Computing and Artificial Intelligence 2025, 3(1), 1489.
https://doi.org/10.59400/cai1489

Article

An automated diagnosis & classification of dengue using advance artificial neural network

Safdar Hayat[1,*], **Rahila Anwar**[2], **Sartaj Aziz**[3]

[1] Information Technology Wing, Prime Minister Secretariat, Muzaffarabad 13100, Pakistan
[2] Computing and Technology Department, Iqra University, Islamabad 44000, Pakistan
[3] Department of Computing and Technology, National University of Sciences & Technology, Islamabad 44000, Pakistan
*** Corresponding author:** Safdar Hayat, safdarhayat@ajk.gov.pk

CITATION

Hayat S, Anwar R, Aziz S. An automated diagnosis & classification of dengue using advance artificial neural network. Computing and Artificial Intelligence. 2025; 3(1): 1489.
https://doi.org/10.59400/cai1489

ARTICLE INFO

Received: 1 July 2024
Accepted: 26 November 2024
Available online: 11 December 2024

Abstract: In this research, an advanced artificial neural network (ANN)-based approach for prognosis and classification of dengue disease is presented. Dengue diagnosis usually relies on clinical assessment; subsequently, there might be a high probability of misdiagnoses due to the complex hodgepodge of symptoms of dengue with other vector-borne diseases. It is needed to develop a system that can help doctors to identify dengue disease much faster than the manual system, which takes longer time and more cost to detect the diseases. Such a system may help users to take an early action before it becomes serious. The study involved three phases: pre-processing, neural network processing, and post-processing. In the pre-processing phase, data were gathered from three high-severity dengue outbreak sites in Pakistan (Benazir Bhutto Hospital, CITI Lab Rawalpindi, and Meo Hospital Lahore) where the dengue outbreak severity was high during the year of 2011. After cleaning and normalizing, 768 samples were obtained, split into 560 for training and 208 for testing. Nineteen critical parameters were selected with input from physicians, medical staff, and prior research. This study presents a supervised feed-forward neural network (FFNN) with two hidden layers, trained using backpropagation and optimized with the Levenberg-Marquardt algorithm, achieving nearly 100% accuracy, minimal runtime, and a very low MSE (0.00000000000032521). The model reached 100% sensitivity, 99.8% precision, and 98.7% specificity, surpassing prior results in dengue diagnosis. The findings support improved diagnostic accuracy and confidence, providing a framework for physicians. Key factors in achieving optimal results include careful selection of architecture, data normalization, parameter selection, and critical evaluation.

Keywords: classification; multilayer perceptron; feed-forward back propagation; dengue expert system; dengue fever; diagnostic; advance artificial neural network (ANN)

1. Introduction

Mosquito-borne diseases are the most popular vector-borne diseases in Pakistan, especially dengue fever and DHF. A lot of research has been done on the diagnosis of dengue, but the fact is that it's a bit difficult to differentiate dengue from other mosquito-borne diseases like malaria, flavivirus, chikungunya, etc., because of intermixed symptoms. Dengue diagnosis usually relies on clinical assessment; subsequently, there might be a high probability of misdiagnoses due to the complex hodgepodge of symptoms of dengue with other vector-borne diseases. So, the physician mostly needs to confirm the dengue by using laboratory tests, for example, enzyme-linked immunosorbent assay (ELISA), PCR, and viral culture. But these serological tests are much more expensive and take so much time that it's unaffordable for middle- and lower-class families in Pakistan to conduct these tests. However, the clinical diagnosis generally resides on the foundation of dengue diagnosis and disease management [1].

Expert systems have become a major area of research in artificial intelligence and the most popular subject in artificial intelligence. Artificial intelligence techniques have been tried out in undertaking research. An expert system is ascertained as a solution to overcome the problem of correctly diagnosing and classifying dengue by identifying the complex intermixed symptoms with high accuracy. The undertaking expert system uses clinical assessments with a combination of specific vital laboratory tests, i.e., CBC enlist platelets count, WBC count, tourniquet test, and hematocrit (HCT % age), instead of serological tests, for early prognosis so that the earlier stage patient can survive and get an early treatment.

Mainly, this system can provide a clear record of the knowledge in early diagnosis of mosquito-borne disease dengue and make it easier to understand than conventional programming. The decision can be made quickly because the system is fully trained with 100% accuracy. By using it, the users can easily take an early action before the disease becomes serious by answering the questions based on the symptoms of the disease in addition to four laboratory findings, adequately without the expert. Moreover, for the diagnosis, the user does not need any serological test; he just needs a CBC test for entering the laboratory parameters, which may be conducted within minutes and also has a very low cost in Pakistan. The system provides a helping framework to the doctors to diagnose the victim correctly and provides the right treatment. Due to the intermixed manifestations of dengue with other mosquito-borne diseases, it is very difficult, even sometimes impossible, for the physicians to diagnose and classify dengue disease correctly and precisely, so this system facilitates the doctors as a right-hand helping tool to differentiate dengue disease among other vector-borne diseases like malaria, yellow fever, chikungunya, etc.

The paper is organized as follows. In the next section, we elaborated on the review of related work in dengue diagnosis. In Section III, we investigate the way of dengue diagnostics in which the methodology of expert system development life cycle (ESDLC) is explained. In Section IV, a proposed design and implementation phase of the dengue expert system is presented. Simulation results are presented in Section V. In this section the performance analysis and comparison of results with previous studies are evaluated. Finally, Section VI presents the conclusion.

1.1. Motivation

A human can easily recognize objects or patterns but be unsuccessful when probabilities have to be assigned to interpretation and observation. Research exposed that around half of the findings aren’t right due to not blocking the self-heart of a few doctors [2]. An individual dependably commits errors, and as a result of his restriction, mistakes do happen amid finding [3]. It implies the value of a conclusion is totally reliant on the ability and encounters of the specialist. Passionate issues, negligence, and weakness debase the specialist's execution [4].

Early research uncovers these fundamental realities in the medicinal zone. The aftereffects of undertaking analysis were clearly as follows [5]:

- Capability of expert human in diagnosis: 79.7%.

- Diagnosis by computer using an expert database: 82.2%.
- PC using nearly 600 patient samples: 91.1%.

It has been proved that in some cases computer-assisted diagnosis is more accurate than those by clinicians [6]. The conclusion is obvious: humans cannot evaluate complex data without errors. Then now what to do? Can neural networks get rid of us from this situation? The above scenario motivates us to develop an automated diagnosis & classification of dengue using advanced artificial neural networks.

1.2. Research objectives

Following are the main objectives concerning our research:

- To evaluate the concise and highly robust clinical symptoms that have a higher probability of association with the dengue virus.
- To develop an expert system in support of the medical diagnostic process of dengue disease, particularly in Pakistan.
- To review the advanced artificial neural network literature in expert systems and estimate the expert system's model that fits in the field of medication.
- To focus only on specific categories of diseases belonging to the specific disease domain and provide a decision support system tool to assist doctors in the early treatment process.
- To help the doctors in addition to their own expert knowledge and analysis to diagnose and classify the dengue precisely.
- To provide high accuracy in dengue diagnosis and classification as compared to previous work.
- To help users to take an early action in diagnosing vector-borne disease before it becomes more serious.
- To differentiate dengue fever from other vector-borne diseases by using expert system techniques.

2. Literature review

A human can easily recognize objects or patterns but be unsuccessful when probabilities have to be assigned to interpretation and observation. Research exposed that around half of the findings aren't right due to not blocking the self-heart of a few doctors [2]. An individual dependably commits errors, and as a result of his restriction, mistakes do happen amid finding [3]. It implies the value of a conclusion is totally reliant on the ability and encounters of the specialist. Passionate issues, negligence, and weakness debase the specialist's execution [4].

A lot of successful applications illustrate that the diagnostic capabilities of ANNs are considerably better than the diagnostic system of humans [7]. Here the short review of the prognostic capabilities of neural networks is discussed. There are lots of winning applications in medicine using neural networks, for example [8–12].

In 2010, Faisal et al. [13] constructed a non-invasive ANN problem-solving system for the assistance of doctors for categorization of the possibility in patients of dengue. They have used samples of 210 patients of dengue from the first day of fever (defervescence day) to the 5th day. They utilized the parameters, clinical manifestations, physiology measurements, and BIA measurements to create a model for prognosis. They train a multi-layer perceptron network using two LMA and SCGA algorithms with an activation function of hyperbolic tangent sigmoid in the hidden layer's neurons, while in the output layer's neuron they have used a transfer function of sigmoid. Their experiments using the Scaled Conjugate Gradient Algorithm achieved prediction accuracy of only 75%. On the other hand, using the Levenberg-Marquardt Algorithm, they could achieve the accuracy of 70.7% only.

Karim and Suryaningsih [14] have also developed a prototype of an expert system for diagnosing dengue fever using a decision tree and decision table as a knowledge base in asking 25 questions to patients to find out about the possibility of being infected by dengue fever. The system generated a sophisticated percentage of accuracy, but the conclusion was that there is still a possibility that the diagnostic result from the expert system has a low certainty factor.

Ibrahim et al. [15] uncovered a tenet base master framework to group hazards in dengue contaminations utilizing 17 bioelectrical impedance examinations (BIA) parameters. The information of 209 dengue-affirmed cases and 223 solid persons from University Kebangsaan Malaysia Hospital (HUKM) was gathered and used to build up a framework to characterize DF and DHF. Visual Basic 6.0 was utilized as an apparatus to plan and customize the framework. In conclusion, the framework could order the danger in dengue patients non-intrusively into three sorts of danger gatherings: higher danger, lower hazard, and no danger or sound gatherings with aggregate characterization precision of just 66.7%.

Rachata et al. in 2008 sent a managed neural system to assess the execution of "a programmed forecast arrangement of Dengue Haemorrhagic-Fever flare-up danger by utilizing entropy strategy and simulated neural system." The entropy idea was coordinated to change an information example to a component vector, which was utilized to prepare an administered encourage forward neural system. The 8 years of information from January 1999 until December 2007, about the state of climate and dengue hemorrhagic fever cases, were utilized. The information parameters—most extreme, least, and normal temperature; precipitation; and relative moistness—were utilized. They inferred that with the entropy system, the most elevated precision of 85.92% happens when they utilized four hidden neurons and the traingdx weight-modifying function [12]. However, without the entropy technique, they achieved a lower percentage of accuracy as compared to using the entropy technique.

During the year 2008, Nor Azura Husin and his team compared the results of "Nonlinear Regression Model (NLRM) and Neural Network Model (NNM) for early prediction of dengue outbreak" using different architectures and parameters, including time series, location, and rainfall [16]. Their research proposed the formula (2n + 1) for the selection of the optimal quantity of hidden layer neurons, where n is the number of inputs. They also recommended the appropriate learning rate, i.e., it should be 0.9, while the best momentum rate should be 0.5 for prediction. They evaluated the overall performance of architecture and concluded that ANN gives superior results than the nonlinear regression model. Their investigation concludes that the good architecture model is a key to better results of prediction.

Another rule-based research for dengue outbreak using comparatively rough set classifiers, decision trees, associative classifiers, and naive Bayes was analyzed by Bakar et al. [17]. They used data of 8505 dengue patients from the years 2003 to 2009 with 134 attributes in their experiments. Rough Set (RS), J48, CBA, and Naïve Bayes (DTNB) algorithms for each of the concerning classifiers were tested and measured the results. Their experiments revealed that multiple classifiers produced superior precision as compared to the single classifier. In conclusion, they succeeded in attaining the highest accuracy of 93.87% by using RS. Later on in 2015, Kamran Shaukat, et al. also analyzed the data mining approach to predict dengue fever [6]. Here they used NBRT, REP Tree, J48, and SOM techniques to analyze the result accuracy for classification. According to their experiments, the highest accuracy was 92% yielded by Naive Bayes, while using J48, they achieved an accuracy of 88%. Moreover, Kashish Ara Shakil and her team [18] also utilized data mining techniques, i.e., the WEKA tool, for the prediction of the survivability of dengue disease in 2015. Their experiments supported the conclusive theory of Shaukat et al. [6] as the accuracy of Naïve Bayes was 100% with 99 correctly classified instances and J48 given 99.70% as compared to REP tree and Random tree.

Furthermore, research [19] illustrates the prediction in early diagnosis of different grading stages (Grade-I, Grade-II, Grade-III, and Grade-IV) of DHF by using a statistical approach. The experiment was made by using the BIA parameters of 210 DHF-confirmed Malaysian patients of the years 2001–2002. The research signified that the multiple linear regressions are not a better method to predict the grading stages of DHF. They were only able to achieve the maximum diagnostic accuracy of 52.1%.

Raúl Beltrán Ramírez and his companions [20] moved one step ahead and developed an expert system as an Android-based mobile application for early diagnosis of influenza and dengue fever using fuzzy logic. The name of their application was BioDnX. Their system was not too sophisticated with respect to time complexity as compared to other techniques because it took approximately 2 to 4 minutes to calculate the probabilities of disease diagnosis, which is probably enough time in expert system diagnosis.

In 2013, M. Naresh Kumar utilized "an alternating decision tree" with a boosting technique to diagnose dengue and compared their performance with the C4.5 algorithm. They have trained the system by using clinical characteristics with a combination of some laboratory measurements. They achieved the highest prediction accuracy of only 89% using an alternating decision tree, while the C4.5-based decision tree was able to classify with an accuracy of only 78%. They conclude that an alternating decision tree with a boosting technique is much better as compared to the C4.5 base decision tree [3].

The students (Wajeeha et al.) of the University of Lahore [21] have performed two experiments for the classification of dengue using the decision tree approach and compared the results. The first general experiment pruned the attributes and classified the disease on behalf of values in the dataset and produced 99.44% accuracy, while in the second experiment they produced the accuracy of 98.62% by introducing the expert-weighted attributes and classified on behalf of minimum cost and resource availability. Their study concluded that we can compromise on the accuracy, specificity, and sensitivity if we have a lower average of Type II errors.

Baker et al. [22] analyze a weather-related dataset to predict the number of illness cases per week in the cities of San Juan and Iquitos by using several machine learning regression algorithms. They utilized and compared different machine learning regression techniques; the performance is evaluated using the Mean Absolute Error (MAE). Consequently, the Poisson Regression Model achieved the best ratios and the lowest mean absolute error ratio of 25.6%. The mean absolute error was used to evaluate these different models. The Poisson regression model obtained the lowest error ratio of 25.6. The paper contributed to testing a large number of machine learning models, obtaining varying error rates. Another research of Mayrose et al. [23] described a method that can support clinicians during dengue diagnosis. It was proposed to automate the peripheral blood smear (PBS) examination using artificial intelligence (AI) to aid dengue diagnosis. A machine learning-based technique was proposed to detect dengue from the images of PBS based on the lymphocyte nucleus. Ten features were extracted, including six morphological and four Gray Level Spatial Dependence Matrix (GLSDM) features, out of the lymphocyte nucleus of normal and dengue cases. Narrow Multilayer Perceptron (MLP) was used to differentiate dengue-infected and normal smears. ReLU was utilized as the non-linear activation function. K-fold cross-validation, with K = 10, was used to build and assess the predictive potential of the classifiers. The best performance was attained using an SVM classifier with 95.74% accuracy. SVM and MLP classifiers, respectively, obtained the best and second-best results.

Gupta et al. [24] proposed a machine learning method to forecast dengue fever. They are employing ML classifiers to evaluate the output according to the accuracy (mean) of disease prediction. They used a variety of machine learning classifiers, beginning with KNN and moving on to decision trees, random forests, Gaussian neighbor boundaries, and support vector classifiers. They utilized k-fold cross-validation to partition the data into ten equal portions for the purpose of classification. As a direct consequence of this, the mean value obtained after ten iterations was shown. They concluded that the random forest classifier was the one that turned out to be the best one, with a mean score of 8.72.

Another research of Ming et al. in 2022 included data of 8100 patients; in total, 2240 (27.7%) patients were diagnosed with dengue infection. Data was randomly split 80/20% into a training and hold-out set. Cross-validation and holdout set testing were used. They concluded that supervised machine learning models are able to discriminate between dengue and OFI diagnoses in patients presenting with an early undifferentiated febrile illness. The median performance in cross-validation of the optimal model was as follows: AUROC of 0.86 (interquartile range 0.84–0.86), specificity 0.92, sensitivity 0.56, positive predictive value (PPV) 0.73, and negative predictive value (NPV) 0.84. The model output used was probabilistic in nature, and isotonic calibration with 10-fold cross-validation was used, and the Brier score was 0.13, demonstrating good calibration. This model was able to maintain the negative predictive value of the model over the study period at a high (> 90%) level when validated, with direct implications for patient care [25].

Ho et al. [26] worked on comparing machine learning with case-control models to identify confirmed dengue cases. In this experiment, they set the dimension of the network to either 16 or 64 and the number of layers to 3, 10, or 100. The performance delivered by the three types of prediction models, i.e., DT, DNN, and LR models, was evaluated based on the area under the receiver operating characteristic (ROC) curve. They observed that for pediatric patients, the DT model delivered higher sensitivity but lower specificity (95.5% sensitivity and 54.8% specificity) than the DNN model, which delivered 87.1% sensitivity and 73.5% specificity.

In 2023, Thirugnanam and Hussain explored machine learning algorithms for the prediction of dengue. An array of machine learning algorithms, including logistic regression, K-nearest neighbor, support vector machine (SVM), decision tree, artificial neural network, and Naive Bayes classifier, were analyzed. Out of the 41 survey articles, it was discovered that 20% of research publications used SVM, 7% used ANN, 7% used KNN, and 5% used LR techniques. The remaining 15% DT, 20% NB, and "boosting" ensemble approach 12% and 5% CNN approaches. They concluded that the Naive Bayes algorithm was found to quickly generate predictions with a precision value of 99.1%. However, the SVM model outperformed all others with a cross-validation score of 98.5%, K-Fold validation of 97.5%, precision of 98.2%, and an F1 Score of 98.0%, thereby enhancing the overall performance of the predictive model [27].

In 2024, Bohm et al. [28] worked on the utilization of machine learning for dengue case screening. The mutual information technique was utilized to assess which variables were most related to laboratory-confirmed dengue cases. Next, a random selection of 10,000 confirmed cases and 10,000 discarded cases was performed, and the dataset was divided into training (70%) and testing (30%). Machine learning models were then tested to classify the cases. It was found that the logistic regression model with 10 variables and the decision tree and multilayer perceptron (MLP) models achieved the best results in decision metrics, with an accuracy of 98%. It was also illustrated that the decision tree achieved the best values of ACC and F1, while the MLP demonstrated the best performance in the AUC metric. A decision tree model achieved the best classification results (90.64%).

In 2020, Huang et al. developed prognostic models for the prediction of dengue severity in patients using machine learning. Machine learning models were trained and tested using the patient dataset that included demographic information and qualitative laboratory test results. To validate the prognosis performance of the models, they used a stratified 10-fold cross-validation approach with training and testing datasets. The discriminative ability of the artificial neural network exhibited good performance for severe dengue prognosis. The artificial neural network showed the highest average discrimination area under the receiver operating characteristic curve (0.8324 ± 0.0268) and balance accuracy (0.7523 ± 0.0273) [29].

Another work in 2021, by Salim et al. [30], illustrated climate data, such as humidity, rainfall, temperature, and wind speed, in the prediction of dengue outbreaks in Selangor, Malaysia, using machine learning techniques. The predictive modelling was performed using several data mining models, namely Decision Trees (CART), Artificial Neural Network (MLP), SVM (LINEAR, POLYNOMIAL, RBF), and Bayes Network (TAN). The models were evaluated based on classification accuracy, sensitivity, specificity, and precision. Based on testing sample results, the SVM Linear model (Accuracy = 70%, Sensitivity = 14%, Specificity = 95%, Precision = 56%) performed better than CART (Accuracy = 63%, Sensitivity = 12%, Specificity = 86%, Precision = 28%), Tree Augmented Naïve Bayes (Accuracy = 62%, Sensitivity = 27%, Specificity = 76%, Precision = 33%) and ANN (Accuracy = 66%, Sensitivity = 14%, Specificity = 95%, Precision = 37%). They concluded that the application of the machine learning models for prediction of dengue outbreaks can provide vital information to healthcare authorities so that they can better prepare for dengue fever outbreaks.

Yang et al. [31] in 2024 proposed a hybrid vision transformer model (Hybrid ViT) to jointly learn global and local information of HSI, including a convolution block and a vision transformer block. They concluded that low-level detailed features can be extracted via CNN architecture. Hybrid ViT is proposed to simultaneously extract small-scale fine-grained information and large-scale global distribution. Hybrid ViT can outperform other methods with higher classification accuracy and finer classification maps. Their experimental results proved that the proposed model can gain higher accuracy and clearer classification maps compared to other methods [31].

From all of the above-mentioned literature surveys, we can say that except for some sort of experiments and techniques, there is no research that is able to explicitly yield 100% results in prognosis and classification of dengue with minimum MSE, high precision, and high sensitivity. So still there is an exigent need to develop a precise and accurate expert system to diagnose the dengue at an early stage. An automated diagnostic and classification model is ascertained as a solution to overcome the problem of correctly diagnosing and classifying dengue by identifying the complex intermixed symptoms with high accuracy. Artificial intelligence techniques have been tried out in undertaking research. The undertaking expert system uses clinical assessments with a combination of specific vital laboratory tests. The following **Table 1** shows the brief summary of the literature review.

Table 1. Summary table of literature review.

Ref.	Year	Techniques	Datasets and Results
[13]	2010	MLP (Levenberg Marquardt) network	Used samples of 210 patients of dengue. Using Scaled Conjugate Gradient Algorithm achieved Prediction accuracy of only 75%. Using Levenberg-Marquardt Algorithm they could achieve the accuracy of 70.7% only.
[14]	2007	Decision tree and decision table as a knowledge base by asking 25 questions.	The system generated sophisticated percentage of accuracy but the conclusion was that there is still a possibility that the diagnose result from the expert system has low certainty factor.
[15]	2007	Rule base expert system	Utilized 17 bioelectrical impedance examinations (BIA) parameters of 209 dengue confirmed cases and 223 healthy persons. Gathered precision of just 66.7%.
[12]	2008	Entropy technique and artificial neural network	The 8 years information from January 1999 until December 2007, about state of climate and Dengue Hemorrhagic Fever cases were utilized. The most elevated precision of 85.92% happens when there utilized four hidden neurons and traingdx weight modifying function.
[17]	2011	Rough Set (RS)	They used data of 8505 dengue patient of the years 2003 to 2009 with 134 attributes their experiments. They succeeded to attain the highest accuracy of 93.87% by using RS.
[6]	2015	NBRT, REP Tree, J48 and SOM techniques	Highest accuracy was 92 % yielded by Nave byes while using J48 they achieved an accuracy of 88%.
[18]	2015	Data mining technique WEKA tool (Naïve Bayes)	100% with 99 correct classified instances.

Table 1. (*Continued*).

Ref.	Year	Techniques	Datasets and Results
[19]	2009	statistical approach	Used the BIA parameters of 210 DHF confirmed Malaysian Patients of year 2001–2002. Able to achieve the maximum diagnostic accuracy of 52.1%.
[20]	2013	"An alternating decision tree" with boosting technique	Used clinical characteristics with combination of some laboratory measurements. They achieved the highest prediction accuracy of only 89% using alternating decision tree while C4.5 based decision tree was able to classify with accuracy of 78% only.
[32]	2016	Multilayer perceptron and support vector machine	Dataset contained 303 samples suffering dengue fever and 346 unaffected patients. The entire dengue dataset is critically analyzed with the primary criteria for analysis being Platelet count (PLT) and Haematocrit (HCT). Dempster-Shafer based classifier fusion strategy theory gives as best as 96.02% with 4 MLPs and 3 SVMs while that of selected features using Fishers Score gives 96.56% accuracy.
[33]	2017	Greedy forward selection method has been applied to select most promising genes to identify the DF, DHF and normal (Either convalescent or healthy controlled) patients. The proposed system performance was compared to the multilayer perceptron feed-forward neural network (MLP-FFN) classifier.	The dataset contains gene expression data for DF, DHF, convalescent and healthy control patients of total 56 subjects. Results proved the dominance of the proposed method with achieved accuracy of 90.91 %.
[34]	2018	Used Data mining technique ANN, Weka and Netbeans IDE	In this study, the ANN model gets 5 inputs x(n) from rainfall, humidity rate, temperature rate, number of dengue cases from the previous month dataset. Bar graph and the training samples were used to test the accuracy of result and provided 95% of accuracy.
[35]	2019	Model used Support Vector Machine (SVM) algorithm to find the optimal loci classification subset, and then an Artificial Neural Network (ANN) was used to classify patients into dengue fever (DF) or severe dengue (SD).	One hundred and two Brazilian dengue patients and controls were genotyped for 322 innate immunity Single Nucleotide Polymorphisms (SNPs). The ANN trained on 13 key immune SNPs selected under dominant or recessive models produced median values of accuracy greater than 86%, and sensitivity and specificity over 98% and 51%, respectively.
[30]	2021	SVM Linear model,	Used climate data, such as humidity, rainfall, temperature and wind speed. (Accuracy = 70%, Sensitivity = 14%, Specificity = 95%, Precision = 56%)
[23]	2023	A Machine Learning based technique Narrow Multilayer Perceptron (MLP) was proposed to detect dengue from the images of PBS based on the lymphocyte nucleus.	Ten features were extracted, including six morphological and four Gray Level Spatial Dependence Matrix (GLSDM) features, out of the lymphocyte nucleus of normal and dengue cases. The best performance was attained using an SVM classifier with 95.74% accuracy.
[28]	2024	logistic regression model with 10 variables, the Decision Tree and Multilayer Perceptron (MLP) models were used	Random selection of 10,000 confirmed cases and 10,000 discarded cases was performed. A decision tree model, achieved the best classification results (90.64%)

3. Research methodology

The undertaking research had been developed using the following standard of Expert System Development Life Cycle (ESDLC). In general, the diagnosis process of the overall research consisted of the following three main steps:

3.1. First step: Pre-processing

The preprocessing stage includes the identification of the problem statement, knowledge acquisition, and knowledge representation, normalization of data, verification, and validation. In the pre-processing stage, the problem was identified firstly, and then all the related knowledge to overcome the problem was gathered. Then we represented the acquired knowledge according to that specific format, which could be used in the implementation phase. The one-year data of 768 patients were collected from Banzir Bhutto Hospital Rawalpindi, CITI Lab Rawalpindi, and Meo Hospital Lahore in the year of 2011. The whole sample was further separated into 560 training sets and 208 testing samples. From a total of 560 patients in the training set, 342 patients were serologically non-reactive while 218 patients were confirmed reactive. From all 218 dengue-confirmed patients, 143 suffered from dengue fever while 75 patients suffered from dengue hemorrhagic fever. The 560 training and 208 testing examples, along with their proper target sets, were successfully generated from the knowledge acquisition. The targets were logically represented into three main classes, i.e., No Dengue (class 00), Dengue Fever (class 01), and Dengue Hemorrhagic Fever (class 10). The input parameters are converted into numeric values for ANN processing. The collected data was then normalized, and the targets were set for every training example. After this, according to the ratio of 70% and 30%, the whole data was separated into training and testing sets. After the normalization and representation of the training and testing sets, we make sure the symptoms that are selected as final input examples are the correct symptoms to diagnose dengue. We also verified whether the targets that are set for each instance correctly represent the right class or not. After verification and validation, the data was then prepared for the neural net.

3.2. Second step: Artificial neural net processing

This is actually an implementation stage. The total training instances along with their target sets were presented to the ANN using the MATLAB tool. Then the trained ANN diagnosed and classified the dengue disease correctly.

3.3. Third step: Post-processing

In the step of post-processing, the outcomes of the artificial neural net were evaluated and transformed into a further comprehensible shape. The results were compared by giving the test samples to the trained ANN and then calculating the errors, accuracy, precision, sensitivity, and specificity of diagnosis and classification of dengue disease.

1) Working of proposed model: The network receives a total of 19 inputs (16 input symptoms and 3 laboratory findings) through input layer neurons, and the output of the network was given by output layer neurons. The backpropagation algorithms commonly use supervised learning. It means that we supply the algorithm with input instances and target sets; after that, the inaccuracy (variation between actual target set and expected output) is calculated as follows. The goal of the network was actually to diminish the inaccuracy, till the learning of training data by ANN.

$$ej(n) = dj(n) - yj(n) \tag{1}$$

where *ej* is the error of j (output neuron), once the neuron's activation, *dj* is the desired outcome of neuron *j* while the *yj* is the yield created by the network after every epoch. To train the network here, we used the Levenberg-Marquardt backpropagation algorithm, which used error correction as a learning rule. The training starts with changing weights, and the object is to regulate the weight so that the error will be minimum. The weights of the first layer have come from the input, then optimum bias is added and shifts the calculated output to the 2nd layer and so on till the last layer. The biases are there at layers. The network output is on the last layer. The weights and biases of each layer are initiated with "initnw". In this system the adaptation is done with "adaptwb", which updated weights with the gradient descent weight and bias learning function. The following flow chart in **Figure 1** illustrates the expert system development life cycle that is followed in our research methodology.

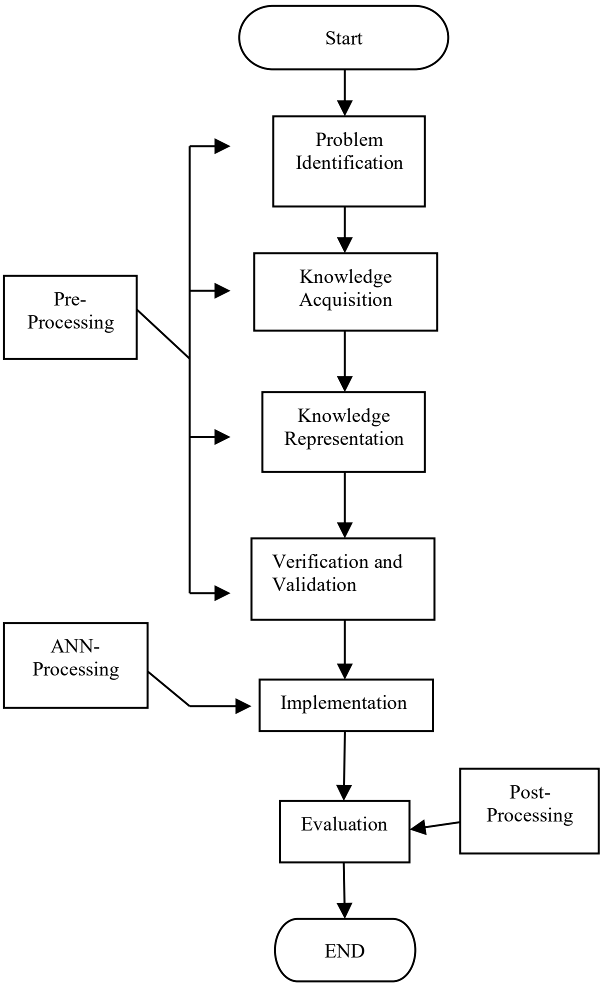


Figure 1. Flow chart of (ESDLC) using in research methodology.

In order to minimize a cost function here in backpropagation learning, it involves adjusting the weights and biases of the network. An error term is always included in the cost function. In order to attain the expected results, computed error is then sent in reverse to the network to renew the weights. The performance of the network is measured by the use of mean square error (MSE) according to the following Equation:

$$Performance = \frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2 \quad (2)$$

Here "n" represent number of total training examples, y_j represents target or expected output while $\hat{y}_j$ represent produced or actual output.

After each training epoch, the generated output is tested by giving the test set to the trained network and checking the accuracy. When the system gave maximum accuracy according to the target sets with minimum MSE, we stopped the training and saved the network.

2) Proposed architecture of artificial neural networks: The subsequent **Figure 2** depicts the final proposed architecture of the neural network, which was used in undertaking research implementation.

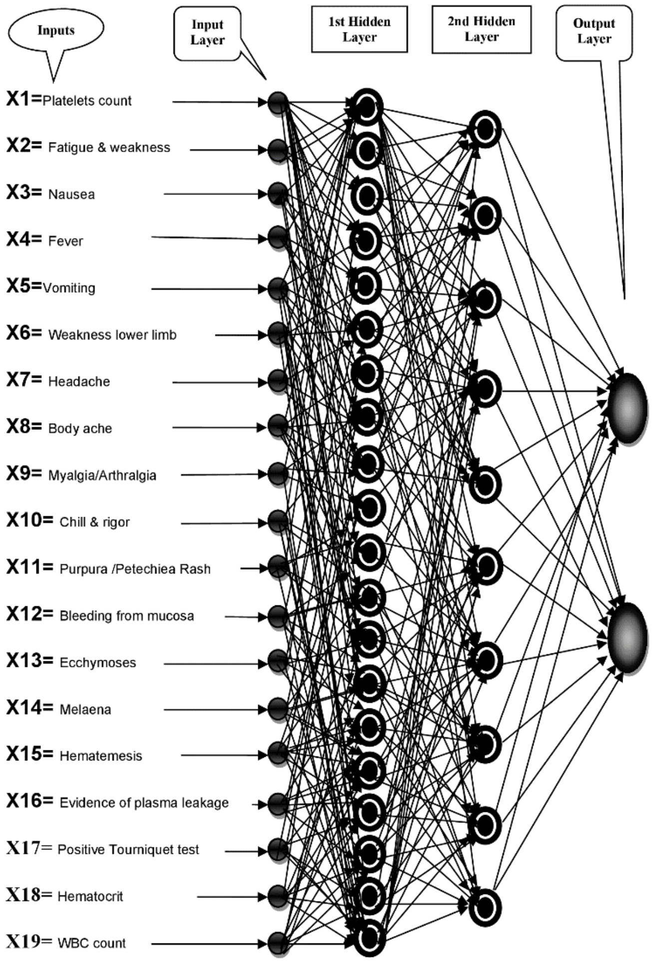


Figure 2. Architecture of ANN for prognosis and classification of dengue.

3) Computation:
   a) Phase 1: Presented the complete set (19 symptoms) of input to the generic network.
   b) Phase 2: The net outputs and inputs of the neurons of the "*j*" hidden layer can be computed as follows:

$$net_j^h = \sum_{t=1}^{N+1} w_{ji} x_i \tag{3}$$

$$y_j = f(net_j^h) \tag{4}$$

c) Phase 3: Compute the net outputs and inputs of the neurons of "k" output layer are:

$$net_k^0 = \sum_{t=1}^{j+1} V_{kj}\, y_j \tag{5}$$

$$Z_k = f = (net_k^0) \tag{6}$$

d) Phase 4: Modify the output layer's weights (for all k, j pairs):

$$V_{kj} \leftarrow V_{kj} + C\lambda(d_k - Z_k)Z_k(1 - Z_k)y_j \tag{7}$$

e) Phase 5: Modify the hidden layer's weights (for all i, j pairs):

$$w_{ji} \leftarrow w_{ji} + C\lambda^2 y_j\big(1 - y_j\big)x_i\Big(\sum_{k=1}^{k}(d_k - Z_k)Z_k(1 - Z_k)V_{kj}\Big) \tag{8}$$

f) Phase 6: Updated the term error

$$E \leftarrow E + \sum_{k=1}^{k}(d_k - Z_k)^2 \tag{9}$$

and reiterate from Phase 2 till the all patterns of input have been presented (This is epoch-1). If the error rate is lower to some predefined level of tolerance, then bring to an end. Else, reset the error $E = 0$, and do again from phase 2 for a new epoch.

4) Algorithm for the system:

Step 1: Allocate all training pattern (input and output).

Step 2: Initialization of all weights with small random numbers, usually (In between −1 and 1)

Step 3: Repeat step 4 to step 14.

Step 4: Present the pattern to the network for every example in training set.

Step 5: Spread the input forward through the network (for each layer in network and for every node in a layer)

Step 6: Calculate the weight sum of the outputs to the node.

Step 7: Add the threshold to the sum.

Step 8: Calculate the activation for the node.

Step 9: Spread the errors backward through the network (for every node in the output layer and Calculate the error signal)

Step 10: Calculate the node's signal error (for all hidden layers and for every node in the layer)

Step 11: Update each node's weight in the network (for all hidden layers and for every node in the layer)

Step 12: Compute Global error.

Step 13: Compute the error function.

Step 14: Do step 4 to step 13 While ((maximum number of iterations < than specified) AND (error function is > than specified))

Step 15: End

5) System implementation: Formerly, the implementation of the system was performed after the system design had been specified. The system was developed using the Neural Network toolbox in MATLAB. The total 19 numeric parameters of each instance were deployed according to the syntax of neural network training. From all of them, 16 parameters represented the signs & symptoms of the disease, while 3 parameters represented the clinical laboratory findings of the patients. The utilized network was created using two hidden layers of feedforward with a backpropagation algorithm. To train the network, it was used back propagation learning. The network was analyzed by using different training and learning functions, also by changing the hidden neurons, number of epochs, momentum, and learning rate. The network's performance was measured using the MSE performance function along with a confusion matrix, and the network was tested by deploying the 208 test sets, which were not part of the training set. The trained ANN then diagnosed and classified the dengue disease. The flow of activities for prognosis and classification of dengue is shown in the following **Figure 3**.

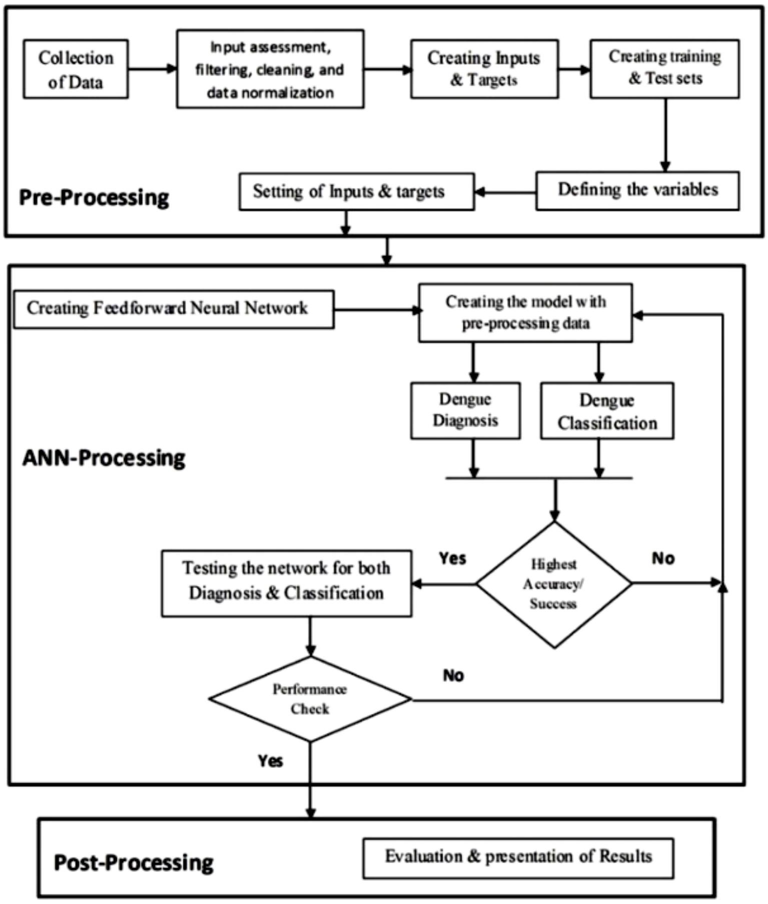


Figure 3. Flow diagram for prognosis and classification of dengue.

4. Design and implementation

The datasets contained one year of real data of 768 patients, which were collected from Banzir Bhutto Hospital Rawalpindi, CITI Lab Rawalpindi, and Meo Hospital Lahore in the year of 2011. The whole sample was further separated into 560 training sets and 208 testing samples. From a total of 560 patients in the training set, 342 patients were serologically non-reactive while 218 patients were confirmed reactive. From all 218 dengue-confirmed patients, 143 suffered from dengue fever while 75 patients suffered from dengue hemorrhagic fever. The 560 training and 208 testing examples, along with their proper target sets, were successfully generated from the knowledge acquisition.

4.1. Input parameters

We found a total of 768 samples after filtering the original data, which were further separated into testing and training sets. The parameter (attribute) of each sample (instance) shows either symptoms or laboratory factors. According to the standard, the data was to be error-free in nature. In order to finalize the input parameters, help was taken from medical experts, medical staff, and published work in dengue cases [1,13,14,36–44]. After careful scrutiny, the final parameters selected to train the neural net are shown in the following **Table 2**:

Table 2. Final selected attributes and their values.

Input Parameters		
	(Clinical Symptoms)	
S. No	Attributes	Values
1	Fatigue & weakness	No, Yes
2	Nausea	No, Yes
3	Fever	No, Yes
4	Vomiting	No, Yes
5	Weakness lower limb	No, Yes
6	Headache	No, Yes
7	Body ache	No, Yes
8	Myalgia/Arthralgia (Muscle & joint pain)	No, Yes
9	Chill and rigor	No, Yes
10	Purpura/Petechiea Rash	No, Yes
11	Bleeding from mucosa	No, Yes
12	Ecchymoses	No, Yes
13	Melaena (Gastrointestinal bleed)	No, Yes
14	Hematemesis (Vomiting blood)	No, Yes
15	Evidence of plasma leakage	No, Yes
16	Positive Tourniquet Test	No (Negative), Yes (Positive)
(Clinical Laboratory attributes)		
17	Platelets count/Thrombocytopenia	10,000–450,000
18	WBC Count/Leucopenia	2500–11,000
19	hematocrit/HCT %age	36–69

4.2. Target set/output classes

The targets were logically represented in the following classes, i.e., No Dengue (00), Dengue Fever (01), and Dengue Hemorrhagic Fever (10).

4.3. Data distribution

The whole dataset of 768 patients was divided into two main sets; set-1 was utilized for the training of the neural network, whereas set-2 was utilized to test the correctness of diagnosis and classification of the trained network. The training set consisted of 560 instances, while the test set consisted of 208 instances. From all of the 560 training samples, 342 samples belonged to group ND, 143 samples belonged to group DF, and 75 samples belonged to group DHF. On the other hand, from a total of 208 testing samples, 111 samples belonged to group ND, 72 samples belonged to category DF, and 25 samples belonged to DHF. The test samples were not included in the training part of NNs; they were only utilized to test the network's performance. Following **Table 3** demonstrates the division of test and training samples.

Table 3. Samples used to test and train the ANNs.

Dataset	**Total Samples**	**Categories**		
		ND	**DF**	**DHF**
Training Set	560	342	143	75
Test set	208	111	72	25

5. Comparison of simulation results

Achieving high accuracy in dengue diagnosis and classification with the minimum possible error was the actual goal of the research. Different architectures of feedforward backpropagation neural networks (**Table 4**) have been utilized for the evaluation of the network's performance. We have used a bottom-up approach for the selection of attributes of hidden layers, hidden neurons, neurons on the output layer, and values of momentum and learning rate for these architectures. Basically, the momentum term is used to "persuade" the changes of weight to carry on in a similar way with larger steps so that the learning process is accelerated and also to prevent the learning process from settling in local minima. After that we have step-by-step increased the number of hidden neurons, hidden layers, momentum term, and learning rate using different learning and training algorithms. Also, by changing the activation functions, we have tested the network for classification accuracy.

The number of hidden layers and hidden neurons were selected using different techniques recommended by previous studies [45]. For example:

- The size of neurons on hidden layer should lie in between the number of the input nodes and output nodes.
- Hidden neurons = $\frac{2}{3}$ × input layer size + size of output layer.
- Number of hidden neurons < 2 × size of the input layer.
- Hidden layer neurons = sqrt (n × m) where n represent input neurons while m represents output neurons.
- Hidden layer neurons = (2n + 1) here n represents number of inputs.

Table 4. Analysis of some best of all selected architectures with different types of activation functions.

Arch.	**Activation Functions**		**T. A**	**Correctly classified**		**Incorrectly classified**	
	H. L (s)	**O. L**		**Training**	**Testing**	**Training**	**Testing**
A	Tansig, logsig	Tansig	trainlm	560	208	0	0
B	Tansig, logsig	Tansig	trainlm	560	208	0	0
C	Logsig, Purelin	Logsig	trainlm	485	183	75	25
D	Tansig, logsig	Tansig	trainlm	560	208	0	0
E	Tansig, logsig	Tansig	trainlm	560	208	0	0
F	Tansig, logsig	Tansig	trainlm	560	208	0	0
G	Tansig, logsig	Tansig	trainlm	560	204	0	4
H	Tansig, logsig	Tansig	trainlm	509	190	51	18
I	Tansig, logsig	Tansig	trainlm	560	208	0	0
J	Tansig, logsig	Tansig	trainlm	560	208	0	0
K	Tansig, logsig	Tansig	trainlm	485	178	75	30
L	Logsig	Tansig	trainlm	560	208	0	0
M	Tansig	Purelin	trainlm	486	78	74	83
N	Tansig	Logsig	trainlm	560	208	0	0
O	Logsig	Tansig	trainlm	560	207	0	1
P	Tansig	Logsig	trainlm	560	208	0	0
Q	Tansig	Logsig	trainlm	560	208	0	0
R	Tansig	Logsig	trainlm	560	208	0	0
S	Logsig	Tansig	trainlm	553	196	07	12
T	Logsig	Tansig	trainlm	560	208	0	0
U	Logsig	Tansig	trainlm	560	208	0	0
V	Tansig	Logsig	trainlm	560	208	0	0

Hints: Arch. = architecture, H. L (s) = Hidden layer/Hidden layers, O.L = Output Layer, T. A = Training Algorithm.

The above **Table 4** shows the brief analysis of correctly classified and misclassified training and testing instances of different architectures with different types of activation functions using the Levenberg-Marquardt algorithm.

From the analysis in above table, it can be seen that the use of tangent sigmoid and log sigmoid transfer functions on first and second hidden layers provides the best architectures for classification accuracy in dengue disease. We have utilized the Levenberg-Marquardt algorithm to train the network because it is considered the best training algorithm for classification problems.

The results of some of the best architectures are presented below in **Table 5**.

Table 5. Results of the ANNs with different architectures system used in proposed system.

FFNN. Arch.	**Epochs**	**Training Time**	**MSE**	**Results**
A	72	00:00:15	0.000000000010444	100%
B	17	00:00:02	0.00000000000032521	100%
C	12	00:00:01	0.00000000012197	87.29395604%

Table 5. (*Continued*).

FFNN. Arch.	Epochs	Training Time	MSE	Results
D	17	00:00:02	0.00000000000063824	100%
E	22	00:00:02	0.0000000000012278	100%
F	228	00:00:17	0.0000000000030713	100%
G	31	00:00:02	0.00000000000433	99.03846154%
H	16	00:00:01	0.00000000054149	91.11950549%
I	15	00:00:00	0.0000000000054149	100%
J	23	00:00:01	0.0000000000056316	100%
K	3000	00:00:41	0.054914	86.09203297%
L	21	00:00:11	0.000000000044486	100%
M	13	00:00:15	0.011455	62.14285714%
N	20	00:00:05	0.000000000014054	100%
O	19	00:00:02	0.000000000024487	99.75961538%
P	13	00:00:01	0.000000000028912	100%
Q	13	00:00:01	0.0000000000060134	100%
R	14	00:00:01	0.0000000000033482	100%
S	3000	00:00:07	0.00385	96.49038462%
T	195	00:00:06	0.000000000015849	100%
U	15	00:00:00	0.000000000022949	100%
V	30	00:00:01	0.000000000042466	100%

FFNN. Arch. = Feedforward Neural Network architecture.

It can be seen clearly that architecture (B) has the highest efficiency because it has the lowest MSE, minimum running time, and highest accuracy. In conclusion, we have reported that the feedforward backpropagation network with two hidden layers, i.e., 20 neurons on the first hidden layer using the tangent sigmoid activation function and 10 neurons on the 2nd hidden layer using the log sigmoid activation function, trained with the Levenberg-Marquardt algorithm, created the $\approx$ 100% results in dengue diagnoses and classification with the lowest MSE of 0.00000000000032521, Precision (PPV) 99.8%, NPV 100%, Sensitivity 100%, and Specificity 98.7%. The values 0.9 and 0.5 were reported as the best values for the learning rate and momentum term for the underlying architecture analyzed by the hit-and-try method.

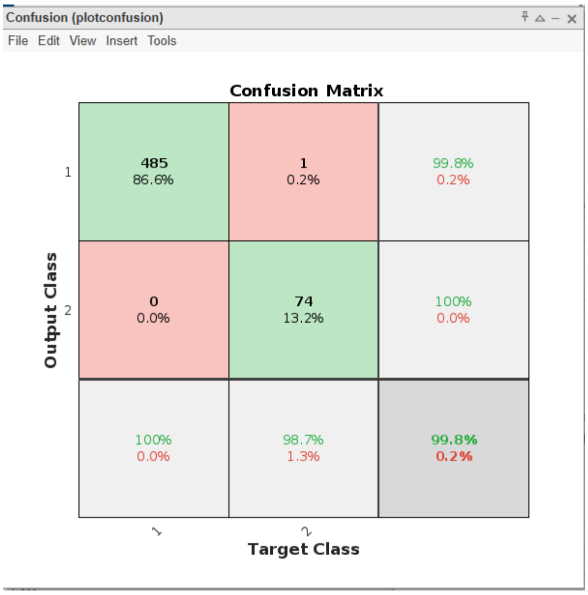
It can also be noted that the two-layer architecture of the feedforward backpropagation network produced good results as compared to the single-layer network in terms of efficiency and accuracy.

The algorithm's performance was evaluated using a standard confusion matrix. The discrepancies between the predicted and actual binary outcomes were quantified by calculating the counts of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). Here, the sum of TP + TN + FP + FN = n represents the total number of observations. Based on these values, standard performance metrics were applied to assess the model's accuracy.

- Accuracy/Correct Classification Rate (A): This metric calculates the fraction of observations accurately classified across all categories, given by A = (TP + TN)/n.
- Positive Predictive Value (PPV) or Precision: Precision measures the accuracy in identifying positive cases among all predicted positives, expressed as PPV = TP/(TP + FP).
- Negative Predictive Value (NPV): This metric indicates the proportion of true negatives among all cases predicted as negative, calculated by NPV = TN/(TN + FN).
- Sensitivity (Sn), also known as Recall or True Positive Rate**:** Sensitivity reflects the rate at which actual positive cases are correctly identified, and it is calculated as Sn = TP/(TP + FN).
- Specificity (Sp), or True Negative Rate: Specificity denotes the rate at which actual negative cases are correctly identified, given by Sp = TN/(TN + FP).

Figure 4 below in the confusion matrix, green cells indicate True Positive (TP) and True Negative (TN) values, while red cells mark False Positive (FP) and False Negative (FN) values. The four gray cells (from the top right, moving clockwise) correspond to Positive Predictive Value (PPV) or Precision, Negative Predictive Value (NPV), Sensitivity, and Specificity. The dark gray cell in the bottom right corner represents the overall accuracy or correct classification rate (A). For color interpretation in this figure, please refer to the online version of this article.

The following plots in **Figures 5–8** expressed a bird's-eye view of the training state, regression, performance, and architecture of the proposed system, respectively.



	1	2	
1	485 86.6%	1 0.2%	99.8% 0.2%
2	0 0.0%	74 13.2%	100% 0.0%
	100% 0.0%	98.7% 1.3%	**99.8%** **0.2%**

Figure 4. Confusion matrix.

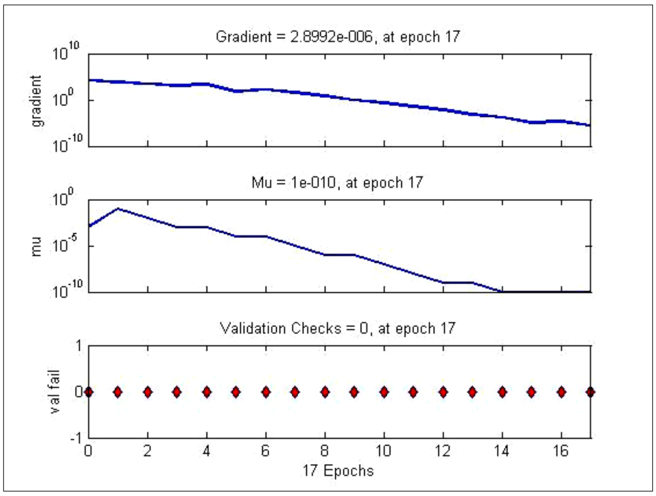


Figure 5. Training state.

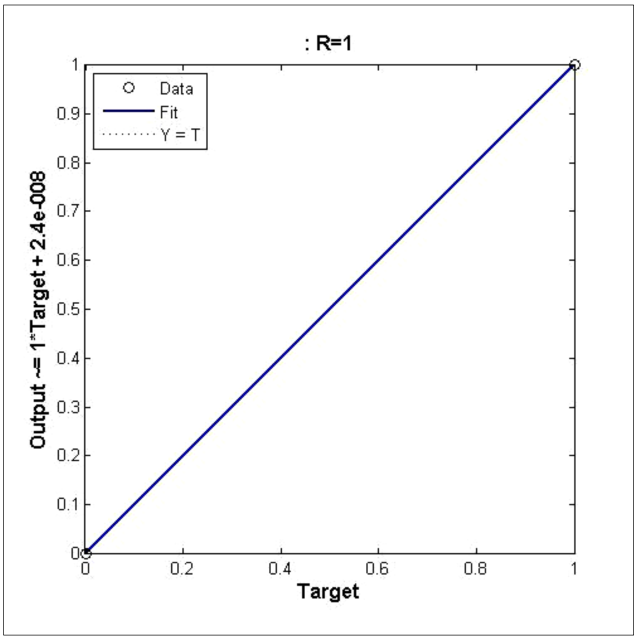


Figure 6. Regression plot.

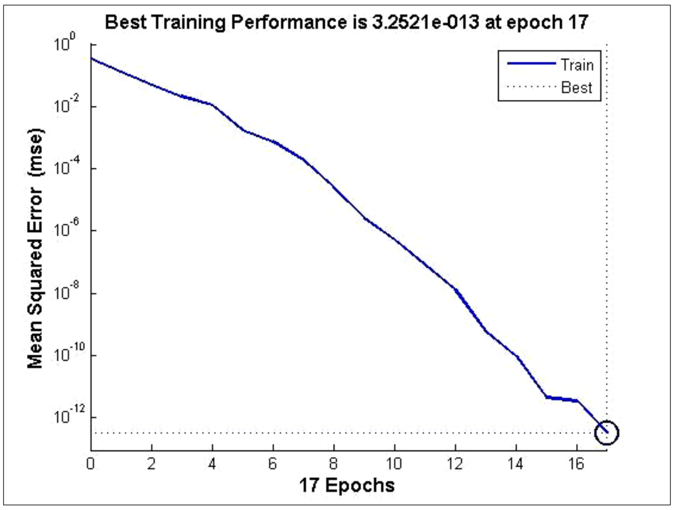


Figure 7. Performance plot of architecture "B".

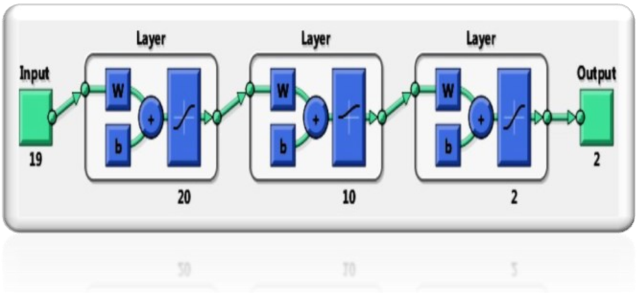


Figure 8. FFNN Architecture for proposed system.

5.1. Performance evaluation

The system was developed using the neural network toolbox in MATLAB R2011b to evaluate the performance of the proposed networks. The utilized network was created using two hidden layers with a feedforward and backpropagation algorithm with 19 inputs and 20 sigmoid hidden neurons on the first layer and 10 pure linear hidden neurons on the second layer. To train the network, it was used with Levenberg-Marquardt backpropagation learning. The network was analyzed by using different training and learning functions, also by changing the hidden neurons, number of epochs, momentum, and learning rate. Training automatically stops when generalization stops improving, as indicated by an increase in the mean square error (MSE) of the validation samples or either minimum gradient or maximum momentum reached. The network's performance was measured using the MSE performance function, and the network was tested by deploying the 208 test sets, which were not part of the training set. The trained ANN then diagnosed and classified the dengue disease with 100% accuracy.

5.2. K-fold cross validation

In our study, the K-fold cross-validation technique was utilized for the course of validation. This practice has been used to minimize the bias connections with the random sampling of the training and testing data in order that the robustness of the whole diagnostic model's performance is preserved [31,32].

The K-fold cross-validation process is employed by randomly distributing the data set into K sets of equal size. On each measurement, one of the K sets is used as a testing set while the other K-1 sets are coupled collectively to form a training set. The overall training procedure is repeated K times. For estimating and analyzing the model's performance, the average of mean squared error (MSE) and average accuracy from all K experiments are evaluated.

We used a cross-validation-based experimental design in finding the "optimal" number of hidden neurons. Specifically, we started from a small number of neurons, measured the performance of the trained neural net on test samples, steadily increased the number of neurons in the first and second hidden layers, and at a point where the performance of the trained neural net on test samples started to decline due to overtraining, we took that as the sign of the "optimal" number of neurons for the hidden layers of the neural network. With the intention to decrease the bias connected with the random splitting of the dataset into training and testing, here we used 10-fold cross-validation. The table above shows the analysis of 10-fold cross-validation performance.

From the analysis in **Table 6**, it is stated that the performance of the network in dengue diagnosis is ≈ 100% in terms of accuracy with a minimum MSE of 0.0000000000032521, precision (PPV) of 99.8%, NPV of 100%, sensitivity of 100%, and specificity of 98.7% when we distribute our data randomly in two parts, i.e., 75% of the data was used as a training set while the network was tested on the rest of the 25% of the data. For the sake of validation, we applied the 10-fold cross-validation technique and measured the performance of each fold. After the analysis of all folds, we found an average accuracy of 98.2984% with an average mean square error (MSE) of 0.00486586. The average rate of misclassified examples was reported as 1.3.

Table 6. Shows the analysis of 10 fold cross validation performance.

Folds #	Mean Squared Error (MSE)	Total testing cases	Correct classified	Misclassified	accuracy % across all *K* trials
1	0.00501	78	76	2	97.43589744
2	0.0050074	76	74	2	97.36842105
3	0.005749	76	75	1	98.68421053
4	0.0050215	76	74	2	97.36842105
5	0.0073436	78	77	1	98.71794872
6	0.0049575	70	69	1	98.57142857
7	0.005	76	75	1	98.68421053
8	0.00522	78	76	2	97.43589744
9	0.00533	78	77	1	98.71794872
10	0.0000196	87	87	0	100
Avg. MSE				**Avg. Misclassified**	**Avg. accuracy across all K trials**
0.00486586				1.3	98.2984384

Table 7 shows the overall results comparison of the proposed system with previous studies in dengue prognosis and classification. It can be easily observed that using feedforward with backpropagation techniques, the results are much better than other techniques previously used for the diagnosis and classification of dengue disease. We can improve the accuracy of dengue diagnosis by selecting appropriate parameters, hidden layers, and hidden neurons. It can also be analyzed that the selection of training and activation functions is also a critical task to generate results with the highest accuracy.

In this study, we present an indicative rather than definitive comparison of methods, acknowledging that differences in datasets across studies may influence the results. Due to variations in data characteristics, direct comparisons are inherently limited. To provide a more rigorous and conclusive assessment, further evaluation on standardized, publicly available datasets would be beneficial. Such testing would allow for a more accurate benchmark of the methods' performances, mitigating the variability introduced by dataset-specific factors and offering a more robust foundation for comparative analysis.

The core intent of this paper is a step in the direction of prognosis and dengue infection's classification using a feedforward backpropagation neural network with high accuracy. The proposed technique has concluded that FFNN using the Levenberg-Marquardt training algorithm is the preeminent classifier practice by way of the fact that it has achieved ≈ 100% accuracy, gets less time to execute, and has the smallest error rate.

Table 7. Comparison of results with previous studies.

Authors	**Methodology**	**Result Accuracy (%)**	**Year**
Ibrahim et al. [15]	Rule base expert system	66.7%	2006
Napa Rachata et al. [12]	Entropy technique and artificial neural network	85.92%	2008
Abdul Rahim et al. [19]	Statistical approach (multiple linear regressions)	52.1%	2009
Tarig Faisal et al. [13]	MLP (Levenberg Marquardt)	75%	2010
Bakar et al. [17]	Rough Set (RS)	93.87%	2011
Kumar [3]	Alternating decision tree with boosting	89%	2013
Farooqi et al. [21]	Decision tree	99.44%	2014
Shaukat et al. [6]	Nave byes classifier	92 %	2015
Shakil et al. [18]	WEKA tool (Naïve Bayes)	100% with 99 correct classified instances	2015
Balasaravanan et al. [34]	Used Data mining technique ANN, Weka and Netbeans IDE	95% of accuracy	2018
Davi et al. [35]	Support Vector Machine (SVM) algorithm & ANN	Models produced median values of accuracy greater than 86%, and sensitivity and specificity over 98% and 51%, respectively.	2019
Salim et al. [30]	SVM Linear model,	Accuracy = 70%, Sensitivity = 14%, Specificity = 95%, Precision = 56%	2021
Mayrose et al. [23]	A Machine Learning based technique Narrow Multilayer Perceptron (MLP)	The best performance was attained using an SVM classifier with 95.74% accuracy.	2023
Bohm et al. [28]	logistic regression model with 10 variables, the Decision Tree and Multilayer Perceptron (MLP)	A decision tree model, achieved the best classification results (90.64%)	2024
Proposed Method	FFNN with back propagation	≈ 100% in term of accuracy with minimum MSE of 0.00000000000032521, Precision (PPV) 99.8%, NPV 100%, Sensitivity 100% and Specificity 98.7% with average misclassified of 1.3	2024

6. Conclusions and future work

This research presented a framework for prognosis & classification of dengue virus using artificial neural networks. The utilized network was a two-hidden-layer feedforward neural network, which belongs to the category of (MLP) classifiers. The backpropagation learning is used for its training. The proposed underlying system was capable of attaining a high level of success. 100% achievement rate was attained along with 99.8% precision (PPV), 100% NPV, 100% sensitivity, and 98.7% specificity. The underlying system is capable of classifying the patients into three groups: no dengue, dengue fever, and dengue hemorrhagic fever groups. This infers that the artificial neural network approach is an efficient and proficient means for employing diagnostic issues. The dengue expert system is built for prognosis & classification of the dengue virus to help users to diagnose the mosquito-borne disease that occurs to the patient.

In the post-processing phase, we have successfully presented the result gathered by the Hit-and-Trail method with the accuracy rate of ≈ 100%, which has never been presented before yet in any dengue diagnosis and classification system using ANN. The purpose of this research is to provide dialectical reasoning and a synthesis of the published research in this area and stimulate further research interests and attempts in the identified topics. The research also provided a helping framework to the physicians for enhancing their experience and helps them to increase the surety of their diagnosis and classification. In short, the better results of the diagnosis of dengue rely upon an excellent model of architecture, so appropriate data, normalization, cleaning, selection of the most probable parameters, hidden layers, learning algorithms, hidden neurons, and also the critical evaluation must be considered.

6.1. Recommendations

The undertaking research effort is recommended to human experts and pathologists who specialize in diagnostics, analyzing, and treatment of dengue virus and associated diseases. Human experts will find it helpful as an assist in the decision-making process and verification of suspected cases. Medical practitioners can also rely on the system for support, which operates in those areas where no specialists (pathologists) are available. This study can also proceed as a plinth for the progress of research and investigation in applications of neural networks in the area of medical diagnostic research. The following project is also recommended for developers of decision-support systems and also for research scholars to explore the new ideas in the related field.

6.2. Future work

The system is used only for prognosis and classification of dengue based on symptomatic and CBC-containing laboratory parameters, so it would be better to include the results of serological methods to improve in confirming dengue diagnosis. Besides this, the system only focuses on a specific mosquito-borne disease, i.e., dengue diagnosis, so in the future we will work on a compact expert system that will be able to cover the diagnosis of other mosquito-borne diseases (malaria, flavivirus, chikungunya, yellow fever, etc.) along with dengue. Also in the future we will work on further classification of dengue hemorrhagic fever into its subclasses like Grade I, Grade II, Grade III, Grade IV, and dengue shock syndrome (DSS).

Author contributions: Conceptualization, SH; methodology, SH; software, SH; validation, SH, RA and SA; formal analysis, SH; investigation, SH; resources, SH, RA and SA; data curation, SH and RA; writing—original draft preparation, SH; writing—review and editing, SH; visualization, SH; supervision, SH and RA; project administration, SH. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: First of all, we would like to express our gratitude to The ALLAH Almighty is the most kind, merciful and compassionate. His benevolence and blessings enabled us to accomplish this task. This research was made possible by the support of many dedicated people. We would like to thank those who shared their ideas and materials. This research would not have been possible without the unrelenting support and encouragement of respected Saba Bashir, Aihab Khan, Rahila Anwar and Fahad khan, special thanks our all-family members for their insistent support.

Data availability: The data that support the findings of this study are available from Ministry of National Health Services, Regulations and Coordination Govt. of Pakistan, but restrictions apply to the availability of these data, which were used under licence for the current study and so are not publicly available. The data are, however, available from the authors upon reasonable request and with the permission of Ministry of National Health Services, Regulations and Coordination Govt. of Pakistan.

Code availability: The code for all analyses reported in the manuscript is available on request.

Conflict of interest: The authors declare no conflict of interest.

References

1. World Health Organization (WHO). Update on the Principles and Use of Rapid Tests in Dengue. WHO; 2009.
2. Brause RW. Medical Analysis and Diagnosis by Neural Networks. Lecture Notes in Computer Science. 2001; 1(13).
3. Kumar MN. Alternating Decision Trees for Early Diagnosis of Dengue Fever. arXiv. 2013; 1305.
4. Khan IY, Zope PH, Suralkar RR. Importance of Artificial Neural Network in Medical Diagnosis of Diseases like Acute Nephritis Disease and Heart Disease. Int. J. Eng. Sci. Innov. Technol. 2013; 2(2).
5. De Dombal FT, Leaper DJ, Staniland JR, et al. Acute Abdominal Pain. Brit. Med. J. 1972; 2: 9–13.
6. Shaukat K, Masood N, Mehreen S, et al. Dengue Fever Prediction: A Data Mining Problem. J. Data Mining Genomics Proteomics. 2015; 6: 181. doi: 10.4172/2153-0602.1000181
7. Dehariya A, Chaudhary VK, Khan I, et al. An Effective Approach for Medical Diagnosis Preceded by Artificial Neural Network Ensemble. In: Proceedings of the 3rd International Conference on Electronics Computer Technology; 8-10 April 2011; Kanyakumari, India. pp. 1–5.
8. Heden B, Edenbrandt L, Haisty Jr WK, et al. Artificial Neural Networks for the Electrocardiographic Diagnosis of Healed Myocardial Infarction. Am. J. Cardiol. 1994; 74(1): 5–8.
9. Snow PB, Smith DS. Artificial Neural Networks in the Diagnosis and Prognosis of Prostate Cancer: A Pilot Study. J. Urology. 1994; 152(5): 1923–1926.
10. Ansari S, Shafi I, Ahmad J, et al. Determination of Hepatotropic Virus in Human Metabolism Using Artificial Neural Networks. In: Proceedings of the International Conference on Emerging Technologies; 15 November 2010; Islamabad, Pakistan.
11. Cetiner BG, Murat S, Aburas HM. Recognition of Dengue Disease Patterns Using Artificial Neural Networks. In: Proceedings of the 5th International Advanced Technologies Symposium (IATS); 13–15 May 2009; Karabuk, Turkey.
12. Rachata N, Charoenkwan P, Yooyativong T, et al. Automatic Prediction System of Dengue Haemorrhagic-Fever Outbreak Risk by Using Entropy and Artificial Neural Network. In: Proceedings of the International Symposium on Communications and Information Technologies (ISCIT); 21–23 October 2008; Vientiane, Laos.
13. Faisal T, Ibrahim F, Taib M. Neural Network Diagnostic System for Dengue Patients Risk Classification. J. Med. Syst. 2010. doi: 10.1007/s10916-010-9532-x
14. Karim S, Suryaningsih H. Expert System for Diagnosing Dengue Fever. In: Proceedings of the Seminar Nasional Aplikasi Teknologi Informasi; 16 Jun 2007; Yogyakarta, Indonesia.
15. Ibrahim F, Osman NAB, Usman J, et al. Classification of Risk in Dengue Fever and Dengue Haemorrhagic Fever Using Rule-Based Expert System. In: Proceedings of the 3rd Kuala Lumpur International Conference on Biomedical Engineering; 11–14 December 2006; Kuala Lumpur, Malaysia. pp. 50–53.
16. Husin NA, Salim N, Ahmad AR, et al. Modeling of Dengue Outbreak Prediction in Malaysia: A Comparison of Neural Network and Nonlinear Regression Model. In: Proceedings of the International Symposium on Information Technology; 26–28 August 2008; Santa Fe, NM, USA. pp. 1–6.

17. Bakar AA, Kefil Z, Abdullah S, et al. Predictive Models for Dengue Outbreak Using Multiple Rule-Based Classifiers. in Proc. In: Proceedings of the International Conference on Electrical Engineering and Informatics (ICEEI); 17–19 July 2011; Bandung, Indonesia.
18. Shakil KA, Anis S, Khan S, et al. Dengue Disease Prediction Using Weka Data Mining Tool. In: Proceedings of the International Conference (ICCI-SEM-2K17); 18 February 2015; New Delhi, India.
19. Rahim HA, Arshad S, Taib MN. (2009). Early diagnostic the grading stages of dengue haemorrhagic fever. 2009 5th International Colloquium on Signal Processing & Its Applications, 375-377. Early Diagnostic the Grading Stages of Dengue Haemorrhagic Fever. In: Proceedings of the 5th International Colloquium on Signal Processing and Its Applications (CSPA); 6 March 2009; Kuala Lumpur, Malaysia.
20. Ramírez RB, Arellano RM, Sandoval GC, et al. An Expert System Oriented Towards the Detection of Influenza and Dengue Developed on Mobile Platforms. J. Softw. Eng. Appl. 2025; 8: 295–301.
21. Farooqi W, Ali S, Wahaba A. Classification of Dengue Fever Using Decision Tree. Comput. Sci. 2014; 3(2).
22. Baker QB, Faraj D, Algazo A. Forecasting Dengue Fever Using Machine Learning Regression Techniques. In Proceedings of the 12th International Conference on Information and Communication Systems (ICICS); 24–26 May 2021; Valencia, Spain.
23. Mayrose H, Bairy GM, Sampathila N, et al. Machine Learning-Based Detection of Dengue from Blood Smear Images Utilizing Platelet and Lymphocyte Characteristics. Diagnostics. 2023; 13(220).
24. Gupta G, Khan S, Guleria V, et al. DDPM: A Dengue Disease Prediction and Diagnosis Model Using Sentiment Analysis and Machine Learning Algorithms. Diagnostics. 2023; 13(1093).
25. Ming DK, Tuan NM, Hernandez B, et al. The Diagnosis of Dengue in Patients Presenting With Acute Febrile Illness Using Supervised Machine Learning and Impact of Seasonality. Front. Digit. Health. 2022; 4. doi: 10.3389/fdgth.2022.849641.
26. Ho TS, Weng TC, Wang JD, et al. Comparing Machine Learning with Case-Control Models to Identify Confirmed Dengue Cases. PLoS Negl. Trop. Dis. 2020; 14(11).
27. Thirugnanam A, Hussain FBJ. Exploring Machine Learning Algorithms for the Prediction of Dengue: A Comprehensive Review. RIA J. 2023; 37(5).
28. Bohm BC, de Melo Borges FE, Silva SCM, et al. Utilization of Machine Learning for Dengue Case Screening. BMC Public Health. 2024; 24.
29. Huang SW, Tsai HP, Huang SJ, et al. Assessing the Risk of Dengue Severity Using Demographic Information and Laboratory Test Results with Machine Learning. PLoS Negl. Trop. Dis. 2020; 14(12): e0008960.
30. Salim NAM, Wah YB, Reeves C, et al. Prediction of Dengue Outbreak in Selangor Malaysia Using Machine Learning Techniques. Sci. Rep. 2021; 11(939).
31. Yang J, Du B, Wu C. Hybrid Vision Transformer Model for Hyperspectral Image Classification. In Proceedings of the EEE International Symposium on Geoscience and Remote Sensing (IGARSS); 7–12 July 2024; Athens, Greece.
32. Saha S, Saha S. Combined Committee Machine for Classifying Dengue Fever. In: Proceedings of the International Conference on Microelectronics, Computing and Communications (MicroCom); 23–25 January 2016; Durgapur, India.
33. Chatterjee S, Hore S, Dey D, Chakraborty S. Dengue Fever Classification Using Gene Expression Data: A PSO-Based Artificial Neural Network Approach. Springer International Publishing; 2017.
34. Balasaravanan K, Prakash, M. Detection of Dengue Disease Using Artificial Neural Network Based Classification Technique. Int. J. Eng. Technol. 2018; 7(3): 13–15.
35. Davi C, Pastor A, Oliveira T, et al. Severe Dengue Prognosis Using Human Genome Data and Machine Learning. IEEE Trans. Biomed. Eng. 2019; doi: 10.1109/TBME.2019.2897285.
36. Mallhi TM, Khan AH, Adnan AS, et al. Clinico-laboratory spectrum of dengue viral infection and risk factors associated with dengue hemorrhagic fever: A retrospective study. BMC Infect. Dis. 2015; 15.
37. Buchy P. Laboratory Tests for the Diagnosis of Dengue Virus Infection. Working paper for the Scientific Working Group on Dengue Research, Special Programme for Research and Training in Tropical Diseases. 2006.
38. World Health Organization. Comprehensive Guidelines for Prevention and Control of Dengue and Dengue Haemorrhagic Fever. WHO; 2011.
39. World Health Organization (WHO). Dengue Guidelines for Diagnosis, Treatment, Prevention and Control. WHO; 2009.
40. Dengue guideline 082012, grade DHF, dengue. Available online: https://www.slideshare.net/slideshow/dengue-guideli (accessed on 2 June 2024).

41. World Health Organization Pakistan. National Guidelines for Clinical Case Management of Dengue Fever/Dengue Hemorrhagic Fever/Dengue Shock Syndrome. World Health Organization Pakistan; 2011.
42. Ibrahim F, Thio THG, Faisal T, et al. A novel dengue fever (DF) and dengue haemorrhagic fever (DHF) analysis using artificial neural network (ANN). Comput. Methods Programs Biomed. 2005; 79: 273–281.
43. Faisal T, Ibrahim F, Taib MN. Analysis of significant factors for dengue infection prognosis using the self-organizing map. In: Proceedings of the 30th Annu. 30th annual international conference of the IEEE Engineering in Medicine and Biology Society; 20–25 August 2008; Vancouver, Canada.
44. Ibrahim F, Taib MN, Abas WABW, et al. Dengue Fever (DF) and Dengue Haemorrhagic Fever (DHF) Symptoms Analysis from an Expert System Perspective. Computer Methods and Programs in Biomedicine. 2005; 79: 273.
45. Panchal FS, Panchal M. Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network. IJCSMC. 2014; 3(11): 455–464.