Forecasting deep learning-based risk assessment of vector-borne diseases using hybrid methodology

Ashok Kumar Nanda[a,*], R. Thilagavathy[b], G.S.K. Gayatri Devi[c], Abhay Chaturvedi[d],
Chaitra Sai Jalda[a] and Syed Inthiyaz[e]
[a]*Department of Computer Science and Engineering, B V Raju Institute of Technology, Narsapur, India*
[b]*Department of Computing Technologies, College of Engineering and Technology, SRM Institute of
Science and Technology, Kattankulathur, Chennai, India*
[c]*Department of Electronics and Communication Engineering, Malla Reddy Engineering College,
Hyderabad, India*
[d]*Department of Electronics and Communication Engineering, GLA University, Mathura, India*
[e]*Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education
Foundation, Vaddeswaram, India*

Abstract.
BACKGROUND: Dengue fever is rapidly becoming Malaysia's most pressing health concern, as the reported cases have nearly doubled over the past decade. Without efficacious antiviral medications, vector control remains the primary strategy for battling dengue, while the recently introduced tetravalent immunization is being evaluated. The most significant and dangerous risk increasing recently is vector-borne illnesses. These illnesses induce significant human sickness and are transmitted by blood-feeding arthropods such as fleas, parasites, and mosquitos. A thorough grasp of various factors is necessary to improve prediction accuracy and typically generate inaccurate and unstable predictions, as well as machine learning (ML) models, weather-driven mechanisms, and numerical time series.
OBJECTIVE: In this research, we propose a novel method for forecasting vector-borne disease risk using Radial Basis Function Networks (RBFNs) and the Darts Game Optimizer (DGO) algorithm.
METHODS: The proposed approach entails training the RBFNs with historical disease data and enhancing their parameters with the DGO algorithm. To prepare the RBFNs, we used a massive dataset of vector-borne disease incidences, climate variables, and geographical data. The DGO algorithm proficiently searches the RBFN parameter space, fine-tuning the model's architecture to increase forecast accuracy.
RESULTS: RBFN-DGO provides a potential method for predicting vector-borne disease risk. This study advances predictive demonstrating in public health by shedding light on effectively controlling vector-borne diseases to protect human populations. We conducted extensive testing to evaluate the performance of the proposed method to standard optimization methods and alternative forecasting methods.
CONCLUSION: According to the findings, the RBFN-DGO model beats others in terms of accuracy and robustness in predicting the likelihood of vector-borne illness occurrences.

Keywords: Deep learning (DL), dengue fever, radial basis function networks, darts game optimizer, vector-borne disease

---

*Corresponding author: Ashok Kumar Nanda, Department of Computer Science and Engineering, B V Raju Institute of Technology, Narsapur, India. E-mail: ashokkumarnanda@yahoo.com.

1. Introduction

In the world's dissimilar tropical and subtropical regions, the lethal viral disease dengue fever severely threatens public health. In 2017, there were 104 million dengue virus cases worldwide, up from 23 million in 1990, according to a recent study. Even though dengue is most often found in developing and underdeveloped countries, the effects of globalization and ongoing climate change are visible in the spread of the illness to previously unaffected areas. The severity of dengue infection presentations ranges from mild fever with symptoms such as increased body temperature, eye pain, aching muscles and joints, skin eruptions, and headaches to more hemorrhagic severe fever and shock syndrome [1]. In the absence of particular antiviral medications or licensed vaccinations, public health officials need to rely on supportive therapies and vector control approaches, such as insecticides, to stop the transmission of dengue infection. The primary way of controlling dengue is vector control. Still, it has limitations due to complicated transmission dynamics, the ineffectiveness of existing treatments, behavioral and social variables, environmental impacts, resource shortages, and gaps in research and innovation. To tackle these problems, cross-sector cooperation, consistent funding for R&D, and a multidisciplinary strategy involving players from many fields are required. Public health professionals employ vector control strategies to manage dengue infections, such as source reduction, larval and adult mosquito control, biological control, community involvement, and integrated vector management. Dengue cases and mosquito numbers are tracked via surveillance systems.

Although the disease is widespread over numerous parts of Southeast Asia, the western Pacific, and Africa, there is still a danger of infection in these locations due to increased mosquito amounts near places where people live. The rapidity at which the dengue virus spreads has been discovered to be determined by temperature, rainfall, and relative humidity. These circumstances favor rapid population growth because Aedes mosquitos mature and reproduce more quickly [2]. Artificial neural networks having input, hidden, and output layers are called radial basis function networks (RBFNs). Gaussian functions quantify how similar the input data and prototype vectors are. Using supervised learning algorithms, training entails optimizing weights and parameters. RBFNs provide interpretability and adaptability in machine learning applications.

Healthcare systems are under unprecedented strain due to the increasing incidence of dengue infections resulting from climate change and unplanned urbanization. According to research, essential occurrences detected and handled early can significantly lower mortality risk and health problems. Precise estimates of sick people's occurrence are also mandatory for resource planning and allocation. It is then possible to increase readiness through educational and awareness movements once the expected number of cases and their timing are known. This could entail shifting resources to high-risk locations and employing medical reinstruction to improve symptom detection and treatment before the peak transmission period. Early warning systems are dangerous for public health professionals in dengue-endemic areas to deploy available resources optimally. Numerous studies have attempted to forecast the beginning and spread of dengue, with varying degrees of success [3]. Numerous academics have endeavored to predict the occurrence and epidemics of dengue, with varying degrees of success. This is critical since developing pre-epidemic preparedness measures depends on accurate dengue predictions. These attempts aimed to create a reliable early warning system alerting medical personnel of an impending outbreak. The models employed to forecast dengue cases differed considerably. Statistical models, ensemble forecasting systems, data-driven machine-learning methods, and mechanical models have all been used in preceding studies. Due to the considerable amount of data required for model parameterization, automatic models based on deterministic differential equations and a thorough understanding of dengue transmission

typically fail to provide precise projections. Another study [4] used internet searches and social media data to forecast dengue prevalence. For this aim, forecasters have successfully used statistical and ML models more than mechanical models, primarily depending on dengue cases and environmental factors [5].

Any blood-sucking vector, like a mosquito, sandfly, or other insect, can spread infectious diseases known as vector-borne diseases (VBD). These ailments include malaria, leishmaniasis, dengue fever, chikungunya, zika, and chikungunya. The global illness burden is rising as a result of these infections. The rise of the VBDs, which claim 700,000 deaths annually, has put the lives of more than 80% of people in the globe in jeopardy. Pakistan is a subtropical country, and this is where VBDs are most prevalent. Seasonal illnesses include VBDs. Numerous such deadly diseases are commonplace in Pakistan. Some of these include leishmaniasis, dengue, and malaria [6].

This section will provide a novel and creative approach to VBD risk assessment that combines DL with a hybrid methodology based on Radial Basis Function Networks (RBFN). This research developed an established and effective forecasting model by combining the strengths of DL algorithms in dealing with complex, high-dimensional data with the interpretability and adaptability of RBFN. DGO is an innovative optimization that takes cues from the planned decisions made by darts players. This exclusive hybrid practice intends to overcome the limits of standard forecasting methods by combining the benefits of deep learning and the efficiency of DGO.

The primary contributions of this work:

- To develop a novel method for forecasting vector-borne illness risk employing the RBFN-DGO algorithm.
- The suggested method is divided into two stages: RBFNs using historical disease data and optimizing their parameters with the DGO algorithm.
- To train the RBFNs, we use an enormous dataset of vector-borne disease incidences, climate variables, and geographical data.
- The DGO algorithm efficiently searches the RBFN parameter space and refinement of the model's architecture to improve forecast accuracy.
- This study increases public health predictive modeling by offering light on effectively controlling vector-borne diseases to protect human populations.

2. Literature survey

Munirathinam et al. [7] describe a unique technique for anticipating vector-borne illnesses using medical data. The method combines a hybrid Support Vector Machine-based Random Forest (SVM-RF) and a Spotted Hyena-based Chimp Optimisation Algorithm (SSC). The proposed hybrid SVM-RF classifier method accurately recognizes vector-borne diseases. The suggested SSC-optimized mixed SVM-RF method performed better on the vector-borne Dataset than the competing algorithms, with 92%, 93.25%, 92.53%, and 91.52%, respectively. Ali et al. [8] proposed that the spatial layout of water bodies substantially impacts the occurrence of vector-borne diseases. Data augmentation, feature engineering, transfer learning, regularisation, ensemble learning, and sensitivity analysis are some tactics used by the RBFN-DGO algorithm to address data scarcity in vector-borne illness forecasting. These steps address the lack of data while enhancing the precision and dependability of risk assessments for vector-borne diseases. They considered the amount of wetness on the surface, the temperature there, how far they were from trash bins, and how frequently dengue and malaria cases had been reported. This recent discovery highlights the critical role of geospatial methods in categorizing disease epidemic hotspots and offers capable chances for use in numerous technical sectors.

Carvalho et al. [9] proposed a compartment model for mosquito life cycles involving six stages (eggs, larvae, pupae, and adults) and hypothesized that environmental variables such as temperature, humidity, rainfall, and ease of access to breeding sites affected the rate of oviposition. Different control tactics were evaluated. However, the effect of model parameters was not considered. According to earlier research, Jayaraj et al. [10] claimed that the weather influences the severity and scope of dengue occurrence. In Tawau, Sabah, a prediction model is still a cutting-edge approach to scientific forecasting. The technique used in this study shows potential for forecasting likely dengue epidemics 1–4 months in advance. These definitions are similar to earlier studies carried out in many countries, which consistently showed that temperature and humidity act as essential indicators for the prevalence of dengue.

Chakraborty et al. [11] introduce a novel hybrid model integrating ARIMA and neural network autoregressive (NNAR) model to account for both linearity and nonlinearity in the data sets. After the ARIMA model removes linear tendencies from the data, the NNAR model is given the residual values. The hybrid methodology we suggest shows better predicting accuracy than the current state-of-the-art dengue time-series forecasting approaches when applied to three different dengue time-series datasets. The consequences of this study designate that the proposed hybrid methodology can accurately predict dengue cases over a sufficiently long-time horizon.

Jagadesh et al. [12] plotted the five distinct disease categories as risk patterns and regionalized the leishmaniasis species according to the environmental circumstances in French Guiana. Climate variables such as temperature, precipitation, and others impact the leishmaniasis species. Taking down trees damages the environment significantly as well. The study looked into the effects of several leishmaniasis epidemics on temperature, precipitation, humidity, and deforestation. Stacked forms distribution models (SSDMs) can accurately identify the areas most vulnerable to contracting leishmaniasis. SSDMs could be used in future studies to examine the regional distribution of leishmaniasis species.

In northeastern Thailand, Rahman et al. [13] presented data on household income, urban versus rural domicile, socioeconomic position, knowledge, attitude, and practices about dengue. Additionally, they discussed the geographic distribution and predictive analysis of the two levels of dengue vector prevalence, high and low. This included a review of the use and coverage of the land, considering urbanized areas, wetlands that are always there, a natural forest canopy, rubber plantations, and rice farming. The authors examined various data types by utilizing machine learning approaches. Although temperature, precipitation, and humidity were not employed in this experiment, it is crucial to emphasize that they should be given greater attention because RF gave the most excellent model performance of any data variable.

Using a particular focus on geographic factors, McMahon et al. [14] undertook a thorough investigation into other factors preventing the transmission of malaria. They used remote sensing methods to get insights into the environmental factors influencing the prevalence of malaria in Ethiopia's two separate geographic zones. Geographical variables like temperature, rainfall, plant cover, land use, and the availability of surface water, as well as indicators like the NDMI (Normalised Difference Moisture Index), NDVI (Normalised Difference Vegetation Index), and LST (Land Surface Temperature), affected the frequency of malaria. These variances were linked to the various environmental conditions in different places. This study used boosted regression tree (BRT) models to consider the relationship between malaria incidence and environmental factors. The subdistricts with minor vegetation and the wettest air experienced the most crucial malaria cases during the outbreak season.

Ferdousi et al. [15] reported dengue fever in Espirito Santo, Brazil. The study used time-series prediction approaches, including dengue disease cases, temperature, precipitation (weekly rainy days count), relative humidity, surface-level pressure, and precipitable water. The researchers developed a novel

approach that enriched the temporal viewpoint by combining historical dengue case data from several places. Their strategy combined correlation, spatial proximity, and local incidence measures to prioritize neighboring area incidence data. The study used RNN, LSTM, and GRU models to estimate dengue disease occurrence accurately. The regionally particular Environmental and economic variables will eventually be incorporated into the model, including these aspects in the feature prioritization procedure.

Anggraeni et al. [16] forecasted the outbreak using fuzzy logistic regression using validation data. This prediction model significantly outperforms rival models such as Naive Bayes, Neural Network (NN), and Random Forest (RF). Temperature and the number of dengue cases were the two most essential components in this projection. Good hygiene habits and a high level of social awareness influence prediction outcomes. These factors should be considered while making predictions. The mosquito population variable is justified since it affects outbreaks.

According to Appice et al. [17], Mexico was included in the upcoming year's monthly projection of dengue cases. The data had air temperature and instances of hemorrhagic dengue and dengue fever from certain geographic and temporal regions. Different time-series prediction methods were applied. Time-series based on auto-encoding the techniques utilized were clustering with the nearest neighbor, M5', Autoregressive Integrated Moving Average, Support Vector Regression (SVR), Vector Autoregression, and kNN. The link between the dependent variable (dengue cases) and the independent variable (temperature) was determined using regression analysis.

A spatiotemporal component fusion model (STCFM) is recommended by Li et al. [18] to forecast the risk of dengue. Since mosquitoes spread dengue, we use feature variables to describe temporal trends and spatial distributions separately, considering mosquito abundance and spatiotemporal lags. To better forecast, we use multiscale modeling of temporal interdependence to find patterns of data fluctuation across temporal segments. We evaluate multivariate spatial correlation analyses as novel characteristics and comprehend the interdependence of locations using the ConvLSTM model. Numerous learning strategies are combined to provide forecasts.

2.1. Limitations of existing system

For an illness organization to be successful, it is essential to rely on accurate and current data. Data limitations might hamper the precision of approximation to improve interventions regarding quality, completeness, and timeliness. It might be challenging to incorporate data from numerous sources, such as data on the climate, disease observation, and vector populations. Data preparations and values that are unsuited can produce unlimited or undersupplied analyses. Predictive models' competence to accurately predict disease outbursts and the dynamics of vector populations can have some limits. In complex environmental systems, these models rely on prospects that may only sometimes be true.

2.2. Problem identification

The first problem is when data from different sources becomes standardized and combined. There needs to be more consistency and reliability in the data sources used to analyze socioeconomic characteristics, illness incidence, vector dispersion, and climate. An important task is developing a dataset that is both dependable and cohesive. Seasonality and regional diversity are regular characteristics of vector-borne illnesses. Creating real-time monitoring systems that can change the environment and send prompt declarations to healthcare authorities is problematic. Any disease management method must include community participation and education. A critical issue is ensuring that the hybrid method properly involves and enables local populations to take preventative steps.

3. Proposed system

The section presents a novel method for forecasting vector-borne illness risk employing the RBFN-DGO algorithm. It is advantageous when there is a lack of data since it is designed to fine-tune neural network parameters and improve model performance professionally. RBFNs are the primary determining model in the obtainable approach. The RBFN class of neural networks is an excellent applicant for problems involving the prediction of illness risk because they are excellent at capturing nonlinear connections in the data. Since RBFNs have transparent structure, adaptive learning, local learning, nonlinear mapping, feature transformation, and adaptive learning, they are good at predicting the likelihood of sickness. They can handle massive datasets and are scalable and computationally efficient. Their architecture and training procedure are described in depth to show how well they can be used for epidemiological data. The research shows the efficiency of the RBFN-DGO combination using real-world epidemiological data from diverse vector-borne diseases. In vector-borne disease epidemiology, comparing model predictions with real-world conditions is essential for accuracy evaluation, assumption validation, calibration, decision support, early warning systems, and trust-building. It promotes cooperation in attempts to prevent and manage disease by bringing model assumptions into line with observed disease transmission patterns. The findings demonstrate that the suggested method outperforms established forecasting techniques, producing risk estimates that are more precise and trustworthy. Figure 1 shows the block diagram of RBFN-DGO. Assessing an RBFN's performance and generalization capacity is imperative by testing it on both in-sample and out-of-sample datasets. While out-of-sample testing gauges the model's ability to generalize to new data, in-sample testing measures the model's fit to the training dataset. By ensuring consistency in performance, comparing the findings improves the accuracy and dependability of the model.

3.1. Dataset description

This study counts lab-confirmed dengue cases weekly in San Juan, Iquitos, and Ahmedabad, three tropical and subtropical regions. We look at statistics on how much rain falls each week to determine how climate affects dengue epidemics. The NOAA Repository2, established due to collaboration between the US government, Peruvian government agencies, and US institutions, contains information on the prevalence of dengue in San Juan and Iquitos. The US Climatology Office and NOAA are monitoring the precipitation in these regions. The data on rainfall totals in the Ahmedabad region and dengue prevalence (per 10,000 persons) are taken from [19]. We examine the weekly data for San Juan from the start of 1991 to the end of 2012. In this study, the periods from 2002 to 2012 and 2005 to 2012 for Iquitos and Ahmedabad, respectively, were looked at. In the next part, we talk in-depth about the global features of these datasets.

3.2. Data preprocessing

Using the information on dengue cases, mosquito vectors, and geographic information shown in Fig. 2, the appropriate staff gathered data and submitted it to the monitoring center. The instances were divided by source region using the conventional matching process. Before feature extraction, data cleaning is required because the raw data is crowded. Data cleansing includes grouping based on grid nodes and time slice segmentation based on timestamps. Four streets and more than 400 dwellings were selected to assess the mosquito population. After checking all outdoor water containers, the larval indices were calculated as shown in Eq. (1). Additionally, the variables sum_con and inv_house display the entire number of
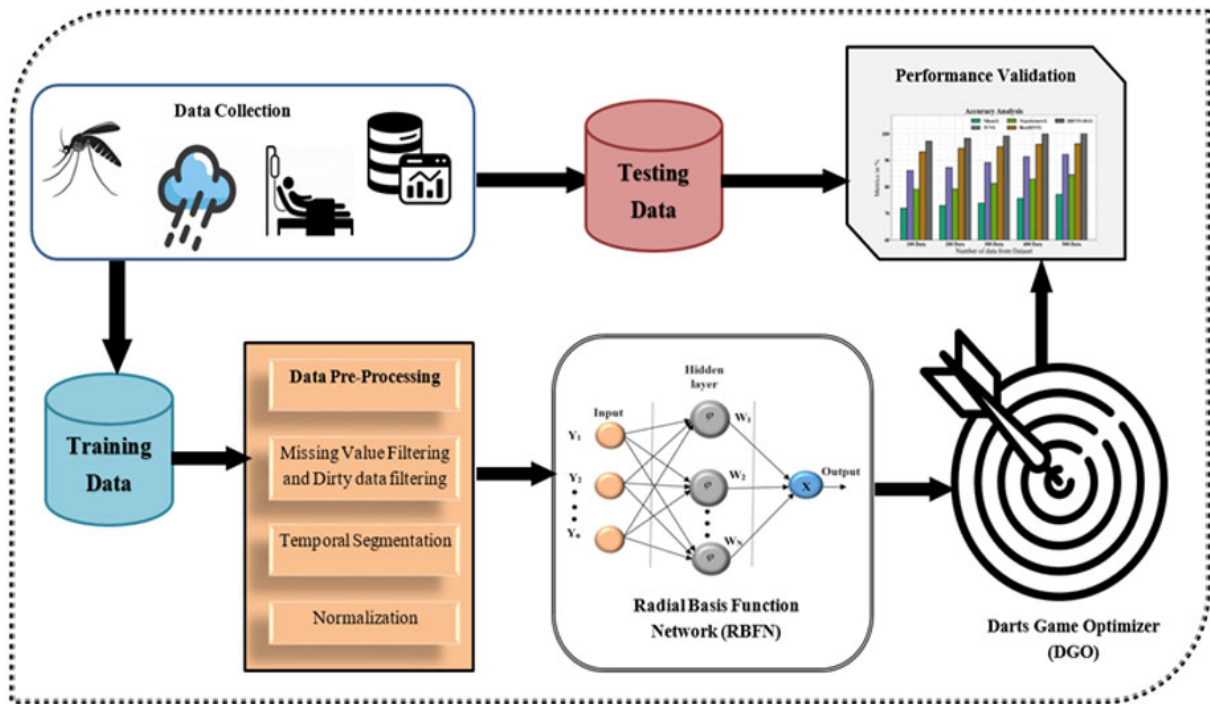
Fig. 1. Proposed diagram for RBFN-DGO.

studied containers and houses, respectively, while pos_con and pos_house display the percentage of containers and homes that contain mosquito larvae. It is crucial to remember that there is a strong link between dengue cases and space's three larval markers (BI, CI, and HI).

$$B = \frac{P_{con}}{In_{hou}} * 100, C = \frac{P_{con}}{S_{con}} * 100, H = \frac{P_{hou}}{In_{hou}} * 100 \tag{1}$$

We go on to data filtering and normalization after data segmentation to eliminate outliers. When data filtering and normalization are performed on time series data, advanced techniques, including interpolation, imputation, time-based imputation, multiple imputations, model-based imputation, normalization, dynamic temporal warping (DTW), and robust statistical measures are employed to handle missing values. Due to the frequency of sampling used for mosquito vector monitoring, time series data must be filled in with missing values. One method for dealing with missing values is to compute the weighted average of nearby data points using the surrounding mean imputation method. Missing values in time series data are addressed using weighted average and forward-filling techniques, which preserve the temporal structure and use nearby observations. By combining neighboring data and considering trends, weighted average computes a weighted average of the closest data points. Forward forward-filling maintains temporal continuity and accommodates brief data outages by propagating the most recent observed value forward. Forward-filling is another strategy that takes advantage of the material connection in time series data and presupposes that missing values correspond to the variable measurements from the previous time step. One computationally effective technique for filling in missing values in time series data is forward-filling, which also preserves temporal patterns. Evaluation is essential because, while it is simple to use and appropriate for linear changes, it may not be suitable for clusters or nonlinear data distributions. Component-by-component feature extraction is finished following the data preprocessing shown in Fig. 2.
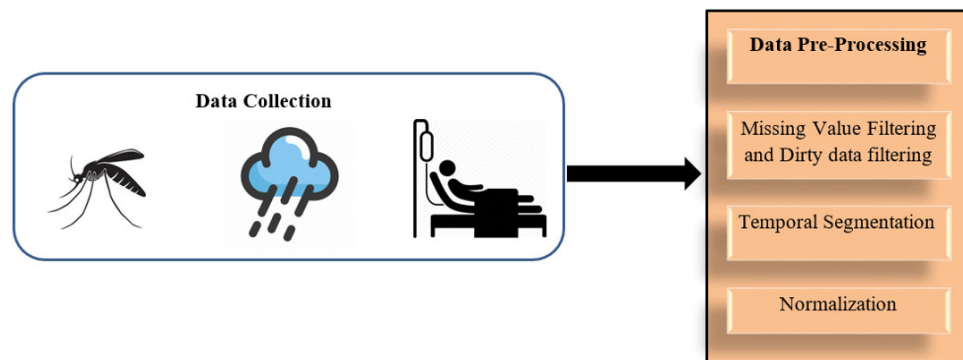
Fig. 2. Steps Involved In Data Preprocessing.

3.3. Radial basis function neural network

The proper set of potential diseases can be found by effectively organizing symptoms using the radial basis network. Radial basis networks are susceptible and clearly distinguish output based on input. During training, a network's number of neurons, spread, gain, and error value is fixed for a particular information set. The network may be used to determine the correct disease from the provided symptoms once it has been trained. An RBFN's ability to identify the right disease based on symptoms after training depends on several factors, including the caliber of the training data, the model's complexity, the feature representation, the training algorithm, the ability to generalize, and the selection of evaluation metrics. The user selects the sensitivity value when choosing how many alternatives are realistic. The network can be introduced again whenever necessary using fresh inputs and parameters. The model may adjust its parameters and better understand the link between input characteristics and target outputs by retraining the radial basis function network (RBFN) with new inputs and settings. Nonlinear modeling and pattern recognition applications of RBFNs include function approximation, time series prediction, control systems, financial forecasting, optimization, and anomaly detection. They are accommodating in engineering, finance, medicine, and telecommunications.

An RBFN, a particular kind of neural network, is utilized for issues including pattern recognition, function approximation, and time series prediction. Community engagement and education are essential to maximize the efficacy and sustainability of disease control programs. Their contributions to the intervention enhance its scalability and durability by ensuring cultural sensitivity, promoting behavioral change, offering social support networks, and providing local expertise and resources. This method guarantees disease management that is both sustainable and suitable for the culture. Although they are less effective when used directly as universal approximators of acceptable precision, RBF networks provide a PPDM solution. A traditional RBF network has layers for input, hidden, and output.

An RBF network's hidden layer can be measured as a function that transforms input patterns from a space that is not linearly distinguishable to one that is. In an RBFN, the hidden layer nonlinearly maps input patterns into a higher-dimensional feature space using radial basis functions, performing a transformational function. With the help of this transformation, the network's predictive power is eventually increased by enabling feature extraction, nonlinear mapping, dimensionality reduction, localized representation, and selective input pattern change. The responses of hidden layer neurons produce new feature vectors that indicate patterns in this altered space. Each output vector represents a collection of input patterns. The architecture of a radial basis network is based on this structure.

Through the association of input patterns with matching output responses, the capture of nonlinear relationships, the compact representation of input characteristics, and the ability to generalize to unseen data, the appropriate link between the output vector and the input dataset in an RBFN promotes pattern recognition. This relationship is the basis for the RBFN's capacity to carry out efficient pattern recognition tasks across various domains. The output vector Y is depicted below with the same size as the input data set X, where X is of size $N x m$:

$$Y = \begin{bmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,m} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ Y_{N,1} & Y_{N,2} & \ldots & Y_{N,m} \end{bmatrix} \tag{2}$$

$$X = (x1, x2, \ldots, x_m) \tag{3}$$

Now, for each row of

$$y_i = y_{i,1} \ y_{i,2} \ldots y_{i,m}$$

targets a row of

$$x_i = x_{i,1} \ x_{i,2} \ldots x_{i,m}$$

We aim to determine which related but unknowable variables Xj has the target function $f(y_i)$ with the lowest prediction error. Finding the weight vector W to locate Y with the least error is equivalent to doing this.

$$W = W_1 \ W_2 \ldots W_3 \tag{4}$$

$$X = f(y) \tag{5}$$

The Radial Basis Network is used to select the function as a radial basis function.

$$f(y) = \sum_{k=1}^{p} w_k \phi_k(y) \tag{6}$$

The radial function can be written as follows:

$$\phi_k(y) = \phi(\|y - y_k\|) \tag{7}$$

Where y_k denotes the neural activation center. Radial basis function's three primary parameters are

– Centre y_k
– Distance Measure $a = \|y - y_k\|$
– The radial basis function's shape

Figure 3 demonstrates how to feed an RBF network with anonymization input and an example function approximator output. It illustrates how the network's output deviates from its input values and instead adopts its course.

The network is created after selecting a training set of data for it.

$$T = \left\{ \left\{ y^k, x^k \right\} \right\}_{k-1}^{p} \tag{8}$$

The objective has been established as

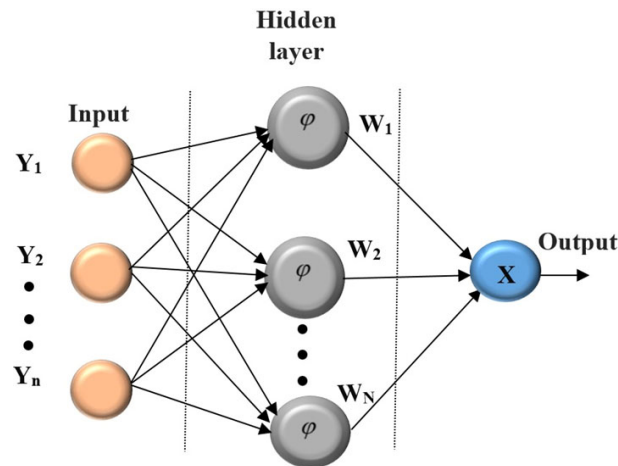$$x^{(k)} \approx f\left(y^{(k)}\right) \tag{9}$$

Fig. 3. RBFN framework.

3.4. Darts game optimizer (DGO)

Everybody, regardless of age or gender, appreciates the sport of darts. Although this game appears straightforward, the players' skill significantly impacts the point total. In darts, accuracy and precision are rewarded by a scoring system that doubles for darts outside the triple ring and triples for those inside. Higher expertise and accuracy are needed for the triple ring, whereas the double ring is said to be less exact. The system encourages Players to aim for accuracy, which rewards more for successful throws in the smaller triple ring and less for those in the wider double ring. Darts and a dartboard are included in the game's equipment [20]. The Darts Game Optimizer employs a sophisticated optimization technique comprising target selection, throw strategy, risk assessment, and feedback loop. By considering player performance, scoring zone layout, and target values, it focuses on optimizing scoring potential and game objectives. The dartboard is divided into 82 sections, each with a different point value. The inner bull receives 50 points for each toss, while the outer bull is awarded 25 points. The results are displayed above for each sector. The score is tripled if the darts land inside the little ring, doubling if they do not. The center of the dartboard is highly significant because it ranks as the fifth-highest-scoring region in the game. This section creates and shows a unique optimizer using the darts game's potential. Players in this game are DGO searcher agents, and the objective of the game is to get the highest score. DGO searcher agents may be used for a variety of optimization tasks because of their properties, which include population-based optimization, crossover operators, local and global search, group formation, and differential evolution.

3.4.1. Mathematical modeling

A matrix represents the population of players, with each row denoting a single individual and each column indicating that player's many traits. The matrix's column count reflects the number of problem variables and their recommended values.

In this case, Z stands for the players' matrix, z_i^b for the ith player's bth dimension, m for the number of variables, and N for the overall number of players.

Equation (10), which shows the data as a result of including z_i in the fitness function, describes the results.

$$F_{best} = \min{(fit)}_{N \times 1} \tag{10}$$

Table 1
Information on different sectors in the dartboard

Sector	Area (mm2)	Angel interval (degree)
Treble ring	278.9734	$0 \leqslant O \leqslant 18$
Outer bull	$2.6702 \times 10 + 3$	$0 \leqslant O \leqslant 360$
Double ring	556.6902	$0 \leqslant O \leqslant 18$
Outer single score	$5.3535 \times 10 + 3$	$0 \leqslant O \leqslant 18$
Inner bull	506.7075	$0 \leqslant O \leqslant 360$
Inner single score	910.1131	$0 \leqslant O \leqslant 18$

$$Z_{best} = Z \left(location of \min (fit), 1:m\right) \tag{11}$$

$$F_{worst} = \max (fit)_{N \times 1} \tag{12}$$

$$Z_{worst} = Z \left(location of \max (fit), 1:m\right) \tag{13}$$

$$F^n = \frac{fit - F_{worst}}{\sum_{j=1}^{N} \left(fit_j - F_{worst}\right)} \tag{14}$$

$$P_i = \frac{F_i^n}{\max (F^n)} \tag{15}$$

The ideal variable values are represented by Z_{best}, while the perfect fitness function value is represented by F_{best}. The least favorable fitness function value is denoted by F_{worst}, and the corresponding unfavorable variable values are indicated by Z_{worst}. P_i represents the probability function for the ith player, while F^n it stands for the normalized fitness function value.

All dartboards have the exact measurements, which are as follows:

– The inner diameters of the double and treble rings are measured at 8 mm
– A bull's inner dimensions: 12.7 mm
– The circumference of the outer bull's exterior diameter: 31.8 mm
– The space between the treble wire and the middle bullseye. 107 mm
– From the innermost to the outermost bullseye wire: 170 mm
– From one side of the dual wire to the opposite side of the dual wire: 340 mm
– A dartboard's entire size: 451 mm

As previously noted, the dartboard has 82 separate spots, and each has been given a unique score. Every player gets three chances to fire darts during each round. Players' skills and random opportunities work together to choose where these darts land on the dartboard.

The six separate sectors of the dartboard are shown in Table 1, each with a different area.

As a result, each participant's throwing score is generated and simulated using Eq. (16) through Eq. (19).

$$H_i = round \left(82 \times (1 - P_i)\right) \tag{16}$$

$$KH_i = \begin{cases} K (1:H), \ rand < P_i \\ K (H+1:82), \ else \end{cases} \tag{17}$$

$$k_i = KH_i (l) \& 1 \leqslant l \leqslant 82 \tag{18}$$

$$k_i^n = \frac{\sum_{throws=1}^{3} k_i^{throws}}{180} \tag{19}$$

The KH_i denotes the ratings of potential players for the player with index "i." K denotes a scoring matrix with scores listed in decreasing order. While k_i^n it stands for the player's score after it has been normalized, it k_i represents the player's score for each throw.

Finally, using Eq. (19), the values of the issue variables and each player's new status are modified.

$$Z_i = Z_i + rand\,(1,m) \times (Z_{best} - 3k_i^n Z_i) \tag{20}$$

3.4.2. Steps of DGO

After the simulation, various optimization issues can be solved using the suggested DGO algorithm. The steps in the optimization process include initiation, assessment, selection, mutation, replacement, recombination, and termination. Iterations are employed to choose parents, develop solutions, spread out the population, and balance exploitation and exploration. Termination criteria identify the optimal course of action and guarantee ongoing development.

Identifying several crucial elements is necessary to define the Digital Growth Objective (DGO). The population of the algorithm consists of 50 players, and the method is stopped after 1000 iterations. A random group of players is created at the start of the process.

A suggested solution to the problem is represented by each participant in the proposed method, which is an m-member vector. The components of this vector represent the variables that make up the issue and are used in the objective function to determine how serious it is. A weighted sum is the objective function used to evaluate the severity of the problem. The weight of each vector component represents its relative relevance, enabling a thorough assessment of the issue's severity.

The various phases for implementing the DGO are as follows:

```
Start DGO
Step1: Creating the initial population of players.
Step2: Calculating the fitness function.
Step3: Updating F_best , Z_best , F_worst , and Z_worst using (2) to (5).
Step4: Updating F^n and P_i using (9) and (12).
Step5: Calculating k_i^n using (14) to (17).
Step6: Updating Z_i using (18).
Step7: Checking the stop condition
Step8: Printing solution.
End DGO
```

4. Result and discussion

4.1. Experimental setup

The radial basis network is implemented using Matlab, and the outcomes are validated. The following procedures should be followed to create a Radial Basis Function Network (RBFN) in MATLAB: data preparation, network design, training, validation, fine-tuning, experimentation with learning algorithms, regularisation approaches, and training parameter optimization. Performance and accuracy are improved by this procedure, which guarantees stable training, generalization, and correct forecasts. The spread,

objectives, and number of neurons in the network are adjusted, and the best result is chosen based on performance indicators. Matlab, which also offers graphical output representation, makes it simple to change the tuning settings for the RBFN network. RBFNs are sophisticated models used to manage nonlinear relationships in epidemiological datasets efficiently. They use interpretability, scalability, clustering techniques, local receptive fields, nonlinear mapping, and adaptive learning to manage high-dimensional datasets. Large datasets and public health decision-making can benefit from these qualities, which offer insightful information on disease patterns, risk factors, and intervention tactics. The first training (in-sample) and testing (out-of-sample) parts are built using the dengue and rainfall datasets. Assessing an RBFN's performance and generalization capacity is imperative by testing it on both in-sample and out-of-sample datasets. While out-of-sample testing gauges the model's ability to generalize to new data, in-sample testing measures the model's fit to the training dataset. By ensuring consistency in performance, comparing the findings improves the accuracy and dependability of the model. The extendable predictive potential of the notion is examined in this study using two separate test sets, each lasting 26 and 52 weeks.

4.2. Performance validation

We evaluate the effectiveness of our proposed RBFN-DGO to several deep learning frameworks, such as NBeatsX [21], TCNX [22], TransformersX [23], and BlockRNNX [24], all of which use exogenous variables. Exogenous variables improve interpretability, generalization, resilience to data changes, and domain-specific adaptability in deep learning frameworks. They also reduce overfitting and facilitate well-informed decision-making across various applications and domains. RBFN-DGO is a deep learning framework that outperforms other frameworks such as NBeatsX, TCNX, TransformersX, and Block-RNNX in forecasting vector-borne illness risk because of its hybrid technique that efficiently captures complicated dynamics in epidemiological datasets. Based on several factors, including prediction accuracy, interpretability, computational efficiency, robustness, generalization, scalability, and domain-specific considerations, the best model among NBeatsX, TCNX, TransformersX, BlockRNNX, and RBFN-DGO is chosen. Setting these criteria in order of importance and carrying out comparison studies assist stakeholders in making an educated choice based on their unique application requirements. Our analysis uses metrics like Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Symmetric Mean Absolute Percent Error (SMAPE), and Mean Absolute Scaled Error (MASE). These metrics' formulas are as follows:

$$RMSE = \sqrt{\frac{1}{h} \sum_{i=1}^{h} (y_i - \hat{y}_i)^2} \tag{21}$$

$$MAE = \frac{1}{h} \sum_{i=1}^{h} |y_i - \hat{y}_i| \tag{22}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{23}$$

Tn, *Fp*, *Tp*, and *Fn*, respectively, are values that stand in for true positives, false positives, true negatives, and false negatives.

$$SMAPE = \frac{1}{h} \sum_{i=1}^{h} \frac{|\hat{y}i - yi|}{(|\hat{y}i| + |y_i|)/2} \times 100 \tag{24}$$

Table 2
MAE Analysis for RBFN-DGO method

Number of data from dataset	NBeatsX	TCNX	TransformersX	BlockRNNX	RBFN-DGO
100	36.55	32.18	29.76	26.87	21.98
200	37.19	32.98	29.12	27.19	22.76
300	37.98	33.18	30.45	27.99	23.18
400	38.19	34.67	30.12	27.87	24.44
500	39.98	35.19	31.76	28.21	25.98
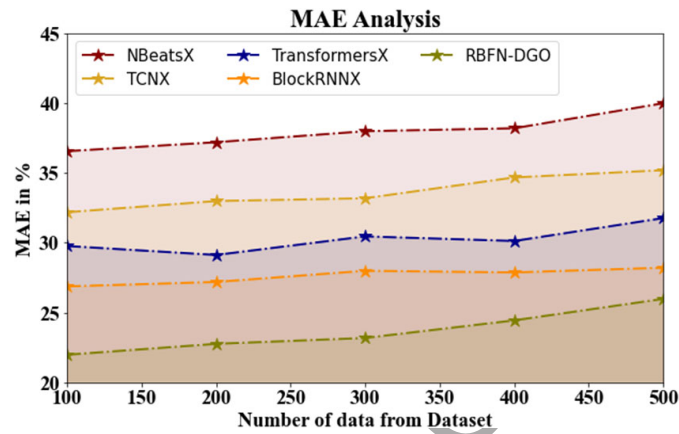
Fig. 4. MAE Analysis for RBFN-DGO method.

$$MASE = \frac{1}{h} \frac{\sum_{i=N=1}^{N+h} |\hat{y}i - yi|}{\frac{1}{N^-}|y_i - y_{i_-}|} \tag{25}$$

Where the forecasted value is yi, and the actual value is yi, N indicates how seasonal the series is, whereas h is the size of the forecasting period. The forecasting period (h) size significantly impacts model assessment, model complexity, and prediction accuracy. More extended periods call for larger models, whereas shorter intervals call for more precise forecasts. The extent of the period is also influenced by the assessment criteria selected, which may introduce prediction horizon effects. By convention, the "best" model is the one that has the lowest numbers for these metrics [25].

4.2.1. MAE analysis

Figure 4 and Table 2 display an MAE study of the RBFN-DGO methodology compared to other existing procedures. The graph demonstrates how the DLarning method improved performance with reduced MAE. In difference, the MAE values for the NBeatsX, TCNX, TransformersX, and BlockRNNX models are 36.55%, 32.18%, 29.76%, and 26.87%, respectively, whereas the RBFN-DGO model has MAE of 21.98% with 100 data. The RBFN-DGO model, however, has performed well with low MAE values for various data. The NBeatsX, TCNX, TransformersX, and BlockRNNX models' respective MAE values are 39.98%, 35.19%, 31.76%, and 28.21%, while the RBFN-DGO model's MAE value is 25.98% with 500 data.

4.2.2. SMAPE analysis

Figure 5 and Table 3 provide a SMAPE comparison of the RBFN-DGO methodology with other methodologies. The graph demonstrates how the DLng technique has improved performance with a

Table 3
SMAPE Analysis for RBFN-DGO method

Number of data from dataset	NBeats	TCNX	TransformersX	BlockRNNX	RBFN-DGO
100	45.19	42.22	38.19	35.19	31.98
200	45.87	42.67	39.98	36.19	31.56
300	46.18	43.18	40.12	36.98	32.28
400	46.88	44.56	40.67	37.12	33.45
500	47.12	44.98	41.18	37.98	34.19

Table 4
RMSE Analysis for RBFN-DGO method

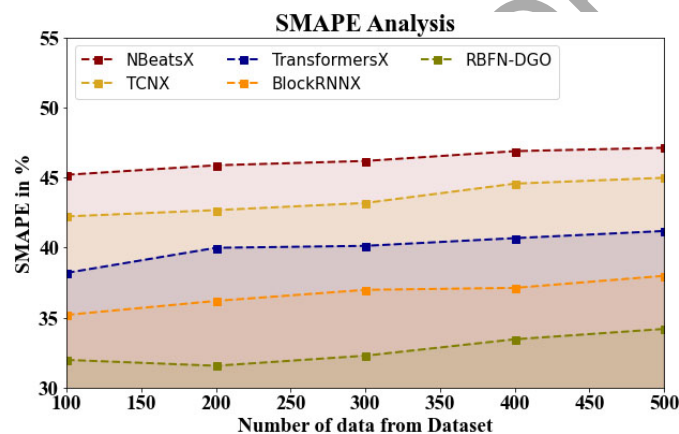| Number of data from dataset	NBeatsX	TCNX	TransformersX	BlockRNNX	RBFN-DGO
100	34.19	38.19	30.65	41.98	27.12
200	34.98	38.67	31.23	42.22	27.98
300	36.19	39.12	31.98	43.76	28.13
400	36.12	39.99	32.87	44.55	28.98
500	37.77	40.12	33.17	45.11	29.12

Fig. 5. SMAPE Analysis for RBFN-DGO method.

decrease in SMAPE. In difference, the SMAPE values for the NBeatsX, TCNX, TransformersX, and BlockRNNX models are 45.19%, 42.22%, 38.19%, and 35.19%, respectively, whereas the RBFN-DGO model has a SMAP of 31.98% with 100 data. The RBFN-DGO model, however, performs well with low SMAPE values for various data. The NBeatsX, TCNX, TransformersX, and BlockRNNX models' respective SMAPE values are 47.12%, 44.98%, 41.18%, and 37.98%, whereas the RBFN-DGO model's SMAPE value for 500 data is 34.19%.

4.2.3. RMSE analysis

In Fig. 6 and Table 4, an RMSE assessment of the RBFN-DGO methodology is shown in contrast to other methods. The graph demonstrates how technology has improved performance with low RMSE. Low RMSE and MASE values signify superior performance over a naïve prediction, predictive solid accuracy, and near alignment between observed and anticipated values. Achieving minimum values increases forecasting effectiveness and decision support confidence. In difference, the RMSE values for the NBeatsX, TCNX, TransformersX, and BlockRNNX models are 34.19%, 38.19%, 30.65%, and

Table 5
Accuracy analysis for RBFN-DGO method

Number of data from dataset	NBeatsX	TCNX	TransformersX	BlockRNNX	RBFN-DGO
100	71.87	86.18	78.98	93.15	97.12
200	72.98	87.23	79.13	94.44	98.23
300	73.76	89.17	81.23	95.12	99.13
400	75.55	91.34	82.87	95.98	99.87
500	77.12	92.17	84.55	96.13	99.98
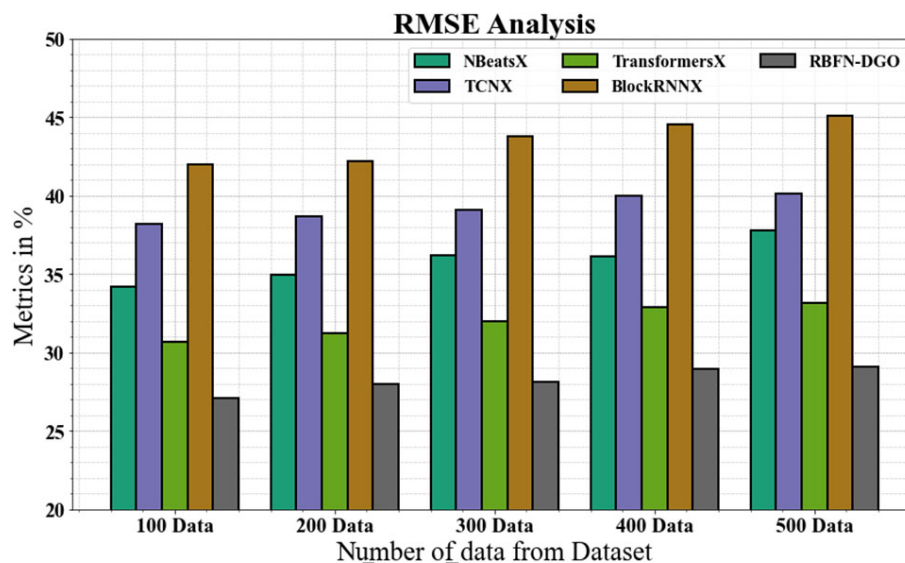
Fig. 6. RMSE analysis for RBFN-DGO method.

41.98%, respectively, whereas the RBFN-DGO model has an RMSE of 27.12% with 100 data. The RBFN-DGO model has, however, performed best for various nodes with low RMSE values. The RMSE value for the RBFN-DGO model with 500 data is 29.12% as opposed to the NBeatsX, TCNX, TransformersX, and BlockRNNX models' respective RMSE values of 37.77%, 40.12%, 33.17%, and 45.11%.

4.2.4. Accuracy analysis

Figure 7 and Table 5 associate the RBFN-DGO methodology's accuracy with other methods. The graph demonstrates the increased accuracy and efficiency of the deep learning approach. For instance, the NBeatsX, TCNX, TransformersX, and BlockRNNX models' respective accuracy values for 100 data are 71.87%, 86.18%, 78.98%, and 93.15%, respectively, as opposed to the RBFN-DGO model's accuracy of 97.12%. However, the RBFN-DGO model has been shown to perform best with a large amount of data. Like this, with 500 data, the RBFN-DGO has an accuracy of 99.98%, while the corresponding accuracy values for NBeatsX, TCNX, TransformersX, and BlockRNNX are 77.12%, 92.17%, 84.55%, and 96.13%.

4.2.5. Execution time analysis

In Table 6 and Fig. 8, the execution time of the proposed RBFN-DGO methodology is compared to that of existing methods, where the RBFN-DGO technique has outperformed all the other methods. The suggested RBFN-DGO approach, for example, took only 1.456 ms to execute 100 data. In contrast, other

Table 6
Execution time analysis for RBFN-DGO method

Number of data from dataset	NBeatsX	TCNX	TransformersX	BlockRNNX	RBFN-DGO
100	14.987	10.567	8.134	5.198	1.456
200	14.543	11.321	8.567	5.987	2.187
300	14.119	11.876	8.456	6.123	2.834
400	15.678	12.876	9.113	6.987	3.987
500	15.998	13.234	9.987	7.145	4.456

Table 7
MASE analysis for RBFN-DGO method

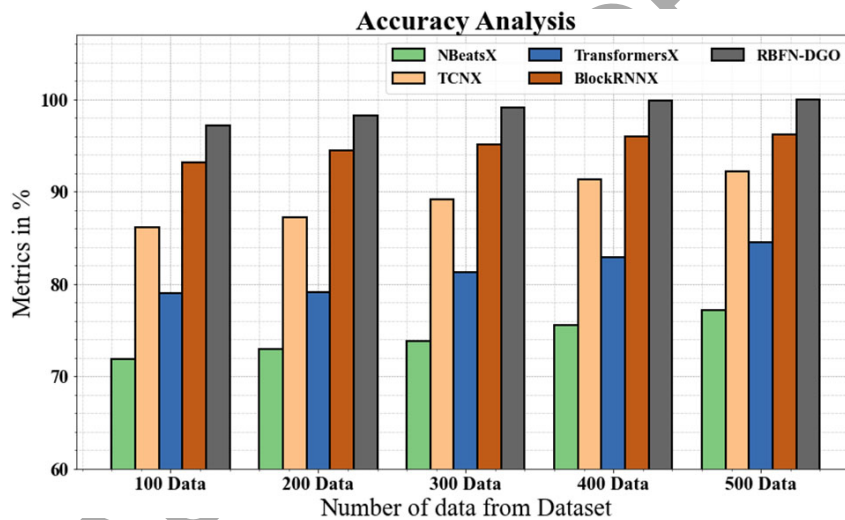| Number of data from dataset	NBeatsX	TCNX	TransformersX	BlockRNNX	RBFN-DGO
100	40.13	43.33	37.19	46.19	34.18
200	41.13	43.87	37.98	46.67	35.19
300	41.76	44.55	38.11	47.17	35.98
400	42.87	44.12	38.65	47.99	36.19
500	42.99	45.66	39.12	48.88	36.98

Fig. 7. Accuracy analysis for RBFN-DGO method.

current methods, such as NBeatsX, TCNX, TransformersX, and BlockRNNX, have taken 14.987 ms, 10.567 ms, 8.134 ms, and 5.198 ms, respectively, as their execution time. Similarly, the suggested RBFN-DGO approach takes 4.456 ms to execute 500 data, while existing techniques like NBeatsX, TCNX, TransformersX, and BlockRNNX have taken 15.998 ms, 13.234 ms, 9.987 ms, and 7.145 ms, respectively.

4.2.6. MASE analysis

Figure 9 and Table 7 demonstrate an MASE comparison of the RBFN-DGO methodology to other methodologies. The performance, stability, seasonality, generalization, and complexity of forecasting models are shown by the MASE comparison. Variations offer complexity trade-offs and flexibility, whereas lower MASE values suggest more accuracy. The graph demonstrates how the DL technology improved performance while decreasing MASE. In difference, the MASE values for the NBeatsX, TCNX, TransformersX, and BlockRNNX models are 40.13%, 43.33%, 37.19%, and 46.19%, respectively,
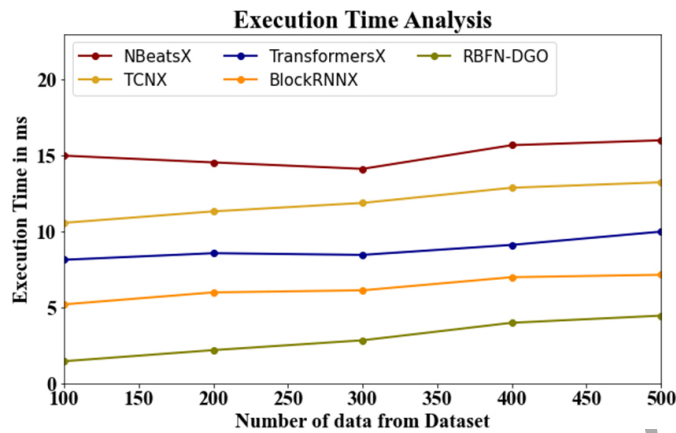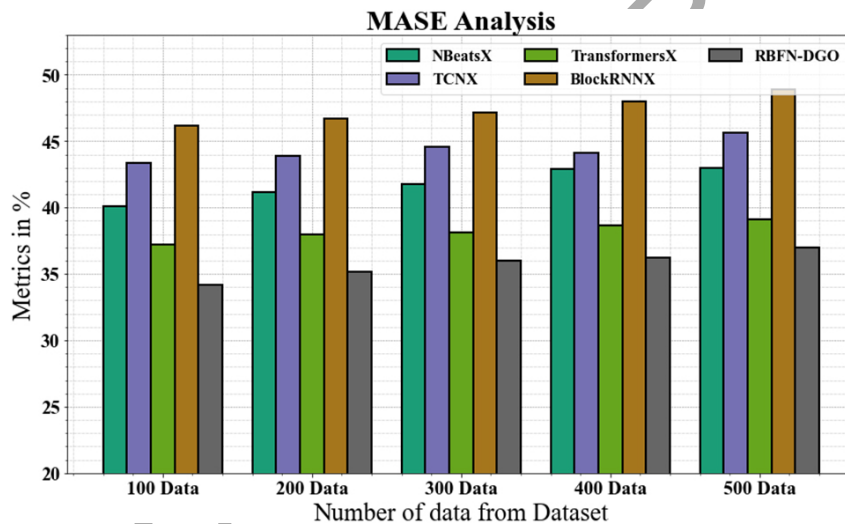
Fig. 8. Execution time analysis for RBFN-DGO method.

Fig. 9. MASE analysis for RBFN-DGO method.

whereas the RBFN-DGO model has an MASE of 34.18% with 100 data. The RBFN-DGO model, however, performs best with low MASE values for various data. The NBeatsX, TCNX, TransformersX, and BlockRNNX models' respective MASE values are 42.99%, 45.66%, 39.12%, and 48.88%, while the RBFN-DGO model's MASE value is 36.98% for 500 data.

4.2.7. Training and testing validation

Training accuracy is a metric used to assess how well the model performed on the training data. The model iteratively modifies its parameters to fit the data better to increase the training accuracy. The model successfully picks up on the patterns and trends in the training data when the training accuracy is high, as shown in Fig. 10.

Testing and training loss analysis of the RBFN-DGO approach is essential for verifying vector-borne illness risk predictions, assessing model performance, and fine-tuning parameters. While testing loss evaluation shows the generalization and dependability of the model, monitoring training loss aids in

Training and Testing Accuracy

Fig. 10. Training and testing accuracy analysis.
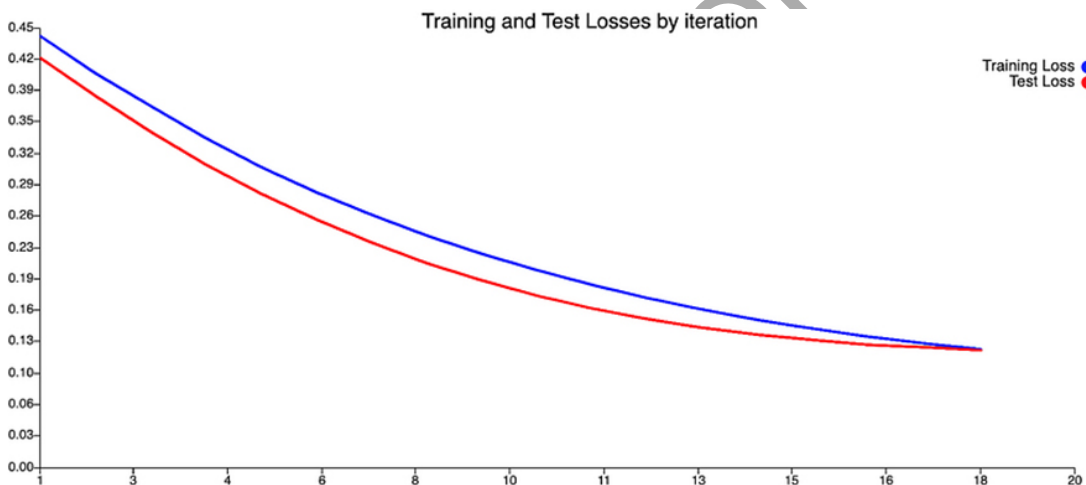
Training and Test Losses by iteration

Fig. 11. Training and testing loss analysis.

the detection of overfitting. Model selection is aided by comparing testing loss among models, and consistently low loss numbers confirm the applicability of the methods in real-world scenarios.

This dataset contains actual observations of the occurrences of vector-borne diseases and pertinent environmental variables. The testing accuracy is determined by comparing the model's predictions with the exact circumstances of vector-borne diseases. The deep learning model used in the hybrid methodology produces predictions on this unobserved data, as given in Fig. 11.

5. Conclusion

In conclusion, the RBFN-DGO to forecasting deep learning-based risk assessment for vector-borne diseases is a significant development in epidemiology and public health. In this work, we developed a hybrid model that performs better in forecasting dengue epidemics. With this novel method, consistent risk

valuation techniques are combined with the strength of deep learning algorithms to provide a prediction model that is more accurate and dependable. Complete acceptance of the factors impacting the occurrence and transmission of vector-borne diseases and the RBFN-DGO integrates numerous data sources, such as climatic data, geographic information, historical illness trends, and vector population dynamics, to provide a detailed. Health authorities and legislatures can conduct battered preventive initiatives to minimize the impact of these diseases on vulnerable groups, more professionally allocate resources, and make conversant decisions due to this level of integration and complexity. The DGO algorithm efficiently searches the RBFN constraint space in our proposed hybrid RBFN-DGO model, which improves the model's structure to increase forecast accuracy. RBFN-DGO offers a probable method for predicting the risk of diseases spread by vectors. This study advances public health predictive modeling by shedding light on the effective control of vector-borne diseases for the security of human populations. Compared to the other existing systems' high performances for our proposed method, it also makes predictions with an accuracy of 99.98% and minimal RMSE and MASE values of 29.12% and 36.98%, respectively. Forecasting deep learning-based risk assessment of diseases transmitted by vectors using a hybrid method is a fascinating new field in disease surveillance, prevention, and management. The research identifies many obstacles to assessing deep learning models on hypothetical data, including sample size, uncertainty estimates, overfitting, generalization problems, and data drift. Utilizing the power of artificial intelligence, we can better understand the dynamics of disease, which will help us work toward a healthier and more secure future for those vulnerable to vector-borne diseases.

Funding

Conflict of interest

The authors do not have any conflicts of interest to report.

Data availability

The data that support the findings of this study are openly available from https://www.kaggle.com/datasets/snehalad/tabular-vector-borne-disease-dataset.

Author contributions

Ashok Kumar Nanda, R Thilagavathy, and G S K Gayatri Devi are responsible for designing the framework, analyzing the performance, validating the results, and writing the article. Abhay Chaturvedi, Chaitra Sai Jalda, and Syed Inthiyaz are responsible for collecting the information required for the framework, providing software, conducting critical reviews, and administering the process.

References

[1]　Zeng Z, Zhan J, Chen L, Chen H, Cheng S. Global, regional, and national dengue burden from 1990 to 2017: A systematic analysis based on the global burden of disease study 2017. E Clinical Medicine. 2021 Feb 1; 32. doi: 10.1016/j.eclinm.2020.100712.

[2] Polwiang S. The time series seasonal patterns of dengue fever and associated weather variables in Bangkok (2003–2017). BMC infectious diseases. 2020 Dec; 20: 1-10. doi: 10.1186/s12879-020-4902-6.

[3] Deb S, Deb S. An ensemble method for early prediction of dengue outbreak. Journal of the Royal Statistical Society Series A: Statistics in Society. 2022 Jan; 185(1): 84-101. doi: 10.1111/rssa.12714.

[4] Rangarajan P, Mody SK, Marathe M. Forecasting dengue and influenza incidences using a sparse representation of Google trends, electronic health records, and time series data. PLoS computational biology. 2019 Nov 21; 15(12): e1007518. doi: 10.1371/journal.pcbi.1007518.

[5] Chakraborty T, Chattopadhyay S, Ghosh I. Forecasting dengue epidemics using a hybrid methodology. Physica A: Statistical Mechanics and its Applications. 2019 Aug 1; 527: 121266. doi: 10.1016/j.physa.2019.121266.

[6] Shaw WR, Catteruccia F. Vector biology meets disease control: using basic research to fight vector-borne diseases. Nature microbiology. 2019 Jan; 4(1): 20-34. doi: 10.1038/s41564-018-0214-7.

[7] Munirathinam B, Aroulanandam VV, Saravanan P. An efficient vector-borne disease prediction using SS optimization-based hybrid support vector random forest model. Signal, Image and Video Processing. 2023 Nov; 17(9): 3943-52. doi: 10.1007/s11760-023-02624-w.

[8] Ali SA, Ahmad A. Spatial susceptibility analysis of vector-borne diseases in KMC using geospatial technique and MCDM approach. Modeling Earth Systems and Environment. 2019 Sep 1; 5: 1135-59. doi: 10.1007/s40808-019-00586-y.

[9] Carvalho SA, da Silva SO, Charret ID. Mathematical modeling of dengue epidemic: control methods and vaccination strategies. Theory in Biosciences. 2019 Nov; 138(3): 223-39. doi: 10.1007/s12064-019-00273-7.

[10] Jayaraj VJ, Avoi R, Gopalakrishnan N, Raja DB, Umasa Y. Developing a dengue prediction model based on climate in Tawau, Malaysia. Acta tropica. 2019 Sep 1; 197: 105055. doi: 10.1016/j.actatropica.2019.105055.

[11] Chakraborty T, Chattopadhyay S, Ghosh I. Forecasting dengue epidemics using a hybrid methodology. Physica A: Statistical Mechanics and its Applications. 2019 Aug 1; 527: 121266. doi: 10.1016/j.physa.2019.121266.

[12] Jagadesh S, Combe M, Ginouvès M, Simon S, Prévot G, Couppié P, Nacher M, Gozlan RE. Spatial variations in Leishmaniasis: A biogeographic approach to mapping the distribution of Leishmania species. One Health. 2021 Dec 1; 13: 100307. doi: 10.1016/j.onehlt.2021.100307.

[13] Rahman MS, Pientong C, Zafar S, Ekalaksananan T, Paul RE, Haque U, Rocklöv J, Overgaard HJ. Mapping the spatial distribution of the dengue vector Aedes aegypti and predicting its abundance in northeastern Thailand using machine-learning approach. One Health. 2021 Dec 1; 13: 100358. doi: 10.1016/j.onehlt.2021.100358.

[14] McMahon A, Mihretie A, Ahmed AA, Lake M, Awoke W, Wimberly MC. Remote sensing of environmental risk factors for malaria in different geographic contexts. International Journal of Health Geographics. 2021 Dec; 20: 1-5. doi: 10.1186/s12942-021-00282-0.

[15] Ferdousi T, Cohnstaedt LW, Scoglio CM. A Windowed correlation-based feature selection method to Improve Time Series Prediction of Dengue Fever cases. IEEE Access. 2021 Oct 14; 9: 141210-22. doi: 10.1109/ACCESS.2021.3120309.

[16] Anggraeni W, Sumpeno S, Yuniarno EM, Rachmadi RF, Gumelar AB, Purnomo MH. Prediction of dengue fever outbreak based on climate factors using fuzzy-logistic regression. In 2020 international seminar on intelligent technology and its applications (ISITIA). 2020 Jul 22 (pp. 199-204). IEEE. doi: 10.1109/ISITIA49792.2020.9163708.

[17] Appice A, Gel YR, Iliev I, Lyubchich V, Malerba D. A multi-stage machine learning approach to predict dengue incidence: a case study in Mexico. Ieee Access. 2020 Mar 13; 8: 52713-25. doi: 10.1109/ACCESS.2020.2980634.

[18] Li L, Fang Z, Zhou H, Tang Y, Wang X, Liang G, Zhang F. Dengue risk forecast with mosquito vector: A multicomponent fusion approach based on spatiotemporal analysis. Computational and Mathematical Methods in Medicine. 2022; 2022(1): 2515432. doi: 10.1155/2022/2515432.

[19] Enduri MK, Jolad S. Estimation of reproduction number and non stationary spectral analysis of dengue epidemic. Mathematical Biosciences. 2017 Jun 1; 288: 140-8. doi: 10.1016/j.mbs.2017.03.007.

[20] Dehghani M, Montazeri Z, Givi H, Guerrero JM, Dhiman G. Darts game optimizer: A new optimization technique based on darts game. International Journal of Intelligent Engineering and Systems. 2020 Oct 1; 13(6): 286-94. doi: 10.22266/ijies2020.1031.26.

[21] Oreshkin BN, Carpov D, Chapados N, Bengio Y. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. arXiv preprint arXiv: 1905.10437. 2019 May 24.

[22] Chen Y, Kang Y, Chen Y, Wang Z. Probabilistic forecasting with temporal convolutional neural network. Neurocomputing. 2020 Jul 25; 399: 491-501. doi: 10.1016/j.neucom.2020.03.011.

[23] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł Polosukhin I. Attention is all you need. Advances in neural information processing systems. 2017; 30.

[24] Herzen J, Lässig F, Piazzetta SG, Neuer T, Tafti L, Raille G, Van Pottelbergh T, Pasieka M, Skrodzki A, Huguenin N, Dumonal M. Darts: User-friendly modern machine learning for time series. Journal of Machine Learning Research. 2022; 23(124): 1-6.

[25] Makridakis S, Spiliotis E, Assimakopoulos V. The M4 Competition: 100,000 time series and 61 forecasting methods. International Journal of Forecasting. 2020 Jan 1; 36(1): 54-74. doi: 10.1016/j.neucom.2020.03.011.