

Design and Implementation of Rainfall Prediction Model using Supervised Machine Learning Data Mining Techniques



Priti Sharma, Deepak Sharma

Abstract: Data mining is a rapidly developing technology that has enriched a lot of field such as business analysis, market analysis, weather forecasting, stock market analysis and many more. It starts with collecting data sets from reliable sources and pre-processing that data. There are some anomalies associated with data collected in large volumes such as outliers, missing values, and duplicated values. Remove these kinds of anomalies is teamed as pre-processing of data. In this paper, collection of weather data and pre-processing it for rainfall prediction model using Rapid Miner tool has been discussed. Also, artificial neural network data mining techniques is used to design a rainfall prediction model. ANN classification techniques is a complex data mining technique results in high accuracy in prediction of rainfall.

Keywords: Data mining, artificial neural network (ANN), Classification techniques, cross validation, Rapid miner

I. INTRODUCTION

Data mining is basically related to a process of mining data for the purpose of extracting meaningful information and patterns using various techniques and well-equipped tools. The term data mining is not appropriate for the process knowledge mining or knowledge discovery would be more suitable because knowledge is to be extracted by this process from large data sets. Data mining helps in deep analysis of collected data in different dimension and perspective with the help of various industrious tools and reliable techniques which ultimately leads to a better decision making and predictive analysis. [1] Data mining is a rapidly developing technology that has enriched a lot of field such as business analysis, market analysis, weather forecasting, stock market analysis and many more. [3] It starts with collecting data sets from reliable sources and pre-processing that data. There are some anomalies associated with data collected in large volumes such as outliers, missing values, and duplicated values. Remove these kind of anomalies is teamed as pre-processing of data. In this paper, collection of weather data and pre-processing it for rainfall prediction model using

Rapid Miner tool has been discussed. [5]

Data mining task broadly divided into two categories which are descriptive data mining and predictive data mining. Descriptive data mining tasks are used to analyze the data and find out the characteristic property of attribute and relation between them.

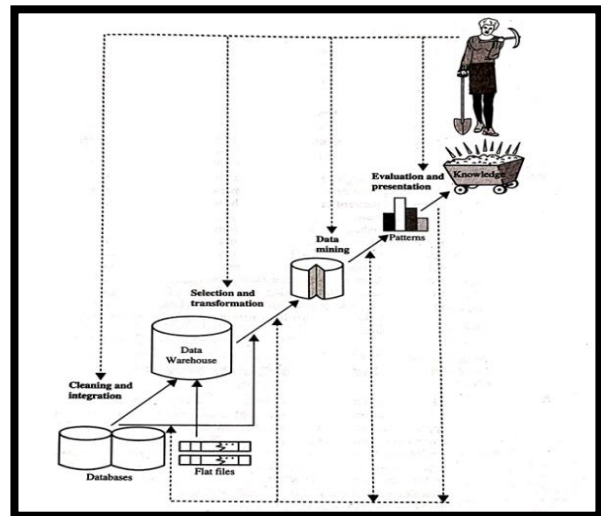


Fig. 1. Process of extracting knowledge

The individual classes are then described in a very concise and well-defined manner called as class descriptions. [12] For class description, data characterization and data discrimination is used. Data characterization congenitally a generalization and summarization of the data of the target class and data discrimination is comparison of target class data with other contrasting classes. [6] The main purpose of doing the analysis and performing several operations is to find out meaningful pattern which one repetitive in nature over a period. Finding out meaningful repetitive patterns ultimately advances towards the revelation of compelling many associations and correlations inside the data. Predictive data mining tasks are used to analyze the given data and perform predictions after performing induction on the current (existing) data. Descriptive data mining tasks includes data characterization, data discrimination, class description, concept description etc. In data mining activities, the data under study is firstly divided into classes or concepts or it is observed that every data entry belongs to a particular class or concepts. Let's take an example of an electronics store, for increasing the profit the owner of the store wants to know which items are bought together so that those items can be placed near each other. [4]

Manuscript received on 12 October 2021 | Revised Manuscript received on 28 October 2021 | Manuscript Accepted on 15 November 2021 | Manuscript published on 30 November 2021.

* Correspondence Author

Deepak Sharma*, Research Scholar, Department of Computer Science and applications, MD University, Rohtak (Haryana), India Email: erdeepaksharmabwn@gmail.com

Dr. Priti Sharma, Assistant Professor, Department of Computer Science and applications, MD University, Rohtak (Haryana), India Email: prish80@yahoo.co.in

© The Authors. Published by Lattice Science Publication (LSP). This is an open access article under the CC-BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Association Rule:-

buys (customer, "Laptop") => buys (customer, "Mouse")
 [Support= 3%, Confidence= 70%]

The above association rule represents a possibility of buying a mouse by a customer who has already brought a laptop. There are two values associated with an association rule i.e. support and confidence. Here in the above association rule "confidence=70%" which shows that there is 70% chance of buying mouse after buying a laptop by a customer. Also "support=3%" means in the data set of transaction under study 3% of the total transaction support the above association rule. [15]

The above association rule is based on a single attribute "buys" so called "single-dimensional association rules".

buys (customer, "Laptop") ^ income (customer, ">50K") => buys (customer, "Speakers")
 [Support= 2%, Confidence= 50%]

The above association rule contains two attributes "buys" and "income" so referred as "Multidimensional association rule". All these tasks discussed are collectively called as Descriptive data mining tasks. The other category is predictive data mining tasks.

Classification and regression techniques are used to assign a class label to an object for which it is not known. In classification, a model is being formulated using the already existing training data. Classification techniques deals with discrete and unordered class labels while regression produce continues values using statistical methods.

Other than classification and regression there are some other predictive data mining tasks such as cluster analysis and outlier analysis. Unlike classification, Cluster analysis doesn't deal with class labels in the beginning. It can be used to generate class labels for a data set. [10]

In clustering the instances are associate in order to maximize the intra class alikeness and minimize the interclass alikeness. Outlier are those instances of the database which do not abide by the data set. Outlier analysis is used to identify these outliers.

In general domain we have two type of prediction problems. One is classification which is to be done with the help of classification data mining techniques such as decision tree and other is numeric prediction in which the model is a continuous valued function in contrast to class label which is to be done by regression analysis (Statistical method).

Classification is the type of data mining which results in a model or classifier representing different classes present in the data under analysis. It helps to understand the data and different classes which are present in the data. [9]

It helps to understand the data and different classes which are present in the data. Some well-known classification techniques which are widely used are Decision tree, Bayesian classifier, rule based classifier, neural network etc. Process of classification is divided into two parts learning part and classification part. In learning part, the training set data has been analyzed with the help of a classification algorithm with the intension to build a classifier. [7] Tuples which are included in training set have already associated class labels which helps the classification algorithm to build association rules by learning from these labels. These tuples are taken from the database by Random sampling method. In a broader perspective, Data mining techniques are divided into

two categories i.e. supervised learning and unsupervised learning. [11]

In supervised learning, the training set is provided to build model in which class label for each tuple is available (classification). On the other hand in case of unsupervised learning, the class labels are not available for tuple in fact of number of classes are not known advance (clustering).

II. METAPHYSICAL INSIGHT

Neural Network is an advanced classification algorithm, which was initially blazed by Neurobiologists. In border way, a neural network is a combinational of three parts i.e. input units, output unit and connections between them.

Every connection has a weight associated to it. Due to presence of these connections this type of learning is called as connectionist learning. Neural network is not simple as decision tree has suffered problem of poor interpretability.

Unlike decision tree, neural network is difficult for humans to interpret especially the hidden layers and weight associated to them and computation process of these weight. Even having problem of poor interpretability, neural network has a huge acceptance due to its high tolerance power towards the noisy data. Unlike previously discussed methods of classification, neural network shows high tolerance towards noisy data and can efficiently recognise and classify the patterns from the noisy data. One more problem associated with neural network other than poor interpretability is that it requires sustainable amount of training time in comparison to other classification algorithm. Best thing about a neural network is that it can also identify those patterns about which it has no previous knowledge from the training. Unlike other classification algorithm, neural network has significant advantage in classifying continuous valued data inputs. Neural networks support parallelism i.e. computer involved in neural network can be enhanced using parallel computing. "Back propagation" neural network algorithm is the most famous ANN algorithm which attain popularity in late 80's.

Multilayer neural network:

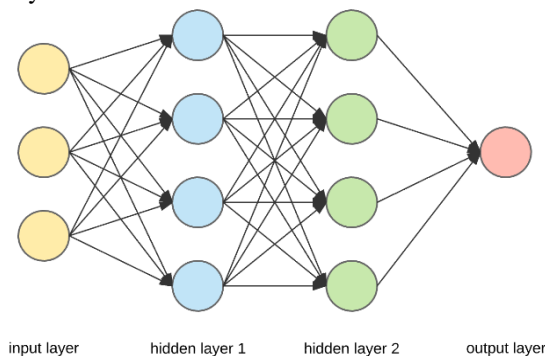


Fig. 2. Multilayer artificial neural Network

Back propagation neural network algorithm is an iterative leaning algorithm in which weights associated to class label predictions are adjusted by learning. The above figure represents a multilayer Feed forward neural network. Multilayer Feed forward neural network is a combination model of one input layer, one output layer and a number of hidden layers. Each layer is a collection of units represented by small circle in the figure.



In the input layer, each units consist of an attribute value is used only to feed the input values. The input layer is only meant to feed the value of attributes as input in the neural network and is not counted as a layer. Only hidden layers and one output layer are count as layers. A neural network can have any number of hidden layer but in general only one hidden layer but in general only one hidden layer is used to minimize the complexity of the model. With increase in hidden layer complexity of the network also increases which ultimately increases the computational requirements. Unit's presents in hidden and output layers are called Neurodes. The multilayer neural network represented in figure is two layer neural network because it has only one hidden layer. If a network has two hidden layer then it is called a three layer neural network. All units of previous layer connects to every node of the next layer and provides value to every node, also all the feed is done in forward direction only from input layer to output layer through hidden layer. This model is called full connected feed forward neural network.

Each edge carries an input from input unit apply a weight on it and gives its value to the associated output unit. From the input unit, each output unit extracts a weighted sum of the output from previous unit. A non-linear activation function is applied to the weight sum of inputs.

Defining a network Topology:

A network topology means the structure of the neural network which is to be defined by the user. Defining a network topology includes processes such as deciding on number of hidden layers, number of units per layer, number of units in the input and output layers etc.

It is “hit and trial” process means getting started with a network topology training that model and performs testing and if found less accurate then changes can be done in the original network topology and weights accordingly.

Neural network is suitable for both classification as well as numeric predictions. It is always well considered to normalize the input data to advance the learning phase.

In general, input values are normalized between 0.0 and 1.0 for continuous valued attributes. For discrete valued attribute, if an attribute has three values A, B, C hen three input units are used to represent three discrete values.

No one can tell about the optimum number of hidden layer for any type of mining as it was already said that defining a network topology is a hit and trail process.

Back propagation algorithm of ANN:

After designing network topology the next step is to train the network. There are many algorithm for training an artificial neural network but back propagation is widely known algorithm in Neural net.

Values of input attribute are feeded in the network with the help of input units and output is predicted. This output can be a numerical value in case of a numeric prediction or a class label in case in case of a classification problem.

This predicted output is compared with the actual output. After comparing with actual output, weights are adjusted in order to reduce the mean- squared error which is to be calculate between and actual value. Weights are adjusted in a backward manner starting from output layer up to the first hidden layer. Weights adjustment is always backward direction and hence the algorithm is called a back propagation. All the weights are initialized to some small random number which ranges from -1.0 to 1.0 or -0.5 to 0.5. Other than input layer units, the hidden layer units and output layer unit have a small value associated with them called as

‘Bias’. Initially these biases are also initialized to small random numbers.

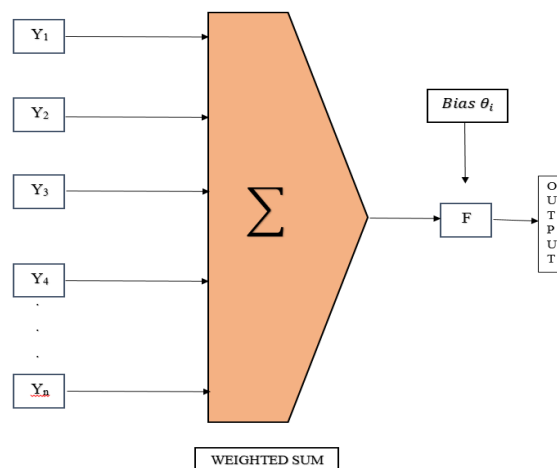


Fig. 3. Back propagation algorithm of ANN

In the above figure a hidden layer or output unit is shown and the process happens on every hidden layer or output layer unit is represented. Output layer unit or hidden layer unit takes input from every unit of the previous layer which is computed as a linear function of the inputs.

$$I_j = \sum_i W_{ij}O_i + \theta_j$$

Where,

I_j : Net input from previous layer unit to the Unit J.

W_{ij} : Weight associated with unit I of previous layer.

θ_j : Bias value of unit j.

O_i : Output of unit i of previous layer.

Output of unit j is calculated using an activation function. A logistic or sigmoid function:

$$O_j = \frac{1}{1 + e^{-I_j}}$$

Output of limit j is computed using a squashing function. It is called a squashing function due to its ability to squash large value to a smaller range i.e. 0 to 1. [12]

The output values for hidden layer units are calculated up to final output layer. After getting all the outputs of the final output layer it’s time to computer the error. Error calculation for unit j of output layer can be done using

$$Err_j = O_j(1 - O_j)(T_j - O_j)$$

Where,

O_j – Actual predicted output of Unit j.

T_j – Target value which was associated with the training tuple originally Errors always computed and propagated in backward direction. For error computations of hidden layers unit j, weighted sum of units to which the hidden layer unit j is connected is considered.

Error computation for a hidden layer unit j is computed by

$$Err_j = O_j(1 - O_j) \sum_k Err_k W_{jk}$$

Here, W_{jk} is the weight of edge connected from j to k and Err_k is error of unit k which is present in next layer.

After calculating error for a unit, Biases and weights are adjusted in accordance to find the correct expected output.



Design and Implementation of Rainfall Prediction Model using Supervised Machine Learning Data Mining Techniques

Change in weight is calculated by

$$\Delta W_{ij} = (l)Err_j O_i$$

$$W_{ij} = W_{ij} + \Delta W_{ij}$$

Here, l stands for learning rate. It is a constant which is used to find the optimum solution. If learning rate l is large in number then it leads to oscillation between in appropriate solution and if l is very low then learning will be very slow. Generally the learning rate (l) has a value between 0.0 and 1.0. Change in biases is calculated by:

$$\Delta \theta_j = (l)Err_j$$

$$\theta_j = \theta_j + \Delta \theta_j$$

This will keep on going up to some pre specified conditions such as:

1. ΔW_{ij} (Change in weight) is very small and reached to some pre specified threshold.
2. Percentage of incorrect classified tuples is very small and reached below to some pre specified threshold.

III. DATA COLLECTION AND PREPROCESSING

Total 35 years of historical meteorological data from year 1988 to 2021 is collected of district “Hisar” of state Haryana. The collected data set is divided into two parts Training set and testing set.

Web Accessible Folder --

<https://www.ncei.noaa.gov/data/global-summary-of-the-day/>
Attribute which are present is the original collected data set and their detailed descriptions are as follows:

Table I. Details of attributes

| S. No | Attribute | Type |
|-------|---------------------------------------|---------|
| 1 | Station Code (STN) | Integer |
| 2 | Date (DATE) | Integer |
| 3 | Temperature (TEMP) | Numeric |
| 4 | Dew point (DEWP) | Numeric |
| 5 | Sea Level Pressure (SLP) | Numeric |
| 6 | Visibility (VISIB) | Real |
| 7 | Wind speed (WDSP) | Numeric |
| 8 | Maximum Sustained Wind sSpeed (MXSPD) | Numeric |
| 9 | Maximum Temperature (MAXT) | Numeric |
| 10 | Minimum Temperature (MINT) | Numeric |
| 11 | Precipitation Amount (PRCP) | Real |

IV. IMPLEMENTATION

Neural Network is an advanced classification algorithm, which was initially blazed by Neurobiologists. In border way, a neural network is a combinational of three parts i.e. input units, output unit and connections between them. Every connection has a weight associated to it. Due to presence of these connections this type of learning is called as connectionist learning. Neural network is not simple as decision tree has suffered problem of poor interpretability. [4] Artificial neural network (ANN) has one input layer, one output layer and a number of hidden layers. Small number of hidden layers leads to less computational cost and computational cost increases with the increase in the hidden layers. Accuracy of the prediction also varies with the increase in hidden layers. The rainfall prediction model of artificial neural network (ANN) classification technique have been implemented using rapid miner data analysis tool. The

designed model is shown in figure 4. After formation of the model the next step is to validate the model using some efficient validation technique. Here in this work we have used cross validation technique to validate all the three models having hidden layer one, two, three respectively.

The models have been validated and compared on five parameters which are Accuracy, Recall, Precision, RMSE (Root mean square error) and computation time. The validation is done using cross validation operator with the help of Rapid miner data analysis tool as shown in figure 9. In cross validation technique, one phase is called training phase which is used to train the model with training data set and other phase is to check the prediction performance called testing phase shown in figure 10.

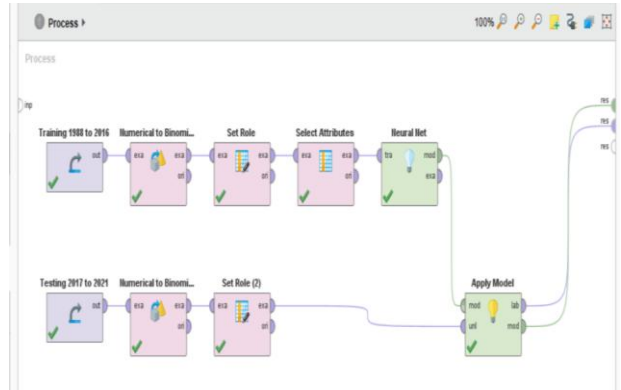


Fig. 4. Rainfall Prediction model using Artificial Neural Network (RPM_ANN)

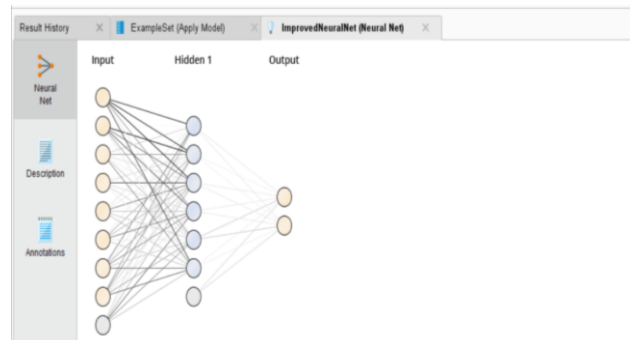


Fig. 5. Artificial Neural Network generated using rapid miner RPM_DT (Hidden layers = 1)

| Row No. | PRCP | prediction(PRCP) | confidence(false) | confidence(true) | STATION | DATE |
|---------|-------|------------------|-------------------|------------------|------------|------------|
| 143 | false | false | 0.984 | 0.016 | 4213109999 | 26-05-2017 |
| 144 | false | false | 0.863 | 0.137 | 4213109999 | 27-05-2017 |
| 145 | false | false | 0.804 | 0.196 | 4213109999 | 28-05-2017 |
| 146 | true | true | 0.348 | 0.652 | 4213109999 | 29-05-2017 |
| 147 | false | false | 0.839 | 0.161 | 4213109999 | 30-05-2017 |
| 148 | true | true | 0.174 | 0.826 | 4213109999 | 31-05-2017 |
| 149 | false | false | 0.958 | 0.042 | 4213109999 | 01-06-2017 |
| 150 | false | false | 0.980 | 0.020 | 4213109999 | 02-06-2017 |
| 151 | false | false | 0.984 | 0.016 | 4213109999 | 03-06-2017 |
| 152 | false | false | 0.984 | 0.016 | 4213109999 | 04-06-2017 |
| 153 | false | false | 0.970 | 0.030 | 4213109999 | 05-06-2017 |
| 154 | false | false | 0.972 | 0.028 | 4213109999 | 06-06-2017 |
| 155 | true | true | 0.381 | 0.619 | 4213109999 | 07-06-2017 |
| 156 | true | true | 0.242 | 0.758 | 4213109999 | 08-06-2017 |

Fig. 6. Predicted value of PRCP (Rainfall) with confidence value (Hidden layers = 1)



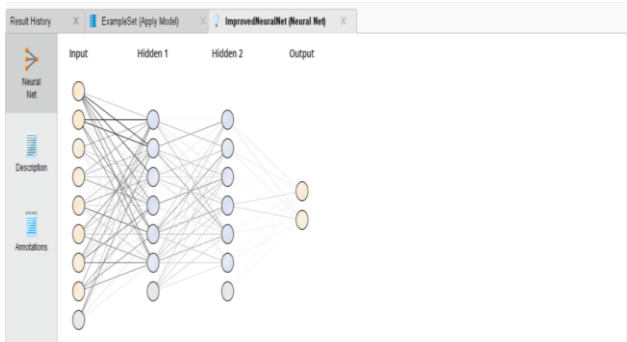


Fig. 7. Artificial Neural Network generated using rapid miner RPM_DT (Hidden layers = 2)

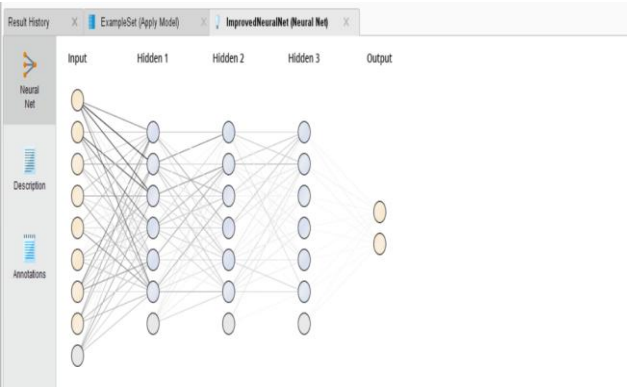


Fig. 8. Artificial Neural Network generated using rapid miner RPM_DT (Hidden layers = 3)

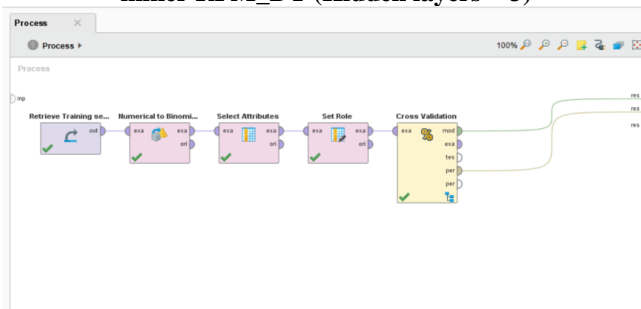


Fig. 9. Cross validation on RPM_ANN

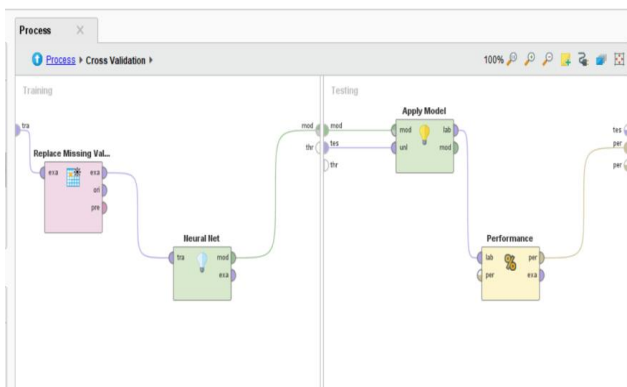


Fig. 10. Training and Testing module of Cross validation on RPM_ANN

V. RESULTS AND ANALYSIS

In this paper, artificial neural network machine learning technique is used to find out the hidden and masked prediction patterns in the 35 years of historical meteorological data. ANN has one input layer, one output layer and a number of hidden layers. Growth in the number

of hidden layers results in the gain of computational complexity of the model. As number of hidden layers increases model formulation time increases. Three models of artificial neural network are represented in figure 5, 7 and 9 having number of hidden layers one, two and three respectively. The accuracy of the prediction varies as well as the other performance parameter with change in hidden layers. Also using data mining in the prediction of rainfall show increase in prediction accuracy as compared to the other methods of predictions.

Table II. Performance analysis and comparison

| MODEL | Precision (%) | RMSE | TIME (sec) |
|------------|---------------|-------|------------|
| RPM_NN_HL1 | 79.35 | 0.289 | 24 |
| RPM_NN_HL2 | 77.76 | 0.293 | 36 |
| RPM_NN_HL3 | 79.03 | 0.291 | 52 |

Table III. Performance analysis and comparison

| MODEL | Accuracy (%) | Recall (%) |
|------------|--------------|------------|
| RPM_NN_HL1 | 88.89% | 71.54 |
| RPM_NN_HL2 | 88.36% | 69.04 |
| RPM_NN_HL3 | 88.74% | 70.22 |



Fig. 11. Comparison (RMSE)

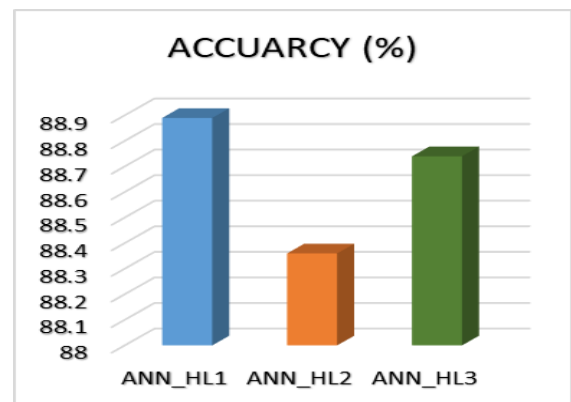


Fig. 12. Comparison (ACCURACY)

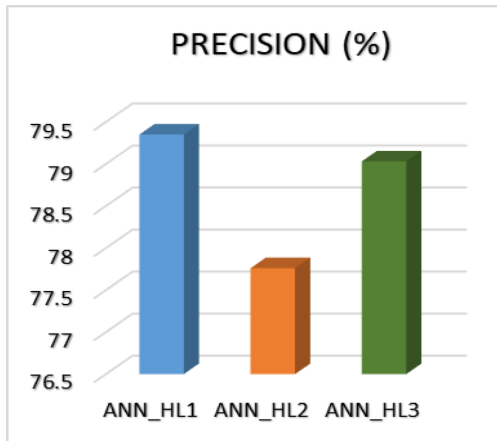


Fig. 13. Comparison (PRECISION)

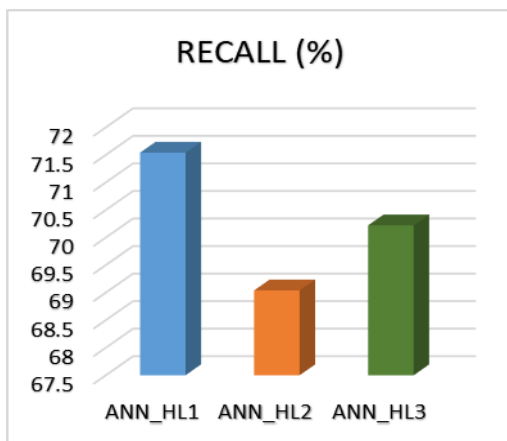


Fig. 14. Comparison (RECALL)

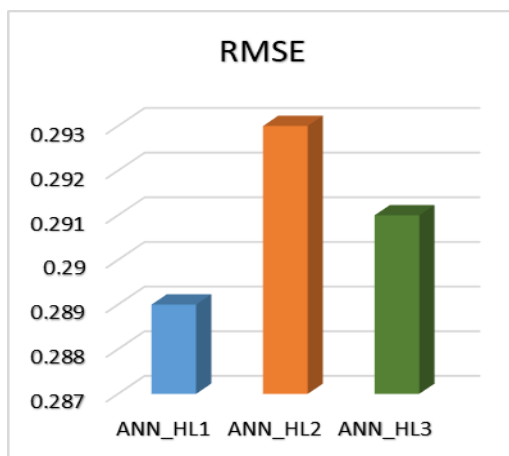


Fig. 15. Comparison (RMSE)

VI. CONCLUSION AND FUTURE SCOPE

Artificial neural network (ANN) is an advanced classification data mining technique which has an advantage in terms of accuracy over simple classification technique such as decision tree. Artificial neural network used in this paper is based on the back propagation algorithm which was discussed in the earlier section of this paper. Artificial neural network is relatively complex in nature as compared to other classification techniques such as decision tree which is less difficult to understand and incorporate. The average accuracy of the artificial neural network rainfall prediction model 88.66 % which is better in case of decision tree 84.31 %. It was clearly observed that with the increase in hidden layers in ANN, computation time also increase to build the

model. Computation taken by the models with hidden layers one, two and three are 24 sec, 36 sec, and 52 sec respectively. Data mining techniques such as logistic regression and linear regression can be used to predict and generate rainfall prediction model.

REFERENCES

1. Tharun V.P, Ramya Prakash, S. Renuga Devi, "Prediction of Rainfall Using Data Mining Techniques", 2nd International Conference on Inventive Communication and Computational Technologies, IEEE-2018. [CrossRef]
2. Abishek.B, R.Priyatharshini, Akash Eswar M, P. Deepika, "Prediction of Effective Rainfall and Crop Water Needs using Data Mining Techniques", International Conference on Technological Innovations in ICT For Agriculture and Rural Development, IEEE-2017. [CrossRef]
3. Fahad Sheikh, S. Karthick, D. Malathi, J. S. Sudarsan, C. Arun, "Analysis of Data Mining Techniques for Weather Prediction", Indian Journal of Science and Technology, Vol 9(38), ISSN (Print): 0974-6846, IJST-2016. [CrossRef]
4. Ramsundram N, Sathya S, Karthikeyan S, "Comparison of Decision Tree Based Rainfall Prediction Model with Data Driven Model Considering Climatic Variables", Irrigation Drainage Sys Eng, an open access journal ISSN: 2168-9768, 2016.
5. Bhaskar Pratap Singh, Pravendra Kumar, Tripti Srivastava, Vijay Kumar Singh, "Estimation of Monsoon Season Rainfall and Sensitivity Analysis Using Artificial Neural Networks", Indian Journal of Ecology (2017) 44 (Special Issue-5): 317-322.
6. Nikhil Sethi, Dr.Kanwal Garg, "Exploiting Data Mining Technique for Rainfall Prediction", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3), 2014, 3982-3984.
7. Chandrasegar Thirumalai, M Lakshmi Deepak, K Sri Harsha, K Chaitanya Krishna, "Heuristic Prediction of Rainfall Using Machine Learning Techniques", International Conference on Trends in Electronics and Informatics - ICEI 2017. [CrossRef]
8. Niketa Gandhi, Owaiz Petkar, Leisa J. Armstrong, "Predicting Rice Crop Yield Using Bayesian Networks", Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016, Jaipur, India.
9. Suvidha Jambekar, Shikha Nema, Zia Saquib, "Prediction of Crop Production in India Using Data Mining Techniques", 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). [CrossRef]
10. Wassamon Phusakulkajorn, Chidchanok Lursinsap, Jack Asavanant, "Wavelet-Transform Based Artificial Neural Network for Daily Rainfall Prediction in Southern Thailand", ISCIT 978-1-4244-4522-6/09 2009 IEEE.
11. N. Tyagi and A. Kumar, "Comparative analysis of backpropagation and RBF neural network on monthly rainfall prediction," Proc. Int. Conf. Inven. Comput. Technol. ICICT 2016, vol. 1, 2017.
12. N. Mishra, H. K. Soni, S. Sharma, and A. K. Upadhyay, "A Comprehensive Survey of Data Mining Techniques on Time Series Data for Rainfall Prediction," J. ICT Res. Appl., vol. 11, no. 2, p. 168, 2017. [CrossRef]
13. Deepak Sharma, Dr. Priti Sharma, "Rain Fall Prediction using Data Mining Techniques with Modernistic Schemes and Well-formed Ideas", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN (Online): 2278-3075, Volume-9 Issue-1, 2019, Page no. 258-263. [CrossRef]
14. Deepak Sharma, "DRI-based implementation for detecting and eliminating cooperative black hole nodes in MANET", International Journal of Recent Technology and Engineering (IJRTE) ISSN (Online): 2277-3878, Vol.8/ Issue 8/ Page No. 8192-8198. [CrossRef]
15. Deepak Sharma & Aaisha Jameel, "Enhancement of security in Flying Ad-hoc Network using a TRUST based routing mechanism" ISSN (Online): 2278-3075, Volume-9 Issue-1, 2019 Page no. 253-257. [CrossRef]
16. Deepak Sharma & Shilpa Singhal, "Detection of fake news on social media using classification data mining techniques", International Journal of Engineering and Advanced Technology (IJTEAT) ISSN (Online): 2249-8958, Volume-9 Issue-1, 2019 Page No. 3132-3138. [CrossRef]

17. Deepak Sharma & Munish Kumar, "Efficient Detection of Black Hole in Mobile Adhoc Networks", International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS) ISSN (Online): 2279-0055, IJETCAS 17-S101/ Special Issue 2017/ Page No 110-115.
18. Deepak Sharma & Lalit Gaur, "Efficient Detection of Brachial Plexus in Ultrasound Images using Machine Learning Algorithms", International Journal of Engineering and Advanced Technology IJEAT), ISSN (Online): 2249-8958, Volume-9 Issue-1, 2019 Page No. 4195-4202. [\[CrossRef\]](#)
19. Deepak Sharma, Dr. Priti Sharma. "Rainfall Prediction Using Classification and Clustering Complex Data Science Models with Geological Significance". International Journal of Computer Science Trends and Technology (IJCTST) V8 (5): Page (39-44) Sep - Oct 2020. ISSN: 2347-8578. www.ijcstjournal.org. Published by Eighth Sense Research Group

AUTHORS PROFILE



Deepak Sharma, has completed his M.tech from C-DAC: Centre for Development of Advanced Computing, Ministry of Communications and Information Technology, Government of India affiliated from Guru Gobind Singh Indraprastha University, Delhi. He is currently pursuing a Ph.D. in Computer Science at M. D. University, Rohtak. His main research areas include Data mining, Mobile Adhoc Network (MANET), wireless sensor network (WSN) and Internet of things (IoT).



Dr. Priti Sharma, MCA, Ph.D. (Computer Science) is working as an Assistant Professor in the Department of Computer Science & Applications, M.D. University, Rohtak. She has published more than 50 publications in various journals/ magazines of national and international repute. She is engaged in teaching and research from the last 12 years. Her area of research includes Data mining, Big data, Software Engineering, Machine Learning.