IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13363

Rainfall Prediction using Hybridized Genetic Algorithm-Based Artificial Neural Network (GA-ANN) and Genetic Algorithm-Based Support Vector Machine (GA-SVM) Models.

Divya Sri.M[1], Venkata Sai Bhargav.K[2], Chandra Kireeti.K[3], Lahari Priya.M[4]

K. Rajeshkumar[5]

CSE, KL University, Andhra Pradesh, India.[1-4]

Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhrapradesh, India[5]

Abstract: Rainfall forecasting is essential for a few industries, including agriculture, water resource management, and flood forecasts. Rainfall prediction is most important, but now-a- days rainfall forecasting has grown to be a difficult issue. The ability to take safeguards is made possible by accurate rainfall predictions. To anticipate the dependent variables temperature, humidity, location, wind speed, and direction the rainfall prediction is dependent on a few constantly shifting factors. The weather calculation also varies according on the location's geographic characteristics and atmospheric variables. To predict rainfall, this study suggests two hybrid models ANN and SVM based on genetic algorithms. To optimize the hyperparameters and boost prediction accuracy, the genetic algorithm (GA) was combined with ANN and SVM, respectively, to create the GA-ANN and GA-SVM models.

The effectiveness of the proposed models was assessed through the application of several statistical metrics, such as Mean Absolute Error (MAE), Correlation Coefficient (CC), and Root Mean Squared Error (RMSE). The results obtained indicate that both GA-ANN and GA-SVM models outperformed the traditional ANN and SVM models, with GA-ANN exhibiting marginally superior performance. Consequently, the suggested hybrid models can be applied to accurately estimate rainfall. The two hybridized models GA-ANN, GA-SVM, are trained and tested using historical rainfall data from a particular region. The study evaluated the performance of both models and compared them to traditional models. The GA-SVM model is a good approach for rainfall prediction, and its accuracy can be further improved by incorporating additional meteorological and environmental factors. The potential of hybridized models that combine genetic algorithms with machine learning techniques to increase the precision of models used to forecast rainfall.

In summary, the suggested models help us to improve the prediction of rainfall with better accuracy and gives us better results in agriculture, water resource management, and flood forecasting.

Keywords: Algorithm, Artificial Neural Network, Genetic Algorithm, Hybridized Models, Rainfall Prediction, Support Vector Machine.

I. INTRODUCTION

Among the most challenging and crucial features of the hydrologic cycle is rainfall forecasting [10]. This is largely due to the variability it exhibits at various temporal and spatial domains. Being the result of heavy rain, flash flooding is a dangerous occurrence. Creating a flood warning and rainfall forecasting system is thought to be a challenging undertaking for typical catchments. The greenhouse effect and global warming make it more difficult to estimate rainfall with accuracy. Due to complicated behaviour, rainfall is largely repressed; hence, rainfall-affecting characteristics are necessary to predict rainfall [12]. Estimating rainfall is significant due to its impact on water supplies, human existence, and water consumption. It is particularly challenging to estimate rainfall that is impacted by regional and geographic differences and features [14].

In the fields of meteorology and hydrology, rainfall prediction is crucial because it ensures appropriate control of water resources and mitigates the negative consequences of droughts and floods. Because machine learning models and algorithms can handle complex and non-linear interactions between meteorological factors, they have become a popular tool for predicting rainfall in recent years.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⧋ Peer-reviewed & Refereed journal ⧋ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13363

Two popular machine learning methods are used for rainfall forecasting are SVM and ANN. To enhance the accuracy prediction of these models, hybridization with optimization techniques such as Genetic Algorithm (GA) has been proposed. This is how the two hybrid models for rainfall prediction the genetic algorithm based on the ANN and genetically altered SVM. Many studies focus on predicting the weather in advance to reduce the amount of harm to people and property. Because of the devastation that heavy rainfall causes and the numerous deaths it takes to cause, one of the most significant climatic variables that must be predicted long in advance is rainfall [8].

The Importance of Rainfall Forecasting and Its Applications are Agriculture: Accurate rainfall forecasting helps farmers plan for their planting and harvesting schedules, and select appropriate crop varieties and optimize water usage, farmers can take precautionary measures to safeguard their crop. Water Resource Management: Rainfall forecasting is crucial for managing water resources, such as reservoirs, rivers, and lakes. It aims in the effective allocation of water for drinking, agriculture, and industrial purposes accurate forecast of rainfall help for prevention of floods and droughts.

The goal of this research article is to forecast the amount of rainfall at a specific place by using user-provided input data, which include location, temperature, humidity, wind direction, and speed [2]. Additionally, GA-ANN and GA-SVM models are used to train these rainfall properties.

The author of [6] conducted a standard survey on every neural network architecture that is now in use and has been heavily used to predict rainfall in recent years. Researchers in [5] used radial basis operations, neural networks, and back propagation to anticipate a monthly collapse. The researchers in [13,11] have presented a general review and conducted a comprehensive analysis of various neural networks for the purpose of forecasting the region's rainfall. Most of the research projects that are suggested have high computation times and relatively low efficiency.

Because ANN can examine and exploiting older climatic records, it is mostly used for forecasting. ANN outperforms other numerical models in terms of precision. BP is the most effective strategy for training ANN. GA is a viable method for Neural Network (NN) optimization [3]. The BPNN model was created by Geetha and Selvaraj (2011) in Chennai, India, to predict the mean monthly rainfall [1]. The NN model approach was created by Chen and Takagi (1993) to estimate rainfall in the open ocean close to Japan's Shikoku. To ascertain the correlation between rainfall intensity and information from the Geostationary Meteorological Satellite (GMS), they employed the neural network (NN) model. The network is trained using the BPNN learning algorithm using the GMS picture as input data [15]. In this thesis, we provide the hybrid Models GA-ANN and GA-SVM which show the promising results in rainfall prediction, and they can be used as tools for decision making in various applications.

II. METHODOLOGY

2.1 Artificial Neural Network:

Research on biological neuron processing influenced a computing architecture known as a neural network. Neural networks come in a wide variety of forms, from very simple to quite complex. The forward-looking neural network Multiple layers of the algorithm components make to a layered feed forward neural network. After performing separate computations on the data, it receives, a layer of processing components passes the results to a higher layer. The succeeding layer conveys the results of its own independent calculations to the layer below. Ultimately, the network's output is determined by layers including one or more processing components [10][9]. Figure 1 shows the feed forward neural network's detailed design.

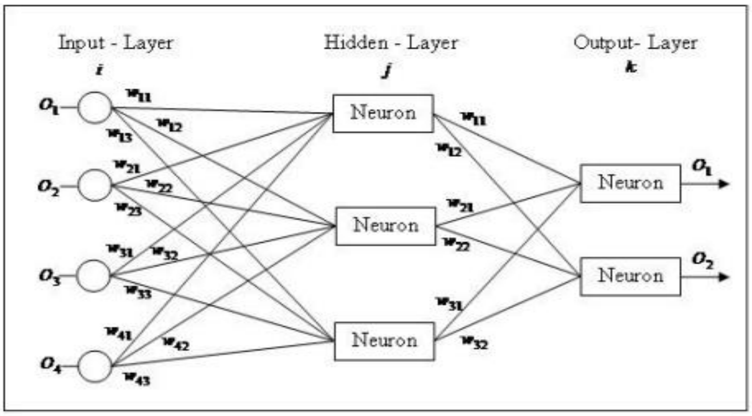

Fig.1—Feed forward neural network architecture

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 13, Issue 3, March 2024
DOI: 10.17148/IJARCCE.2024.13363

Backpropagation, or backward propagation of faults, is a method intended to identify mistakes as they make their way from input to output nodes backward. Backpropagation is a method of learning used by ANN to produce a gradient decline with respect to weight values for the various inputs. The architecture of backpropagation is mentioned in the fig 2.

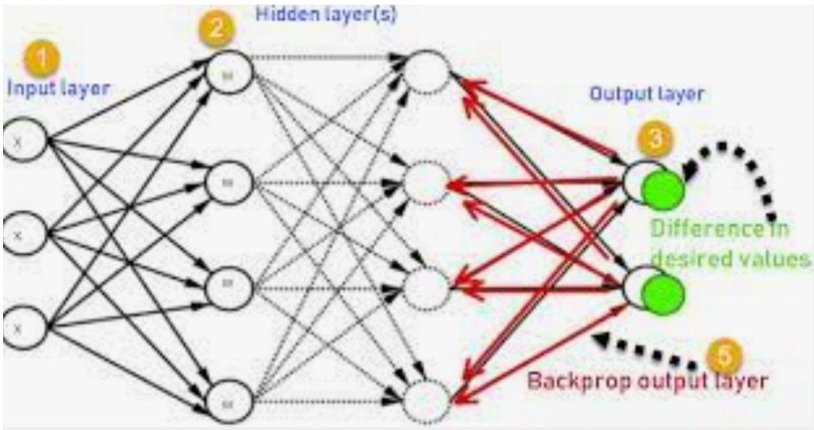

Fig.2 Backpropagation architecture

2.2 Genetic Algorithm:

A genetic algorithm flow is shown in Figure 2. The operators for selection, crossover, and mutation will be used to allocate the fitness. The next prediction will be made if the value is true. These three operators will be used once more to allocate the fitness if the response is false. These three operators each have their own special operations. The values from the rainfall dataset were predicted using this technique. for selecting the top from the chosen area or region. It made it possible for our suggested system to anticipate values more quickly and provide the best performance validation [8].

Fig.3 Flow chart of genetic algorithm

2.3 Support Vector Machine:

Powerful machine learning methods like SVM are frequently employed in classification and regression issues [4]. A hyperplane is generated by the SVM classifier to categorize data instances, by creating the greatest distance between the several occurrences, hyperplanes can be identified. SVMs can receive an instruction in complicated small- to medium-sized datasets, unlike deep learning. SVMs can also perform structural risk minimization, which minimizes the generalization error bound and prevents overfitting [7]. The detail architecture of support vector machine is in fig 4.

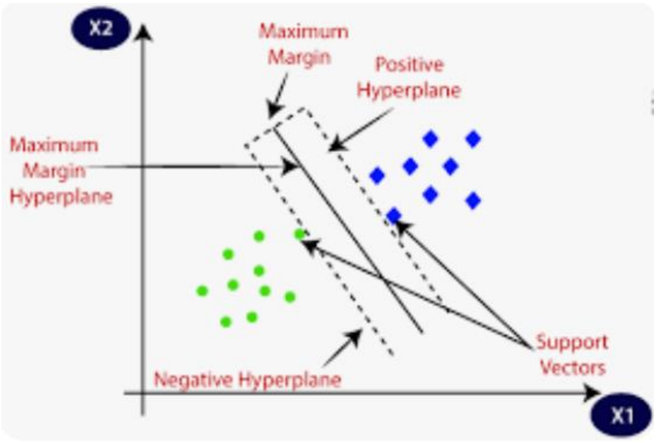

Fig.4 Support Vector Machine architecture

2.4 Data Preprocessing: Handle missing values in the dataset via ascription or evacuation of insufficient entries. Identify and address exceptions that will have an impact on demonstration preparation and expectation exactness. Normalise or standardise the information to make sure every part is model-ready and has the same scale.

2.5 Data Segmentation: Sets for preparation and testing should from the previously processed data. After the models have been developed using the preparation set, their performance will be assessed using the testing set.

2.6 Genetic Algorithm for Hyperparameter Optimisation: Implement a hereditary computation to optimise the hyperparameters of the GA-ANN and the GA-SVM models. The ANN's learning rate, the number of neurons and hidden layers within it,and bit capacities and regularisation parameters in the SVM are examples of hyperparameters. Set the parameters of the hereditary algorithm, such as the population estimate, transformation rate, and end criteria.

2.7 GA-ANN Showcase Improvement: Create an Artificial Neural Network (ANN) for precipitation forecasting. Include input, cover up, and yield layers, and start the arrangement with weights. To train the GA-ANN demonstration, use the preparation data and the hyperparameters optimised by the hereditary calculation. Backpropagation or other appropriate preparatory computations should be used.

2.8 GA-SVM Model Construction: Create a Support Vector Machine (SVM) for precipitation forecasting by selecting an appropriate part work (e.g., direct, spiral premise work, polynomial). Prepare the GA-SVM demonstration using the preparation data and the hyperparameters optimised by the hereditary computation.

2.9 Model Assessment: Use the testing dataset to evaluate the GA-ANN and GA-SVM models' performance. Use built-in evaluation measures like increase forecast accuracy, employ the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). By comparing the model's performance against that of conventional models, decide which one performs the best.

2.10 Forecasting and foreseeing: Use the crossbreed models that have been built (GA-ANN, GA-SVM, or gathering) to generate precipitation projections for upcoming time periods based on current meteorological data.

2.11 Show Support and Checking: Continuously monitor the crossover models' performance and retrain them as necessary using newer data to maintain their accuracy.

2.12 The act of visualising: Create visualisations that separate the expected precipitation values from the actual precipitation data to enable a visual assessment of the model's performance.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⩓ Peer-reviewed & Refereed journal ⩓ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13363

2.13 Detailing and Record Keeping: Plan a thorough report that summarises the whole methodology, including data sources, preprocessing methods, demonstration models, hyperparameter optimisation, evaluation outcomes, and any notable discoveries or pieces of knowledge.

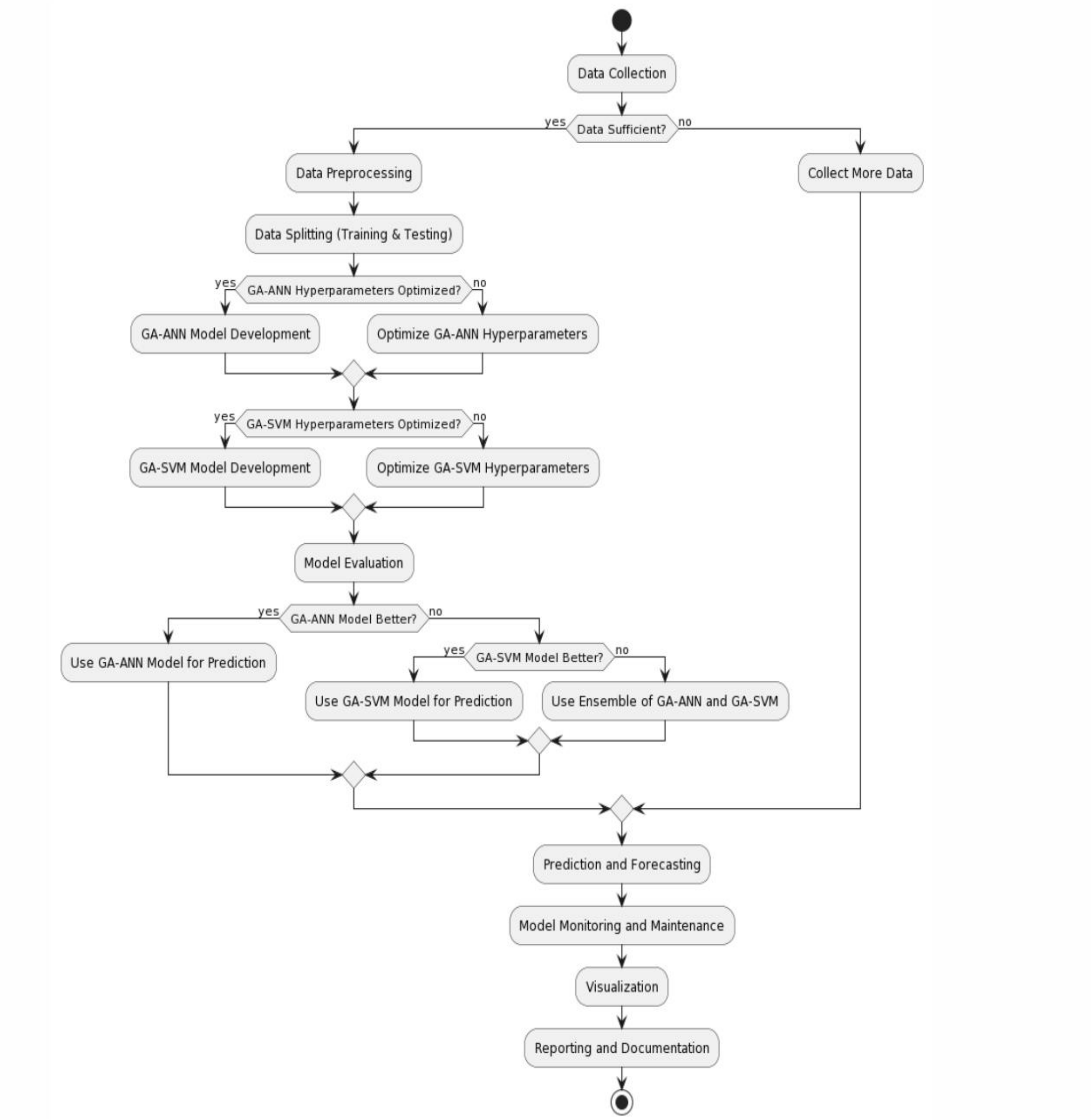

Fig.5 Flowchart of Methodology

III. RESULTS

	RMSE	MSE	MAE	VARIANCE	r
Training	0.130642	0.017067	0.085439	0.017111	0.705014
Validation	0.135717	0.018419	0.084692	0.018882	0.644
Testing	0.119862	0.014367	0.082109	0.014733	0.711715

TABLE 1: shows the error statistics of the GA-ANN model used to forecast rainfall at the Mahanadi River's Tikarapara station. The acronym for mean absolute error is MAE. RMSE stands for root mean square error, MSE for mean square error, and r for correlation coefficient.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ※ Peer-reviewed & Refereed journal ※ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13363

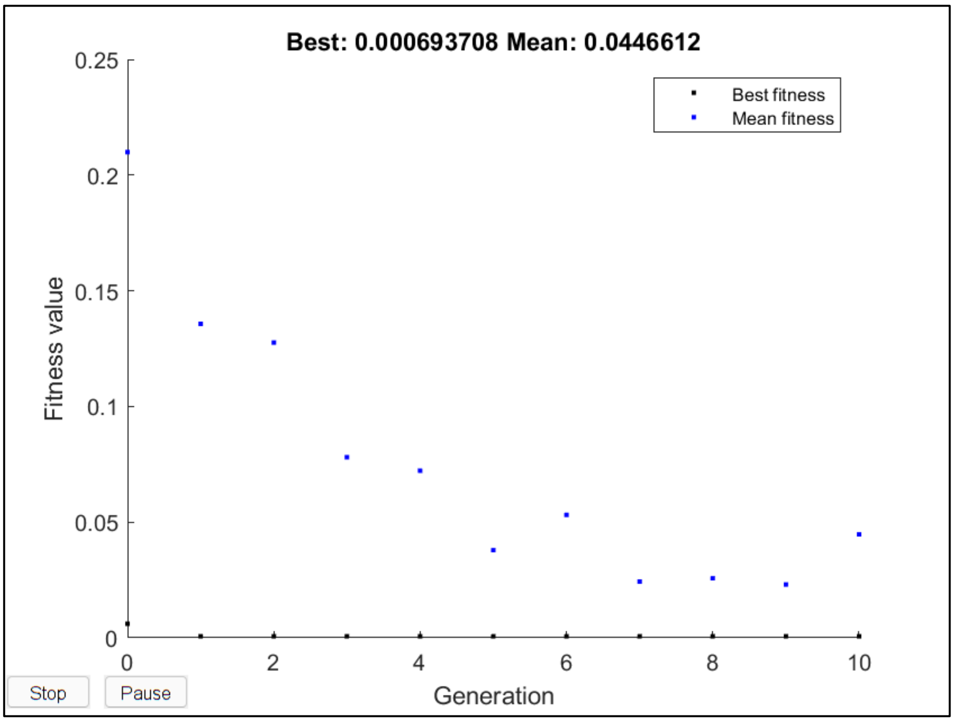

Fig 6. Generation-wise fitness function profile variation during GA-based ANN learning for rainfall forecasting.

Using the test data set, the created GA-ANN model's performance and generalization capacity were assessed. Table 1 displays the rainfall data used in the datasets for testing, validation, and training to perform the statistical analysis of errors. Figure 6 shows that for all three, the coefficient of correlation (r) is quite high and the RMSE is very low.

Additionally, all three have very low and consistent mean absolute errors (MAEs). It is determined based on the RMSE, MAE, and r values, this GA-ANN model has successfully predicted rainfall with a good degree of accuracy. The generated model's ability to generalize was shown by the consistency of all parameters across all three. Through It was shown by testing, validation, and training that this model protects against both excess and insufficient fit due to its elevated r values and low error parameter values.

A set of final solutions at a predetermined stopping criterion matching to maximum generations was supplied by the GA-ANN model. Figure 1 illustrates the fluctuation of the mean fitness and the best fitness value (RMSE) in each generation during the training phase. Out of all the generations, 0.000693708 had the best fitness (Mean: 0.0446612). Additionally, the optimal fitness function of every genetic learning generation was noted.

It was shown by the best chromosomal matching the optimal fitness function that takes into account the requirements of temperature, precipitation, and water outflow chosen to build a neural network model. The outcomes demonstrated that the concealed layer's ideal neuronal count.

The tan-sigmoid activation function, a continuous, monotonically increasing S-shaped curve, was the most suitable option for an activation mechanism of the output and hidden layers. The GA-ANN model determined the optimal value in the LM algorithm model, combination coefficient (μ) is utilized. We did our best to select the initial bias terms and weights of connections. The GA ANN model's ideal solution was the best-fit chromosomal solution from the evolution run.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024
DOI: 10.17148/IJARCCE.2024.13363

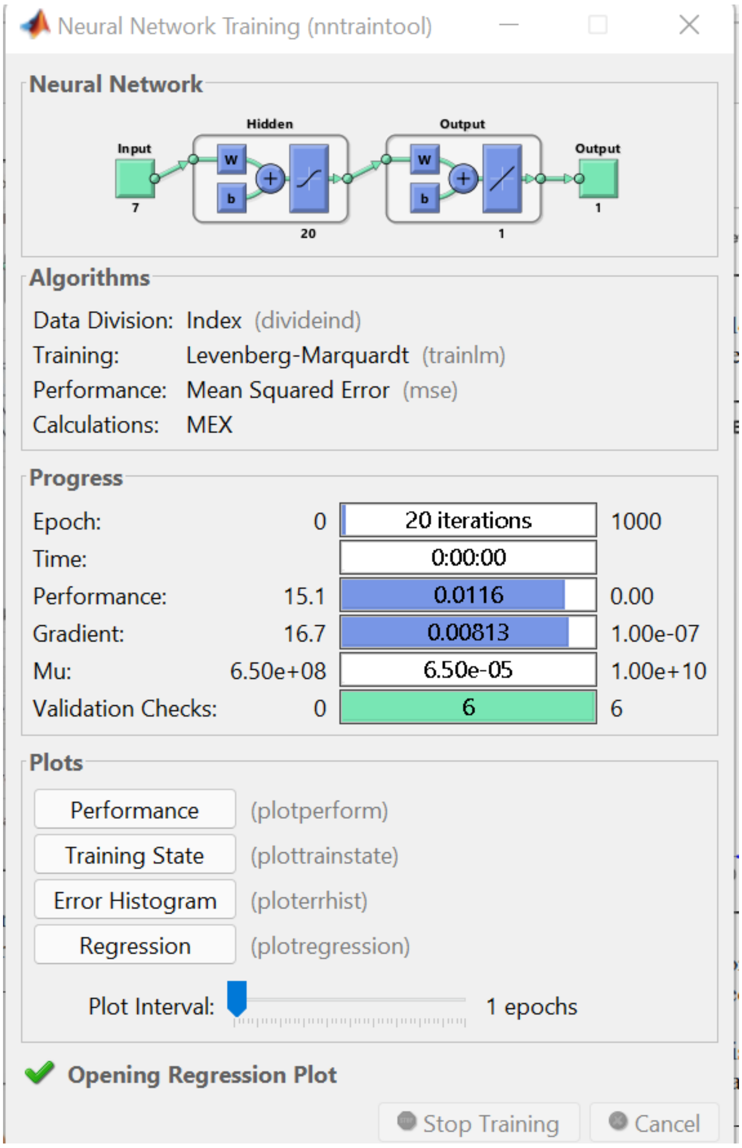

Fig 7. Artificial neural network's training parameter structure.

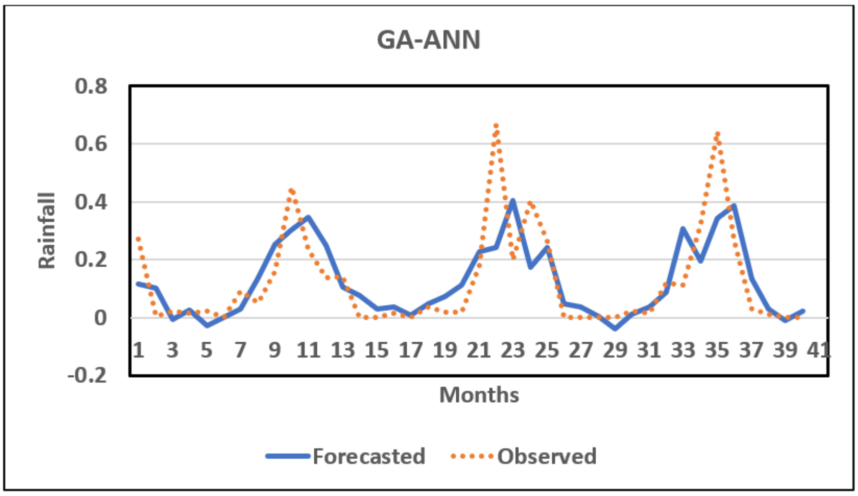

Fig 8. A comparison using testing data between the observed and GA-ANN anticipated rainfall.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 3, March 2024
DOI: 10.17148/IJARCCE.2024.13363

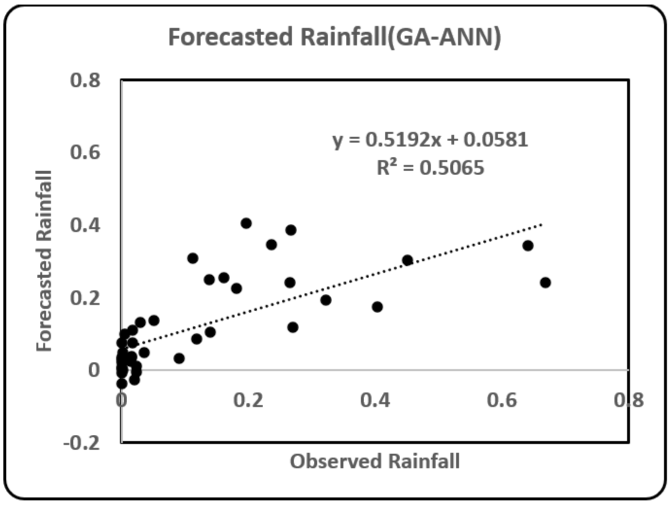

Fig 9. Using testing data, create a scatter plot comparing the observed and GA-ANN anticipated rainfall.

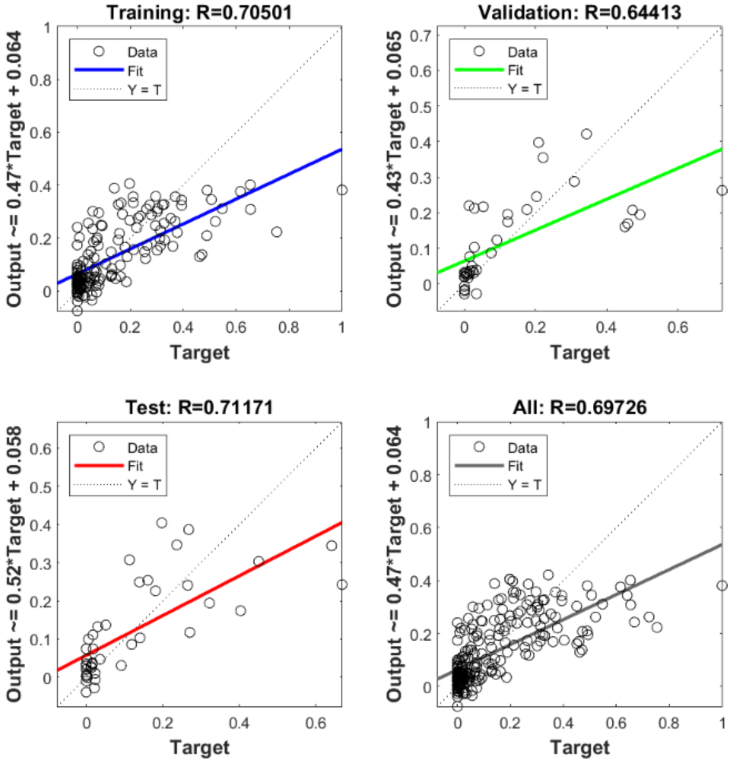

Fig10. A scatter plot that compares the observed rainfall to the Rainfall anticipated by GA-ANN using training, validation, testing, and all combined data.

Water discharge, precipitation fall, and temperature were utilized as the model's input parameters to determine the best GA-ANN model to use in predicting the actual and anticipated rainfall values across various time periods (monthly).

Figures 3 and 4 depict the observed and anticipated rainfalls and clearly illustrate the most times that the rain has fallen is overstated. There were just two months in which it was not fully appreciated. Therefore, it can be deduced that the GA-ANN model yields highly accurate results at the peaks as well.

DOI: 10.17148/IJARCCE.2024.13363

IV. CONCLUSION

Data on temperature, precipitation, sediment yield, water level, and water discharge sets are used as inputs. India, to predict water flow using the ANN and new GA- ANN models. These inputs' parameters possessed been discovered to be the most important parameter for predicting water flow. In the past, several ANN models predicted negative values at times of low water, howeverin this investigation, The water flow values predicted by the ANN and GA-ANN models are all positive. Additionally, the model-predicted water flow is shown to be larger than the observed water flow by the scatter plot and hydro graph comparing the two. At the Tikarapara gauge station in the basin of the Mahanadi River. India, this study proved the calculation of suspended using GA-ANN models with various Sediment yields were computed using inputs of hydroclimatic factors such as water discharge, rainfall, and temperature. The most important regulating characteristics of the Mahanadi River's suspended sediment were found to be water discharge, rainfall, and temperature. Climate change may have an impact on a river's sedimentation process.

For calculating the suspended sediment yield, a new ANN structure was taken into consideration, and this structure was effectively optimized by using the GA. The suggested GA-ANN models use temperature, precipitation, and water discharge as simultaneous inputs and forecast the sediment yield well. The concurrent optimization of input variables and parameters for modelling the prediction of sediment output is one of the work's contributions. Although the method has been studied in other application areas, this research is unique in that it concentrated on predicting sediment output. Furthermore, the study's methodology, data sources, and results are all unique.

REFERENCES

[1]. Abhishek K., Kumar A., Ranjan R., Kumar S., A rainfall prediction model using artificial neural network, IEEE Control and System Graduate Research Colloquium (2012), 82-87.
[2]. Akash D Dubey, Artificial Neural Network Models for Rainfall Prediction in Pondicherry, International Journal of Computer Applications 120(3) (2015).
[3]. Ankit Chaturvedi, Rainfall Prediction using Back-Propagation Feed Forward Network, International Journal of Computer Applications 119(4) (2015).
[4]. A.Geron, ´ Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems." O'Reilly Media, Inc.", 2017.
[5]. B. K. Rani and A. Govardhan, "RAINFALL PREDICTION U SING DATA MINING TECHNIQUES - A SURVEY," pp. 23–30, 2013.
[6]. D. Nayak, A. Mahapatra, and P. Mishra, "A Survey on Rainfall Prediction using Artificial Neural Network," Int. J. Comput. …, vol. 72, no. 16, pp. 32–40, 2013.
[7]. Eslam Hussein, Mehrdad Ghaziasgar, Christopher Thron on Regional Rainfall Prediction Using Support Vector Machine Classification of Large-Scale Precipitation Maps in Virtual conference, Pretoria, South Africa 6 - 9 July 2020.
[8]. J.Refonaa, M. Lakshmi, R S S SrinivasaRao, P Eshwar Prasad on Rainfall Prediction using Genetic Algorithm in International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-2S3, July 2019.
[9]. Kumar Abhishek, Abhay Kumar, Rajeev Ranjan, Sarthak Kumar on A Rainfall Prediction Model using Artificial Neural Network in 2012 IEEE Control and System Graduate Research Colloquium (ICSGRC 2012).
[10]. M.Nasseri, K.Asghari , M.J.Abedini on Optimized scenario for rainfall forecasting using genetic algorithm coupled with artificial neural network. Expert Systems with Applications 35 (2008) 1415–1421.
[11].] Rajalakshmi, V., M. Lakshmi, and V. Maria Anu. "A Complete Privacy Preservation System for Data Mining Using Function Approximation." J. Web Eng. 16, no. 3&4 (2017): 278-293.
[12]. R.Arya and Maya L Pai on Rainfall Prediction Using an Optimised Genetic-Artificial Neural Network Model in 2018 International Journal of Pure and Applied Mathematics.
[13]. R. Venkata Ramana, B. Krishna, S. R. Kumar, and N. G. Pandey, "Monthly Rainfall Prediction Using Wavelet Neural Network Analysis," Water Resour. Manag., vol. 27, no. 10, pp. 3697–3711, 2013.
[14]. Santosh Kumar Nanda, Debi Prasad Tripathy, Simanta Kumar Nayak, Subhasis Mohapatra on Prediction of Rainfall in India using Artificial Neural Network (ANN) Models in 2013 Intelligent Systems and Applications.
[15]. Wong K.W., Wong P.M., Gedeon T.D., Fung C.C., Rainfall Prediction Using Soft Computing Technique, Soft Computing 7 (2003), 434-438.