DS Journal of Digital Science and Technology
Volume 1 Issue 2, 1-10, Oct - Dec 2022
ISSN: 2583-5416/ https://doi.org/10.59232/DST-V1I2P101

Original Article

An Optimal Prediction of Dengue Fever Based on Pso-Optimized Fuzzy-ELM

R.Surendiran[1],

[1] *School of Information Science, Annai College of Arts and Science, Kumbakonam, India.*
ORCID :0000-0003-1596-7874

surendiranmca@gmail.com

Received: 29 October 2022; Revised: 12 November 2022; Accepted: 20 November 2022; Published: 06 December 2022;

Abstract - Dengue is an illness caused by dengue viruses (DENVs) carried by Aedes mosquitoes. Dengue hemorrhagic fever can develop into life-threatening dengue shock syndrome. To provide timely supportive care and therapy, it is necessary to have indispensable practical instruments that accurately differentiate dengue and its types in the early stages of illness advancement. Due to a scarcity of vaccines and medications, early detection of a dengue outbreak is critical to lowering the number of deaths. Several computer vision-based studies have been conducted in recent years to recognize DENV and the stages of fever. However, the existing standard machine learning techniques face challenges of inaccurate prediction, while backpropagation neural yields sufficient results, yet it requires very large training data and has slow consolidation. To overcome the existing challenges, this paper proposes a novel machine learning (ML) model based on PSO-FELM (PSO optimized Fuzzy-ELM) to predict dengue fever. First, the collected data is pre-processed to transform raw data into a beneficial and efficient format (cleaned, standardized, and noise-free). Then Fuzzy-ELM based machine learning framework is presented for optimally predicting dengue fever. Herein fuzzy logic is applied to overcome the imbalance and weighted classification problems. PSO is employed to obtain the optimal parameters, which assist to improves the diagnostic accuracy of FELM. The goal of this study is to help patients diagnose dengue illness on their own. It allows patients to consult with an expert and reduces the medical examiner's workload in advance. When compared to existing methodologies such as ELM, FELM, PSO-ELM, and PSO, the research shows that the suggested model can attain 95 percent accuracy.

Keywords - Dengue, Viruses, Aedes mosquitoes, PSO-FELM, Fuzzy-ELM.

1. Introduction

Dengue is a life-threatening disease caused by a particular variety of mosquitoes known as Aedes mosquitoes in the world's tropical regions. Heavy downpour results in stagnant water, which is suitable for this type of mosquito breed [18]. So, under the perfect climatic conditions, the female mosquitoes known as Aedes aegypti carry the dengue virus (DENV) and inject it into human blood when it comes in contact. Normally the disease is mild and can be cured on its own within a week. However, in some cases, the disease is severe and can be fatal. This can be avoided if the disease is diagnosed at an early stage and treated properly with medications. There is no distinct treatment for dengue, but early detection will help to avoid risk factors and reduce the mortality rate. Detecting the presence of DENV also helps to control the spread of the disease; thus, an epidemic can be avoided, saving many human lives.

The spread of dengue infection in the world has increased fiercely in recent years [1]. According to WHO, dengue fever is recognized as one of the four main contaminants frightening global well-being (World Health Organization 2019b). There were more than 3.1 million cases in the American continent [18]; among that, 25000 were reported as severe. In 2021 there were 1,64,103 reported dengue victims in India, against 2,05,243 occurrences in 2019.

This can be best achieved by detecting the presence of the dengue virus through Machine Learning (ML)[22]. Machine Learning facilitates the processing of a large number of associated variables and allows for the modeling of multiple variables' interactions. [11]. Machine Learning techniques such as SVM and RFC [23] [24] were used previously to detect the disease, but the accuracy is low. These models were applicable for a particular region as the variables and time series are low. Artificial Neural Network (ANN) is adapted in [4] [25] and classifies infected persons by preparing image datasets, and an improved version is proposed in [9][13][15]. Researchers have used Recurrent Neural Network (RNN) method [7], the K-means Clustering algorithm [10], and the Decision Tree technique [8] for the prediction of dengue fever. In later days, Extreme Learning Machine (ELM) [16] [17] is widely used for better recognition in the classification process for dengue prediction [18]. The inclusion of the optimization step [27] gives better output for the big data classification process [26]. However, there were many drawbacks to the above models, like inaccuracy, time consumption, the requirement of a massive training set, and complexity.

In this paper, we propose the PSO-FELM method, an advanced algorithm for automatically detecting the presence of the dengue virus. FELM is a simple, improved method, and it is easy to implement, retaining the merits of ELM. FELM stands out in selecting the exact fuzzy membership for a particular challenge. The PSO is used to enhance the parameters of the FELM approach. Then the improved FELM is used to determine the dengue-infected people. This proposed method has higher accuracy, quicker results, and more reliability, and the overall performance in the diagnosis process is improved. When tackling the identification problem, a fuzzy ELM with a proper fuzzy membership function can effectively limit the effects of outliers.

The remaining structure of the paper is as follows: Section 2 (Related Studies) includes the previous works and their drawbacks. Section 3 describes the proposed method and its algorithm. Section 4 describes the suggested method's analysis and experimental findings. Section 5 makes a conclusion with the final results of the new method and the comparison with the previous algorithms.

2. Related Works

Researchers have used machine learning to detect dengue outbreaks in multiple studies. This section outlines some of the existing benchmarking techniques and also their limitations.
In 2017 Manivannan, PIDP and Devi, P.I. [10] proposed K Means clustering for predicting dengue fever. Dengue serotypes will be used to cluster dengue cases in households based on their age groups. By using K-Means clustering, it provides high proficiency in the output. Yet K-means clustering did not perform efficiently with the global cluster.

In 2017, Gambhir, S., et al. [9] created a PSO-ANN-based diagnostic model for earlier detection of dengue disease. The PSO methodology is used to optimize the weight and bias parameters of the ANN algorithm. The suggested diagnostic model is a capable and strong tool for detecting dengue fever early and more accurately.

In 2018 Anitha, A. and Wise, DJW et al. [8] developed a Data mining technique based on the decision tree. To validate the real-time approach dataset was collected from south Tamil Nadu. For classification and validation, the j48 trimmed tree is employed. The j48 technique is simple to understand, and the acquired knowledge can be used to make early predictions.

In 2019 Chovatiya., et al. [7] presented the prediction of dengue fever using LSTM RNN. The data are collected from various government websites and healthcare centers. Using Google heatmaps, the consequence of the possibility of an epidemic will be illustrated by highlighting the places where dengue could occur. Yet the paper only concentrates on the dengue epidemic disease

In 2019 Sanjudevi, D. and Savitha, D. [6] predicted Dengue fever using ML-based classification techniques. The feature selection is made by forwarding and backward stepwise regression, and for classification, SVM and decision tree are used. The analytical results show that the SVM efficiency was comparatively higher than DT. However, SVM does not operate with a large dataset

In 2020 Sarma, D., et al. [5] presented a framework based on Decision Tree (DT) and Random Forest (RF) to predict dengue fever. A real-time dataset was collected from various Bangladesh hospitals to validate the approach. The quantitative results show that the DT efficiency was comparatively higher than RF. The dataset had a small number of attribute details.

In 2021 Dourjoy, et al. [3] developed a framework based on SVM and Random Forest algorithm to analyze the symptoms and the early prediction of the symptoms. This model improves the prediction intensity of dengue hemorrhagic fever. The experimental data illustrate that SVM efficiency was better than RF efficiency.

In 2021 Swaraj, P.K. and Kiruthiga, G. [4] presented a deep learning model based on ANN to anticipate the infection status of a dengue patient. The real-time dataset was collected from different health centers across India. The validation shows that the ANN offers fast reactions through a quick confirmation process. However, this model requires lots of computation and training time.

In 2022 Amin, et al. [1] presented a unique approach to diagnosing Dengue illness based solely on posts on social media. The proposed technique RNN with LSTM is used for a surveillance system that works automatically and detection of places in which the split is occurring at a frightening rate, as well as guiding healthcare professionals on how to control the spread.

In 2022 Grampurohit, S and Sagarnal, C. [2] implemented the Ensemble ML technique to identify the spread of Dengue illness. SVM, DT, NB, and LR are designed in an ensemble strategy and assessed with acquisition data. The result reveals that the ensemble learning techniques had a higher prediction rate compared to the individual algorithm. The space and time consumption make this model less feasible.

The constraints of the aforementioned strategies include inaccurate prediction, restricted dataset, training time, slow convergence, computing expense, and small attribute data collection from affected individuals. This research provides an optimal Dengue fever prediction based on a PSO-optimised Fuzzy-ELM to get around the obstacles.

3. PSO-FELM Method

In this paper, the data set is collected from people manually. The data set comprises personal details, clinical pieces of information, and also pathological particulars. In the pre-processing step, the data are normalized and fed back to the data set again. The normalized data is given to the fuzzy module, where data is processed and optimized by the PSO module, and the result is compared with the trained model. Here the output is classified as a particular type of dengue fever.

The suggested method consists of four main processes to investigate the given data. 1. Data Collection and pre-processing 2. Data modeling 3. Classification module 4. Assessment Module. Different steps in each module are explained in figure.1.

3.1. Data Collection

The main intention of this method is to identify the infected people with the dengue virus by processing the data set obtained from the public. Here the public information is collected manually and given to the system. The data set includes information like personal data, clinical records, and pathological test reports. Data from 309 patients were collected from the Medical Colleges in Bangladesh and Chittagong. The types of fever are classified into three types: common fever, Ordinary dengue, and critical dengue.

Table 1. Statics of patient dataset

Class	Male	Female	Total
Common Fever	48	40	88
Ordinary Dengue	71	52	123
Critical Dengue	65	33	98
Total	184	125	309

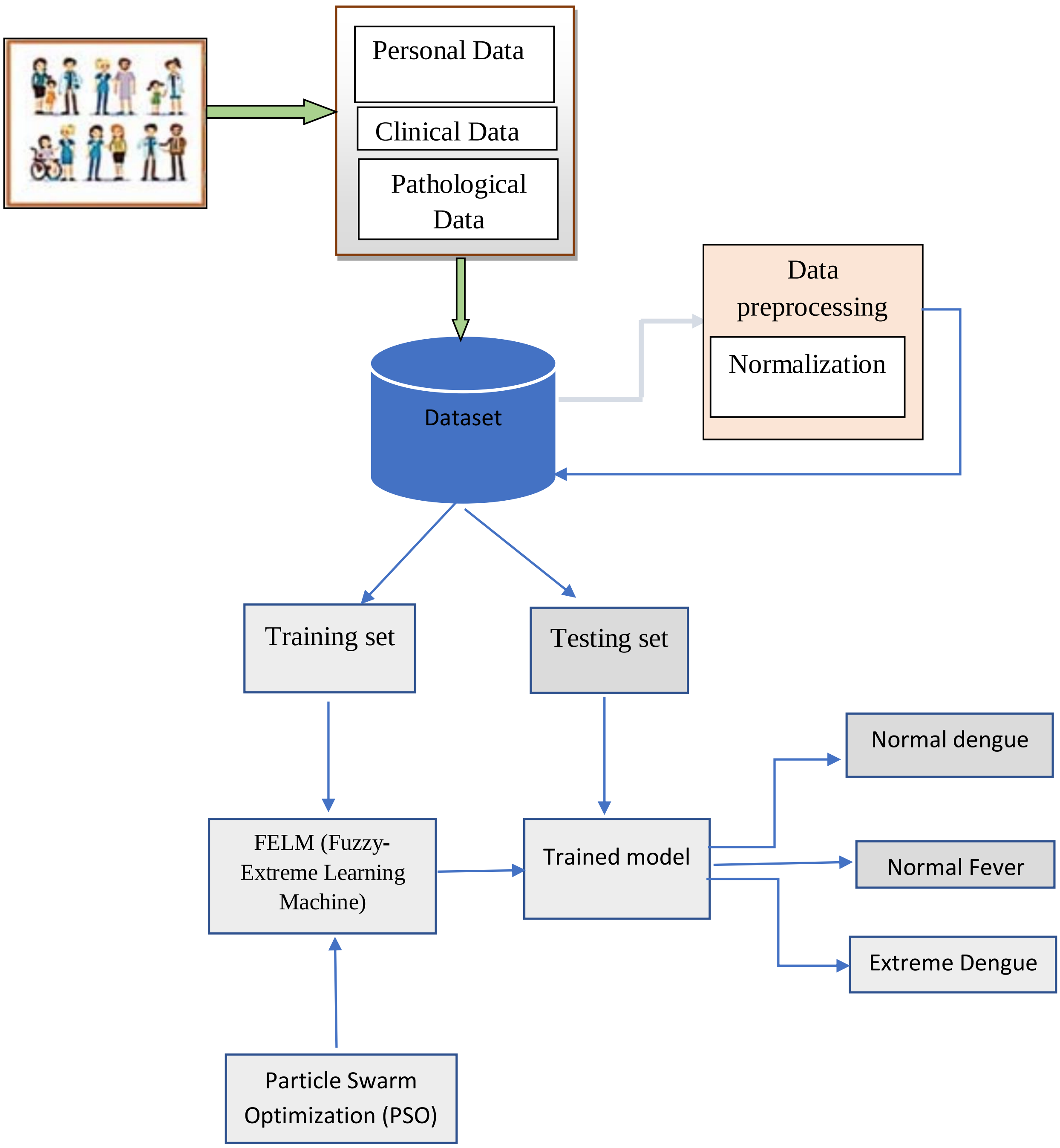

Fig. 1 Structural outline of the proposed PSO-optimized FELM method

3.2. Data Pre-processing

Data Pre-processing is one of the important steps, as the raw data cannot be processed. In this step, the data set is normalized using standard methods. The unwanted data and irrelevant information are removed. The obtained data is cleaned and standardized, and the data is noise-free.

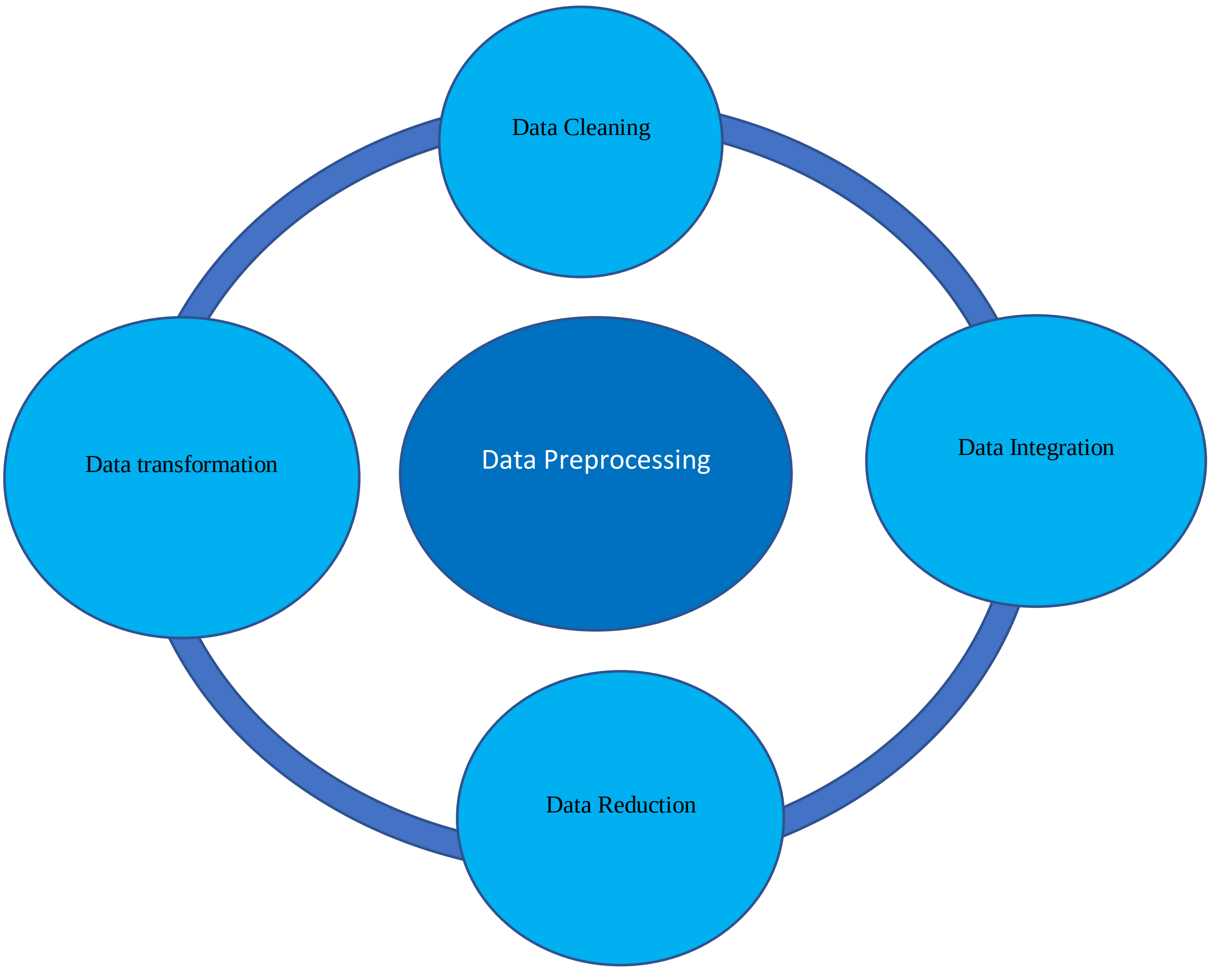

Fig. 2 Tasks in pre-processing

The grade of the data is inspected before applying Machine Learning algorithms. In pre-processing, data quality is assessed against a set of criteria, including accuracy, completeness, consistency, timeliness, believability, and interpretability. The four key responsibilities in data pre-processing are cleaning, integration, reduction, and transformation, as shown in fig 2.

3.3. Data Modeling

Based on their features, the normalized data is divided into two sets, a training set and a testing set. The output of the training set is given to the fuzzy ELM module for classification, whereas the testing set output is given to the training model unit for evaluation.

3.4. PSO Unit

Particle Swarm Optimization is one of the most effective optimization approaches since it uses a limited number of hyperparameters. It is a basic bio-inspired approach for determining a function's maximum and minimum values. Particles travel through a multidimensional cast region in PSO, and each particle represents an attainable solution.

The synopsis of a PSO can be explained in the following steps [13].

Step 1: Design a searching space with a population of a random potential solution.
Step 2: Fitness Calculation of each particle in the swarm
Step 3: The fitness of each particle is compared to its prior best-obtained fitness p_i^k. If the current value is better than the previously acquired pik, update the pik value to the present value and the location to the present location.
Step 4: In comparison, find the best global location with the greatest fitness g_i^k
Step 5: The acceleration towards p_i^k and g_i^k is weighed by a random term. Each particle's new position is calculated by adding the new velocity value
Step 6: Do a repletion of steps from step 2 to step 5 until the convergence is reached.
PSO provides the best possible solution by comparing the whole functional space and interacting socially, yielding an optimized result. Usage of the PSO technique improves the accuracy rate.

3.5. Fuzzy ELM Unit

Extreme Learning Machine (ELM) is a simple and accurate classification programming language used to categorize large data. Figure 3. Explains the basic working of a PSO-ELM having a single hidden layer. ELM can be used when dealing with problems that require real-time network retraining. The different input points may not be precisely assigned to one of the classes in many real-world applications. This issue is beyond the scope of the ELM convention. As a result, we offer Fuzzy ELM, a unique processing technique for this scenario. In addition to the standard ELM, this has a set of fuzzy memberships and a fuzzy matrix. Following that, inputs with varied fuzzy memberships can contribute differently to the learning of output weights. Thus, the above-mentioned problem is solved by Fuzzy ELM. It reduces the effects of outliers and provides more logical results compared to ELM. FELM can produce a more rational conclusion than ELM for weighted classification issues.

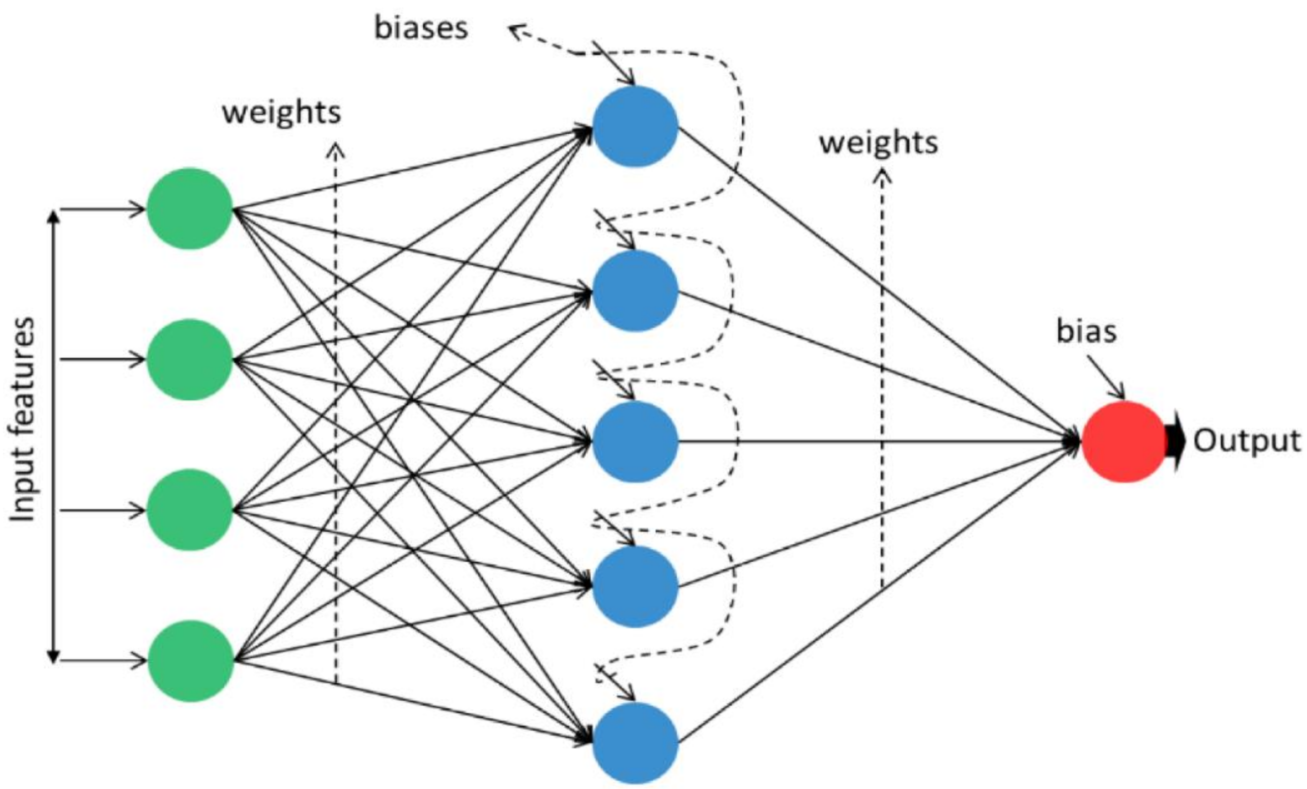

Fig. 3 PSO-ELM model

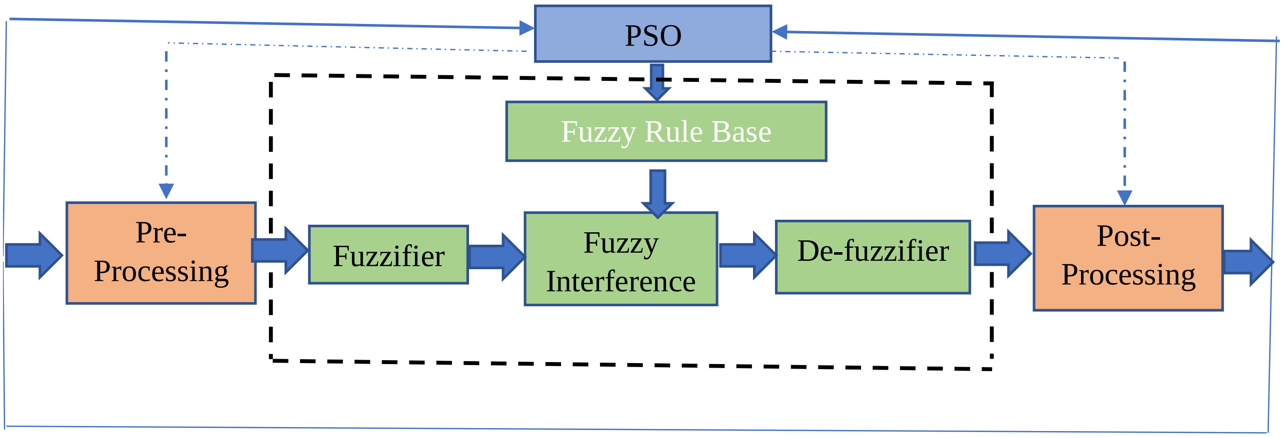

Fig. 4 Flow chart of PSO optimized FELM

The process in a PSO-FELM is explained in fig.4. The PSO-FELM unit consists of the pre-processing unit, PSO unit, fuzzy rule base, fuzzy interference unit, fuzzifier, defuzzifier, and post-processing unit.

3.6. Assessment Module

The classified data is given to the training model and compared with the testing set. On iteration, the given data is classified into any one of the results, namely Normal fever, Normal Dengue, or severe dengue. Thus, a person can evaluate his results through this method and take necessary actions for the same.

4. Results and Discussion

The outcomes in the realm of medical diagnosis should be precise. The accuracy, sensitivity, and specificity of test data sets are calculated to evaluate the classification model's efficiency. The sum of positive instances of a model divided by the total number of instances is the model's accuracy. The accuracy parameter provides the percentage of correctly identified cases. To appropriately diagnose a person with disorders, sensitivity is utilized to determine the degree of the qualities. To appropriately describe the individual without disorders, specificity is employed to determine the degree of the traits.

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

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

Where
TP – no. of correct cases identified as patients,
FP – no. of incorrect cases identified as patients,
TN – no. of correct cases identified as healthy and
FN – no. of incorrect cases identified as healthy.

The Area Under the Curve (AUC) is a critical measure for evaluating diagnostic test performance and determining a condition's prevalence. It is a two-dimensional graphic that analyses the validity of medical tests by comparing sensitivity and specificity.

Table 2. Numerical values of the attributes

Variables	Description	Possible Values (scores)
Sex	Patient Gender	0=Male, 1=Female
Mode of Work	Occupation of the Patient	0=Desk bound, 1=fairly active, 2= Extremely Active
Cough Symptoms	Heavy Cough/ Throat Pain	0=No, 1=Yes
Pulse	Pulse Rate of the Patient	Values in numbers
Fever Scale	Temperature measurement	0=Normal, 1= Medium, 2= High
Headache	Heaviness in head	0=Normal, 1= Medium, 2= High
Muscle pain	Body pain	0=Normal, 1= Medium, 2= High
Skin rash	Rashes in the body	1=Yes, 0=No
Fatigue	Weakness/Tiredness	1=Yes, 0=No

The collected data from the patients are normalized using the numerical values of the attributes like sex, work, pulse, fever, muscle pain, skin rash, etc., as shown in table 2. This dataset is then processed by the proposed PSO-FELM method. The normalized dataset is shown in table:3.

Table 3. Normalized dataset

Age	Sex	Work	Cough	Pulse	Fever	headache	Body pain	Skin Rash	Fatigue
45	0	2	1	68	2	1	2	1	1
38	1	2	0	98	0	0	1	1	1
76	1	0	1	56	1	1	1	1	0
63	0	1	1	77	1	1	1	1	1
48	0	1	0	88	0	1	0	0	0
73	1	0	1	93	1	0	1	1	0
59	1	1	1	67	1	0	1	1	1
64	1	0	0	83	2	1	0	0	1
49	0	1	1	70	1	1	1	1	0
37	1	2	1	95	1	1	1	1	0

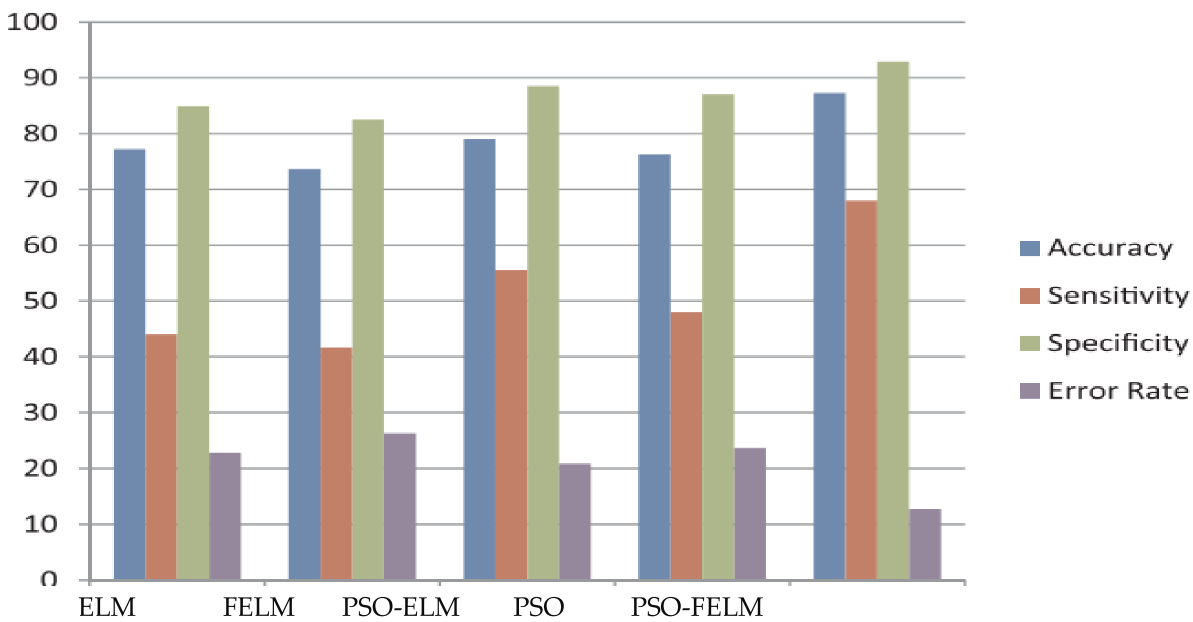

Fig. 5 Performance comparison of different approaches

The performance of different approaches is compared, and the bar diagram showing the results is shown in fig.5. The proposed PSO-FELM method has high accuracy and less error rate when compared with the other methods like ELM, FELM, PSO, and PSO-ELM.

5. Conclusion

In the proposed system PSO - Fuzzy ELM, dengue fever is predicted accurately in a short period, compared with the existing models. The drawbacks in the previous models are overcome in this method. The imbalance problems and weighted classification issues are conquered in this method. The usage of the PSO technique improves accuracy. Now the patient can identify the disease without consulting medical officers and thus reducing the burden on doctors. An accuracy of 95 % is obtained compared with existing approaches like ELM, ANN, DT, NB, and PSO.

References

[1] Samina Amin et al., "Recurrent Neural Networks with TF-IDF Embedding Technique for Detection and Classification in Tweets of Dengue Disease," IEEE Access, vol. 8, pp. 131522-131533, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Sneha Grampurohit, and Chetan Sagarnal, "Disease Prediction Using Machine Learning Algorithms," 2020 International Conference for Emerging Technology, IEEE, pp. 1-7, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Saif Mahmud Khan Dourjoy et al., "A Comparative Study on Prediction of Dengue Fever using a Machine Learning Algorithm," Advances in Distributed Computing and Machine Learning, pp. 501-510, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] P.K. Swaraj, and G. Kiruthiga, "Design and Analysis of Medical Image Classification for Dengue Detection Using an Artificial Neural Network Classifier," ICTACT Journal on Image and Video Processing, vol. 11, no. 3, pp. 2407-2412, 2021. [Google Scholar] [Publisher Link]

[5] Dhiman Sarma et al., "Dengue Prediction using Machine Learning Algorithms," 2020 IEEE 8th R10 Humanitarian Technology Conference, IEEE, pp. 1-6, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] R. Sanjudevi, and D. Savitha, "Dengue Fever Prediction using Classification Techniques," International Research Journal of Engineering and Technology, vol. 6, no. 2, pp. 558-563, 2019. [Google Scholar]

[7] Megha Chovatiya et al., "Prediction of Dengue using Recurrent Neural Network," 2019 3rd International Conference on Trends in Electronics and Informatics, IEEE, pp. 926-929, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[8] A. Anitha, and D.C. Joy Winnie Wise, "Forecasting Dengue Fever using Classification Techniques in Data Mining," 2018 International Conference on Smart Systems and Inventive Technology, IEEE, pp. 398-401, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[9] Shalini Gambhir, Sanjay Kumar Malik, and Yugal Kumar, "A PSO-ANN-Based Diagnostic Model for the Early Detection of Dengue Disease," New Horizons in Translational Medicine, vol. 4, no. 1-4, pp. 1-8, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[10] P. Manivannan, and P. Isakki Devi, "Dengue Fever Prediction using K-Means Clustering Algorithm," 2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing, IEEE, pp. 1-5, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[11] Dhaval Raval et al., "Medical Diagnosis System Using Machine Learning," International Journal of Computer Science & Communication, vol. 7, no. 1, pp. 177-182, 2016. [CrossRef] [Google Scholar]

[12] Kasra Mohammadi et al., "Predicting the Wind Power Density Based Upon Extreme Learning Machine," Energy, vol. 86, pp. 232-239, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[13] Mosbeh R. Kaloop et al., "Particle Swarm Optimization Algorithm-Extreme Learning Machine (PSO-ELM) Model for Predicting Resilient Modulus of Stabilized Aggregate Bases," Applied Sciences, vol. 9, no. 16, p. 3221, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] Yonghe Chu et al., "Fuzzy ELM for Classification Based on Feature Space," Multimedia Tools and Applications, vol. 79, no. 37, pp. 27439-27464, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Bin Zhu et al., "Hybrid Particle Swarm Optimization with Extreme Learning Machine for Daily Reference Evapotranspiration Prediction from Limited Climatic Data," Computers and Electronics in Agriculture, vol. 173, p. 105430, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] S. R. Mugunthan, and T. Vijayakumar, "Design of Improved Version of Sigmoidal Function with Biases for Classification Task in ELM Domain," Journal of Soft Computing Paradigm, vol. 3, no. 2, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Abdullah Toprak, "Extreme Learning Machine (ELM)-Based Classification of Benign and Malignant Cells in Breast Cancer," Medical Science Monitor, vol. 24, pp. 6537-6543, 2018. [CrossRef] [Google Scholar] [Publisher Link]
[18] Hector Estigarribia, Daniel Romero, and Juan Vicente Bogado Machuca, "Application of Extreme Learning Machine to Forecast Outbreaks of Dengue Cases," Research Cate, 326305815, 2018. [Google Scholar]
[19] Weihong Cai et al., "PSO-ELM: A Hybrid Learning Model for Short-Term Traffic Flow Forecasting," IEEE Access, vol. 8, pp. 6505-6514, 2020. [CrossRef] [Google Scholar] [Publisher Link]
[20] Xiujuan Sui, Lei Xue, and Dunyu Li, "Dynamic Expression Recognition Based on Hybrid Features and Optimized Extreme Learning Machine Model," IOP Conference Series: Materials Science and Engineering, CTCE, vol. 719, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[21] Sirote Khunkitti et al., "A Hybrid DA-PSO Optimization Algorithm for Multiobjective Optimal Power Flow Problems," Energies, vol. 11, no. 9, p. 2270, 2018. [CrossRef] [Google Scholar] [Publisher Link]
[22] Iqbal. N., Islam. M., 2017. Machine learning for dengue outbreak prediction: An outlook. ISSN No. 0976-5697
[23] Kamran Kowsari et al., "Text Classification Algorithms: A Survey," Information, vol. 10, no. 4, p. 150, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[24] Harsh Patel, and Purvi Prajapati, "Study and Analysis of Decision Tree-Based Classification Algorithms," International Journal of Computer Sciences and Engineering, vol. 6, no. 10, pp. 74-78, 2018. [CrossRef] [Google Scholar] [Publisher Link]
[25] Mucahid Mustafa Saritas, and Ali Yasar, "Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification," International Journal of Intelligent Systems and Applications in Engineering, vol. 7, no. 2, pp. 88-91, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[26] Punia. S. K., Kumar. M., Stephan. T., Devarajan. G.G., Patan, R., "Performance Analysis of Machine Learning Algorithms for Big Data Classification: ML and AI-Based Algorithms for Big Data Analysis," International Journal of E-Health and Medical Communications, vol. 12, no. 4, 2021. [CrossRef] [Google Scholar] [Publisher Link]
[27] Sawyer D. Campbell et al., "Review of Numerical Optimization Techniques for Meta-Device Design [Invited]," Optical Materials Express, vol. 9, no. 4, pp. 1842-1863, 2019. [CrossRef] [Google Scholar] [Publisher Link]