Rainfall Prediction a Novel Approach

Sheetal Rasal
B. E. Computer Engineering
VIVA Institute of Technology
Mumbai, India
sheetalvijayrasal@gmail.com

Bhavesh Rao
B. E. Computer Engineering
VIVA Institute of Technology
Mumbai, India
bhaveshrao180@gmail.com

Saqlain Chaus
B. E. Computer Engineering
VIVA Institute of Technology
Mumbai, India
saqlainchaus@gmail.com

Prof. Umesh Mohite
Computer Engineering
VIVA Institute of Technology
Mumbai, India
umeshmohite@viva-technology.org

***Abstract* —** The proposed framework is a prescient model for precipitation forecast. It is a framework which will foresee the future climate condition in light of the over a wide span of time climate conditions. The proposed framework makes utilization of more number of characteristics than the current framework which makes the forecast more precise. The framework utilizes the classifier calculation that is Naïve Bayes Algorithm for ordering the datasets. The system makes use of the huge dataset of weather condition of various location. The system uses the classifier algorithm called Decision tree classifier for classifying the dataset into a tree format which is used for the purpose of decision making.

***Keywords*-** *Naive Bayesian Algorithm, Machine Learning, Data Extraction, Genetic Algorithm, BPNN.*

*****

I. INTRODUCTION

Rainfall prediction has always been a very challenging task for the metrological department. The proposed system is a rainfall prediction tool which is used to predict the rain based on the data mining technique. The proposed system is a prescient model for precipitation forecast. It is a system which will predict the future climate condition in light of the over a wide span of time climate conditions.

The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Data mining is the process of extracting or mining knowledge from large amount of data. In other words, Data mining is the efficient discovery of valuable, non-obvious information from a large collection of data. It extracts hidden predictive information from large databases, is a powerful new technology with great potential to help in analysis of data and for decision making [8]. Data mining tasks provide a very useful and accurate knowledge in a form of rules, models, and visual graphs.

Weather prediction, in general, is a complex process and challenging task. It requires various parameters to forecast the weather [4]. Rainfall is one of the most important components of the hydrologic cycle and affects surface water resources. Prediction of rainfall has been one of the most challenging problems as it depends on many local and global parameters [2].

II. RELATED WORK

The existing system predicts the rainfall based on the raw dataset which is static and used again and again. The existing system also uses limited number of parameters for the prediction. The problem of missing predictor category assumes the record with the same category has zero probability. Linear regression is the most basic type of regression and commonly used for data extraction and sorting purpose. The proposed system extracts the dataset and then the prediction is carried out. The output is then used to update the existing dataset for the purpose of learning.

III. PROPOSE SYSTEM

The proposed system prepares a model for prediction which uses Naive Bayesian classifier. Pre-processing on the dataset. The proposed system performs the steps of data extracting, error detection and verification, probability prediction of data. This system will be used for rainfall forecasting. This System will be having a tree like structure to classify the input dataset instances. Each sample is assigned a class label by moving in a top down manner from root and testing the condition at the branch. It is based on greedy approach as best split is used at each step to build the tree. Decision trees are easy to work with and even with low domain knowledge.

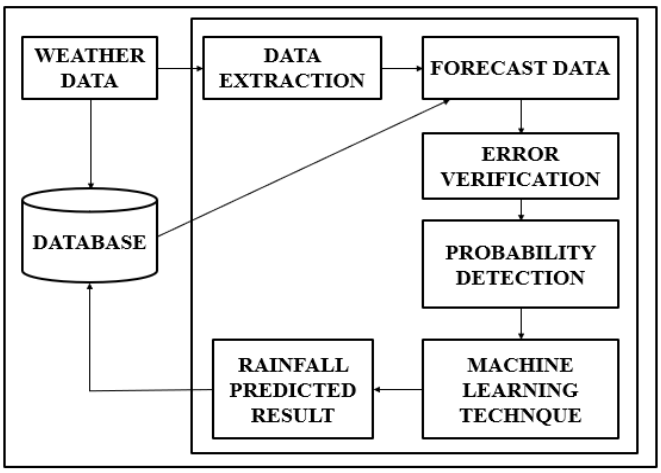

Fig. 1: Block Diagram of Propose System

A. *Weather Data and Data Extraction*

Datasets includes the temperature, rain or snow, wind speed, humidity, pressure, etc. Data extraction is where data is analyzed complete to retrieve a relevant information from database in a specific pattern.

B. *Dynamic Line Rating Algorithm*

Dynamic line ratings (DLR) can increase power transmission capacity of power lines. The use of DLR becomes effective, if DLR is possible to be known beforehand, e.g., a day-ahead for the different hours of the following day.

C. *Genetic Algorithm*

In computer science and operations research, a genetic algorithm (GA) is a metaheuristic inspired by the process of natural selection that belongs to the larger class of evolutionary algorithms (EA). Genetic algorithms are commonly used to generate high-quality solutions to optimization and search problems by relying on bio-inspired operators such as mutation, crossover and selection.

D. *Back Propagation Neural Network*

Backpropagation is a method used in artificial neural networks to calculate the error contribution of each neuron after a batch of data (e.g. in image recognition, multiple images) is processed. This is used by an enveloping optimization algorithm to adjust the weight of each neuron, completing the learning process for that case. Technically, it calculates the gradient of the loss function. It is commonly used in the gradient descent optimization algorithm.

E. *Naïve Bayesian Algorithm*

Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set.

VI. CONCLUSIONS

The proposed system will help to calculate the accurate results using machine learning techniques in real time. With the help of back tracking the system calculate the result using the previous dataset of the rainfall.

Thus, the proposed system implements a weather prediction model using the Decision tree classifier applied on the preprocessed dataset which is being collected for various cities for a specific period. The preprocessing technique includes the data transformation, data selection and reconstruct missing value. The decision tree classification algorithm was used to generate decision trees and rules for classifying weather parameters such as maximum temperature, minimum temperature, rainfall and wind speed etc. in terms of the month and year.

VII. EXPECTED RESULTS

The system will calculate the rainfall by considering the previous data of the rainfall. Even it will check whether the predicted percentage of rainfall is being fallen or not. It will predict the actual number of rainfall fallen and depending on the predicted and previous output it will give the expected result.

The system will help to calculate the accurate results using machine learning techniques in real time. With the help of back tracking, the system calculates the result using the previous dataset of the rainfall.

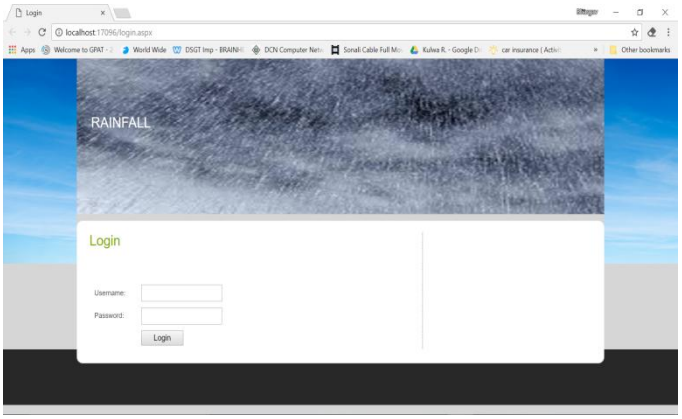

Fig. 2: Login

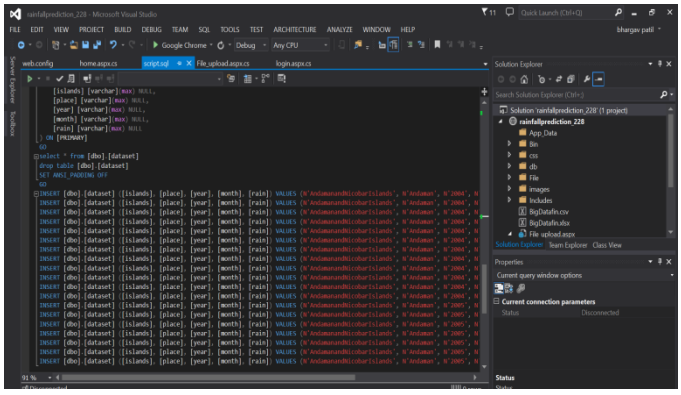

Fig. 3: SQL Script

islands	place	year	month	rain
Andaman	Andaman	2004	January	59.1
Andaman	Andaman	2004	February	109
Andaman	Andaman	2004	March	26
Andaman	Andaman	2004	April	30.2
Andaman	Andaman	2004	May	546.8
Andaman	Andaman	2004	June	456.6
Andaman	Andaman	2004	July	313.9
Andaman	Andaman	2004	August	336.1
Andaman	Andaman	2004	Septembe	270.8
Andaman	Andaman	2004	October	239
Andaman	Andaman	2004	Novembe	177.3
Andaman	Andaman	2004	December	2.2
Andaman	Andaman	2005	January	2.7
Andaman	Andaman	2005	February	0
Andaman	Andaman	2005	March	3.6
Andaman	Andaman	2005	April	36.7
Andaman	Andaman	2005	May	220
Andaman	Andaman	2005	June	482.7
Andaman	Andaman	2005	July	400.5

Fig. 4: Database

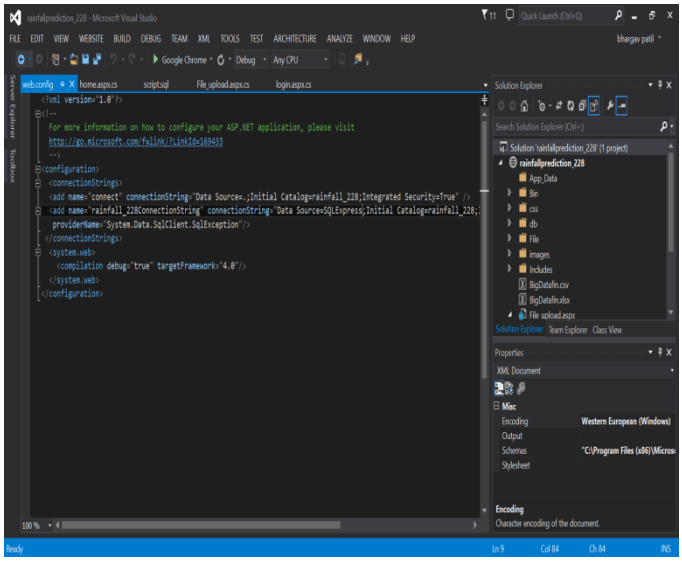

Fig. 5: Web Connection

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to the teaching faculty at Viva Institute of Technology whose timely inputs and suggestions, helped in the completion of the project. We would also like to thanks the Library and Computer Engineering Department of our college for allowing us to carry out our research. Finally, we are thankful for having been given this opportunity to learn something new about the world of technology.

REFERENCES

[1] Nikam, Valmik B., And B. B. Meshram. "Modeling Rainfall Prediction Using Data Mining Method: A Bayesian Approach." In Computational Intelligence, Modelling and Simulation (Cimsim), 2013 Fifth International Conference On, Pp. 132-136. IEEE, 2013.

[2] Sharma, Ashutosh, And Manish Kumar Goyal. "Bayesian Network Model for Monthly Rainfall Forecast." In Research in Computational Intelligence And Communication Networks (Icrcicn), 2015 IEEE International Conference On, Pp. 241-246. IEEE, 2015.

[3] 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, No. 2 (2001): 249-256.

[4] Anandharajan, T. R. V., G. Abhishek Hariharan, K. K. Vignajeth, And R. Jijendiran. "Weather Monitoring Using Artificial Intelligence." In Computational Intelligence and Networks (Cine), 2016 2nd International Conference On, Pp. 106-111. IEEE, 2016.

[5] Liu, Qing, Feng Zhang, Min Liu, And Weiming Shen. "A Fault Prediction Method Based on Modified Genetic Algorithm Using Bp Neural Network Algorithm." In Systems, Man, And Cybernetics (Smc), 2016 IEEE International Conference On, Pp. 004614-004619. IEEE, 2016.

[6] Augusto, Andre Abel, Julio Cesar Stacchini De Souza, And Milton Brown Do Coutto Filho. "Correction of Power Grid Parameters Using Genetic Algorithms." In Intelligent System Application to Power Systems (Isap), 2015 18th International Conference On, Pp. 1-6. IEEE, 2015.

[7] Haviluddin, Rayner Alfred, "A Genetic-Based Backpropagation Neural Network for Forecasting In Time-Series Data", 2015 International Conference On Science In Information Technology (Icsitech).

[8] Tyagi, Nikita, And Ajay Guide Kumar, "Comparative Analysis of Backpropagation And Radial Basis Function Neural Network On Monthly Rainfall Prediction." (2016).

[9] Uski, Sanna. "Dynamic Line Rating Forecastability For Conservative Day-Ahead Line Rating Values." In Industrial Electronics Society, Iecon 2015-41st Annual Conference of The IEEE, Pp. 003738-003742. IEEE, 2015.

[10] Gaikwad, Ganesh P., And V. B. Nikam. "Different Rainfall Prediction Models and General Data Mining Rainfall Prediction Model." Int. J. Eng. Res. Technol 2 (2013): 115-123.

[11] Wahyuni, Ida, And Fitri Utaminingrum. "Error Numerical Analysis for Result of Rainfall Prediction Between Tsukamoto Fis And Hybrid Tsukamoto Fis With Ga." In Advanced Computer Science and Information Systems (Icacsis), 2016 International Conference On, Pp. 365-372. IEEE, 2016.

[12] Luo, Xian, You-Peng Xu, And Jin-Tao Xu. "Regularized Back-Propagation Neural Network for Rainfall-Runoff Modeling." In Network Computing and Information Security (Ncis), 2011 International Conference On, Vol. 2, Pp. 85-88. IEEE, 2011.

[13] Siddharth S. Bhatkande, Roopa G. Hubballi, "Weather Prediction Based on Decision Tree Algorithm Using Data Mining Techniques", International Journal of Advanced

Research in Computer and Communication Engineering Vol. 5, Issue 5, May 2016.

[14] S. Monira Sumi, M. Faisal Zaman, Hideo Hirose, "A Rainfall Forecasting Method Using Machine Learning Models and Its Application to The Fukuoka City Case", Int. J. Appl. Math. Comput. Sci., 2012, Vol. 22, No. 4, 841–854 Doi: 10.2478/V10006-012-0062-1

Alshareef, Almahdi, Azuraliza Abu Bakar, Abdul Razak Hamdan, Sharifah Mastura Syed Abdullah, And Othman Jaafar. "Pattern Discovery Algorithm for Weather Prediction Problem." In Science and Information Conference (Sai), 2015, Pp. 572-577. IEEE, 2015.