INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Rainfall Prediction for North Maharashtra, India Using Advanced Machine Learning Models

Husain H Dawoodi[1]*, Manoj P Patil[2]

1 System Analyst, School of Computer Sciences, Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon, Maharashtra, India
2 Associate Professor, School of Computer Sciences, Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon, Maharashtra, India

OPEN ACCESS

Received: 21-11-2022

Accepted: 27-02-2023

Published: 30-03-2023

Citation: Dawoodi HH, Patil MP (2023) Rainfall Prediction for North Maharashtra, India Using Advanced Machine Learning Models. Indian Journal of Science and Technology 16(13): 956-966. https://doi.org/10.17485/IJST/v16i13.2235

* **Corresponding author**.

hhdawoodi@nmu.ac.in

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: The main goal of this research is to design and analyze the performance of intelligent Machine Learning (ML) algorithms such as Support Vector Machines and Naïve Bayes for the prediction of rainfall in the three districts of Jalgaon, Dhule, and Nandurbar in Maharashtra, India. This research attempts to identify the factors that contribute to the occurrence of rainfall at the regional level. **Methods:** The data used in this study are meteorological variable parameters from the previous ten years for the 21 locations in the study area. The predictive performance of the model was validated using several statistical metrics such as precision, accuracy, and f-score. **Findings:** The experimental results demonstrate that both Machine Learning models provided acceptable predictions of rainfall. However, the Support Vector Machine was found to be the best model, with 93% accuracy, for predicting rainfall. In large datasets, prediction results improve significantly. **Novelty:** The Support Vector Machine Model outperforms the Naïve Bayes and Decision Tree models. The study investigates how the meteorological parameters of an atmospheric pressure of 1008 Mb, with winds flowing from the western ghats from the west, and temperatures ranging from 20 to 30 °C affect the decision-making capability of the Support Vector Machine model for drawing the precise hyperplane. Further, Identifying the rain and non-rain situations precisely.

Keywords: Rainfall Prediction; Machine Learning; Support Vector Machines; Naïve Bayes; Meteorological Parameters

1 Introduction

Rainfall prediction is an important application in the field of meteorology. Rainfall directly affects various fields, including agriculture, water resources, hydroelectricity generation, etc., which in turn has a direct impact on the region's economic, political, and social well-being. The total agricultural production of developing countries like India depends about 65% on monsoon rainfall[1]. Further, crop production at the district level has a strong relationship with rainfall and climate variables[2].

Rainfall is usually predicted using numerical weather prediction models [3]. The satellite-based observation techniques and data observed from radars [4] were used for improving the forecast. Due to the complexities of the atmospheric processes that create rainfall, it is often not possible to anticipate rainfall using a statistical or physically based process model. Despite enormous advancements in information technology, modelling precise hydrological forecasts remains a difficulty for meteorological professionals all around the world. Recently, due to the enormous progress in the field of artificial intelligence, machine learning models are widely used in the prediction of rainfall in various countries [5].

The authors in [6] have used Multiple Linear Regression (MLR) and Support Vector Regression (SVR) for the prediction of rainfall. The data from 1901 – 2015 of the sub-divisions in India was collected for prediction. Principal Component Analysis (PCA) was applied to reduce the feature set. The authors have done comparative studies of the machine learning models. However, the study lacks the details of the dataset used, a detailed description of the features, and the methodology for predicting the rainfall, making it difficult to replicate the results by other researchers. The study does not give directions to the readers as to which machine learning techniques are better in its discussion.

Daily rainfall prediction at Bahir Dar City, Ethiopia, by [7], had used Multiple Linear Regression (MLR), Random Forest (RF), and Extreme Gradient Boost (XBoost). Pearson correlation was applied to select relevant environmental variables for input to machine learning models. The past 20 years' data (1999-2018) was used for training and testing the models. The authors concluded that XGBoost is best suited for daily rainfall prediction. The study does not include meteorological datasets for predicting daily rainfall and is limited to Bahir Dar City.

Day-ahead rainfall forecasting is done by [8] using Logistic Regression (LR), Support Vector Machines (SVM), and Random Forest (RF). The study uses eight features for predicting rainfall; however, it does not specify the study region. The data preprocessing techniques and method for carrying out experiments are missing. The authors have not stated the metrics for evaluating the performance of the various machine learning algorithms. The results do not indicate the best machine-learning algorithm for further research by the researchers.

The authors in [9] have investigated rainfall patterns and their variability and predicted the average rainfall for the next five years for the state of Maharashtra in India. Monthly rainfall from 1901-2015 was collected from the Indian Meteorological Department. The authors have applied Artificial Neural Networks (ANN) for prediction. The study predicts the average rainfall for the next five years for the whole Maharashtra state instead of specifying the time and intensity of the rainfall that will occur.

Rainfall prediction for Lahore, Pakistan, is done by [10] using machine learning models like Decision Tree (DT), Support Vector Machines (SVM), Naïve Bayes (NB), and K-nearest Neighbor (KNN). The study shows that the fusion-based machine learning methods outperform all other techniques. The historical weather data from 2005 to 2017 was collected. The study states that predictions cannot be trusted if the data used for the prediction is compromised.

Maharashtra's northern region frequently experiences droughts, extreme rainfall, and hailstorms. Farmers in this region are in great economic distress due to the unpredictable nature of rainfall, which is one of the influencing factors for agricultural production in the region. Further, to fully comprehend climate change at the regional level, researchers must first identify the contributing elements that influence rainfall forecasting and weather patterns at the regional level. Various meteorological agencies conduct weather predictions at a national, state, or district level; however, precise, and reliable weather forecasting at the regional or taluka level is scarce. As a result, rainfall forecasting in North Maharashtra is not precise.

Hence, it is necessary to investigate the rainfall patterns and their variability in the North Maharashtra region to precisely model the machine learning models for accurate rainfall prediction. To date, no studies have been made at the taluka level in the North Maharashtra region to predict rainfall using hourly meteorological observations using machine learning models. Further, there seems to be no research to identify the parameters affecting rainfall. Hence, efforts have been made to identify a good machine learning model to precisely predict the rainfall in the study region and find out the meteorological parameters that affect the occurrence of the rainfall.

The paper is organized as follows: Section 2 describes the study area, the methods used in the paper, and the methodology; section 3 reports results and discusses them; and finally, section 4 highlights the conclusion of the paper.

1.1 Study area and data source

1.1.1 Description of the study area

Maharashtra is one of the biggest states by population, and area-wise, it ranks second in India. The state was formed on May 1st, 1960. Maharashtra state has 36 districts. The study region of North Maharashtra is geographically vast, having three districts: Jalgaon, Dhule, and Nandurbar. The study area is located at latitudes ranging from $20^{\circ}15'30''$N to $22^{\circ}03'00''$N and longitudes ranging from $73^{\circ}47'00''$E to $76^{\circ}16'00''$E. It lies in the north-western part of the Deccan Plateau in central India. The Satpura range lies to the north, and to the west lies the Western Ghats. The study area is bordered to the east by the Berar (Varhad) region and to the south by the Ajanta Hills (Figure 1).

1.1.2 Meteorological data

Hourly meteorological data was collected from [11] for 21 locations in the three districts of Jalgaon, Dhule, and Nandurbar for this study. In modelling, a total of fourteen meteorological parameters, namely moon illumination, sun hours, temperature, wind speed, wind direction, humidity, visibility, pressure, cloud cover, heat index, dew point, wind chill, wind gust, and feels like, were used as input variables, and rainfall data was used as an output variable for generating training and testing datasets. In total, 17,73,912 data samples were collected during the period from January 01, 2009, to December 31, 2018.

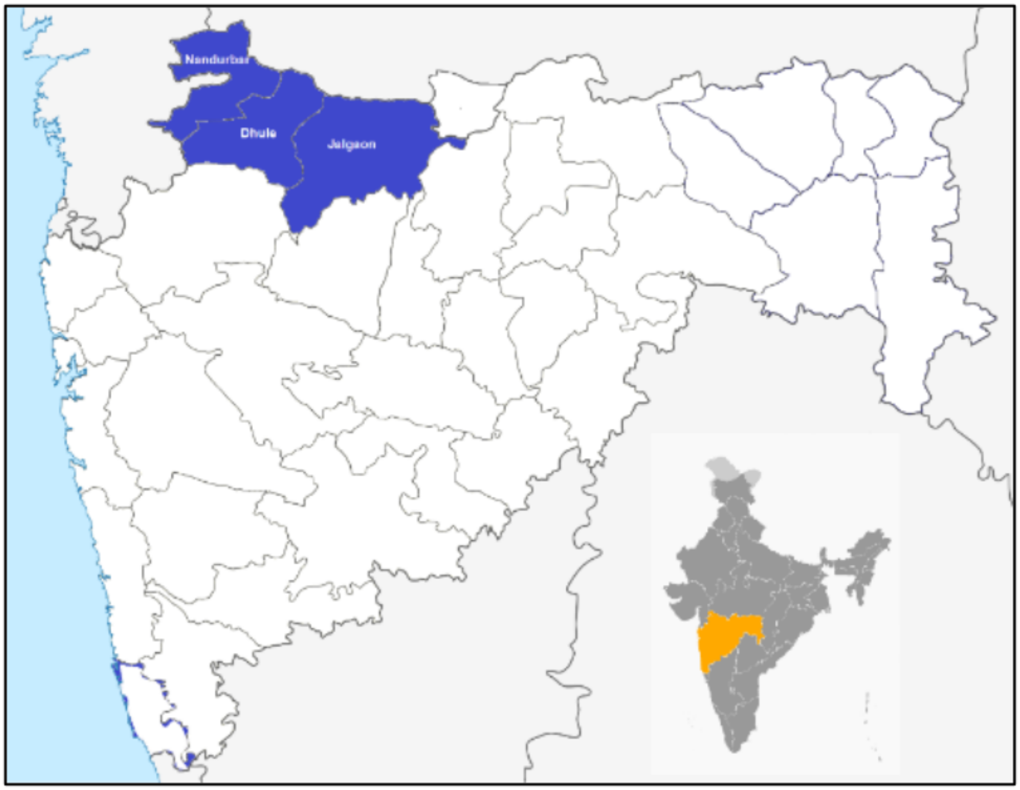


Fig 1. North Maharashtra region having Jalgaon, Dhule, and Nandurbar districts

An initial analysis of the data used is presented in (Table 1). It is observed that the rainfall varies from 0 mm (no rain) to 49.60 mm (rain), with a mean value of 0.03 mm and a standard deviation of 2.04. The temperature varies from 8 °C (winter) to 49 °C (summer), with a mean value of 29.05 °C and a standard deviation of 5.71 °C. Wind speed varies from 0 m/h to 24 m/h with a mean of 7.61 m/h and a standard deviation of 4.01 m/h. Humidity ranges from 7% to 99% with a mean value of 44.77% and a standard deviation of 24.41%. The visibility ranges from 2 miles to 20 miles, with a mean of 9.69 miles and a standard deviation of 1.29 miles.

2 Methodology

2.1 Support Vector Machines (SVM)

A Support Vector Machine is a machine learning algorithm based on supervised learning techniques. It was developed based on the idea of statistical learning theory to perform classification and regression tasks [12]. The main objective of the SVM algorithm is to find a hyperplane that distinctly separates data points in a two-dimensional space by separating them into rainy and non-rainy situations by maximizing the distance between the two of them. In SVM, support vectors are data points that are close to the hyperplane, which determines the position and orientation of the hyperplane. The SVM divides the labeled input data into two classes. A kernel function is used to plot the input space in a higher-dimensional feature space. The linear kernel function separates data linearly.

Table 1. Analysis of meteorological parameters for prediction

Parameters	Unit	Min	Max	Mean	Median	SD
Sun Hours	Hrs.	7.50	13.70	11.70	11.80	1.08
Moonshine	%	0	97	48.32	48	29.87
Temperature	^{o}c	08	49	29.05	29	5.71
Humidity	%	7.00	99.00	44.77	39	24.41
Pressure	mb	996	1020	1008.77	1009	4.54
Visibility	m	2	20	9.69	10	1.29
Wind Speed	m/h	0	24	7.61	7	4.01
Wind Direction	deg	0	360	203.58	255	98.42
Wind Gust	m/h	0	29	10.69	10	4.93
Wind Chill	^{o}c	12	49	29.15	29	5.80
Heat Index	^{o}c	12	55	30.41	30	6.19
Cloud Cover	%	0	100	19.04	5	26.88
Feels Like	^{o}c	12	55	30.34	30	6.27
Dew Point	^{o}c	-7	30	13.25	13	7.93
Precipitation	mm	0.00	49.60	0.03	0	2.04

2.1.1 Naïve Bayes (NB)

Naive Bayes is a simple supervised machine learning algorithm based on Bayes' theorem with the naive assumption of independence between the features to procure results. Bayes' theorem is a simple mathematical formula used for calculating conditional probabilities. Whereas conditional probability is a measure of the likelihood of one event occurring if another event occurs.

$$P(A \mid B) = (P(B \mid A) * P(A))/(P(B)) \tag{1}$$

Where the equation (1) tells how often A happens given that B happens, written P(A|B), is also called posterior probability; when one knows how often B happens given that A happens, written P(B|A), and how likely A is on its own, written P(A), and how likely B is on its own, written P(B).

2.2 Validation criteria

In the present work, Machine Learning models used to predict rainfall occurrence (or probability) are evaluated using the following metrics:

$$f - score = 2 * \frac{precision * recall}{precision + recall} \tag{2}$$

where precision is defined as:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

Further, recall is calculated as:

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

and accuracy is calculated as:

$$Accuracy = \frac{TP + TN}{N} \tag{5}$$

Where TP means true positives, the number of days correctly predicted as a rainy day when it was observed as a rainy day. FP means false positives; the number of days predicted by the model as a rainy day when they were observed as a non-rainy day.

TN means true negatives, the number of days correctly predicted as a non-rainy day when they were observed as a non-rainy day. FN means false negatives, the number of days predicted as a non-rainy day by the model when it was observed as a rainy day, and N is the total number of observations.

Accuracy measures the goodness of the model, and precision measures the probability of the model making a correct prediction. In all these measures, a value of one (1) represents a perfect score, while zero (0) is the minimum value for all these measures. A model that predicts a rainy day when it was a rainy day and misses some of the observed rainy days by predicting a non-rainy day will have high precision but will have a smaller recall value due to the prediction of false negatives. However, a model that predicts a rainy day when it was a rainy day but also predicts a rainy day when it was observed as a non-rainy day will have a higher recall but smaller precision value due to the presence of false positives in the prediction.

Accuracy is a good choice for evaluating an ML model with a balanced dataset, and the f-score metric is useful to compare the model's prediction skills.

2.3 Prediction Modelling

Modelling for the prediction of rainfall using machine learning models was carried out in the following steps (Figure 2), (1) Collection of historical meteorological data, (2) Dataset normalization, (3) Construction of ML models, and (4) Validation of ML models.

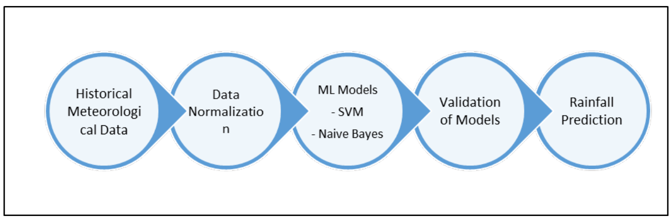


Fig 2. Methodology flowchart of the present study

Step 1 - Historical meteorological data: Hourly data were collected from 21 locations having fourteen meteorological parameters from 2009 to 2018 as described in Section 1.1.2.

Step 2 - Data normalization: A Z-Score normalization approach was used to normalize the hourly data x, collected as given by the equation eq. 6:

$$x^{`} = \frac{x - arithmetic\ mean}{standard\ devination} \quad (6)$$

Where the arithmetic mean and the standard deviation is given by Eq. 7 and 8 respectively:

$$arithmetic\ mean = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (7)$$

$$\text{standard deviation} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \quad (8)$$

Step 3 - Construction of the ML models: The data were randomly split into two parts. 70% of the data in one part was used to create the training dataset. Another part of the remaining 30% of the data was used to create the testing dataset. For the prediction of rainfall, two approaches were used: (i) the Fixed Window size (FW) approach and (ii) the Variable Window size (VW) approach. To construct SVM, a linear kernel function was chosen.

a) The Fixed Window Size (FW) approach: In this approach to the prediction of rainfall, the data for the past year is used to predict the rainfall for the next year. For example, to predict the rainfall for 2010, data of 2009 was used for training purposes.

b) The Variable Window Size (VW) approach: In this approach, the prediction of the rainfall is carried out by providing all the past year's data for training to predict the next year's rainfall. For example, to predict the rainfall for 2016, the data from 2009 to 2015 is used for training. Similarly, for predicting the value for 2015, the values from 2009 to 2014 were used for training.

Step 4 - Validation of ML models: The constructed models were validated using training and testing datasets. Various criteria, namely precision, recall, f-score, and accuracy, were used to validate and evaluate the robustness of the predictive capability of the models.

3 Results and Discussion

The experiments were carried out on an Intel Xeon E5-2667 @ 3.20 GHz 64-core processor with 64 GB of RAM to analyze the performance of the machine learning models for the prediction of rainfall in the North Maharashtra region. Octave, an open-source tool, is used to build the models.

The study of the correlation of the hourly meteorological data collected for the three districts, viz., Jalgaon, Dhule, and Nandurbar, of the North Maharashtra region, shows that the region usually receives seasonal rainfall in August and September. The correlation of various meteorological parameters with the occurrence of rainfall in the North Maharashtra region is shown in Figure 3.

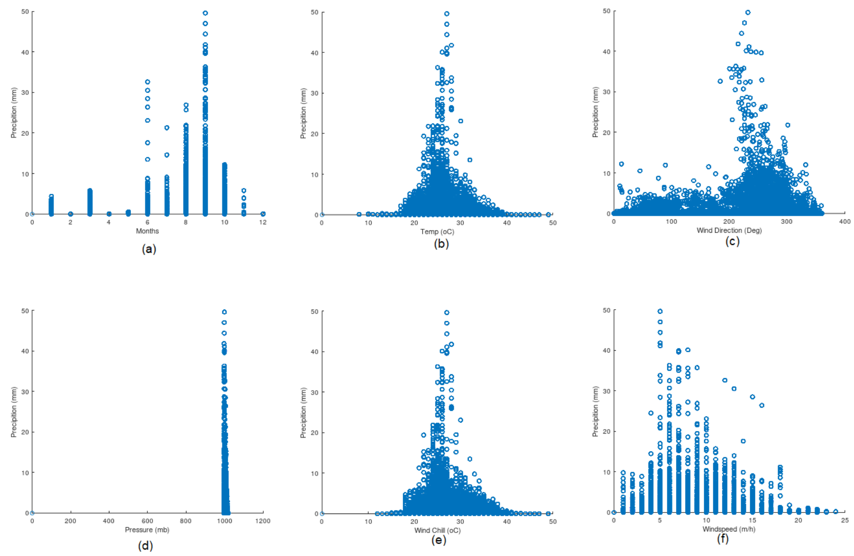


Fig 3. Correlation of meteorological parameters with rainfall (a) Month of Year, (b) Temperature(^{o}C), (c) Wind Direction (Degrees), (d) Pressure (Mb), (e) Wind Chill (^{o}C), (f) Wind speed (m/h)

Further, it is observed that atmospheric pressure plays an important role in the occurrence of rainfall. From Figure 3, it is evident that when there is low air pressure, there is rainfall. Also, temperatures from 20 ^{o}C to 30 ^{o}C are favorable for the occurrence of rainfall in this region. The wind coming from the western ghats, situated to the west of the region, and blowing towards the Satpura Range in the east, cause rainfall.

The prediction evaluations of the two machine learning models, Support Vector Machine and Naïve Bayes, against the observed values for rainfall for the three districts, namely Jalgaon, Dhule, and Nandurbar, are represented in Figures 4, 5 and 6 respectively, for the period of 10 years from 2009 to 2018.

Figure 4 shows that in the study area of the Jalgaon district, the non-rain period typically occurs from January to May, whereas the rain period lies between July and October. Based on the plotted time series rainfall data of the two proposed ML models, SVM exhibits a strong ability to track the behavior of the rain process. The meteorological record indicated that the period from February to April in the year 2015 was rainy. During this period, SVM and NB models also predicted the accurate quantity of rainfall. However, for the year 2015, July to October received a small quantity of rainfall, which is accurately predicted by the proposed ML models.

Figure 5 shows a time series plot of the observed and simulated values of rainy days per year for all ML techniques for the Dhule district. Further, it is observed that the time series of rainfall occurrence is perfectly predicted using machine learning methods, and the prediction is acceptable. Both models could properly predict the extremes observed in 2015 when rainfall occurred from February to April compared to the usual monsoon season, which runs from June to October.

Figure 6 compares the observed and simulated distributions of rainfall intensities in the Nandurbar district. The figure shows that observed values are not well predicted by the ML models for a few years like 2009 and 2017. The models were able to correctly predict the occurrence of rainfall multiple times in a year, as shown in 2005 and 2006, besides the season in which the rainfall normally occurs, thereby proving the model works best even if there are variations in the occurrence of the seasonal rainfall.

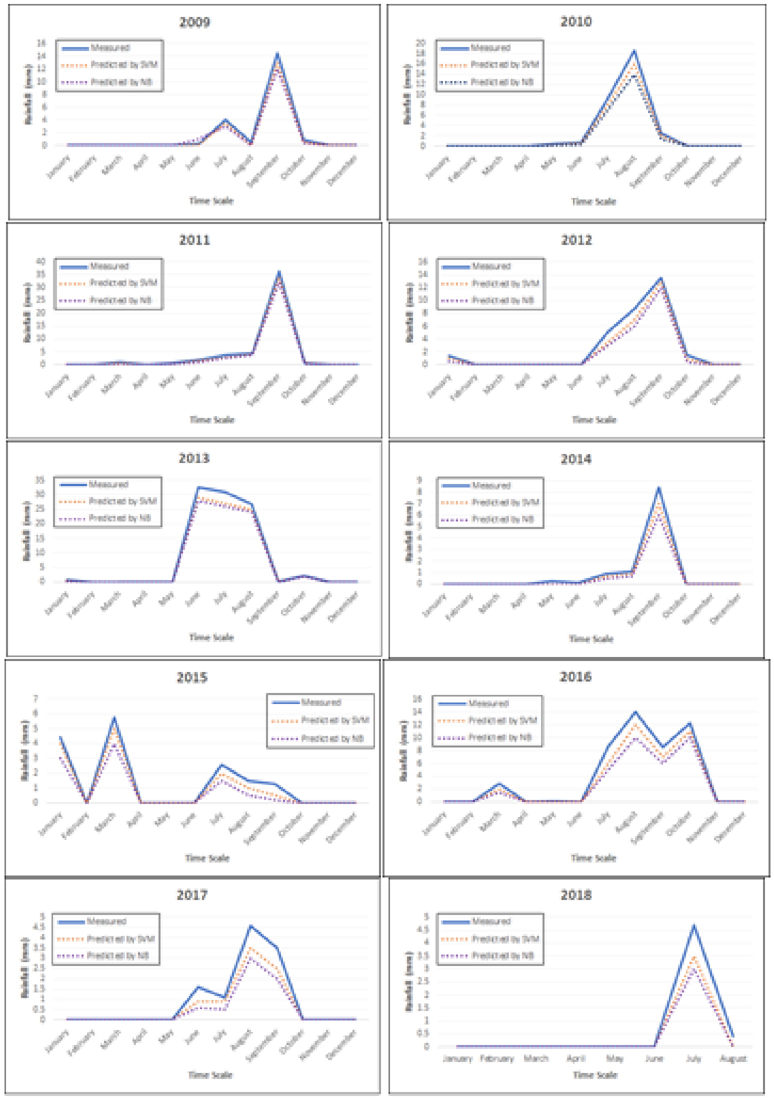


Fig 4. Evaluation of yearly rainfall for 10 years from 2009 to 2018 based on measured data and predicted data of ML models for Jalgaon district

Figure 7 shows that the accuracy value lies between 0.829 and 0.863 for the Naïve Bayes technique for the fixed window size approach and from 0.914 to 0.920 for the variable window size approach. However, the values of accuracy for the Support Vector Machine vary from 0.850 to 0.895 for the fixed window size approach, but the value of accuracy seems to be constant for the variable window size approach. The variance in the value of Naïve Bayes is due to its method of calculating the probability for the rainfall prediction. Whereas the SVM separates the occurrence of rain from no rain with the help of a hyperplane, thereby having a constant accuracy rate.

The precision value for both the FW and VW window size approaches tends to be in the range of 0.7 to 0.9, as precision measures the prediction of the occurrence of rain from the overall observed data. The precision of the SVM algorithm tends to decrease for the VW size approach because, as we increase the size of training data, the fraction of rainy-day occurrences among the overall observed data seems to be lower as compared to non-rainy days. The variability of the rainy days and non-rainy days data for each year plays an important role in the prediction of rainfall. The SVM model can efficiently handle the variance in the yearly data. However, the NB model cannot do so due to its nature of prediction based on probability.

The recall is the ratio of correct rainy days predicted to the total number of rainy days, with Nave Bayes having a value in the range of 0.754 to 0.826 for the fixed window size approach, and the recall value appears to be constant for the variable window

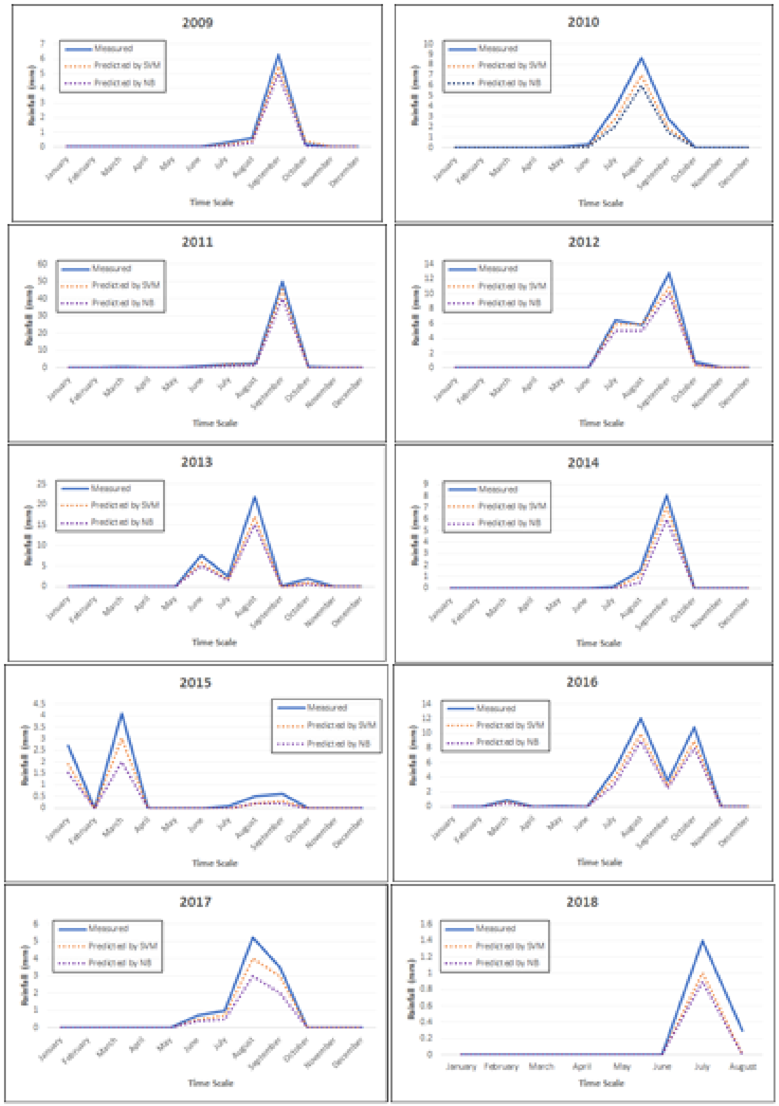


Fig 5. Evaluation of yearly rainfall for 10 years from 2009 to 2018 based on measured data and predicted data of ML models for the Dhule district

approach where the previous 10 years of data are used for training purposes, making the probability of rainy-day prediction quite stable. The recall rate for the SVM algorithm decreases in the variable window size approach as we increase the training data. The probability of predicting a rainy day gets mixed up with the observation of non-rainy days as the number of non-rainy days is greater compared to the number of rainy days.

The F-score values for Naïve Bayes and Support Vector Machines are close to each other, and the f-score seems stable for both techniques. In the fixed window size approach, the number of training records is constant for each year as only the past year's data is used for training purposes. However, the f-score values for the variable window size approach show promising results for the Support Vector Machine. The accuracy of predicting the rainy day or non-rainy day increases with the increase in the no. of observations given for the training. The SVM was able to properly classify between the rainy day when it was a rainy day and the non-rainy day when it was observed, thereby minimizing the type-I and type-II errors in the prediction of the rainfall.

Table 2 shows a comparative analysis of the proposed Support Vector Machine model and Naïve Bayes model with previously published results for rainfall prediction in terms of accuracy. The proposed model is compared with Decision Tree (DT). The

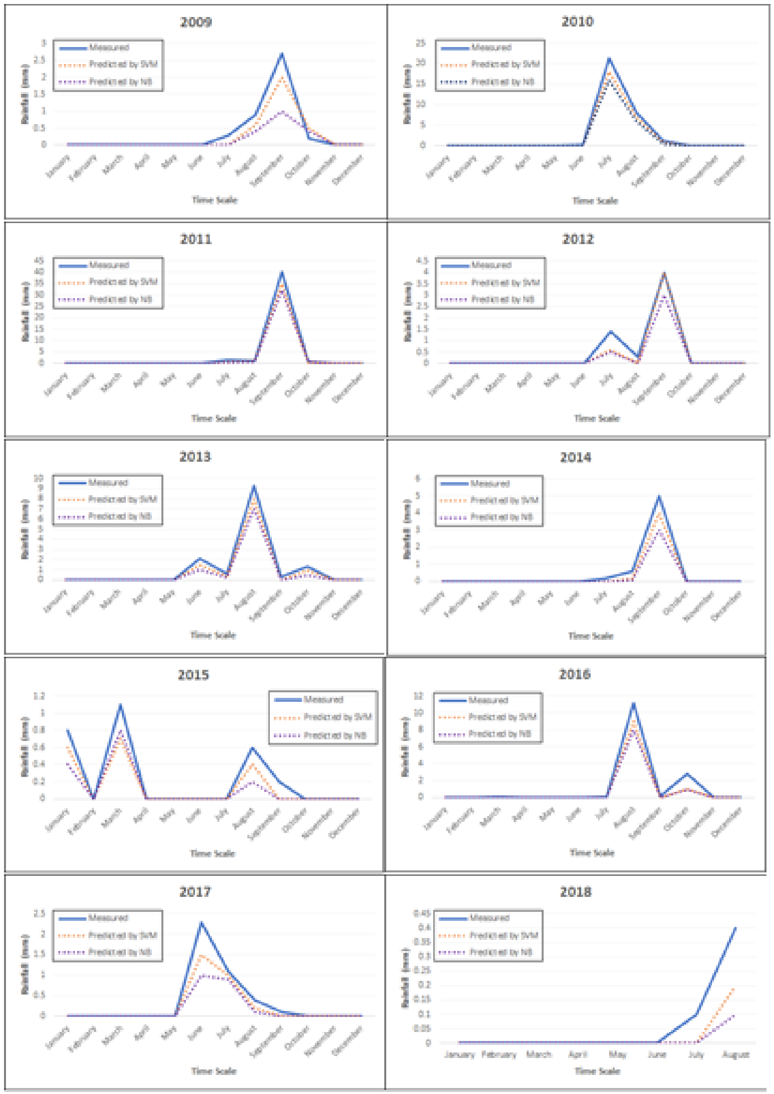
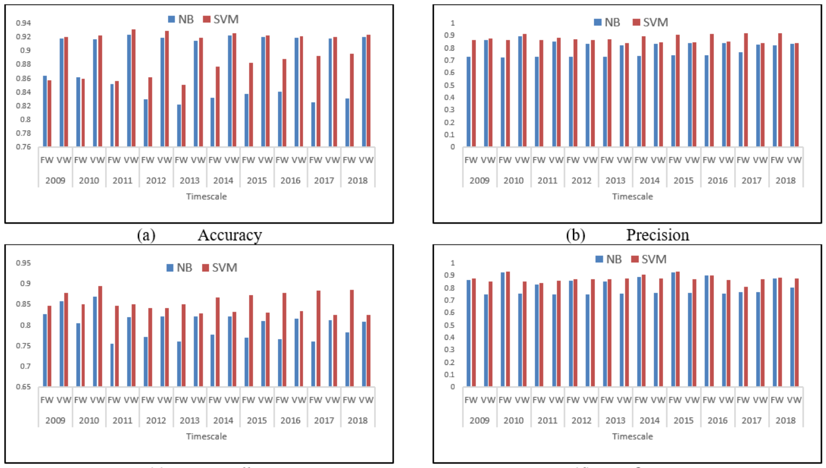


Fig 6. Evaluation of yearly rainfall for 10 years from 2009 to 2018 based on measured data and predicted data of ML models for the Nandurbar district

Table 2. Comparison of proposed Naïve Bayes and Support Vector Machine models with previously published literature

Algorithm	**Accuracy**					
	Yoga IK. et. al. [14]		**Zhang J. et. al.** [13]	**Rahman AU. et. al.** [10]	**Proposed Algorithm**	
Naïve Bayes (NB)	Daily:	67%	—	90%	Fixed Window Size	86%
	Monthly:	88%			Variable Window Size	**92%**
Support Vector Machine (SVM)	—		85.7%	92%	Fixed Window Size	89%
					Variable Window Size	**93%**
Decision Tree	—		85%	91%	—	



(a) Accuracy (b) Precision

(c) Recall (d) f-Score

Fig 7. Evaluation of ML models on criteria such as (a) Accuracy, (b) Precision, (c) Recall, and (d) f-Score

proposed Support Vector Machine model is the best candidate for predicting the rainfall in the North Maharashtra region as compared to the Naïve Bayes and Decision Tree models which have 93% accuracy. The SVM model can predict rainfall with 93% accuracy due to meteorological parameters like pressure and wind direction causing the model to identify the boundary points to precisely draw the hyperplane, distinguishing between rainy and non-rainy days.

4 Conclusion

Rainfall prediction with accuracy is the challenging task of weather prediction. The machine learning algorithms can learn the pattern in the historical data and predict rainfall with accuracy. In the present study, hourly meteorological data for the past ten years from 2009 to 2018 for 21 talukas in the Jalgaon, Dhule, and Nandurbar districts of North Maharashtra was used to predict the rainfall using two Machine Learning (ML) models such as Naive Bayes and Support Vector Machines.

In the present study, it is found that the meteorological parameters having an atmospheric pressure of 1008 Mb with winds flowing from the west and temperatures ranging from 20 to 30 °C affect the decision-making capability of the SVM model for drawing the precise hyperplane.

Upon comparison with the existing Naïve Bayes, Support Vector Machine, and Decision Tree models, the proposed Naïve Bayes and Support Vector Machine models have higher accuracy than the existing approaches. The NB has 92% accuracy, and the SVM has 93% accuracy. As the training samples are increased by increasing the past year's data, the accuracy of the model increases. The higher accuracy in the proposed SVM model is due to parameters like atmospheric pressure, wind speed, and wind direction, which aid in determining the boundary points for accurate rain and no-rain predictions.

Further, this study can be extended in the future by applying an appropriate feature selection technique to make cost-effective predictions. Efforts will be made to incorporate other machine learning techniques, like Artificial Neural Networks.

References

1) Guntukula R. Assessing the impact of climate change on Indian agriculture: Evidence from major crop yields. *Journal of Public Affairs*. 2020;20(1):2040–2040. Available from: https://doi.org/10.1002/pa.2040.
2) Kelkar SM, Kulkarni A, Rao KK. Impact of Climate Variability and Change on Crop Production in Maharashtra, India. *Current Science*. 2020;118(8):1235. Available from: https://www.currentscience.ac.in/Volumes/118/08/1235.pdf.
3) Kimura R. Numerical weather prediction. 2002. Available from: https://doi.org/10.1016/S0167-6105(02)00261-1.
4) Ricciardelli E, Paola FD, Gentile S, Cersosimo A, Cimini D, Gallucci D, et al. Analysis of Livorno Heavy Rainfall Event: Examples of Satellite-Based Observation Techniques in Support of Numerical Weather Prediction. *Remote Sensing*. 2018;10(10):1549. Available from: https://doi.org/10.3390/rs10101549.
5) Hussein EA, Ghaziasgar M, Thron C, Vaccari M, Jafta Y. Rainfall Prediction Using Machine Learning Models: Literature Survey. . Available from: https://doi.org/10.1007/978-3-030-92245-0_4.
6) Mohammed M, Kolapalli R, Golla N, Maturi SS. Prediction of Rainfall Using Machine Learning Techniques. *International Journal of Scientific and Technology Research*. 2020;01(9):3236–3240. Available from: https://www.ijstr.org/final-print/jan2020/Prediction-Of-Rainfall-Using-Machine-Learning-Techniques.pdf.

7) Liyew CM, Melese HA. Machine learning techniques to predict daily rainfall amount. *Journal of Big Data*. 2021;8(1):1–1. Available from: https://doi.org/10.1186/s40537-021-00545-4.
8) Joseph J, Mathai MK, Kottayam K. Rainfall Prediction Using Machine Learning. *Proceedings of the National Conference on Emerging Computer Applications (NCECA)* . 2022;4:265. Available from: https://nceca.in/index/2022NCECA%20(61).pdf.
9) Desai VP, Kamat RK, Oza KS. Rainfall Modeling and Prediction using Neural Networks: A Case Study of Maharashtra. *Disaster Advances*. 2022;15(3):39–43. Available from: https://doi.org/10.25303/1503da3943.
10) Rahman AUU, Abbas S, Gollapalli M, Ahmed R, Aftab S, Ahmad M, et al. Rainfall Prediction System Using Machine Learning Fusion for Smart Cities. *Sensors*. 2022;22(9):3504. Available from: https://doi.org/10.3390/s22093504.
11) World Weather Online. 2020. Available from: https://www.worldweatheronline.com/.
12) Cortes C, Vapnik V. Support-vector networks. 1995. Available from: https://doi.org/10.1007/BF00994018.
13) Zhang J, Qiu X, Li X, Huang Z, Wu M, Dong Y. Support Vector Machine Weather Prediction Technology Based on the Improved Quantum Optimization Algorithm. *Computational Intelligence and Neuroscience*. 2021;2021:1–13. Available from: https://doi.org/10.1155/2021/6653659.
14) Yoga IK, Prasetyowati SS, Sibaroni Y. Prediction and Mapping Rainfall Classification Using Naive Bayes and Simple Kriging. 2022. Available from: https://doi.org/10.29100/jipi.v7i4.3264.