IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 6, June 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006139|

# Novel Approached Flood Forecasting System Using Machine Learning and Big Data

**Prof. Ajay N, Swapna K, Yarramsetty Syona Sree, Namitha B C, Shivadarshini R**

Assistant Professor, Dept. of CSE, S J C Institute of Technology, Chikkaballapur, Bangalore, India

B.E Student, Dept. of CSE, S J C Institute of Technology, Chikkaballapur, Bangalore, India

B.E Student, Dept. of CSE, S J C Institute of Technology, Chikkaballapur, Bangalore, India

B.E Student, Dept. of CSE, S J C Institute of Technology, Chikkaballapur, Bangalore, India

B.E Student, Dept. of CSE, S J C Institute of Technology, Chikkaballapur, Bangalore, India

**ABSTRACT:** Flood is one of the most disruptive natural hazards, responsible for loss of lives and damage to properties. A number of cities are subject to monsoons influences and hence face the disaster almost every year. Early notification of flood incident could benefit the authorities and public to devise both short and long terms preventive measures, to prepare evacuation and rescue mission, and to relieve the flood victims. Geographical locations of affected areas and respective severities, for instances, are among the key determinants in most flood administration. this paper proposes a novel flood forecasting system based on fusing meteorological, hydrological, geospatial, and crowdsource big data in an adaptive machine learning framework. Data intelligence was driven by state of-the-art learning strategies. Subjective and objective evaluations indicated that the developed system was able to forecast flood incidents, happening in specific areas and time frames. It was also later revealed by benchmarking experiments that the system configured with an MLP ANN gave the most effective prediction.

**KEYWORDS:** Flood Forecasting System, Geographical, Meteorological, Hydrological, Geospatial, Crowdsource, MLP ANN, Big Data, Machine Learning.

## I.INTRODUCTION

Natural flood is one of the most recurrent disasters. Unlike stagnant water discharge, occasionally experienced in poorly planned cities, major flood incidents always cause considerable damages to properties and, more often than not, loss of lives. Several Asiancountries, particularly Thailand, are subject to both southwest and northeast monsoons and accordingly facing seasonal deluge almost every year and in most parts of the countries.

Among notable causes, sudden and enduring heavy rain is the most pertinent one in Thailand. Furthermore, overflow from main rivers along shore sides to surrounding basins can greatly spread the damages. Although being located further away from a river, an area with inappropriate land uses are unable to efficiently discharge accumulated precipitation, and hence are inevitably prone to even more frequent floods. Regardless of causes, however, a flood is generally sudden and thus almost formidable for the general public and relevant organization to be adequately prepared for the incident. This is mainly due to the lack of an effective means of anticipating the disaster well in advance.

Despite the recent extensive development of computerized flood forecasting systems, they remainedbased primarily on present precipitation, monitored by rain stations or rain gauges. These facilities are normally owned by a meteorology department or similar organizations. Besides, they are scantly located in a few areas due to costly installation and maintenance. Hence, it is difficult to determine precipitation or predict flood accurately, especially in areas with no such facility. To remedy this issue, precipitation in these areas were typically estimated either by inter- or extrapolation from those with rain stations present. Due to a limited number of these stations and those readings in one area may not be a good representative to others. Therefore, estimated pre imitation was insufficiently accurate to make a realistic forecast. Conventional meteorological readings, e.g., precipitation, temperature, and humidity, etc., took really long

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006139|

time to measure, process, record, and transfer to relevant organizations. Analyses based on past precipitation were known to be associated with several shortcomings.

## II.LITERATURE REVIEW

**Flood Forecasting System Based on Integrated Big and Crowdsource Data by Using Machine LearningTechniques.**This paper proposes a novel flood forecasting system based on fusing meteorological, hydrological, geospatial, and crowdsource big data in an adaptive machine learning framework. Data intelligence was driven by state of-the-art learning strategies. Subjective and objective evaluations indicated that the developed system was able to forecast flood incidents, happening in specific areas and time frames.

**Disadvantages**:- Data considered in this study, such as GLOFAS, were not of intrinsically high spatial resolution.

**Flood Prediction Using Machine Learning Models**

The holistic framework of the IHIP includes five layers (data access, data integration, servicer, functional subsystem, and end-user application) and one database for effectively dealing with flood disasters. The IHIP provides real-time flood-related data, such as rainfall and multi-step-ahead regional flood inundation maps. **Disadvantages:-**Web server deals merely with sending requests and responses between clients and the IHIP.

**Flood Hydrograph Prediction Using Machine Learning Methods**

This study discusses the application of the artificial neural network (ANN), the genetic algorithm (GA), the ant colony optimization (ACO), and the particle swarm optimization (PSO) methods for flood hydrograph predictions.

**Disadvantages**:-Delay should be improved

**Development of Heavy Rain Damage Prediction Model Using Machine Learning Based on Big Data.**

Authors developed a model for the prediction of heavy rain damage based on the big data provided by the Korea Meteorological Administration and machine learning that can maximize the prediction performance of the model, and the model could be used in implementing a proactive disaster management system. However, this study has somelimitations on the number ofdamage data and the use of hydro-meteorological data

## III.METHODOLOGY

This project proposed a novel distributed flood forecasting system, based on integrating meteorological, hydrological, geospatial, and crowd source data. Big data made available by prominent agencies were acquired by means of various cross platform APIs. Forecasting was performed based on these data learned by modern ML strategies. They were decision tree, RF, Naïve Bayes, MLP and RBF ANN, SVM, and fuzzy logics, It was elucidated empirically that the developed system could be used to alert the public and authorities alike of not only a current flood but also future ones. This system also enhanced user experience via responsive graphical interfaces, interoperable on different computing devices including mobiles. This advantage effectively encouraged greater contribution of crowdsource data from the public, enriching data aggregation and hence increasing system accuracy and reliability.

**Objectives**

- To collect the flood forecasting data such as 1. Geospatial, 2. Meteorological and hydrological, obtained from GLOFAS, 3. Crowdsource (or volunteer) data.

- Normalization of data and make a common coordinate frame, a pre-processing need to be done. Data integration preprocessed data.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006139|

• Prediction and evaluating the results using MLP ANN, RF, SVM.

## IV.PROPOSED SYSTEM

To elevate the limitations stated above, the proposed system thus analyzed and designed a flood forecasting system that improved over the current ones. The aspects considered herein were supports of responsive web technology, automation of key processes,and availability and usability of the system. To this end, the proposed system was developed by using both meteorological and hydrological models in forecasting accumulated precipitation from data obtained from TMD big data and GLOFAS, and ML models in forecasting flood situations in given areas. The analyses were made based on meteorological, hydrological, geospatial and crowd sourcing data. Our novel flood forecasting system based on fusing meteorological, hydrological, geospatial, as well as crowd sourcing data, and integrating them into an ML framework. These data were compiled from various big data platforms, by using online application programming interfaces (API).

The forecasting mechanism was driven by a machine learning strategy. To determine the most suitable one for the task, several state-of-the-art MLs, i.e., decision tree, random forest, naïve Bayes, artificial neural networks, support vector machine

**ADVANTAGES**

• Forecasting was performed based on these data learned by modern ML strategies. They were decision tree, RF, Naïve Bayes, MLP and RBF ANN, SVM

• The proposed system also enhanced user experience via responsive graphical interfaces, interoperable on different computing devices including mobiles.

• This advantage effectively encouraged greater contribution of crowd source data from the public, enriching data aggregation and hence increasing system accuracy and reliability

## V. IMPLEMENTATION

Data used in flood forecasting could be categorized into four main groups. They were geospatial,meteorological and hydrological, obtained from GLOFAS , hourly rainfalls prediction from TMD Big Data platform , crowd source (or volunteer) data. They were stored in geo-database and then processed by modern ML strategies.

The key elements in this module were data acquisition and interchange between the system and respective sources and their intelligence via MLs. Thematic data acquisitions were divided into four groups, i.e., meteorological and hydrological data, hourly precipitation data, area specific geospatial data, and crowdsource data.

Unlike a physical based approach, data driven ML does not focus on insights into functional models, but intrinsic relationships between flood relevant factors and corresponding outcomes, learned from the past events. The design of ML models adopted in this framework and their characteristics are described as follow: Firstly, MLP is a configuration of ANN with multiple layers and is suitable for complicate learning tasks.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006139|

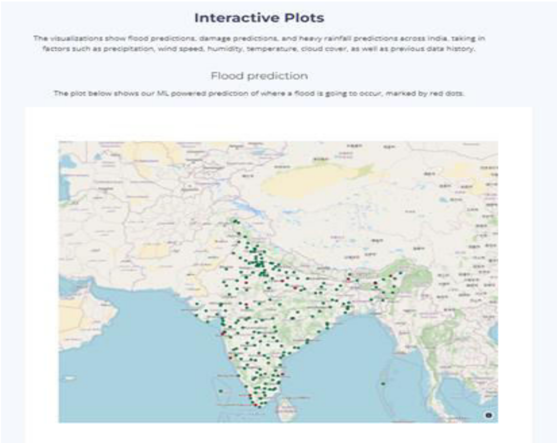


**Figure5.1: Interactive plots**

The above figure shows the Flood Prediction , the plots shows ML powered prediction of where a flood is going to occurred, which are marked by red dots.


**Figure 5.2: Damage Analysis**

The above figure shows the damage analysis which is based on the flood risk prediction. The color scale of the heatmap indicates the extent of predicted monetary damage, measured in INR.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006139|

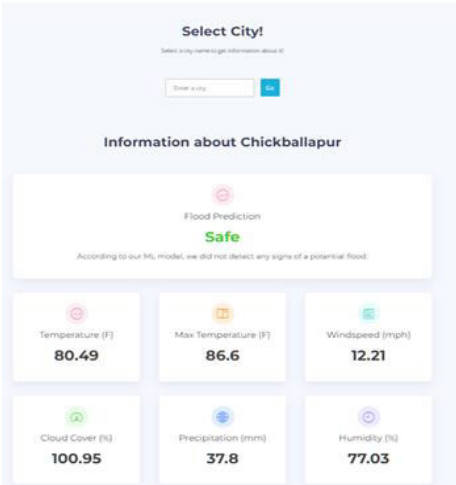


**Figure5.3: Flood Prediction for a city**

The above figure shows the prediction of the particular place in India about the Flood Prediction whether the particular place is safe or not from the floods. It consists of parameters like Temperature, Wind speed, Cloud cover, Precipitation and Humidity.

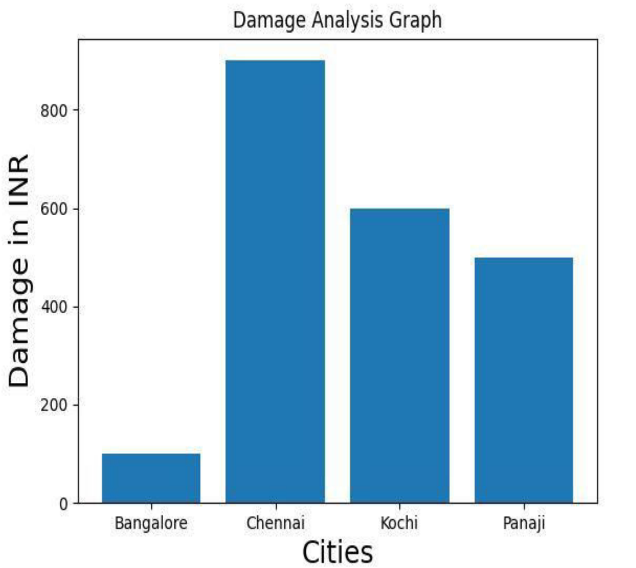


**Figure 5.4: Damage Analysis Graph**

The above figure is the prediction of the damage analysis graph of few cities which will affect from floods.

## VII.CONCLUSION

The main applications of this proposed approach is employing crowd source data, not only in training but also in verification. The system obtained data from https://india.mongabay.com/, to which public users could send notifications of flood situations. It is worth noted that, without dedicate equipment, accurate quantification of relevant data is prohibitive in practice, especially when provided by the public. Instead of requesting an explicit number, the proposed system relied on GUI, by which a participant could provide the experienced factors. Particularly, they could choose, for instance, one out of four different rainfall levels, i.e., none, low, moderate, or heavy, as they had actually

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 6, June 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1006139|

experienced. The Crowd source data were represented by pin icons. Each pin is coded in three different colors with respect to reported flood levels. Flood forecasting was rendered based on learning of thematic data by an ML method. The resultant forecast was displayed on a web browser. With this platform, users may access this information from various devices.

## REFERENCES

[1] Natural Disaster Data Book, ADRC, Hyogo, Japan, 2012. ¯

[2] Natural Disaster Data Book, ADRC, Hyogo, Japan, 2015. ¯

[3] D. Komori, S. Nakamura, M. Kiguchi, A. Nishijima, D. Yamazaki, S. Suzuki, A. Kawasaki, K. Oki, and T. Oki, ''Characteristics of the 2011 Chao Phraya River flood in central Thailand,'' Hydrolog. Res. Lett., vol. 6, no. 6, pp. 41–46, 2012.

[4] P. Promchote, S.-Y. S. Wang, and P. G. Johnson, ''The 2011 great flood in Thailand: Climate diagnostics and implications from climate change,'' J. Climate, vol. 29, no. 1, pp. 367–379, Jan. 2016.

[5] S. Puttinaovarat, P. Horkaew, K. Khaimook, and W. Polnigongit, ''Adaptive hydrological flow field modeling based on water body extraction and surface information,'' J. Appl. Remote Sens, vol. 9, no. 1, Dec. 2015, Art. no. 095041.

[6] S. K. Jain, P. Mani, S. K. Jain, P. Prakash, V. P. Singh, D. Tullos, S. Kumar, S. P. Agarwal, and A. P. Dimri, ''A Brief review of flood forecasting techniques and their applications,'' Int. J. River Basin Manage., vol. 16, no. 3, pp. 329–344, Jul. 2018.

[7] N. Belabid, F. Zhao, L. Brocca, Y. Huang, and Y. Tan, ''Near-real-time flood forecasting based on satellite precipitation products,'' Remote Sens., vol. 11, no. 3, p. 252, Jan. 2019.

[8] A. K. Kar, A. Lohani, N. Goel, and G. Roy, ''Rain gauge network design for flood forecasting using multi-criteria decision analysis and clustering techniques in lower Mahanadi river basin, India,'' J. Hydrol., Regional Stud., vol. 4, pp. 313–332, Sep. 2015.

[9] M. Dembélé and S. J. Zwart, ''Evaluation and comparison of satellitebased rainfall products in Burkina Faso, West Africa,'' Int. J. Remote Sens., vol. 37, no. 17, pp. 3995–4014, Sep. 2016.

[10] A. Foehn, J. G. Hernández, B. Schaefli, and G. De Cesare, ''Spatial interpolation of precipitation from multiple rain gauge networks and weather radar data for operational applications in Alpine catchments,'' J. Hydrol., vol. 563, pp. 1092–1110, Aug. 2018.

[11] F. Cecinati, A. Moreno-Ródenas, M. Rico-Ramirez, M.-C. Ten Veldhuis, and J. Langeveld, ''Considering rain gauge uncertainty using kriging for uncertain data,'' Atmosphere, vol. 9, no. 11, p. 446, Nov. 2018.

[12] S. Stisen and M. Tumbo, ''Interpolation of daily raingauge data for hydrological modelling in data sparse regions using pattern information from satellite data,'' Hydrolog. Sci. J., vol. 60, no. 11, pp. 1911–1926, Sep. 2015.

[13] Q. Hu, Z. Li, L. Wang, Y. Huang, Y. Wang, and L. Li, ''Rainfall spatial estimations: A review from spatial interpolation to multi–source data merging,'' Water, vol. 11, no. 3, p. 579, Mar. 2019.



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING