ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 2, February 2023

A Review Paper on Performance Analysis of Machine Learning Methods for Flood Forecasting

[1] Pallavi Bangare, [2] Unnati Bokade*, [3] Mayur Dhamankar, [4] Chetan Dhawale, [5] Dr. S. W. Mohod

[1] Department of Computer Engineering, Bapurao Deshmukh College of Engineering Sevagram, Maharashtra, India.
Corresponding Author Email: [1] unnatibokade2018@gmail.com

Abstract*— Floods are one of the common natural disasters that affect all life. This is due to the rapid renewal of fresh water, which acts outside their natural limits and causes widespread destruction of life and infrastructure. Flooding in an area depends on many factors, including average rainfall, rainfall, proximity to bodies of water, and vegetation. This project's goal is to build a database dependent on the factors and analyses the data using a variety of models: Each classification model's accuracy was tested using a combination of regression model, support vector machines, decision trees, and random forests. The results can be used efficiently by climatic organizations and teams for disaster response inform actions in flood-prone region. In this project, we try to test their success in detecting floods and simulating water flow in a given area.

Index Terms*— Flood Detection, Rainfall , flooding parameters, Forecasting model, SVM algorithm, Naive Bayes algorithm.

I. PROBLEM STATEMENT

In recent decades, algorithms and machine learning techniques have provided promising solutions to many of the problems we face every day. Certainly, machine learning algorithms can be used to identify and predict heavy floods in a specific region. By using machine learning is important because it can process large amounts of data by feeding appropriate machine learning algorithms and applying both supervised and unsupervised learning. The purpose of this project is to prepare and collect data. An important feature of accurate forecasting is information on average rainfall, vegetation cover, rainstorm, and topographic factors that affect flooding in a specific areas over time.

In addition to topographic factors, wetness and air pressure are also factors that affect regional precipitation. Since these factors indirectly lead to the increase of the flood level, these factors should also be considered as parameters. Using live data, officials can learn first-hand the details of areas likely to be affected by floods. Algorithms used to successfully train and build random forest models are support vector machines, decision trees, and linear regression. Since our projects are selected established on accurate flood forecasting in the regions, it was necessary to select the appropriate algorithm as well as to clean and pre-process the collected data. In addition to use different classification model, the goal is to use the HEC-RAS tool to simulate water flow. This can cause flooding. Simulations are never included alongside predictions. This framework aims to achieve this goal. To simulatethe complex mathematical representation of physical floods, machine learning (ML) techniques have significantly contributed to the evolution of forecasting system over the last two decades, leading to better performance and cost-effective solution.

Accurate flood forecasting is necessary to manage environmental and water resource systems, reduce flood hazards, evacuate residents from flooding prone regions, establish insurance rates, and control flood risk. The most popular model in the modern era may be artificial neural network model which are among the content methodology that have attracted a lot of awareness recently for forecasting floods. It is essential to choose important input variables when using ANN modeling to predict floods. attention The CE method uses nonlinear statistical dependence measurement and makes no functional form assumptions. Then successful case conducted a study from the Yangtze River basin in China are used to illustrate how the new CE method can be applied to flood forecasting. The discussion of the range of forecasting unpredictability for flooding showed that it is not always appropriate to apply the normal distribution assumption.

II. MACHINE LEARNING CLASSIFIER FOR FLODD PREDICTION OVERVIEW

Following classifiers we use for flood prediction:

SVM Classifier:

Flood is a devastating phenomenon that threatens human lives, damages houses and has many economic effects. Anticipation is essential to planning and implementing effective mitigation strategies for when and where floods will occur. This study proposes a Support Vector Machine (SVM) approach to river forecasting based on a combination of rain gauge and stream flow measurements from the Don River Basin in the UK. The purpose of this work is to demonstrate the potential of SVM methods for predicting future floods.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 2, February 2023

Naive Bayes Classifier:

Naive Bayes classifier are a set of classifying algorithms depend on Bayes rule. It is not a single algorithm, but a collection of algorithms that share a frequent principle. That means each classification symbols not depend on each other.

Multilayer Perceptron Classifier:

Floods are a powerful and destructive natural disaster, and their renovation is a dangerous undertaking. The development of flood forecasting systems helps to reduce flood-related risks, support policy, reduce loss of life, and reduce property damage. Over the past two decades, neural network approaches have greatly improved flood forecasting by incorporating dynamic formulations of physical floods.

III. PROPOSED METHODOLOGY

With the growth of machine learning, computing is divided into traditional methods and machine learning methods. This section discusses related work in flood forecasting and how machine learning methods outperform traditional methods. Machine learning (ML) methods greatly contribute to the development of predictive systems that provide better performance and more cost-effective solutions. With the growth of machine learning, computing is divided into traditional methods and machine learning methods. In this section, we discuss related work in flood forecasting and how machine learning techniques out perform traditional techniques.

Following are the architecture diagram of flood prediction system:

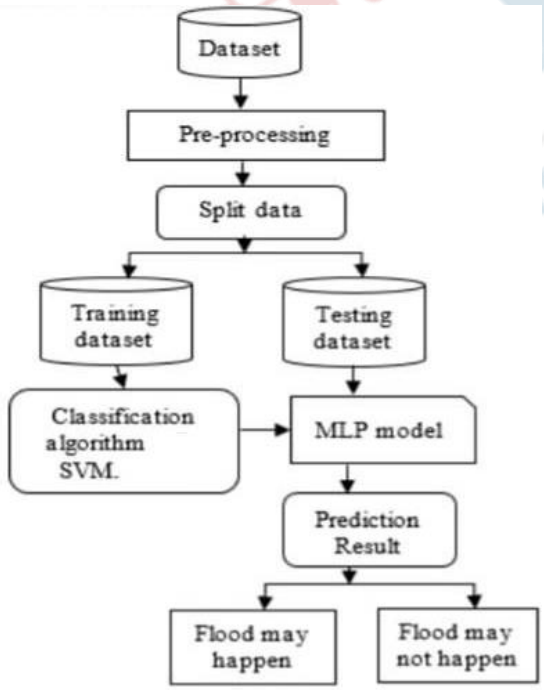

Figure1: Block Diagram of Flood Prediction System

Pre – Processing Data:

There are following steps we will used in our projects:

Step1: In proposed work , we will collect dataset based on rainfall data from kaggle.in website.

Step2: we will apply pre-processing method labelling and labelling of the data.

Step3: Then we will apply training and testing method.

Step4: Then we will apply SVM algorithm for classification of model.

Step5: Then we will fetch live data of rainfall ranges and testing method will applied on this and accuracy will compared.

Step6: Then, we will predict the chances of flood for the designated area using RF algorithm flood may occurred or not. But it requires a lot of memory and the result is not accurate.

Training and Testing Data:

There are two separate dataset sections: a training dataset and a testing dataset, each of which contains 80% training and 20% testing, while models learn from data and made forecasting. Training and testing the precision of model are differentiate by creating a error matrix for algorithm.

IV. RESULT AND ANALYSIS FOR EXISTING SYSTEM

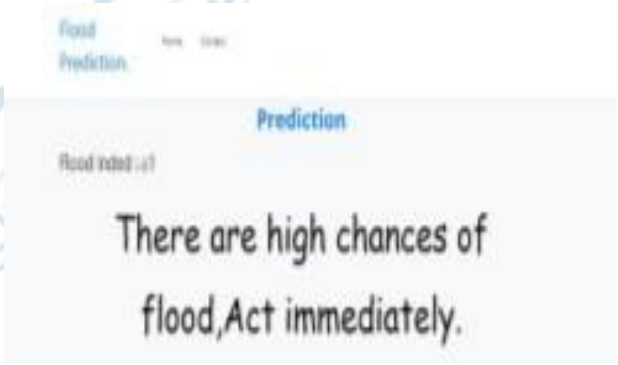

Fig-2 Consequences of predicting the probability of flooding.

It is a binary output where 1 indicates flood and 0 indicates no flood.

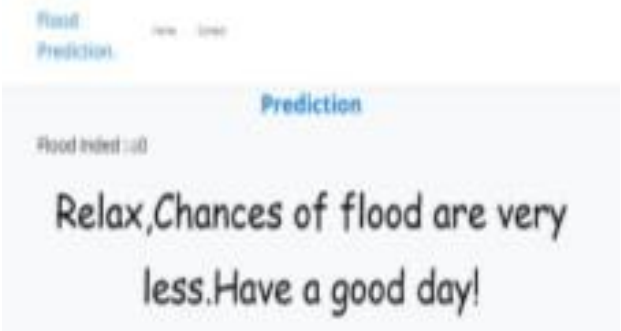

Fig-3 Consequences of not predicting floods

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 2, February 2023

Fig-4 A summary of forecast data interpreted from live data.

V. CONCLUSION

We conclude that a systematic process begins with the process of data cleaning, missing value management, exploratory analysis, and finally building a model for evaluation. The performance and accuracy of the test data are considered and the model with the best performance and accuracy is implemented in the machine learning model. Performance factors such as precision, recall, F1 score, sensitivity and specificity are calculated for each method. In addition, a conclusion matrix including TP, TN, FP and FN is calculated.

REFERENCES:

[1] Ajay Katti, K.V Ashish, Aditya Loke, Kranti Bade, A Pluvial Flood Detection Model Using Machine Learning Techniques and Simulate The Flow of Water, Proceedings of the Fifth International Conference on Communication and Electronics Systems (ICCES 2020) IEEE .

[2] Chinmayee Kinage, Abhishek Kalgutkar, Amruta Parab[7], Sejal Mandora, Sunita Sahu, Machine Learning Based Algorithms for Flood Prediction and Model for Real Time Flood Prediction, 2019 5th International Conference On Computing, Communication, Control And Automation(ICCUBEA) IEEE.

[3] D Han, L.Chan and N Zhu: FLOOD FORECASTING USING SUPPORT VECTOR MACHINES, in 2007 in Journal of Hydroinformatics.

[4] Amir Gosavi, Pinar Ozturk,Kwok-wing Chau: "FLOOD PREDICTION USING MACHINE LEARNING MODELS, 27 October 2018, MDPI Journal.

[5] P.C Nayak, K.P Sudheer, D.M Rangan, K.S Ramasastri: "SHORT TERM FLOOD FORECASTING WITH A NEUROFUZZY MODEL, in 5 April 2005, Water Resources Research, Volume-41.

[6] A Yohan Felix, T. Sasiprabha: FLOOD DETECTION USING GRADIENT BOOST MACHINE LEARNING APPROACH, in December 11-12 2019, International Conference on Computational Intelligence and Knowledge Economy.

[7] Abhishek Bhatt, S.K Ghosh, Anil Kumar: AUTOMATED CHANGE DETECTION IN SATELLITE IMAGES USING MACHINE LEARNING ALGORITHMS FOR DELHI, INDIA.

[8] Supatra Puttinaoporavat, Paramate Horkaew: FLOOD FORECASTING SYSTEM BASED ON INTEGRATED BIG AND CROWDSOURCE DATA BY USING MACHINE LEARNING TECHNIQUES, 10.1109/ \PRP6105021450.