responsibility of Dean, Examinations and Dean, Students' Welfare in Centurion University of Technology and Management, Odisha. He is a Passionate Researcher and experienced Associate Professor with a demonstrated history of working in the higher education industry. He is skilled in social services, students' welfare, teaching, research, data analysis, and higher education. He is a strong education professional having multidimensional approaches such as engineering, technology, management and law. Social responsibility initiative is his passion always. His research interests and activities include data mining, big data, artificial intelligence, image processing, data analytics, machine learning, cloud computing, IoT, waste management and social development. He has published two patents, two copyrights, more than thirty technical papers in different SCI, WoS, Scopus indexed and UGC approved journals. He has presented more than ten technical papers in different national and international conferences. He has translated and published seven Odia language books on chemotherapy, breast cancer, lung cancer, ovary cancer, brain cancer, blood cancer, and solid waste management.