International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.072 ISSN: 2581-4621

Downfall Prediction Using Machine Learning

S.Prathap[1], G.Sandhya[2], T.Sai Vineela Reddy[3], A.Sneha[4], M.Venkata tharun[5]
Assistant Professor, Student, Department of CSE, AITS, Tirupati

Abstract

Accurate and timely rainfall prediction can be very helpful to take effective security measures in advance regarding ongoing construction projects, transportation activities, agricultural tasks, flight operations and flood situation, etc. The Downpour prediction helps people to take preventive measures. Two types of rainfall predictions- long term predictions and short-term predictions are performed. Accuracy of rainfall statement has nice importance for countries like India whose economy is basically dependent on agriculture. Rainfall prediction is one of the challenging tasks in weather forecasting. This system is designed to work on long term predictions of rainfall. The main motive behind the development of this model is to predict the amount of rainfall in a particular division or state well in advance.

Keywords: Forecasting, security, projects, prediction, long term, short term, machine learning techniques

1. Introduction

Data which are collected with time stamps in a specific pattern are called time series data [3–5]. This type of time-oriented data is collected with a specific time interval, such as on an hourly, daily, or weekly basis. Time series data can be utilized effectively to make predictions in various areas and domains, including foreign currency rates, stock and security measures in advance for flight operations, agricultural tasks, water reservoir systems, and constructions and transportation activities [2,8,9]. A red alert in advance in the case market trends, energy consumption estimations, and climate change. There are many hardware devices for predicting rainfall by using the weather conditions like temperature, humidity, and pressure. These traditional methods cannot work in an efficient way, so by using machine learning techniques we can produce accurate results. We will just do it by having the historical data of rainfall and other parameters like temperature, pressure, humidity, cloud, wind can predict the rainfall for future seasons. We can apply many techniques like classification, and regression according to the requirements, and also, we can calculate the error between the actual and prediction and also the accuracy. Different techniques produce different accuracies, so it is important to choose the right algorithm and model it according to the requirements.

Machine learning and data mining techniques can be utilized to extract the hidden patterns from historical data in order to forecast the future trend [1,2,5,6]. Machine learning techniques tend to predict future weather conditions by using hidden patterns and relations among the features of historical weather data [2]. The real-time weather data are collected from multiple sensors located in various vital locations of the city. Classification techniques are used in the proposed framework for fusion, including Naïve Bayes (NB). To achieve high accuracy, a fuzzy logic-based layer is included in the proposed framework, which integrates the predictive performance of used classification techniques. These algorithms belong to a supervised class of data mining, in which training is required first with pre-classified data, where classification rules are built and then applied to the input dataset (test data). A weather forecasting website is used to extract the relevant data.

2. Related Work

Rainfall prediction using Machine Learning Techniques - Prediction of rainfall is a very important aspect in the country that leads to the economy of the country too and can help the people by preventing some serious natural disasters. some areas in India are economically dependent on rainfall as agriculture is dependent on rainfall and agriculture is the primary occupation of many states. This

Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.072 ISSN: 2581-4621

prediction helps to identify the crop patterns and proper management of water resources for crops. For this, linear and non-linear models are very commonly used for seasonal rainfall prediction. In this paper, we have performed qualitative and accurate analysis using some classification algorithms like Decision Trees and Logistic regression. Overall, we analyze algorithms that are efficient to be used in order to qualitatively predict rainfall. Logistic regression methods are used. ANN was able to yield an accuracy of 87% with usage of huge volume of data to train the model and it consumes more time. According to Aakash Parmar et.al [16] predicting heavy rainfall is a big concern for meteorological departments since it is so directly linked to the economy and human existence. It is the cause of natural disasters such as floods and droughts that affect people all over the world every year. Rainfall forecasting accuracy is critical for countries like India, whose economy is heavily reliant on agriculture. Statistical techniques fail to provide good accuracy for rainfall forecasting due to the dynamic character of the atmosphere. Because of the nonlinearity of rainfall data, Artificial Neural Networks are a superior technique. for rainfall.

According to Kaushik Dutta et.al [17] rain prediction models based primarily on artificial neural networks have been proposed in India. This paper compares two rainfall forecast methods and determines which one is the most accurate. Statistical and numerical methods are now in use, but they do not operate accurately when there is a non-linear trend. When the complexity of the datasets containing historical rainfall rises, the existing approach fails. To determine the best technique to predict rainfall, researchers will look at both machine learning and neural networks, and the algorithm that delivers the greatest results will be employed in prediction. Weather forecasting is the simplest and quickest way to get a greater outreach.

3. Dataset Attributes

Attribute Name	Attribute Type	Measurement
Temperature	Continuous	Degrees Celsius
Visibility	Continuous	Kilometers
Dew Point Temperature	Continuous	Degrees Celsius
Dew Point Temperature	Continuous	Degrees Celsius
Atmospheric Pressure (sea level)	Continuous	Millimeters of Mercury
Atmospheric Pressure (weather station)	Continuous	Millimeters of Mercury
Relative Humidity	Continuous	Percentage
Pressure Tendency	Continuous	Millimeters of Mercury
Maximum Temperature	Continuous	Degrees Celsius
Minimum Temperature	Continuous	Degrees Celsius

DOI:10.46647/ijetms.2023.v07si01.072 ISSN: 2581-4621

Mean Wind Speed	Continuous	Meters per Second

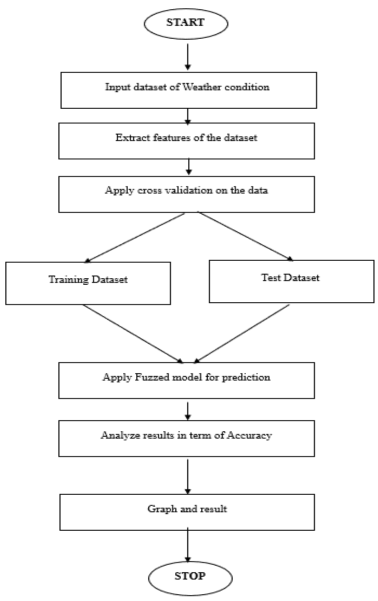
Table 1: Dataset Attributes

4. Methodology

The proposed framework is composed of two layers, namely the training and testing layers, which are further subdivided into multiple stages. In the first stage of the training layer, the weather attributes are extracted from advanced sensors in the smart city. The pre-processing stage involves three activities, namely cleaning, normalization, and splitting. The cleaning process is aimed at removing missing values in the dataset through mean imputation, while normalization brings attribute values within a specific range to aid the classifiers in achieving maximum accuracy. The cleaned and normalized data is then split into two subsets, namely the training data and test data, using a 70:30 ratio of class split rule.

To prepare the data for machine learning algorithms, missing data was analysed and separated by cities to avoid discarding a large amount of data. Eliminating samples with missing data in any of their variables would have resulted in the removal of approximately 50% of the samples. This helps to avoid issues related to variables with large magnitudes that may have a greater influence on the application of machine learning algorithms. For the purpose of testing, 30% of the data was reserved for testing while 70% of the data was used for training. The dependent variable for this study is "Rain", while the other parameters such as Minimum temperature, Maximum temperature, Pressure, Humidity, Wind, and Cloud are independent variables

Figure 1: Flow Diagram



4. Models

4.1 Naïve Bayes The NB classification technique was used for training and testing the rainfall prediction model. The achieved accuracy for training with NB was 90.7%, with a miss rate of 9.3%. The accuracy achieved for testing with NB was also 90.7%, with a miss rate of 9.3%, when the predicted output was compared with the expected output.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.072 ISSN: 2581-4621

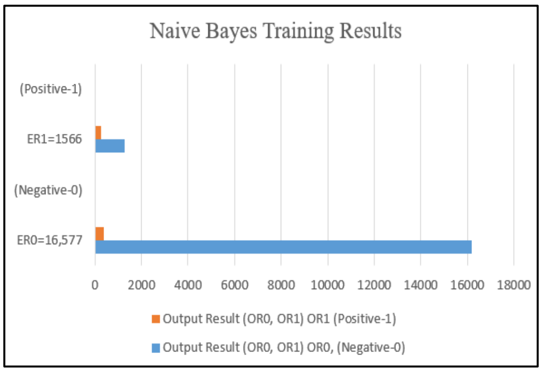
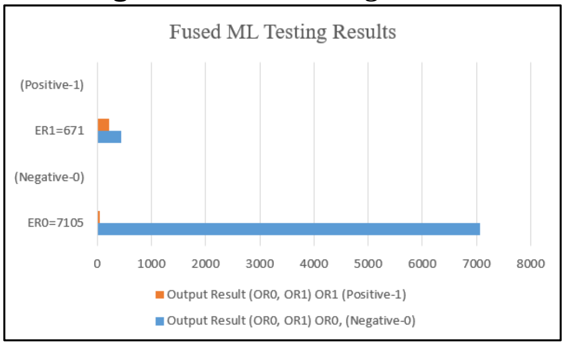


Figure 2: NB Training Results



4.2 Fuzzed Model

The fuzzy system was used to make the ultimate prediction on all the test data. The input for the fuzzy system consisted of the test data and the predictions from the previous classifiers. The testing data is fed into the fuzzy system for final prediction, along with the output class and predictions from the classifiers used. The proposed machine learning-based fuzzy system accurately classified 7063 out of 7105 negative instances and 228 out of 671 positive instances, resulting in an overall accuracy of 94% and a miss rate of 6%. In comparison to the machine learning techniques, the proposed fused technique performed the best, as demonstrated. The given table presents a comparative analysis of the proposed fused model with other published techniques for rainfall prediction in terms of accuracy and miss rate, indicating its superiority. This proposed framework can be integrated into smart cities for more precise rainfall prediction using advanced weather sensors. The sensors will provide continuous weather data for real-time prediction.

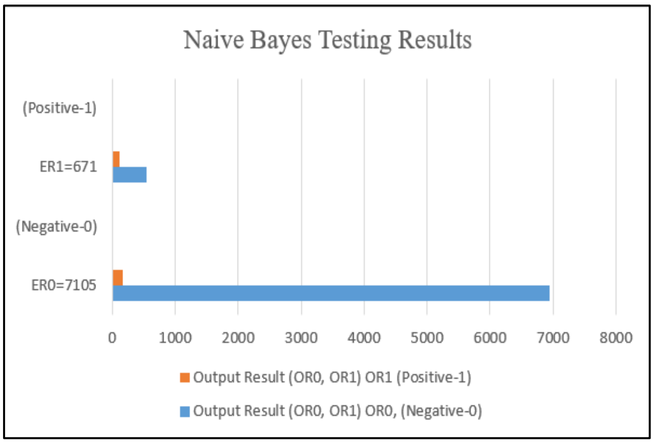


Figure 4: Fuzzed ML Testing Results

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.072 ISSN: 2581-4621

5. Result:

Rainfall prediction is one of the challenging tasks in weather forecasting process. Accurate rainfall prediction is now more difficult than before due to the extreme climate variations. Machine learning techniques can predict rainfall by extracting hidden patterns from historical weather data. The activities of the pre-processing stage, including cleaning and normalization, are performed on the rainfall dataset before the classification stage. To predict, classification techniques are used: Naïve Bayes. Classification technique is optimized iteratively until maximum accuracy is achieved.
Table displays a comparative analysis of the proposed fused machine learning technique with previously published techniques for rainfall prediction in terms of accuracy and miss rate. The proposed model was compared with Naïve Bayes. The results show that the proposed fused model performed better than the other technique.
This proposed machine learning fusion framework can be implemented in smart cities to achieve accurate rainfall predictions. The system will be linked to highly sensitive weather sensors that will provide real-time data for continuous predictions.

ML Algorithm	Task	Spe	LRP	LRN	PPV	NPV	AY
Naïve Bayes	Training	0.98	7.39	0.84	0.41	0.93	0.9
	Testing	0.98	7.31	0.85	0.41	0.93	0.9
Proposed Fuzzed ML	Testing	0.99	57.48	0.66	0.84	0.94	0.94

Table 2: Comparison of ML models

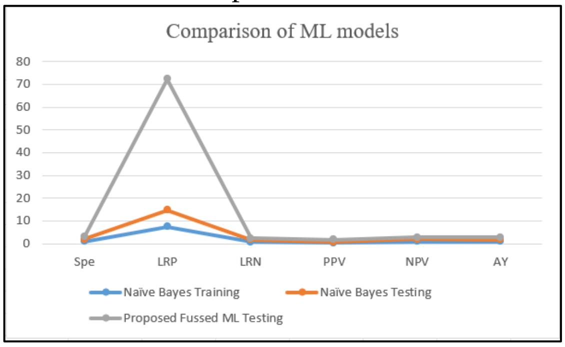


Figure 5: Comparison of ML models

6. Conclusion:

The task of accurately predicting rainfall is a difficult aspect of weather forecasting, but recent developments in machine learning techniques have increased the precision of such systems. This study proposes a novel real-time rainfall prediction system for smart cities that uses a machine learning fusion approach. The framework obtains real-time feature-based weather data from advanced and sensitive weather sensors for accurate predictions. The proposed system integrates the prediction accuracy of supervised machine learning technique, namely Naïve Bayes. The study extracted 12 years of historical weather data for Lahore from a weather forecasting website and performed pre-processing activities like cleaning and normalization to improve classification and prediction accuracy. The goal of this paper is to make the techniques and approaches used in this study accessible to non-experts.
Results show that the proposed system is highly effective compared to other modern techniques, although it is subject to limitations. For example, compromised data or malfunctioning weather sensors could reduce prediction accuracy. A monitoring system and information security system are

Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.072 ISSN: 2581-4621

incorporated into the framework to ensure data integrity until it is used for prediction. Future work will explore the fusion of ensemble machine learning techniques on more diverse datasets, and feature selection techniques to reduce costs, and will incorporate other types of Artificial Neural Networks for weather forecasting. The level of accuracy and prediction highly depends on the data being used as input for classification and prediction. For our weather dataset, it was concluded after analyzing various models of supervised learning that the usage of Fuzzed model has an appreciable level of accuracy and acceptance.

7. References:

1. Ahmad, M.; Aftab, S.; Muhammad, S.S. Machine Learning Techniques for Sentiment Analysis: A Review. Int. J. Multidiscip. Sci. Eng. 2017, 8, 27.
2. Available online: http://ru8.rp5.ru/ (accessed on 30 January 2021).
3. Sivapragasam, C.; Liong, S.-Y.; Aftab, S.; Ahmad, M.; Hameed, N.; Salman, M.; Ali, I.; Nawaz, Z. Rainfall Prediction in Lahore City using Data Mining Techniques. Int. J. Adv. Comput. Sci. Appl. 2018, 9, 254–260.
4. Aftab, S.; Ahmad, M.; Hameed, N.; Salman, M.; Ali, I.; Nawaz, Z. Rainfall Prediction using Data Mining Techniques: A Systematic Literature Review. Int. J. Adv. Comput. Sci. Appl. 2018, 9, 143–150.
5. Nayak, M.A.; Ghosh, S. Prediction of extreme rainfall event using weather pattern recognition and support vector machine classifier. Arch. Meteorol. Geophys. Bioclimatol. Ser. B 2013, 114, 583–603.
6. Yue, T.; Zhang, S.; Zhang, J.; Zhang, B.; Li, R. Variation of representative rainfall time series length for rainwater harvesting modelling in different climatic zones. J. Environ. Manag. 2020, 269, 110731.
7. Mishra, N.; Soni, H.K.; Sharma, S.; Upadhyay, A. A Comprehensive Survey of Data Mining Techniques on Time Series Data for Rainfall Prediction. J. ICT Res. Appl. 2017, 11, 168.
8. Gupta, D.; Ghose, U. A comparative study of classification algorithms for forecasting rainfall. In Proceedings of the 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) Trends and Future Directions, Noida, India, 2–4 September 2015; pp. 1–6.