International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Predictive Analysis Using a Novel Deep Learning Algorithm "Geonet" for Flood Likelihood Monitoring

Ms. Anjali P. Kadam[1], Prof. Dr. Prakash Devale[2]

Submitted:05/10/2024 **Revised**:18/11/2024 **Accepted**:01/12/2024

Abstract: In alignment with the United Nations' Sustainable Development Goals, India is committing itself to a comprehensive framework of 17 goals and 179 targeted objectives that address social, economic, and environmental challenges, aiming to achieve these by 2030. Against the backdrop of recurring flood events in Kolhapur, Maharashtra, our research takes on a critical role: predicting the likelihood of flooding by analyzing key weather parameters. The proposed innovative approach involves meticulous data collection, thorough pre-processing, and the application of the groundbreaking "GeoNet" deep learning algorithm. This state-of-the-art algorithm is meticulously designed to classify and predict flood-prone conditions by analyzing vital parameters such as pressure, maximum temperature, actual temperature, minimum temperature, dew point, humidity, cloud cover, sunshine, wind direction, and wind speed. By processing daily data over an entire year, the proposed study creates a robust model that offers actionable insights. Furthermore, the representation of weather variables through the proposed model is informative and visually compelling. As demonstrated by the proposed comprehensive analysis of the confusion matrix, the results reveal that the GeoNet algorithm significantly outperforms existing machine learning classifiers, achieving an impressive accuracy rate of 85%. This advancement strengthens our understanding of flood dynamics and enhances our ability to implement timely and effective mitigation strategies.

Keywords: *Machine learning, Deep learning, Geological AI, Flood prediction, LSTM*

1. Introduction

In previous seasons, global environment transformation and the rising occurrence of intense climate incidents have formed inundating catastrophes among the critical concerns experienced by different countries worldwide. In this circumstance, the quick and legitimate acquisition of details of flooding devastation is essential for authorities and relevant institutions to develop well-timed alleviation and avoidance actions [1]. The risk of surges from tropical stormy weather increases because of weather swaps and humane development. Maps of previous flood extents can help in future planning and mitigation initiatives to reduce flood chances [2].

Nevertheless, day-by-day parametric analysis is essential to monitor a water surge in vulnerable spaces. Because of the surges, there were certainly many cutbacks in the areas of Kolhapur and Sangli in 2019. Rainwater provokes dangerous situations like a water surge, and so for the predictive analysis, the authors acknowledged the rainfall data at Gaganbawada, Kolhapur, Lanja, and Miraj-Kolhapur section near Rukadi village, Kolhapur [3]. Fig. 1 shows the Panchaganga River basin view.

Research Scholar,
Department of Computer Engineering,
BVDU, COE, Pune, India
Research Supervisor,
Department of Information Technology,
Bharati Vidyapeeth (Deemed to be) University,
College of Engineering, Pune, Maharashtra, India

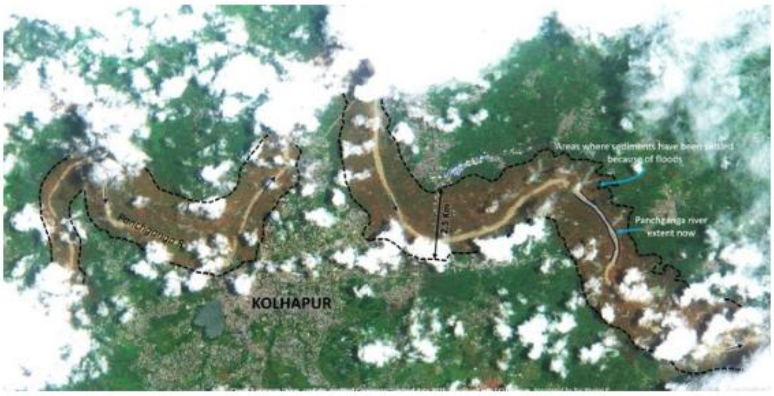

Fig.1: Flood Site: Panchaganga River, Kolhapur Maharashtra, India (Source: Maps of India)

A generalized flooding occurrence is water overflow in dried-out sectors due to natural or man-made triggers. Flash surges are recognized by all their abrupt starting point. Classifications differ, as they frequently stress the rapid lag between major rainfalls and surging or an unexpected spike in water levels [4]. Traditional machine learning models, due to their difficulty in recording intricate spatial activities and non-linear associations, often struggle to accurately predict flooding. This limitation hampers their predictive precision, making it challenging to forecast floods with high accuracy. To address this, various analyses have presented hybrid machine-learning models. Reliable flooding risk evaluation and prediction are significant for increasing metropolitan flood administration and disaster response and building up metropolitan strength. Consequently, the requirement for precision improvement is further advised by previous research [5, 6].

One of the most intellectually stimulating tasks in environmental science is predicting the pattern of rainfall in the climate forecasting process. With the increasing intensity of weather variations, this task has become more complex than ever. Predictive machine learning techniques, by uncovering hidden patterns in historical climate data, can provide a solution. However, the selection of an appropriate categorization strategy for prediction remains a complex challenge, underscoring the need for innovative solutions in this field [7].

Rainfall, a fundamental aspect of natural processes, is a key resource for agriculture, energy, and other industries and is therefore vital for the country's economy. Rainfall prediction facilitates the management of water resources, flood alerts, air travel, transportation, and construction activities. The growing global awareness of the economic and social benefits of accurate rainfall prediction is a testament to its positive impact. In a country like India, where agriculture is a major economic driver, precise rainfall prediction is crucial for managing water resources and mitigating extreme hydrological conditions [8,9].

2. Literature Review

The author tailored the investigation region of Fars province, Iran, to investigate the features of ensemble machine learning methods in forecasting rainwater attributes, integrating the monthly rainwater volume level, maximal day-by-day rainfall, and the number of rainy days. The author formulated a hybrid system using random forest, tandem perturbs, combined trees, and instance-based K-nearest neighbor models for the rainfall forecasts [10].

The author implemented a network of 26 SML sensors over a study area equivalent to a catchment area of 140 km^2, and the data accumulated from it was applied to train two machine learning (ML) algorithms and then a statistical approach to forecasting wet/dry circumstances. According to the results, ML methods are useful alternatives for real-time rainwater monitoring, possibly in an extensive sensor network. The author applied the voting procedure and observed that it enhances the auguration of poor precipitation at the outlay of larger false positive ranges [11].

As per the author, a trained DNN model employing a landslide range of usual rainfall years is competent in forecasting landslides at the time of intense rainfall events. Adding of the landslides that took place within the latest EREs (since 2018) into the existing landslide array gives a considerably more legitimate and enhance augmentation of landslide susceptibility that helps risk-informed landscaping forecasting and development of the location [12]. However, it is important to predict rainfall/floods based on daily data recordings.

The author applied a periodic accrued gridded rainfall dataset and a recurring proportioned daily mean temperature dataset from 1901 to 2021 to explore and augment yearly rainwater activities throughout India with a city-precise strategy. The author likened the deep learning (DL) algorithms like Long Short Term Memory (LSTM), Bi-directional LSTM (BiLSTM), Gated Recurrent Unit (GRU), as well as, Convolution 1D LSTM (Conv1DLSTM) worked for the long-term rainwater auguration across various locations of India. [13] However, the author introduced rainfall percentage and temperature as analysis parameters, and certainly, there is a need to analyze other environmental parameters like wind, humidity, dew point, etc.

In another study, the author investigated 12 hybrid deep learning and machine learning models to forecast daily rainfall, applying meteorological parameters such as optimum humidness, least temperatures, the upper limit of temperatures, and rainfall. The author discovered that the PSO-BiLSTMANN II model obtained the most effective capabilities and outperformed the PSO-SVR model by 6.4%. The PSO-BiLSTM-ANN II model likewise mandates fewer cells in the hidden layer than other models and converges with the minimum epochs. The outcomes demonstrate the benefit of conjoining the ANN layer in the RNN, LSTM, and BiLSTM models, and this study presents a benchmark model for forecasting rainwater in the research area [14]. However, a new deep learning algorithm can be developed to correlate the existing models to enhance prediction accuracy.

The author designed the hybrid deep learning models CNN-LSTM and LSTM-LSTM. The previous hybrid model removes the spatiotemporal aspects, and the other model incorporates this functionality as a remembrance. The other model forecasts the traffic stream parameters based on the former features and provisory input. The author noted 78-82% accuracy for rainy weather [15]. Additionally, the new algorithm can be developed and tested on a real-time flood prediction dataset using rainfall data and other environmental parameters.

The author looks at the study of ensemble learning strategies applied in rainfall prediction, considering that many researchers' focus was obtained by earlier research for rainfall auguration with several exciting issues. The accumulated data needs to be assessed and trained effectively to be examined by ensemble learning approaches that are more effective in getting the forecasted results with minimal error among the assessed value and the typical set [16].

3. Methodology

We applied the Regional Flood Frequency Analyses (RFFA) model, and the proposed methodology is shown in Fig. 2. We used the flood dataset for Kolhapur City, Maharashtra, for training and validation. We used Python and the Anaconda server to develop the proposed algorithm.

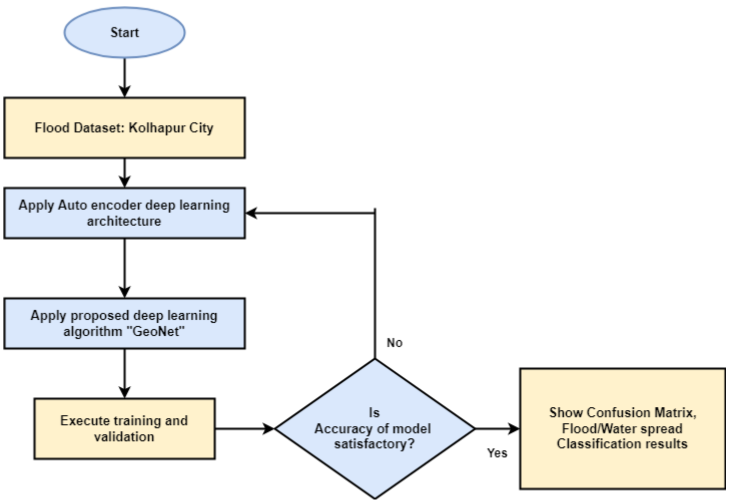

Fig. 2: Proposed system block diagram

For representation of the confusion matrix, we will calculate the accuracy, precision, recall and F-1 score using proposed algorithm “GeoNet”.

$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$ ………………………………..……..... (1)

$\text{Precision} = \frac{TP}{TP+FP}$ …………..…………………………………………... (2)

$\text{Recall} = \frac{TP}{TP+FN}$ ……………………………………………………… (3)

$\text{F1 Score} = 2 * (\text{Recall} * \text{Precision})/(\text{Recall} + \text{Precision})$ …..…. (4)

Where ‘TP’ is True Positive, ‘FP’ is False Positive, and ‘FN’ is False Negative. Precision is the accuracy of optimistic predictions, and ‘Recall’ defines the percentage of cases/instances you have caught. ‘F1-score’ depicts the optimistic predictions that are correct. The accuracy will be compared to existing machine learning algorithms.

3.1 ‘GeoNet’ Algorithm pseudo code

IMPORT Sklearn, SCciKit python libraries

IMPORT Dataset

PRE-PROCESSING to verify missing values and replace/remove particular row is Not-a-Number

BALANCE the database

SET Auto encoder for incoming data validation

IDENTIFY Data Splitting Random_State

IDENTIFY Classifiers and configure function calls

SET Deep Learning Model sequence of operation

SET INPUT PARAMETERS: Pressure, maximum temperature, actual temperature, minimum temperature, dew point, humidity, cloud status, sunshine, wind direction, wind speed

FOR each input value,

RECORD State-of-rainfall based on INPUT PARAMETERS

STORE Prediction values and generate distance plots and box plots for each parameter

SET K-fold for classifier execution

SET parametric analysis matrix

GET Confusion Matrix for all clubbed classifies

GENERATE accuracy, precision, recall, F1-score and support value

3.2 Mathematical Model

Deterministic modeling is developed for rain/flood routing and represented by a mathematical model for predicting the changing magnitude of input parameters and shape of a parameter wave as a function of time (hours, days, etc.), i.e., the graph for a specified duration. The proposed study uses daily data records for input parameters (Refer to Fig. 6).

A mathematical model determines the state of dynamic system parameters using daily observations of the system. The flood prediction step is updated by recurring observation parameter input. The prediction phase incorporates the system's state transition model to make predictions about the system's recent state based on its previous state, and the final decision is represented by the graph.

Further, the matrix of state transitions, which relates the present state to the previous state, is given by:

$\overline{X}_k = F_k \overline{X}_{k-1} + B_k u_{k-1}$ …………………………………….(5)

Where, $\overline{X}_k$ is measured value of parameter that control input matrix. $F_k \overline{X}_{k-1}$ is an error covariance matrix,

$Q_k = B_k u_{k-1}$ ……………………………...(6)

Where, Q_k is an updated parameter value obtained as a second stage reading for parameters and is computed during the prediction stage, which represents the uncertainty in the predicted state estimate which is given by:

The predicted output $\overline{P}_k$ is given by:

$\overline{P}_k = F_k \overline{P}_{k-1} F_k^t + Q_k$ …………………………(7)

Where F_k^t is predictive error from day-1 till day ‘n’ which is counted as initial time ‘k’ to‘t’.

The measurements and predicted state estimate are combined in the final update stage to provide the updated state estimate and error covariance matrix.

$X_k = \overline{X}_k + \overline{P}_k + \cdots + Q_k$ …………………………..8)

The parametric observation matrix represents the measurements for flood prediction. The weight assigned to the predicted state estimate and the measurements in the updated state estimate are determined by the GeoNet gain matrix. The error covariance matrix is also updated:

$$\overline{P}_k = (I - H_k) + \overline{P}_k \quad \text{……………………………….(9)}$$

Where, ‘I’ is the final parametric identity matrix which keeps adding updated parameters H_k is the last update about the each parameter. Further section 4 discusses results for flood prediction.

4. Results

The following Fig. 3 shows the prediction of flood based on the parametric study by “GeoNet” algorithm execution.

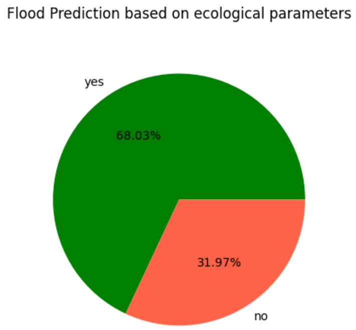

Fig. 3: Flood Prediction for Kolhapur dataset

The specific parameters/weather variables are analyzed and hydrograph represented in Fig. 4 below. The box plot is shown in Fig. 5.

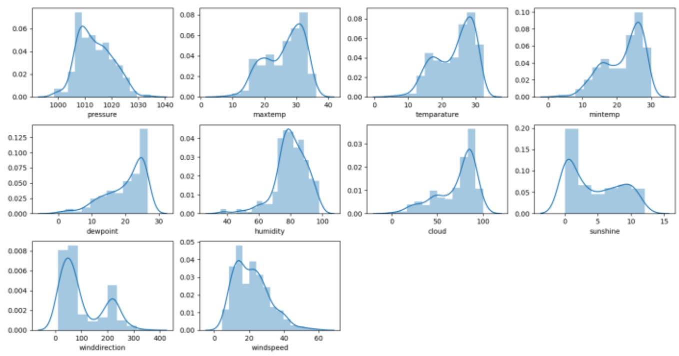

Fig. 4: Analysis of weather variables using the proposed model

Further, Fig. 5 shows box plotting for all weather variables.

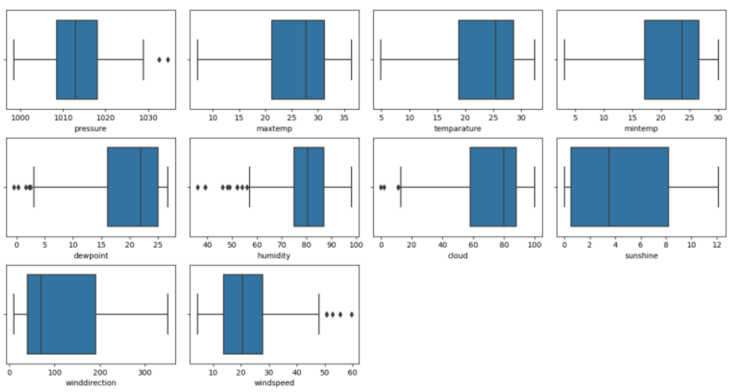

Fig. 5: Box plotting for weather variables using proposed model

Based on the analysis of weather variables, the parametric mapping has been done to identify most influential variables contributing for rainfall leading to flood situation. Fig. 6 shows all parameters mapping each other and sub-matrix shows the high impact weather variables.

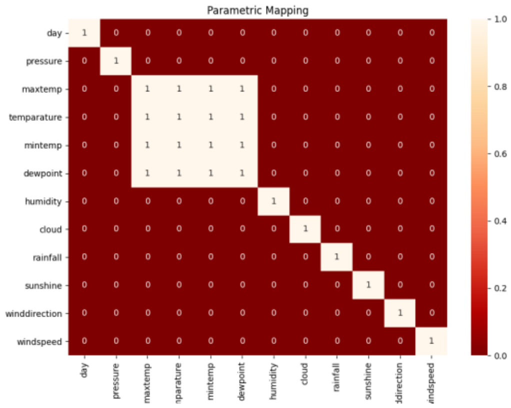

Fig. 6: Parametric mapping for most influential parameter identification

Based on the proposed "GeoNet" deep learning algorithm, Fig. 7 represents the confusion matrix.

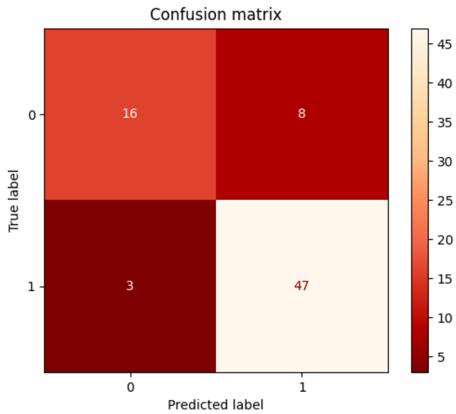

Fig. 7: Proposed system confusion matrix representation

The proposed 'GeoNet' algorithm execution and optimization results are shown in Fig. 8 with the accuracy of 85%.

	precision	recall	f1-score	support
0	0.84	0.67	0.74	24
1	0.85	0.94	0.90	50
			0.85	74
macro avg	0.85	0.80	0.82	74
weighted avg	0.85	0.85	0.85	74

Fig. 8: Proposed model execution results

Further, the proposed 'GeoNet' algorithm will be compared with existing classifiers to conduct comparative analysis for accuracy.

5. Conclusion

There are several reasons for flood/flash floods, and consequently, looking at the environmental effect, the proposed research was executed to predict the rainfall-causing water surge scenarios. The proposed research data reveals that Pressure, maximum temperature, actual temperature, minimal temperature, dew point, humidity, cloudiness, sunlight, wind direction, and wind speed are key contributor parameters of flooding. Hence, the proposed deep learning hybrid algorithm “GeoNet” focuses on classifying rainfall possibilities founded on parametric analysis. The dataset is comprised of daily readings for all parameters. The proposed model execution reveals that the classification accuracy is 85%, outperforming the existing classifiers. The confusion matrix values are promising. The future development can be focused on the comparative analysis between the proposed model and existing classifiers. The model is useful for daily data analysis and early augmentation of rainfall/flood scenarios.

References:

[1]Guo, Qichen, et al. "Assessment of urban flood disaster responses and causal analysis at different temporal scales based on social media data and machine learning algorithms." International Journal of Disaster Risk Reduction, Elsevier, 117 (2025): 105170.

[2] Composto, Rebecca W., et al. "Quantifying urban flood extent using satellite imagery and machine learning." Natural Hazards, Springer, 121.1 (2025): 175-199.

[3] Srishailam, C., N. Vivekanandan, and R. G. Patil. "Trend Analysis of Panchganga Catchment Rainfall by Statistical." Hydrology and Hydrologic Modelling: Proceedings of HYDRO 2023 410 (2025): 493.

[4] Chowdhury, Md Enayet, et al. "A machine learning-based approach for flash flood susceptibility mapping considering rainfall extremes in the northeast region of Bangladesh." Advances in Space Research, Elsevier, 75.2 (2025): 1990-2017.

[5] Wang, Yongyang, et al. "Toward explainable flood risk prediction: Integrating a novel hybrid machine learning model." Sustainable Cities and Society, Elsevier, 120 (2025): 106140.

[6] Yang, Haidong, Ting Zou, and Biyu Liu. "Assessment of urban flood susceptibility based on a novel integrated machine learning method." Environmental Monitoring and Assessment, Springer, 197.1 (2025): 1-27.

[7] Dhannaseth, Nilesh, and Sanjay Yedey. "Machine learning based model for rainfall pattern prediction: a comprehensive overview." Technological Innovations & Applications in Industry 4.0, Taylorfrancis, (2025): 112-115.

[8] Sohana, C. A., et al. "Analysing trends and forecasting rainfall in subregions of Karnataka using machine learning models." AIP Conference Proceedings. Vol. 3162. No. 1. AIP Publishing, 2025.

[9] Bhunia, Syamantak, and Ujjwal Saha. "Long-range forecasting of daily rainfall using machine learning techniques." International Journal of Hydrology Science and Technology, Inderscience, 19.2 (2025): 117-151.

[10] Khosravi, Khabat, et al. "Prediction of three vital rainfall characteristics: Advanced hybrid tree-or lazy-based learner?." Results in Engineering, Elsevier, 25 (2025): 103840.

[11] Gianoglio, Christian, et al. "Rainfall Classification in Genoa: Machine Learning vs. Adaptive Statistical Models Using Satellite Microwave Links." IEEE Access (2024).

[12] Achu, A. L., et al. "Redefining landslide susceptibility under extreme rainfall events using deep learning." Geomorphology 448 (2024): 109033.

[13] Singh, Sanjeev, et al. "Analysis and Forecasting of Temporal Rainfall Variability Over Hundred Indian Cities Using Deep Learning Approaches." Earth Systems and Environment, Springer, 8.3 (2024): 599-625.

[14] Zoremsanga, Chawngthu, and Jamal Hussain. "Particle swarm optimized deep learning models for rainfall prediction: a case study in Aizawl, Mizoram." IEEE Access (2024).

[15] Nigam, Archana, and Sanjay Srivastava. "Hybrid deep learning models for traffic stream variables prediction during rainfall." Multimodal Transportation, Elsevier, 2.1 (2023): 100052.

[16] Kundu, Saranagata, et al. "A review on rainfall forecasting using ensemble learning techniques." e-Prime-Advances in Electrical Engineering, Electronics and Energy, Elsevier, 6 (2023): 100296.