RESEARCH REVIEW JOURNALS
Volume-04
Issue-05
May-2019

ISSN: 2455-3085 (Online)
RESEARCH REVIEW International Journal of Multidisciplinary
www.rrjournals.com [UGC Listed Journal no : 44945]

A Novel Approach for Weather Prognosis Using Naive Bayes Modeling Technique over Srinagar, J&K, India

*1Azad Abeena & 2Dixit Sarvottam

*1 *Mewar University, Ph. D Scholar, Department of Computer Application, Gangrar, Chittorgarh (Rajasthan), India.*
2 *Mewar University, Professor, Computer Science and Engineering, Gangrar, Chittorgarh (Rajasthan),India.*

ARTICLE DETAILS

Article History
Published Online: 25 May 2019

Keywords
Weather prognosis, rainfall, precipitation prediction, Data Mining, Naive Bayes Classifier, Confusion matrix, precision, ROC

***Corresponding Author**
Email: a.abeena@gmail.com

ABSTRACT

Weather prognosis has become a very important in terms of global economy of the country, agriculture and industries. The accurate prediction of precipitation can sense the heavy rainfall and can provide warnings and information regarding the disasters that could happen in future. Various techniques were proposed and implemented to forecast precipitation, but did not attain much accuracy due to changing weather conditions. In this research work we have tried to implemented Naive Bayes algorithm to build rainfall forecast model which will forecast the precipitation with significant accuracy. Historical weather data set of Srinagar, J&K, India, is gathered from November 2015 to November 2016 from http:///www.wundergrounds.com website. From nine (9) available set of attributes; Temperature, Humidity, sea level pressure, wind speed and Events attributes are considered for this work to achieve better results. Experimental outcomes of various performance measures illustrate that Naive Bayes approach to forecast precipitation has appreciable accurateness and credibility.

1. Introduction

Weather forecast remained the most exhilarating and challenging job for the scientists, meteorologists and government agencies since ages. In this era of information processing and sharing, weather prognosis has become the most challenging and important thing which helps us to forecast the atmosphere of a location [22]. Prognosis of weather accurately helps people to be ready for any natural disasters that could happen in future. Forecasting weather precisely requires knowledge in multiple domains and also involves computing complex mathematical formulas. With the rapid advancement in computer science and technology, scientists have enabled to forecast the weather more accurately. Precise forecast of weather parameters is a complex task due to the dynamic nature of climate. Weather forecasting is the process of collecting data on atmospheric conditions, which records the temperature, humidity, rainfall, wind speed and its direction using high speed computers, wired and wireless sensors, satellites and weather radars [1]. With the progression in information technology complex mathematical and statistical models has enabled us to attain better weather forecasting. The precise prediction of precipitation is the major issue in Indian Summer Monsoon Rainfall (ISMR) and it seeks to be mostly preferred. It is a big issue in the growing economy. Different methods, such as statistical, empirically, and numerical modeling are adapted for predicting the rainfall most accurately. However, soft computing does not yield much over the forecast mechanisms [27]. From the past few decades, data mining techniques are most widely used for weather prediction and have shown a remarkable level of accuracy and applicability in prediction [24] [25] [26] [28]. A number of weather prediction models are proposed by researchers to improve accuracy of the models. Various weather research and forecasting models (WRF), numerical weather prediction models (NWP), and automatic weather stations (AWS) are developed and implemented to forecast weather more accurately [5]. The prediction of precipitation is most complex due to its complex physical properties and changing parameters, which results in precipitation. It seems to be a too noisy and disorganized natural phenomenon [23]. Huge quantity of weather data set is available in metrological centers that contain large volume of weather data which is used for the weather prediction [2] [3] [4]. Data mining is the study of how to determine underlying patterns in the data to help make optimal decisions on computers when the database involved is voluminous, hard to characterize accurately, and constantly changing [23] . It deploys techniques based on machine learning, alongside the conventional methods. These techniques can produce decision or prediction models, based on the large volumes of actual historical data. Thus, they represent true evidence-based decision support [6]. Various classification machine learning algorithms are implemented to forecast the weather conditions. Naive Bayes Networks, Support Vector Machines (SVM), Decision Trees, Artificial Neural Networks (ANN), Fuzzy Logic, Genetic Algorithms are some of the frequently used data mining techniques that are predominantly implemented for weather prognosis..

In this work, we have used classification technique in which the information from the predictors or independent variables is used to categorize the data samples into two or more distinct classes. Based on various atmospheric variables, we have implemented Naïve Bayes approach to build rainfall prediction model which will predict the rainfall with appreciable accuracy. Naive Bayes algorithm has been implemented to predict the Rainfall to class label YES or NO. The performance of the developed rainfall prediction model is done by using various performance measures like Recall, precision, accuracy, ROC etc. The prognosis of accurate weather conditions particularly forecasting rainfall is very vital, as rainfall is directly linked to extreme weather conditions like crop production, aviation, thunderstorms, hurricanes, flash floods and droughts. Thus, predicting accurate and in-time rainfall is quite a big challenge and requirement for an agricultural country like India to have sustainable growth and development.

2. Related Work

In [23], Razeef et al, proposed SALM-NARX rainfall prediction model by training the Self Adaptive Levenberg-Marquardt (Self Adaptive LM) algorithm that selects the weight optimally. Furthermore, LM algorithm is modified with learning rate to make it more accurate for predicting rainfall. The rainfall data obtained from the regions of Jammu and Kashmir and India is collected, for which the proposed SALM-NARX performs the rainfall prediction. The effectiveness of the proposed SALM-NARX is checked with MSE and PRD values and is evaluated to be the best when compared to other existing techniques with least MSE value as 0.008 and PRD value as 1.721%.

In [11], Sharma et al, developed a rainfall prediction model based on Bayesian network. For this research, the authors collected the monthly weather data of 20 years from 1981 to 2000 of 21 stations in Assam, India. K2 algorithm is implemented on the weather data set and conditional probability is found using maximum likelihood approximations. Five different atmospheric parameters viz. Temperature, Cloud cover, Relative humidity, Wind speed and Southern Oscillation Index (SOI) are used. Experimental results showed that temperature is most efficient and wind speed least one. Southern Oscillation Index is also found important in improving the results. Some station got efficiency above 95% whereas other station got satisfactory results.

In [8], Mehmet Tektaş, implemented Adaptive Network Based Fuzzy Inference System (ANFIS) and Auto Regressive Moving Average (ARIMA) models fro weather prognosis. Nine years weather data of Göztepe, İstanbul, Turkey from 2000-2008 is used for this work which showed that ANFIS has superior prognostic potential than ARIMA.

In [7], Liu et al, proposed an improved Naïve Bayes classifier (INCB) model for rainfall prognosis by using genetic algorithms (GAs) for feature selection. Weather data of Hong Kong is used for comparisons of genetic algorithms, C4.5, and INBC with relative frequency or initial probability density. The outcome od experiments showed that the performance of the proposed INBC model is much superior other methods with 90% accuracy rate on the rain/no-rain (Rain) classification problems. The experimental results also revealed that INBC has reasonable predictive capability to predict rainfall range levels with 65% to 70% accuracy rate.

In [10], KavithaRani et al, proposed a novel rainfall prediction model using hybrid classifier implementing artificial bee colony algorithm is collaboration with the genetic algorithm for training performance and predictive capability than Artificial Bee Colony with Neural Network.

In [9], Santhanam et al, proposed Neural network based model for weather prediction. Authors implemented and compared the performances of propagation neural network (BPN) and radical basis functioned neural network (RBF). Whole weather data of ten years is collected from meteorological department, Kanyakumari, Tamil Nadu, India. The experimental results showed that radical basis functioned neural network (RBF) has improved accuracy and is faster and more reliable for weather forecasting. The prognostic accuracy of RBF was 88.49% which makes it more useful for fast real time weather forecasting.

3. Data Collection and preprocessing

Preparing the data set which will be consumed for the data mining task to yield better results is the most time-consuming part of the data mining process. Very rarely data are available in the form required by the data mining algorithms [13]. Poor data quality and selection is the main challenge in weather forecast. For this reason we pre-processed data carefully to obtain accurate and correct prediction results. Weather data of Srinagar, India from November 2015 to November 2016 is collected for this work from http:///www.wundergrounds.com website to develop a prediction model that predicts rainfall based on historical weather data [12]. The obtained raw weather data consists of nine measured attributes which are date, temperature , Dew point , Humidity, sea level pressure, visibility, wind speed, precipitation, Events. Out of these 9 attributes we have used the temperature, Humidity, sea level pressure, wind speed and Events attributes as shown in table 1. In order to achieve better prediction and accuracy, we have ignored less relevant attributes in the data set. In this phase of data mining, unwanted data or noise is removed from the collected data set. Missing values in the data set is another major issue which can produce poor results and prediction. Various techniques are available which are used to fill the missing values. In this work we have implemented mean and modes methods based on existing data. Adding the missing values provides a more complete dataset for the classifiers to be trained on [14].

Table 1
Table 1: Weather data description

Attribute	Type	Description
Temperature	Numerical	Temp is in deg. C
Humidity	Numerical	Humidity in Percentage
Sea Level Pressure	Numerical	Sea Level Pressure in hpa
Windy Speed	Numerical	Wind Speed in Kmph
Events	Numerical	Rainfall in mm

The obtained weather data set contains numerical attribute values; however our proposed model requires categorical values, so we implemented equal width binning method (discretization) for converting the numerical attribute variables into categorical counterparts. The data is divided 'K' intervals of equal size. The width (w) of the interval is given by:

$$w = \frac{\max value - \min value}{k}$$

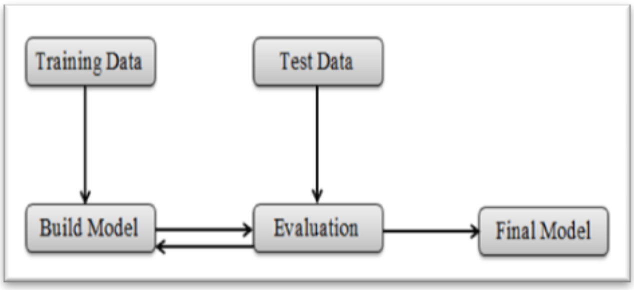
and interval boundaries are:

$$\min value + w, \min value + 2w, \ldots \min value + (k-1)w$$

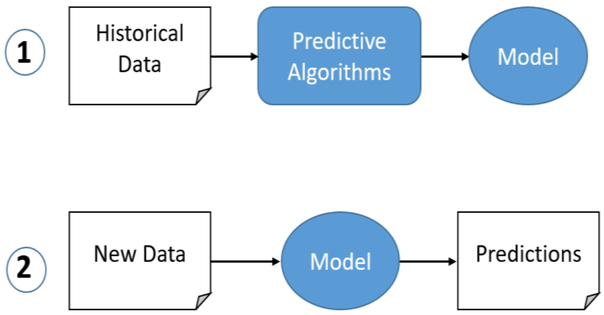
4. Proposed Rainfall Prediction Model

A model is the abstract representation of the data and its relationships in a given data set [26]. Data mining models can be classified into the following categories: Classification, Regression, Neural Network, Association Analysis, Clustering, And Outlier or Anomaly Detection. Each category has many different algorithms; each takes a slightly diverse approach to solve the problem [13]. Classification and Regression approaches are predictive techniques because they predict an outcome variable based on one or more input variables. Prognostic algorithms need a known prior data set to “learn” the model [13]. The percentage of correctly classified instances by the classifier (model) known as classification accuracy gives us the performance measure of the classifier (model). In this work we will build a classifier (model) that will predict the rainfall for future event will appreciable accuracy.

Figure 1
Steps of building Prediction Model.



(a)



(b)

4.1 Steps for Building Classifier (Model)

1) To develop classifier (model), pre-processed and cleaned data set is used with all chosen attributes, including the target class attribute (Event). This data set is called training data set which is used to create a model. Another data set known as the test data set or validation set is needed to check the validity and accuracy of the developed model. The overall data set is divided into training data set and a test data set or test data set is applied externally and is not necessarily the part of training data set. A standard thumb rule is that two-thirds of the data becomes training and one-third to go to the test data set.
2) A classification algorithm is applied to training data set with known class labels which produces a classifier (model).
3) To check the validity of the model, a test data set or a validation data set which is already known is applied to the classifier (model).
4) The developed model can predict the target class attribute as well.

5. Methodology

In this work, we will use a simple Naive-Bayes classification method to model the data set and generalize the relationship between the selected attributes and the target attribute (Event). We will use the prepared weather data set to create the prediction model using Naive-Bayes algorithm which will be then used to predict precipitation. We will also calculate the accuracy of the developed rainfall prediction model by using various accuracy measures.

5.1 Naive Bayes Classifier

Naive Bayes classifier is probabilistic classifier based on Bayes’ theorem with an assumption of independence among the predictors, in other words knowing the value of one attribute does not tell us anything about the other attribute. The Naive Bayesian classifier was first described in [15] in 1973 and then in [16] in 1992. , Naive Bayes is robust to noise and irrelevant attributes and the learnt theories are easy for domain experts to understand. A Naive Bayes classifier structure, when represented as a Bayesian network shows the independence assumption among all features in a data instance. A Naive Bayesian Model is easy to build as no complicated parameter estimation is needed which makes it useful for large data sets. Naive Bayes often outperforms some sophisticated classification algorithms

6. Experimental Study Area And Results

The data set used in this research is weather data of Srinagar, India from November 2015 to November 2016 [12] which we first pre-possessed and cleaned by applying the data mining process model. The experiments are conducted in order to predict the rainfall using Naive Bayes algorithm. In our collected weather data set, EVENT is predicted variable which tells whether it will rain on a particular day or not. By applying Naive Bayes algorithm on the cleaned data set a model is generated which is also known as classifier. The percentage of correctly classified instances by the classifier (model) known as classification accuracy gives us the performance measure of the classifier (model). There are total 540 records in dataset. Each record has 5 attributes including the last attribute defines the class label of the record, whether it will rain or not. The Naive Bayes classification algorithm is applied to data set with known class labels using 10- fold cross validation test which produces a classifier (model). The model produces 411 correctly classified instances which are 76.11% and 129 incorrectly classified instances which are 23.89%. The developed prediction model is used to predict the future events.

7. Performance Evaluation Measures

Performance evaluation of a model (classifier) is an integral part of model development process. Model evaluation helps to find the better model for our data and also reveals how well the chosen model will perform in future. In order to evaluate the performance of developed model, a number of performance measure are used which are based on confusion matrix.

7.1 Confusion Matrix

Confusion matrix is the matrix visualization of outcome of machine learning model. A basic confusion matrix is traditionally arranged as a 2 x 2 matrix for binary classification problem which contains information about actual and predicted classifications done by a classification algorithm as shown in table 2. Performance of such systems is commonly evaluated using the data in the matrix [17].

Table 2
A sample confusion matrix

Predicted by Model		Actual Labels	
		YES	NO
	YES	True-Positive(TP)	False-Positive(FP)
	NO	False-Negative (FN)	True-Negative(TN)

The predicted classes are arranged horizontally in rows and the actual classes are arranged vertically in columns, although sometimes this order is reversed [18]. Basic confusion matrix for a binary or binomial classification can have two classes (say, Yes or No). The four different possible outcomes of a single prediction for binary classification problem are:

1) True-Positive (TP): the number of instances which are actually positive and are also predicted positive by the model.
2) True-Negative (TN): the number of instances which are actually negative and are also predicted negative by the model.
3) False-Positive (FP): the number of instances which are actually negative and are predicted positive by the model.
4) False-Negative (FN): the number of instances which are actually positive and are predicted negative by the model.

Table 3 shows the confusion matrix that is produced after applying Naive Bayes algorithm to Srinagar weather data set.

Table 3
Confusion Matrix of weather data set

Predicted by Model		Actual Labels	
		YES	NO
	YES	255 (TP)	100 (FP)
	NO	29 (FN)	156 (TN)

There are many performance measures for classification algorithms. In this work we have implemented Confusion Matrix of weather data set with following performance measures: Accuracy, Precision, Recall, F-measure, Receiver Operating Characteristic (ROC), root mean square error (RMSE), and mean absolute error (MAE). These performance measures are based on the values of confusion matrix.

1) Sensitivity or Recall: measure the ratio of number of instances which are actually positive and are also predicted positive by the model. It is also known as True Positive Rate and is measured as:

True Positive Rate / Recall/ Sensitivity:

$$= \frac{TP}{TP+FN} = \frac{255}{255+29} = \frac{255}{284} = 0.9$$

2) False-Positive Rate (FPR): measures the ratio of instances which are actually negative and are predicted positive by the model.

False Positive rate $= \frac{FP}{FP+TN} = \frac{100}{100+156} = \frac{100}{256} = 0.4$

3) Precision: is defined as the proportion of cases found that were actually relevant [11]. It is the ratio of instances that are positive cases and are correctly identified by model.

Precision/ predictive Positive Rate:

$$= \frac{TP}{TP+FP} = \frac{255}{255+100} = \frac{255}{355} = 0.7$$

4) F-measure: It is the harmonic mean of precision and recall. F-measure has been widely used in information retrieval [12].

F-measure $= \frac{2 \text{ x Precision x } Recall}{\text{Precision} + \text{Recall}} = \frac{2\times0.7\times0.9}{0.7+0.9} = \frac{1.3}{1.6} = 0.8$

5) Specificity: It is the ability of a classifier to reject all the cases that need to be rejected. In other words, it will have no false positives. In reality, any classifier will select some cases that need to be rejected and thus have some false positives [13].

Specificity is expressed as a ratio (or percentage), calculated as follows:

True Negative rate/Specificity $= \frac{TN}{TN+FP} = \frac{156}{156+100} = \frac{156}{256} = 0.6$

6) Accuracy: It is one the most widely used classification performance metrics. Accuracy is defined as the ability of the classifier to select all cases that need to be selected and reject all cases that need to be rejected. For a classifier with 100% accuracy, this would imply that FN = FP = 0 [13].

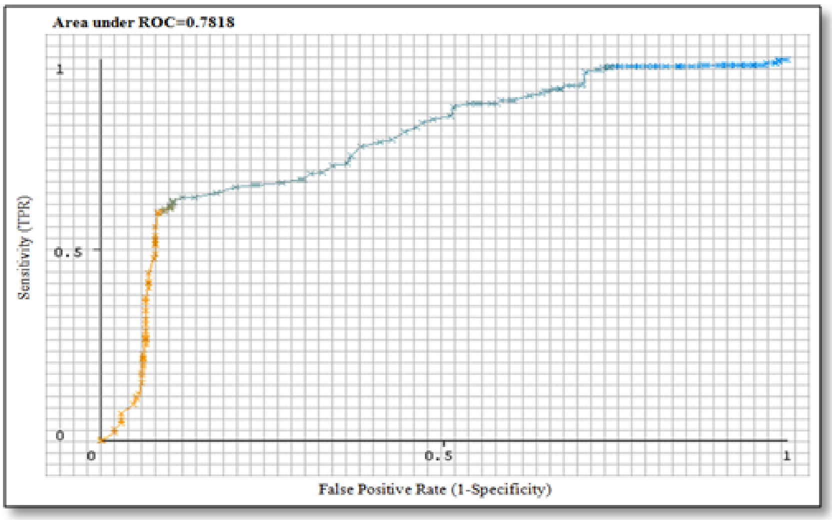
Overall Accuracy $= \frac{TN+TP}{TP+FP+FN+TN} = \frac{255+156}{156+100+29+255} = \frac{411}{540} = 0.8$

Finally, error is simply the complement of accuracy, measured by (1 – accuracy).

7) Receiver Operating Characteristic (ROC) Curve:

ROC curve offer a more robust framework for evaluating classifier performance than the traditional accuracy measure and have become increasingly popular and useful in machine learning [19]. The receiver operating characteristic (ROC) curve is a two-dimensional measure of classification performance. The area under the ROC curve (AUC) is a scalar measure gauging one facet of performance [20]. ROC curve is similar to gain or lift charts. The gain or lift charts are used to compare the classification models while ROC curve reveals how good a model is. ROC curves were originally developed in the field of signal detection [21]. A ROC curve is created by plotting the fraction of true positives (TP rate) versus the fraction of false positives (FP rate). When we generate a table of such values, we can plot the FP rate on the horizontal axis (x-axis) and the TP rate (same as sensitivity or recall) on the vertical axis (y-axix). The FP can also be expressed as (1 – specificity) or TN rate [13].

Figure 2
ROC curve for the Srinagar weather data set

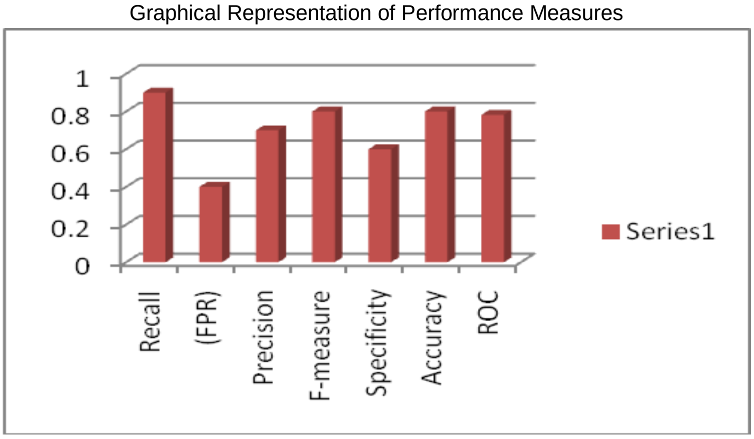


As shown in figure 2 the ROC curve climbs quickly towards top left corner which means model correctly predicts the classes. Area under ROC curve is often used as a measure of quality of the classification model. A random classifier has an area under the curve of 0.5, while area under ROC for a perfect classifier is 1. In our experimental work, Area under ROC curve (AUC) is 0.7818 (78.18%) which is quite reasonable and acceptable for prediction. Figure 3(a) and 3 (b) shows the graphical representation of various performance measures.

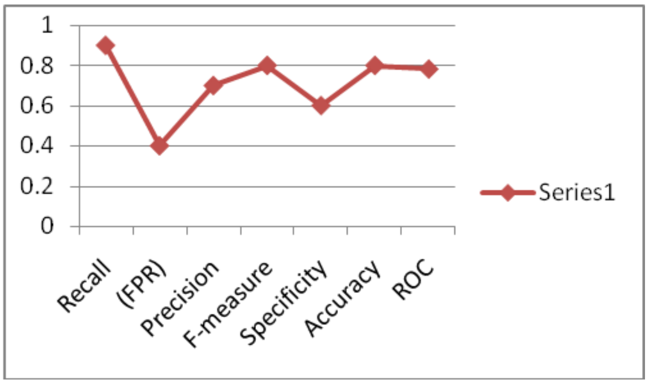
Table 4
Performance Measures

S.No	Performance Measures	Values
1	Sensitivity or Recall	0.9
2	False-Positive Rate (FPR)	0.4
3	Precision	0.7
4	F-measure	0.8
5	True Negative rate/Specificity	0.6
6	Overall Accuracy	0.8
7	ROC	0.7818

Figure 3
Graphical Representation of Performance Measures



(a)



(b)

8. Conclusion and Future Scope

Naive Bayes is a probability-based classification method, which assumes that attributes are conditionally mutually independent given the class label. In this work we have developed the rainfall prediction model (classifier) that predicts the rainfall using Naive Bayes approach. The weather data for this work is taken from Srinagar, India. The developed model's performance is calculated using various performance measures like precision, F-measure, Recall, Accuracy, ROC etc. The prediction model is constructed using 10-fold cross validation test which produces 411 correctly classified instances which are 76.11% and 129 incorrectly classified instances which are 23.89%. The quality of the model is measured by constructing ROC which is 78.18%. The experimental observations show that Naïve Bayes approach to predict rainfall has good level of prediction and is quite satisfactory in rainfall prognosis. Based on our work, future research can be proposed by implementing amalgam of two or more prediction algorithms to develop the model with better accuracy rate.

Acknowledgement

We would like express our gratitude to Mewar University, Gangrar, Chittorgarh, Rajasthan (India) for granting us the valuable opportunities and facilities to work on this study. We are also highly grateful to Director and Head of Department of Computer Applications for providing all the facilities and valuable suggestions during the present research work.

References

[1] Sawaitul, Sanjay D., K. P. Wagh, and P. N. Chatur. "Classification and prediction of future weather by using back propagation algorithm-an approach." International Journal of Emerging Technology and Advanced Engineering 2.1 (2012): 110-113.

[2] R. Sallehuddin, et al., "Forecasting Time Series data using Hybrid Grey Relational Artificail Neural Network and Auto Regressive Integrated Moving Average," Journal of Applied Artificial Intelligence, vol. 23,

[3] M. Tektaş, "Weather forecasting using ANFIS and ARIMA models. A case study for Istanbul," Environmental Research, Engineering and Management, vol. 51, pp. 5-10, 2010.

[4] Binu Thomas, Raju G. and Sonam Wangmo, "A Modified Fuzzy C-Means Algorithm for Natural Data Exploration," World academy of Science, Engineering and Technology 49 2009.

[5] Geetha, A., and G. M. Nasira. "Data mining for meteorological applications: Decision trees for modeling rainfall prediction." Computational Intelligence and Computing Research (ICCIC), 2014 IEEE International Conference on. IEEE, 2014.

[6] Liu, James NK, Bavy NL Li, and Tharam S. Dillon. "An improved naive Bayesian classifier technique coupled with a novel input solution method [rainfall prediction]." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 31.2 (2001): 249-256.

[7] Liu, James NK, Bavy NL Li, and Tharam S. Dillon. "An improved naive Bayesian classifier technique coupled with a novel input solution method [rainfall prediction]." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 31.2 (2001): 249-256.

[8] Tektaş, Mehmet. "Weather forecasting using ANFIS and ARIMA models." Environmental Research, Engineering and Management 51.1 (2010): 5-10.

[9] Santhanam, Tiruvenkadam, and A. C. Subhajini. "An efficient weather forecasting system using radial basis function neural network." Journal of Computer Science 7.7 (2011): 962.

[10] KavithaRani, B., and A. Govardhan. "Effective Features and Hybrid Classifier for Rainfall Prediction." International Journal of Computational Intelligence Systems 7.5 (2014): 937-951.

[11] Sharma, Ashutosh, and Manish Kumar Goyal. "Bayesian network model for monthly rainfall forecast." Research in Computational Intelligence and Communication Networks (ICRCICN), 2015 IEEE International Conference on. IEEE, 2015.

[12]https://www.wunderground.com/history/airport/VISR/2015/11/6/CustomHistory.html?dayend=6&monthend=11&yearend=2016&req_city=&req_state=&req_statename=&reqdb.zip=&reqdb.magic=&reqdb.wmo=

[13] Kotu, Vijay, and Bala Deshpande. Predictive analytics and data mining: concepts and practice with rapidminer. Morgan Kaufmann, 2014.

[14] Ahmed, Bilal. "Predictive capacity of meteorological data: Will it rain tomorrow?." Science and Information Conference (SAI), 2015. IEEE, 2015.

[15] R. O. Duda and P. E. Hart, Pattern classification and scene analysis, John Wiley and Sons, 1973.

[16] P. Langley, W. Iba and K. Thompson, "An analysis of Bayesian Classifiers.," in Proceedings of the Tenth National Conference on Artificial Intelligence, San Jose, CA, 1992.

[17] N. Friedman, D. Geiger and M. Goldszmidt, "Bayesian Network Classifiers.," Machine Learning, vol. 29, pp. 131-163, 1997.

[18] [10] Kohavi, Ron, and Foster Provost. "Glossary of terms." Machine Learning 30.2-3 (1998): 271-274.
[19] Qin, Zeng-Chang. "ROC analysis for predictions made by probabilistic classifiers." Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on. Vol. 5. IEEE, 2005.
[20] Marzban, Caren. "The ROC curve and the area under it as performance measures." Weather and Forecasting 19.6 (2004): 1106-1114.
[21] Green, D. S. (1966). Signal Detection Theory and Psychophysics. New York: John Wiley and Sons.
[22] Razeef Mohd, Muheet Ahmed Butt, Majid Zaman, “Comparative Study of Rainfall Prediction Modeling: A Case Study on Srinagar, J&K, India”, Asian Journal of Computer Science and Technology, (AJCST), Volume 6 Issue 1, September 2018. ISSN 2249- 0701. [UGC Approved Journal].
[23] Razeef Mohd, Muheet Ahmed, Majid Zaman,” SALM-NARX: Self Adaptive LM- based NARX model for the prediction of rainfall”, International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2018) organized by SCAD Institute of Technology, Palladam, Tamilnadu, India and technically sponsored by IEEE.
[24] R. Mohammad, M. Butt Ahmed, and M. Baba Zaman, “Tools for Predictive Analytics : An Overview,” vol. 6, no. 7, pp. 748–750, 2017.
[25] R. M. Shah, M. A. Butt, and M. Z. Baba, “Predictive Analytic Modeling: A Walkthrough,” Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 7, no. 6, pp. 421–426, Jun. 2017.
[26] R. M. Shah, M. A. Butt, and M. Z. Baba “Review of Predictive Analytic Modeling Techniques,” vol. 6, no. 4, pp. 58–62, 2017.
[27] Y. Dash, S. K. Mishra, S. Sahany, and B. K. Panigrahi, “Indian summer monsoon rainfall prediction: A comparison of iterative and non-iterative approaches,” Appl. Soft Comput. J., 2017.
[28] R. Mohammad, M. Butt Ahmed, and M. Baba Zaman,“Predictive Analytics: An Application Perspective, ” vol. XI, issue VIII, Aug. 2017.

