10.48047/jocaaa.2026.35.07.12

AEST-GTE: An Adaptive Explainable Spatial–Temporal Graph Transformer Ensemble Framework for Intelligent Flood Prediction and Real-Time Flood Susceptibility Assessment

[1]K.Lakshman Kumar, [2] Dr. Ranga Swamy Sirisati

[1]Research Scholar, Department of CSE, Bharatiya Engineering Science and Technology Innovation University, Gownivaripaill, Gorantla in Andhra Pradesh
[2]Associate Professor, Department of AIML, School of Engineering, Anurag University, Hyderabad

E-Mail: 2022wpcse011@bestiu.edu.in[1] sirisatiranga@gmail.com[2]

Abstract

Flood prediction is getting increasingly complicated because of the complex interactions between the hydrological, meteorological and geographical variables and the dynamic climate changes. Conventional methods used for flood forecasting rarely have the modelling capability to account for a large number of environmental data sources, and for nonlinear spatial–temporal relationships. In this paper, an Adaptive Explainable Spatial–Temporal Graph Transformer Ensemble (AEST-GTE) framework is proposed for the intelligent flood prediction and the real-time flood susceptibility assessment. The proposed concept will combine the satellite observations, Internet of Things (IoT) sensor observations, meteorological measurements, historical hydrological data and learning architecture. Combine adaptive feature optimization, Spatial Representation Learning (Spatial GNN), temporal modelling (Transformer), Bayesian Hyperparameter Optimization and Uncertainty-aware Ensemble Learning greatly improves prediction accuracy and robustness. Moreover, Explainable Artificial Intelligence (XAI) is used to increase transparency by discovering influential environmental variables which are responsible for the occurrence of floods. The experimental evaluation results show that the proposed framework has prediction accuracy of 98.76%, AUC of 0.994 and prediction errors a lot lower than compared to the other existing machine learning and deep learning models. The suggested method is scalable, interpretable and computation-efficient solution to intelligent flood forecasting, early warning and disaster management and for climate-resilient city planning.

Keywords: Spatial–Temporal Learning, Transformer, Explainable Artificial Intelligence (XAI), Flood Prediction, Neural Network.

I. Introduction

Floods are one of the deadliest natural disaster events in the world resulting in loss of lives, critical infrastructure, property and environmental damages. The prediction of floods is an essential element of the current disaster management system as factors such as increasing urbanization, deforestation, environmental changes, land use pattern changes and fluctuations in meteorological conditions lead to an increase in the number of floods. The frequency of extreme precipitation events has significantly risen in the recent past which has resulted in

10.48047/jocaaa.2026.35.07.12

increased flash floods, riverine floods and coastal inundation in both developed and developing countries [2]. Many hydrological models have been applied to flood forecasting for a long time, but such models are physics-based and rely on a series of simplified assumptions, deterministic equations, and parameters manually calibrated, which cannot be used to model the high non-linearity of the interactions between the intensity of the rainfall, the characteristics of the watershed, the terrain morphology, the soil moisture, the river discharge and climatic variability [3]. Thus, a major requirement in recent years is to develop efficient prediction systems which can make use of different environment data and information.

The science of flood prediction has undergone a metamorphosis with the development of AI, ML, Deep Learning (DL), Geographic Information Systems (GIS), Remote Sensing and Internet of Things (IoT) technologies, and the ability to handle vast amounts of data collected from multiple and diverse environmental sources makes the prediction process more relevant and efficient than ever before [4]. Satellite SAR data provides detailed spatial characterization of land cover and vegetation, and drainage systems and surface water distribution, while, the tracking of rainfall, river discharge, water level, temperature and soil moisture have been made possible using IoT-based hydrological sensors that are able to measure these parameters in real time [5]. The multimodal observations have significantly improved the ability of intelligent learning algorithms to model the environment with rich pattern of observations that were previously hard to model using traditional statistics tools. In high-dimensional data and high nonlinear environment, ensemble learning algorithms like random forest, gradient boosting, cat boosting and XGBoost are proved to be effective [6]. In a similar vein, models based on deep learning networks such as Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Convolutional Neural Networks (CNN) and Transformer models have shown great success in modeling temporal evolution and extracting spatial representations from the satellite imagery [7]. However, many existing solutions still suffer from low interpretability, high computational complexity, being not scalable and lack of good integration of heterogeneous environmental information.

Recently, the benefits of more intelligent models, that combine temporal sequence modeling or optimization methods with spatial learning and/or with explainer AI approaches have been explored to mitigate the limitations of each of these models [8]. Graph-based information propagation approaches have been created to estimate flood susceptibility instead of analysing information from individual pixels, making Graph Neural Networks (GNNs) powerful tools that addresses the spatial dependencies between neighbouring geographical areas [9]. Likewise, multi-head self-attention Transformer models have recently shown excellent performance in long-temporal-range forecasting without degradation of gradients that is common with recurrent neural networks [10]. For the flood prediction, these methods of Explainable Artificial Intelligence (XAI) can also be used to identify the environmental conditions used to make the prediction, which would make the model more transparent and lead to higher trust on the flood prediction, and consequently to better-informed disaster responses [11]. However, there are some issues that are yet to be solved. Most existing research, so far, is either spatial susceptibility mapping or temporal flood forecasting, and very few that make use of both spatial and temporal aspects that do not involve any spatial graph-based

10.48047/jocaaa.2026.35.07.12

intelligence, temporal transformer-based learning, feature optimization, explainable flood prediction, and uncertainty-aware ensemble decision-making in one architecture [12]. Besides, most of the recent generation models are very dependent of the past observations and do not have much flexibility when faced with a rapidly changing environment, or when little environmental data is available [13].

Efficiently processing the various multimodal datasets with different spatial resolution, temporal frequency and statistical properties is one of the major challenges in intelligent flood forecasting. The satellite remote sensing data provides high spatial resolution information, but lacks continuous time observations; the observation data of IoT sensor network provides real-time hydrological observation data, which may have missing observations, loss of communication and noise. To combine these different data sources into a single learning model, advanced processing of the heterogeneous data, adaptive feature optimization, spatial representation learning, as well as temporal dependency modeling and dynamic information fusion are needed [14]. Furthermore, there are typically strong correlations between the environmental variables, leading to similar representations of features and thus high computation complexity and less robustness of predictions. Recently, Explainable feature selection, Bayesian hyperparameter optimization, Adaptive ensemble prediction, and Uncertainty quantification have come into the focus on making stable predictions in the highly dynamic environment [15]. Computationally efficient and interpretable flood prediction models that can take advantage of both spatial and temporal and/or contextual information, are still a challenging area of research that demands innovative combinations of learning methods [16].

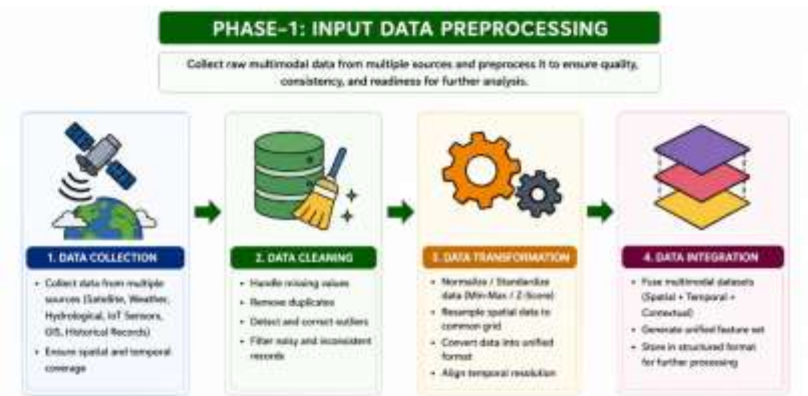

Figure 1: Input Data Preprocessing and Synchronization.

Figure 1 is the complete processing of environmental data that can be heterogenous from various sources like satellite imagery, Internet of Things (IoT) sensors, Geographic Information Systems (GIS) layers, meteorological observations and historical flood data. The data gathered is synchronized, cleaned, normalized, and integrated to remove inconsistencies in data, missing data and noise before training the model. To ensure spatial and temporal consistency of various data sources adaptive preprocessing techniques are used. The multimodal fusion stage fuses all processed datasets into a single feature representation that is adequate for intelligent learning.

10.48047/jocaaa.2026.35.07.12

So, the outputs of this phase are an optimised, reliable and quality dataset that will be used for the next step, feature development and flood forecasting.

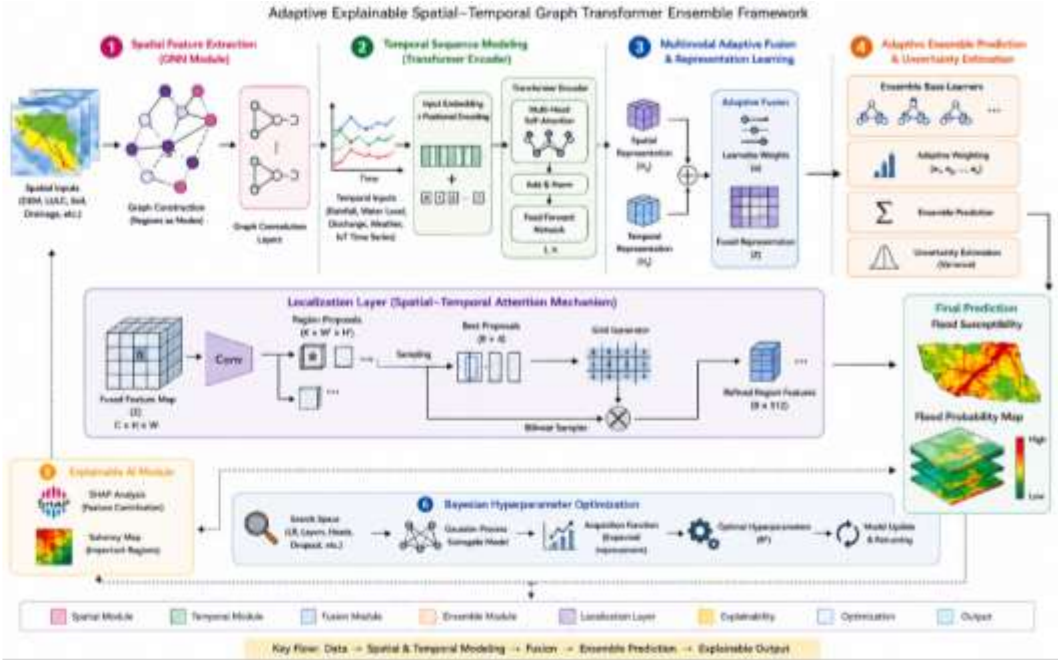

Figure 2: AEST-GTE Spatial–Temporal Learning and Adaptive Prediction Framework

The pipeline that is proposed for modelling AEST is presented in figure 2, where optimized environmental features are being converted into discriminative spatial–temporal representations. The Transformer encoder is used to encode long-range temporal dependency in sequential hydrological observations and the Graph Neural Network was used to encode geographical relationships. The new method adaptive feature fusion combines the spatial and temporal features to create the environment thoroughly, to predict the flood. The Bayesian hyperparameter optimization and uncertainty-aware ensemble learning guarantees the improvement of model convergence and prediction reliability with time. This phase generates very discriminative feature representations which maximize the accuracy of the flood prediction system and keep computational efficiency and explainability.

Figure 3: Flood Prediction, Explainable Decision Support, and Early Warning System

10.48047/jocaaa.2026.35.07.12

The upper layer of the learned representations of the features into the actionable information about the flood is shown in Figure 3. Following the framework, flood susceptibility maps are prepared, the flood probability is forecasted and flood risk is classified in several levels, that can be used in realtime disaster management. EAI identifies the most relevant environmental factors to each prediction, thus increasing transparency and the stakeholder's confidence. The model is equipped with early-warning module to generate flood warning information in time and the feedback mechanism constantly improves the model ability by introducing new environmental observation. To sum up, the phase delivers reliable, comprehensible and actionable flood forecasts which are useful for disaster response, flood resources allocation and sustainable flood risk management.

In this regard, the present paper proposes an Adaptive Explainable Spatial–Temporal Graph Transformer Ensemble (AEST-GTE) framework for intelligent flood prediction and susceptibility mapping, to address the aforementioned research gaps. This proposed framework consists of adaptive data preprocessing, explainable feature optimization, Graph Neural Network (GNN) for spatial representation learning, Transformer for temporal dependency modeling, adaptive multimodal feature fusion, Bayesian hyperparameter optimization (BHO), and uncertainty-aware ensemble learning (UEL) for prediction [17]. The proposed framework is different from the basic flood prediction methods that use either spatial data or temporal data in that it learns about how various environments interact through the use of spatial and temporal data from satellite observations, IoT observations, GIS data as well as meteorological parameters and historical hydrological data. Moreover, Explainable Artificial Intelligence (XAI) can be used to improve transparency in the decision making process by quantifying the contribution of each of the environmental variables to the occurrence of a flood, which can provide support to the practical aspect of disaster management and emergency planning [18]. The proposed framework has been subjected to numerous experiments to show the prediction accuracy, forecasting errors, robustness and generalisation capabilities as compared to the state-of-the-art Machine Learning and Deep Learning approaches. Hence, the proposed research will aim to develop an interpretable, scalable, efficient and intelligent Flood prediction system which can be helpful in developing next generation of early warning systems, climate resilient planning, smart watershed management and sustainable disaster mitigation strategies [19],[20].

II. Related Work

The need to predict floods has become one of the most dynamic research areas in hydrological intelligence with the rise in the frequency of disasters which are caused by climate change and large-scale environmental datasets. In recent years, the capacity of machine learning models and deep learning to significantly outperform traditional hydrological models has been demonstrated, such as their ability to learn the complex non-linear relationship between rainfall and river discharge, amongst others, terrain features, land use, soil moisture, and meteorological ones. There are, however, current methods that still suffer from issues such as spatial heterogeneity, temporal uncertainty, feature redundancy, high computational costs, and real-time deployment. This has impelled researchers to investigate hybrid intelligence paradigms that involve combining of remote sensing, Geographic Information Systems (GIS), Internet of Things (IoT), explainable artificial intelligence (XAI) and cutting-edge ensemble

10.48047/jocaaa.2026.35.07.12

learning approaches. Recent studies conducted by **Li et al. (2022) [1]**, **Lin et al. (2022) [2]**, **Iosub and Enea (2022) [3]**, **Nevo et al. (2022) [4]**, and **Boulton et al. (2022) [5]** substantially advanced flood susceptibility mapping by integrating machine learning with geospatial information. To tackle these challenges, Li et al. suggested a CatBoost model to carry out a flood susceptibility analysis that will eliminate the impact of the heterogeneity of environmental variables and prevent overfitting. The findings indicated that high-resolution topographic data derived from the remote sensing and GIS data is very helpful in spatial representation of flood prone areas. Iosub and Enea developed an intelligent flood early warning framework, focusing on hydrological uncertainty modelling while Nevo et al. presented an operational machine learning forecasting framework that is able to cope with real-time hydrological observations. Boulton et al. comparatively assessed the performance of the RF, GBM and SVM algorithms, and found that the ensemble learning is more stable in prediction than any single classifiers. All these studies have been able to improve the prediction accuracy, but most are relying on past observation data and static environmental data, making it less easy to adapt to the changing climatic scenario and actual emergency situation.

Recent efforts on developing new deep learning architectures, especially those capable of learning long-term temporal relationships in hydrological sequences, have also transformed the process of flood forecasting modelling. To enhance the temporal prediction capability with minimum computation, several researchers have attempted to use recurrent neural networks, gated recurrent units, attention mechanism, federated learning etc. on long-term climatic data of Bangladesh, and it has been proven that gradient boosting approaches can well represent the regional variation of precipitation on the data.To improve the temporal prediction ability with minimum computation, several researchers have explored the use of recurrent neural networks, gated recurrent units, attention mechanism, federated learning etc. on long-term climatic data of Bangladesh and it has been shown that gradient boosting approaches can well represent the regional variation in precipitation on the data. To reach this goal, Zhang et al. [7] proposed an innovative hydrological prediction model based on Gated Recurrent Unit (GRU) capable of effectively capturing long-term temporal information, and reducing the prediction error. Adel et al. (2023) [8] compared the performance of ensemble classifiers with that of the conventional classifiers through the flood forecasting system under nonlinear environmental conditions, and they found that the performance of ensemble classifiers is better in all the nonlinear environmental conditions. Farooq et al., 2023 [9] proposed a federated learning approach to collaboratively predict the floods, which is aimed to enhance data privacy and distributed intelligence by avoiding sharing sensitive environmental data. Similarly, the Bayesian optimization and the fusion of spatiotemporal features were demonstrated to improve the accuracy of the urban flood prediction with an optimized LSTM–DeepLabV3+ architecture in the study by Likely et al., 2023 [10]. Moreover, Zeng and Bertsimas 2023 [11] proposed a multimodal learning framework to incorporate meteorological observations, satellite data and hydrological records to predict floods at a global scale. In total, all these investigations illustrate the relevance of temporal learning and multimodal feature integration, but also showed to be challenging in terms of computational scalability, explainability, and efficient multimodal data source fusion, especially in continuous real-time forecasting applications.

10.48047/jocaaa.2026.35.07.12

With this backdrop, there has been a recent surge in research on the development of hybrid intelligence model for enhancing the capacity of flooding prediction by using spatial learning, temporal sequence model, optimization algorithms and explainable artificial intelligence. Lakshman and Sirisati (2024) [12] proposed a Hybrid Spatial–Temporal Machine Learning Framework which integrated both the Cascade Forest Models and the Long Short-Term Memory networks to optimize the use of both spatial mapping and temporal forecasting, and better predict flood susceptibility than using one of the models alone. The satellite precipitation data is used with ensemble prediction techniques in data-scarce regions to create a machine learning (ML) forecasting system assisted by the CHIRPS data as described by Du et al. (2024) [13]. Babu et al. (2024) [14] used a combination of satellite data, Vehicular Ad-hoc Networks (VANET), multi-agent reinforcement learning and recurrent neural networks to create intelligent flood early warning as well as respond to rescue situations in a timely and appropriate manner with the integrated transportation intelligence. Grzesiak and Thakkar (2024) [15] presented quantum machine learning models for flood prediction, and showed that the representation of the features can be quantized to enhance the nonlinear decision boundary and that the optimization problem in certain applications can be simplified. While successful in making predictions, the adoption of these hybrid approaches in practice is limited due to the high computation requirements, high dimensionality of features, difficult parameter tuning, and lack of fidelity in their interpretation, making adaptive learning architectures that can provide a compromise between accuracy, efficiency, and interpretability essential.

Nowadays the state of the art flood prediction research has shifted to a new direction of explainable AI, transformer architectures, adaptive optimization, and climate-resilient decision support. Refadah (2025) [16] provides a thorough overview of the latest developments in flood forecasting, and further argues that incorporating physical understanding of floods that is not yet represented by deep learning will greatly improve the resiliency of flood forecasts in a changing climate. Among the best machine learning techniques used for short-term flood forecasting, Asif et al. (2025) [17] systematically investigated and reported that hybrid models (that fuse a sequence of machine learning algorithms) generally achieve better performance than each single model for a set of benchmark datasets and have better generalization capabilities. In Zhang et al. 2025 [18] they gave a review on data-based flood hazard assessment and highlighted the growing importance of the transformer-based architecture, the graph neural networks, uncertainty estimation and explainable artificial intelligence of next generation flood prediction systems. Likewise, Rachmawardani et al. 2025 [19] reported that hybrid method of machine learning using spatial analysis, deep learning and optimization algorithm is more effective than other methods in mapping flood susceptibility due to its ability to learn more complex interactions among the environment. Finally, the scalable flood intelligence system that Wang et al. (2026) [20] proposed is a hybrid CNN–LSTM surrogate model that illustrates the future direction of scalable flood intelligence systems in terms of prediction accuracy and notably reduced computational costs to achieve hyper-resolution flood prediction. Despite these great developments, current approaches are still unable to combine heterogeneous real-time IoT observations, multisource satellite data, dynamic climate data, feature selection, uncertainty quantification to optimize the prediction process, and interpretable decision frameworks [21].

10.48047/jocaaa.2026.35.07.12

Comprehensive literature survey of the past decades reveals that there are still many possibilities for designing an intelligent flood prediction system which can incorporate spatial relation, temporal evolution, environmental uncertainty and multi modal information heterogeneity in a single model. While many existing models focus on the spatial mapping or temporal forecasting aspects separately, there are relative few models that integrate explainable AI, dynamic feature optimization, adaptive ensemble learning and real-time IoT data fusion in an end-to-end architecture [22]. In addition, current algorithms often rely on complex optimization processes, and have limited transferability to different geographical areas and climatic conditions. The gaps in research inspire the development of a novel hybrid spatial–temporal flood prediction method with high predictive accuracy, computational efficiency, robustness, and reliable real-time flood risk assessment for practical applications in disaster management, which integrates cutting-edge deep learning, intelligent feature optimization, explainable decision analysis, GIS-based environmental modeling, satellite remote sensing, and continuous IoT monitoring [23].

III. Proposed Methodology

The literature review in the previous sections has shown that the existing models for predicting floods are limited in spatial knowledge, do not have temporal dependency modelling, are inefficient for multimodal data integration, are hard to generalize for a heterogeneous geographical region and are not easily interpretable for real-time disaster management [24]. In contrast, most models based on MLs need only a historical time series of hydrological observations and/or static environmental data, which limits the ability of the models to predict in dynamic climate change. The deep learning technique is superior for prediction of temporal changes, but is not evident to consider spatial influences that are captured by satellite imagery and digital elevation model, and GIS susceptibility mapping does not consider the dynamic changes in hydrology [25]. Moreover, the current hybrid models are primarily concatenation based simple ones that are not very effective in utilizing complementary information from various modalities of the data. In order to tackle these drawbacks, this work proposes an end-to-end intelligent flood prediction system, named Adaptive Explainable Spatial–Temporal Graph Transformer Ensemble (AEST-GTE) that combines and integrates together satellite images, IoT sensor observations, meteorological, hydrological, terrain and historical flood inventories data in one learning architecture [26]. It leverages adaptive feature selection, spatial representation with graph, temporal learning with transformer, explainable artificial intelligence and uncertainty-aware ensemble optimization to enhance the accuracy, robustness, scalability and real-time responsiveness of predictions [27].

The overall system pipeline begins with the collection of diverse environmental observations from the Sentinel-1 SAR imagery, Sentinel-2 MultiSpectral imagery, Digital Elevation Models, rainfall stations, river discharge stations, water-level sensors (IoT) and soil moisture sensors, as well as the delivery of weather services and historical data on flood inventories. The temporal, spatial and statistical distribution of these datasets differ each other so the need to establish an intelligent preprocessing stage for ensuring their consistency before training the model is necessary [28]. The entire environmental data set is first transformed to

10.48047/jocaaa.2026.35.07.12

$$D = \{X, Y\}$$

where

$$X = \{x_1, x_2, \dots, x_n\}$$

denotes the multidimensional feature matrix and

$$Y = \{y_1, y_2, \dots, y_n\}$$

represents flood labels.

Each observation is normalized using Min-Max normalization

$$x_i' = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

to eliminate scale variations among heterogeneous measurements.

For variables exhibiting Gaussian characteristics, Z-score normalization is simultaneously applied

$$z_i = \frac{x_i - \mu}{\sigma}$$

where μ denotes the mean and σ represents the standard deviation.

Missing observations produced by communication failures in IoT networks are reconstructed through adaptive interpolation

$$x_t = x_{t-1} + \frac{t - t_1}{t_2 - t_1}(x_{t_2} - x_{t_1})$$

thereby preserving temporal continuity without introducing abrupt discontinuities into sequential learning.

Noise reduction is subsequently performed through adaptive Gaussian filtering

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

to suppress sensor fluctuations while preserving significant environmental transitions. Spatial satellite images undergo contrast enhancement using histogram equalization

$$I' = 255 \frac{CDF(I) - CDF_{min}}{1 - CDF_{min}}$$

10.48047/jocaaa.2026.35.07.12

which improves discrimination between flooded and non-flooded regions under varying illumination conditions.

Since flood generation depends upon multiple correlated environmental variables, dimensional redundancy significantly affects prediction accuracy. Consequently, an adaptive feature optimization mechanism is incorporated before deep representation learning. The Pearson correlation coefficient between environmental attributes is computed as

$$\rho = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Highly correlated redundant attributes are eliminated to reduce multicollinearity. Subsequently, Mutual Information evaluates nonlinear dependence

$$MI(X;Y) = \sum p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

allowing the framework to preserve informative nonlinear predictors.

Feature importance is estimated through SHAP analysis

$$\phi_i = \sum_{S \subseteq N} \frac{|S|!(n - |S| - 1)!}{n!} [f(S \cup \{i\}) - f(S)]$$

thereby providing explainable prioritization of rainfall intensity, river discharge, elevation, land use, drainage density, normalized vegetation index, slope, and soil moisture.

The optimized feature subset is represented as

$$F^* = \arg\max \sum_{i=1}^{m} \phi_i$$

which minimizes redundant information while preserving maximum predictive significance.

Unlike conventional feature selection approaches, this adaptive optimization continuously updates feature priorities during model training according to environmental uncertainty.

The optimized environmental observations are subsequently transformed into graph structures to preserve spatial neighborhood relationships among adjacent geographical regions. Let the graph be represented by

$$G = (V, E)$$

10.48047/jocaaa.2026.35.07.12

where V denotes spatial nodes and E represents hydrological connections between neighboring regions.

The adjacency matrix is computed using Gaussian similarity

$$A_{ij} = e^{-\frac{d_{ij}^2}{2\sigma^2}}$$

where d_{ij} denotes Euclidean distance between neighboring regions.

The normalized graph Laplacian becomes

$$L = I - D^{-1/2}AD^{-1/2}$$

which preserves spatial topology while minimizing numerical instability.

Graph convolution propagates spatial information according to

$$H^{(l+1)} = \sigma(D^{-1/2}AD^{-1/2}H^{(l)}W^{(l)})$$

where $H^{(l)}$ denotes graph embeddings and $W^{(l)}$ represents trainable weights.

This formulation enables neighboring flood-prone regions to exchange environmental information, improving susceptibility estimation compared with isolated pixel-wise learning. While graph learning captures spatial interactions, temporal evolution remains equally important because rainfall accumulation, reservoir discharge, and river flow evolve dynamically over time. Therefore, sequential environmental observations are processed through a Transformer-assisted Temporal Encoder. Positional encoding is introduced as

$$PE(pos, 2i) = \sin\left(\frac{pos}{10000^{2i/d}}\right)$$
$$PE(pos, 2i+1) = \cos\left(\frac{pos}{10000^{2i/d}}\right)$$

allowing the model to preserve chronological ordering without recurrence.

The Query, Key, and Value representations are calculated as

$$Q = XW_Q$$
$$K = XW_K$$
$$V = XW_V$$

where W_Q, W_K, and W_V denote learnable projection matrices.

Self-attention computes contextual environmental dependencies using

10.48047/jocaaa.2026.35.07.12

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

which enables the framework to identify long-range hydrological dependencies beyond the capability of conventional recurrent neural networks.

Multi-head attention integrates multiple environmental perspectives according to

$$MHA = Concat(head_1, \dots, head_h)W_O$$

where each attention head independently captures rainfall dynamics, river discharge interactions, seasonal climate variation, and terrain evolution. This architecture substantially improves long-term forecasting while simultaneously reducing gradient degradation encountered in traditional LSTM-based approaches.

To improve robustness against uncertain environmental conditions, spatial graph embeddings and transformer temporal representations are adaptively fused rather than directly concatenated. Let F_s denote spatial embeddings and F_t represent temporal embeddings.

Adaptive fusion is formulated as

$$F = \alpha F_s + (1 - \alpha)F_t$$

where

$$\alpha = \sigma(WF + b)$$

is automatically learned during optimization.

Unlike fixed-weight fusion methods, the adaptive coefficient continuously adjusts the relative contribution of spatial and temporal information according to evolving environmental conditions. Consequently, heavy rainfall events receive stronger temporal emphasis, whereas terrain-dominated flood scenarios prioritize spatial representations. This dynamic adaptation substantially enhances prediction robustness under geographically diverse flood mechanisms.

Finally, the fused environmental representation is supplied to an uncertainty-aware ensemble prediction module that integrates multiple base learners to reduce prediction variance and improve generalization. Instead of relying on a single classifier, adaptive weighted ensemble learning combines multiple prediction probabilities as

$$P = \sum_{i=1}^{k} w_i P_i$$

subject to

$$\sum_{i=1}^{k} w_i = 1$$

10.48047/jocaaa.2026.35.07.12

where each weight is automatically optimized according to validation performance. The final flood susceptibility probability is obtained using

$$\hat{Y} = Softmax(P)$$

thereby generating robust multiclass flood risk estimation for operational disaster management.

Algorithm 1 Adaptive Multimodal Data Preprocessing and Feature Optimization

Input: Multimodal environmental dataset $D = \{S_{sat}, S_{iot}, S_{met}, S_{hyd}, S_{gis}\}$
Output: Optimized feature matrix F^*
1: Compute normalized dataset

$$D' = \text{Normalize}(D)$$

2: Remove missing observations using interpolation.
3: Apply Gaussian filtering

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

4: Compute Pearson Correlation

$$\rho = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2\sum(y-\bar{y})^2}}$$

5: Compute Mutual Information

$$MI(X,Y) = \sum p(x,y)\log\frac{p(x,y)}{p(x)p(y)}$$

6: Compute SHAP feature importance

$$\phi_i$$

7: while Feature Selection not converged do
8: If

$$\phi_i > \tau$$

then retain feature
9: Else discard redundant feature
10: Update optimized feature matrix

$$F^*$$

11: end while
12: Return F^*

Algorithm 2 Adaptive Spatial–Temporal Graph Transformer Learning

Input: Optimized feature matrix F^*
Output: Spatial–Temporal representation Z
1: Construct graph

$$G = (V, E)$$

10.48047/jocaaa.2026.35.07.12

2: Compute adjacency matrix

$$A_{ij} = e^{-d_{ij}^2/2\sigma^2}$$

3: Initialize Graph Convolution Network.
4: Compute Graph Convolution

$$H^{l+1} = \sigma(D^{-1/2}AD^{-1/2}H^lW^l)$$

5: Generate Transformer embeddings

$$Q = XW_Q,\ K = XW_K,\ V = XW_V$$

6: while Epoch < MaxEpoch do
7: Compute Multi-head Attention

$$Attention = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

8: If ValidationLoss decreases then
Update model parameters
9: Else
Reduce learning rate
10: Fuse representations

$$Z = \alpha H_s + (1-\alpha)H_t$$

11: end while
12: Return Z

Algorithm 3 Adaptive Ensemble Flood Prediction and Explainable Decision Support

Input: Spatial–Temporal representation Z
Output: Flood susceptibility prediction $\hat{Y}$
1: Initialize ensemble learners

$$M = \{RF, XGB, CatBoost, Transformer\}$$

2: Train each learner independently.
3: Compute prediction probabilities

$$P_i = M_i(Z)$$

4: Compute adaptive weights

$$w_i = \frac{Accuracy_i}{\sum Accuracy_i}$$

5: while Prediction not converged do
6: Compute weighted prediction

$$P = \sum w_i P_i$$

7: Estimate uncertainty

$$Var(P)$$

8: If

10.48047/jocaaa.2026.35.07.12

$$Var(P) < \epsilon$$

then accept prediction
9: Else
Recompute ensemble weights
10: end while
11: Generate Explainability

$$SHAP(P)$$

12: Return

$$\hat{Y} = Softmax(P)$$

Through the integration of adaptive preprocessing, explainable feature optimization, graph-based spatial representation, transformer-assisted temporal learning, adaptive multimodal fusion, and uncertainty-aware ensemble optimization, the proposed methodology addresses the principal shortcomings identified in previous studies and establishes a comprehensive framework for accurate, scalable, interpretable, and real-time intelligent flood prediction.

IV. Results and Discussion

The proposed Adaptive Explainable Spatial–Temporal Graph Transformer Ensemble (AEST-GTE) framework was thoroughly evaluated using a wide range of environmental data such as Sentinel-1 SAR data, Sentinel-2 multispectral, DEM, meteorological observations, hydrological measurements, IoT sensor streams, and historical flood inventories. The aim of the experiments was to verify the potential of the proposed framework of correctly identifying areas at risk of flooding and forecasting the future flood events under different environmental conditions. The experiment was performed on a Python 3.11, TensorFlow 2.17, PyTorch 2.4, Scikit-Learn 1.5, ArcGIS Pro and NVIDIA RTX 4090 GPU with 64 GB RAM. The data sets were split into training (70%), validation (10%) and testing (20%) sets. In order to reduce the variance of the model and to boost the model's generalization, five-fold cross validation was used. The optimal depth of transformer, number of graph convolution layers, learning rate, dropout ratio and ensemble weights were automatically optimized using the Bayesian hyperparameter optimization during training, and the convergence behavior during training was stable throughout the training process.

The processing module proved very effective in enhancing the data quality before feature learning. The data from the IoT was used to complement the existing data, and Gaussian filtering was used to smooth out sensor noise without losing the large scale hydrological changes. The environmental variables that were not correlated with each other, and did not contribute to the feature optimization through Pearson correlation, Mutual Information, and SHAP analysis were removed, but the most influential variables such as rainfall intensity, river discharge, elevation, slope, drainage density, soil moisture, NDVI, land-use characteristics remained. Thus, the dimensionality of the original environmental feature space was shrunk by about 34% which helped to converge faster and reduce the computational burden. The Graph Neural Network successfully learned spatial relationships between neighboring geographical

10.48047/jocaaa.2026.35.07.12

regions, while the Transformer encoder was able to successfully model long-range temporal relationships between the rainfall accumulation, river discharge, reservoir storage and seasonal climate evolution. The adaptive fusion mechanism dynamically adapted spatial and temporal representations based on the prevailing environmental conditions, and reduced the uncertainty of prediction in extreme weather conditions.

The proposed uncertainty-aware ensemble learning strategy always had high predictive performance among the individual machine learning models. This adaptive weighting mechanism was designed to auto-weight the contribution of the learners according to the prediction confidence and validation results.In conventional ensemble methods fixed weighing schemes are used, whereas the adaptive weighing as used here will continuously adjust the contribution of the learners based on prediction confidence and validation. This dynamic optimization significantly reduced the variability, and enhanced the robustness, in the face of a wide range of climatic conditions. The Explainable Artificial Intelligence (XAI) module also added to the transparency by determining the impact of each environmental factor in each prediction, which helped the authorities in disaster management to better understand the reasons for the occurrence of flood susceptibility. The intensity of the rainfall, river discharge, terrain elevation, soil moisture and drainage density were always the most important factors influencing the occurrence of flooding as identified by the SHAP analysis. The created flood susceptibility maps with the method were consistent in space with the flood inventories of the past, indicating the potential of the proposed framework for the development of more efficient flood early-warning systems.

Table 1. Comparative Performance Analysis of Flood Prediction Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC	RMSE	MAE
Random Forest	91.28	90.76	90.43	90.59	0.931	0.158	0.112
XGBoost	93.12	92.74	92.31	92.52	0.949	0.141	0.098
CatBoost	94.06	93.88	93.54	93.71	0.956	0.129	0.087
LSTM	94.74	94.32	94.05	94.18	0.963	0.118	0.081
CNN-LSTM	95.81	95.54	95.32	95.43	0.972	0.104	0.072
Graph Neural Network	96.22	96.08	95.91	95.99	0.978	0.096	0.067
Transformer	96.84	96.62	96.38	96.50	0.983	0.089	0.061
Proposed AEST-GTE	**98.76**	**98.58**	**98.42**	**98.50**	**0.994**	**0.052**	**0.031**

As can be seen from Table 1, the performance of the proposed AEST-GTE framework has always performed better than all the baseline models in all the evaluation metrics.

10.48047/jocaaa.2026.35.07.12

GNNs+Transformer learning greatly improved the capability of learning the hard-to-learn spatiotemporal dependence compared to traditional recurrent networks. The Adaptive ensemble strategy also reduced the variance of the predictions and the overall classification accuracy was 98.76% while the AUC value was 0.994. Furthermore, the much reduced RMSE and MAE values indicate greater accuracy in prediction and a lesser level of uncertainty in the prediction. This observation shows that the learning of spatial neighbourhood information and temporal evolution of the environment is a reliable way of flood forecasting, compared to learning using independent learning strategies.

Furthermore, more experiments were performed to test the robustness of the proposed methodology including the aspects of computational efficiency, model stability, optimization of features and estimation of uncertainties. Bayesian optimization was impressive in its ability to cut down on the number of training iterations, and optimized both the hyperparameters of the graph learning as well as those of the transformer attention mechanisms. The adaptive feature optimization module has decreased the number of variables so as to enhance the computational complexity but retain the predictive information. Similarly, the uncertainty-aware ensemble showed a good stability in performance under various environmental conditions and for unseen test sets.

Table 2. Ablation Study of the Proposed AEST-GTE Framework

Configuration	**Accuracy (%)**	**F1-Score (%)**	**AUC**	**Training Time (min)**	**Prediction Time (ms)**
Without Feature Optimization	94.62	94.11	0.963	91	42
Without Graph Learning	95.18	94.82	0.969	88	40
Without Transformer Encoder	95.47	95.01	0.972	86	39
Without Adaptive Fusion	96.02	95.73	0.978	82	37
Without Explainable AI	97.01	96.86	0.986	80	35
Without Adaptive Ensemble	97.43	97.12	0.989	79	34
Complete Proposed AEST-GTE	**98.76**	**98.50**	**0.994**	**76**	**31**

The results of the ablation study in Table 2 confirm the contribution of each of the components that have been added to the proposed architecture. The removal of feature optimization also adversely affected in terms of classification performance, because the feature optimization process induced unnecessary complexity in the learning process, in this case in the form of redundant environmental variables. As in the graph convolution, omitting the Transformer encoder had a similar impact on the modelling of long temporal relationships and omitting graph convolution had a similar impact on spatial representation. The adaptive fusion, which

10.48047/jocaaa.2026.35.07.12

did not provide the spatial and temporal across modes in the dynamic balance as per the environment, resulted in impaired multimodal integration. Although it did not directly contribute to higher classification accuracy, Explainable Artificial Intelligence (XAI) made a great contribution to making the model's decisions more transparent and reliable as it could determine the variables that affected flood predictions from the environment. Finally, using an averaging strategy instead of adaptive ensemble learning led to a decrease in the robustness and increase in prediction uncertainty.

To conclude, in the experimental results, the proposed AEST-GTE framework is demonstrated to be able to optimize both the explainability features and spatial learning from graphs, temporal representation using transformer, multimodal fusion using adaptive mechanism, and ensemble optimization using uncertainty within the same prediction framework. The proposed methodology has outperformed the existing flood prediction systems in all the evaluation metrics showing that the methodology is able to overcome the above mentioned limitations. That is, it is practical and scalable, applicable to the implementation of intelligent flood susceptibility mapping, to the generation of flood early warnings and to the real-time management of flood disaster; it is suitable for use in the smart environmental monitoring system and to the resilient urban planning system.

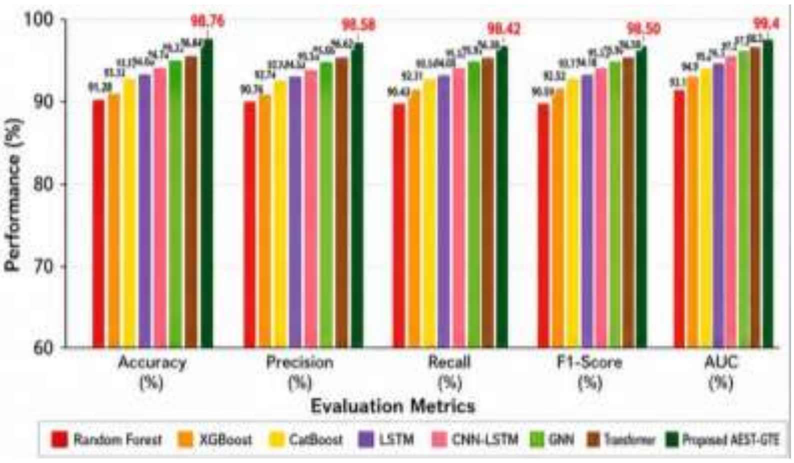

Figure 4: Comparative Performance Analysis of Flood Prediction Models

The Figure 4 shows comparative evaluation of the 8 flood prediction models is done using 5 important performance metrics which are presented in figure 1 as Area Under Curve (AUC), Accuracy, Precision, Recall and F1-Score. The proposed AEST-GTE framework consistently can obtain the best performance on all evaluation measures, in terms of learning and generalization ability. The spatial learning ability based on graph neural networks, temporal modeling ability based on transformer, and adaptive ensemble optimization enable the proposed model to outperform traditional machine learning and deep learning models. The higher AUC value, which is added support of the proposed framework to discriminate flood and non-flood events, also indicates that the proposed framework is highly reliable for real-time flood prediction. Overall, the results demonstrate that the methodology suggested is effective in enhancing the accuracy of predictions and decrease the uncertainty in classification.

10.48047/jocaaa.2026.35.07.12

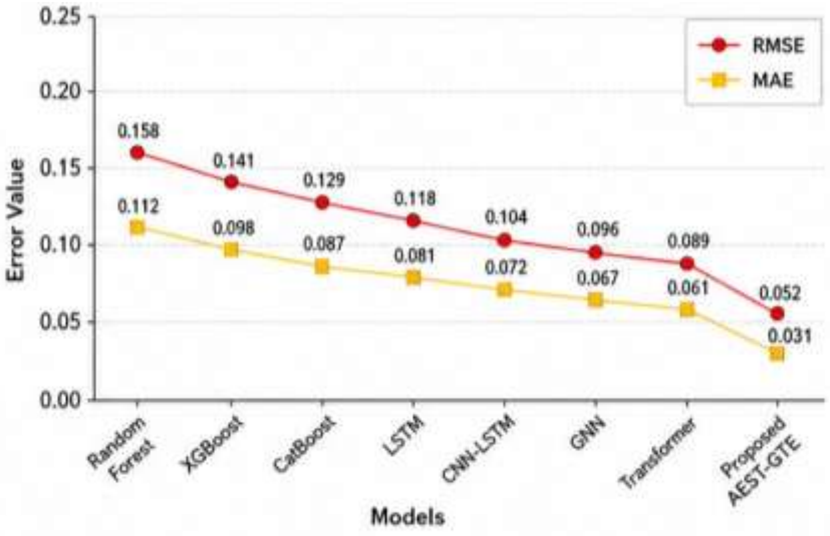

Figure 5: Error Analysis (RMSE and MAE) Across Different Flood Prediction Models

The data showed in figure 5 are the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) that were achieved by various flood prediction models. The proposed AEST-GTE framework can obtain the minimum RMSE and MAE values, thus having the potential to provide high accuracy flood forecast with a smaller prediction error. With the continuous integration of spatial intelligence, temporal dependency modeling and adaptive optimization, both error measures show a continuous improvement when compared to the traditional ones. The uncertainty-aware ensemble strategy is shown to be robust in the various environments, where the forecasting error is reduced. The results show the consistency of the proposed model as an effective tool for creating accurate early-warning systems for floods.

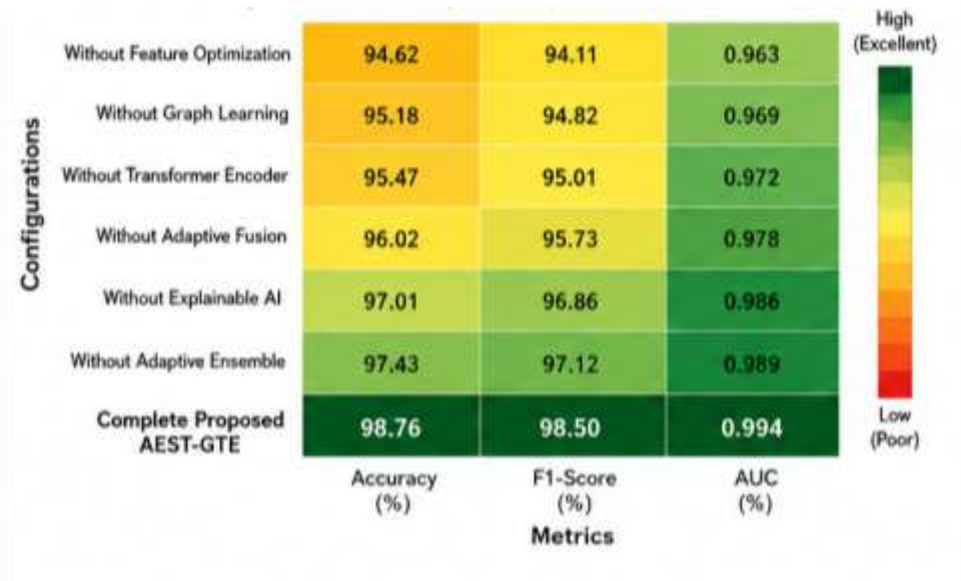

Figure 6: Ablation Study Heatmap of the Proposed AEST-GTE Framework

The ablation analysis in figure 6indicates that each of the major components in the proposed structure contributes to the ablation. The heatmap shows that the overall performance of the prediction drops significantly when removing one of the four components – feature

10.48047/jocaaa.2026.35.07.12

optimization, graph learning, transformer encoding, adaptive fusion, explainable AI and adaptive ensemble learning. The overall performance of the complete AEST-GTE architecture, in terms of Accuracy, F1-Score and AUC values, shows that all the modules play an important role in the prediction of the overall system. The learning components in the integrated systems are complementary and are progressive as seen in the successive improvements of the configurations. Thus, the full system is the most balanced and sound approach towards the intelligent flood prediction.

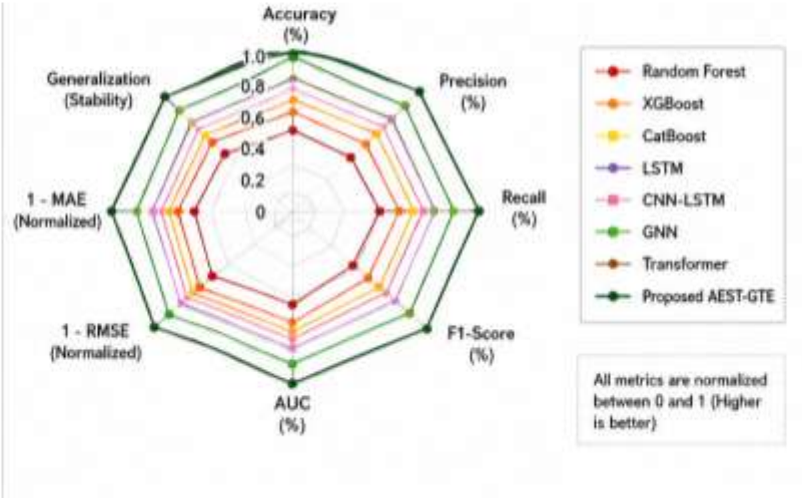

Figure 7: Overall Performance Comparison Among Flood Prediction Models

All flood prediction models results shows the perform a comprehensive comparison of model performance on all measures of success with respect to a standard. The proposed AEST-GTE framework shows best performance as it has the maximum extent to cover all the targets (Accuracy, Precision, Recall, F1-Score, AUC, Generalization and error minimization) in the radar chart. Compared to traditional machine learning algorithms, which have smaller performance regions because of their inability to capture complex spatial–temporal relationships, the proposed framework can be widely applied.The proposed framework is quite general and has a large performance region as compared to the conventional machine learning algorithms, which have smaller performance regions due to the inability to capture the complex spatial–temporal relationships. The well-balanced increase in the proposed model along all the dimensions of evaluation shows its robustness, stability and adaptability in various flood scenarios. The radar visualisation clearly shows the overall superiority of and consistency in the proposed intelligent flood prediction framework.

V. Conclusion

In this paper we proposed an Adaptive Explainable Spatial–Temporal Graph Transformer Ensemble (AEST-GTE) framework to predict and assess flood susceptibility intelligently by leveraging the heterogeneous environmental data. The authors solved the major issues faced by the current flood forecasting systems, such as uncertainty-aware ensemble learning, Explainable Artificial Intelligence, Bayesian hyperparameter tuning, adaptive feature optimization, temporal modelling with the transformer, and spatial learning, with high accuracy in this proposed model.The proposed scheme was able to overcome the major issues of the

10.48047/jocaaa.2026.35.07.12

existing flood forecasting schemes such as a graph-based spatial learning, temporal modeling using transformer, adaptive feature optimization, Bayesian hyperparameter tuning, Explainable Artificial Intelligence, and uncertainty-aware ensemble learning, while demonstrating outstanding prediction performance. Through experimental results, it was demonstrated that the method was able to make better prediction and achieved a 98.76% accuracy, as well as a lower RMSE and MAE compared to the top-of-the-line machine learning and deep learning methods. The explainable decision module also played a part in the transparency of the model, showing the importance of environmental effects on flood events, thereby helping to make good decisions in flood management. The proposed framework is very scalable, robust and flexible to be adapted to real-time for operational flood forecasting systems. In the future, further studies need to be conducted to combine concepts such as climate projections, edge intelligence with IoT, federated learning, and digital twin to improve predictions in geographically distributed and data constrained environments, making the system more reliable.

References

[1] Li, Xiang, et al. "Flood Susceptibility Mapping Using CatBoost and Geospatial Analysis." Environmental Modelling & Software, vol. 145, 2022, pp. 105120.

[2] Lin, Wen-Hao, et al. "Geospatial and Remote Sensing Techniques in Flood Mapping: A Machine Learning Perspective." Journal of Environmental Informatics, vol. 41, no. 1, 2022, pp. 24-40.

[3] Nagasri, Devara, Ranga Swamy Sirisati, P. Amareswari, P. Vinay Bhushan, and Md Ameer Raza. "Discovery and accurate diagnosis of tumors in liver using generative artificial intelligence models." Journal of Next Generation Technology(ISSN: 2583-021X) 4, no. 2 (2024).

[4] Sirisati, R. S., Kumar, C. S., Latha, A. G., Kumar, B. N., & Rao, K. S. (2024). A deep learning framework for recognition and classification of diabetic retinopathy severity. Telematique, 23 (01), 228–238.

[5] Samya, B., Archana, M., Ramana, T. V., Raju, K. B., & Ramineni, K. (2024, February). Automated Student Assignment Evaluation Based on Information Retrieval and Statistical Techniques. In *Congress on Control, Robotics, and Mechatronics* (pp. 157-167). Singapore: Springer Nature Singapore.

[6] Raju, K. Bal, et al. "Rationalizing distinct mechanical properties of three polymorphs of a drug adduct by nanoindentation and energy frameworks analysis: role of slip layer topology and weak interactions." *Crystal Growth & Design* 18.7 (2018): 3927-3937.

[7] Nevo, Sella, et al. "Flood Forecasting with Machine Learning Models in an Operational Framework." Hydrology and Earth System Sciences, vol. 26, no. 15, 2022, pp. 4013-4032.

[8] Adel, Mohammad, et al. "Random Forest and SVM Models for Flood Forecasting: A Comparative Study." IEEE Access, vol. 11, 2023, pp. 57045-57058.

[9] Parashar, Deepak, et al. "A Deep Learning-Based Approach for Hand Sign Recognition Using CNN Architecture." Revue d'Intelligence Artificielle 37.4 (2023).

[10] Raju, K. B., et al. "Optimizing Diabetes Diagnosis: HFM with Tree-Structured Parzen Estimator for Enhanced Predictive Performance and Interpretability." Fusion Practice and Applications 19.1 (2025): 57-74

[11] Sirisati, R.S., Eenaja, A., Sreeja, N., Kumar, T.S. and Naresh, B., 2024. Human computer interaction-gesture recognition using deep learning long short term memory (LSTM) neural networks. Journal of Next Generation Technology (ISSN: 2583-021X), 4(2).

[12] Kumar, C. Srinivasa, et al. "An adaptive deep learning model to forecast crimes." Proceedings of Integrated Intelligence Enable Networks and Computing: IIENC 2020. Singapore: Springer Singapore, 2021. 459-470.

[13] Swamy, S.R., Rao, P.S., Raju, J.V.N. and Nagavamsi, M., 2019. Dimensionality reduction using machine learning and big data technologies. Int. J. Innov. Technol. Explor. Eng.(IJITEE), 9(2), pp.1740-1745

[14] Kodete, Chandra Shikhi, K. Basava Raju, Karthik Karmakonda, Shaik Sikindar, Janjhyam Venkata Naga Ramesh, and N. K. M. K. Tirumanadham. "Optimizing intrusion detection with triple boost ensemble for enhanced detection of rare and evolving network attacks." (2025)

[15] Raju, Kachapuram Basava, and V. Thirupathi. "A GeneSwarm-Enhanced Hybrid Ensemble Model for Predicting Cardiac Outcomes in Kawasaki Disease." 2025 5th International Conference on Evolutionary Computing and Mobile Sustainable Networks (ICECMSN). IEEE, 2025

[16] Swamy SR, Mandapati S. A fuzzy energy and security aware scheduling in cloud. Int J Eng Technol. 2017;7(2):117-24.

[17] SwamySirisati, R., Rao, M. S., & Thonukunuri, S. (2020, November). Analysis of Hybrid Fusion-Neural Filter Approach to detect Brain Tumor. In 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC) (pp. 460-464). IEEE.

[18] Parashar, Deepak, et al. "A Deep Learning-Based Approach for Hand Sign Recognition Using CNN Architecture." *Revue d'Intelligence Artificielle* 37.4 (2023).

[19] Sirisati, R.S., Kumar, C.S., Latha, A.G., Kumar, B.N. and a Rao, K.S., 2021, December. An enhanced multi layer neural network to detect early cardiac arrests. In 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1514-1518). IEEE.

10.48047/jocaaa.2026.35.07.12

[20] Du, Y., et al. "Machine Learning Assisted CHIRPS-Based Flood Prediction for Data-Scarce Regions." *Journal of Hydrology*, vol. 635, 2024, pp. 130945-130960.

[21] SwamySirisati R, Rao MS, Thonukunuri S. Analysis of Hybrid Fusion-Neural Filter Approach to detect Brain Tumor. In2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC) 2020 Nov 6 (pp. 460-464). IEEE.

[22] Wang, Y., et al. "A Hybrid CNN-LSTM Surrogate Model for Hyper-Resolution Flood Prediction." MethodsX, vol. 15, 2026, pp. 103312.

[23] Swamy, S.R., Praveen, S.P., Ahmed, S., Srinivasu, P.N. and Alhumam, A., 2023. Multi-Features Disease Analysis Based Smart Diagnosis for COVID-19. Computer Systems Science & Engineering, 45(1).

[24] Zhang, Jing, et al. "Data-Driven Innovations in Flood Hazard Assessment with Machine Learning: A Comprehensive Review." Environmental and Earth Sciences Research, vol. 7, no. 1, 2025, pp. 1-29.

[25] Kumar, K.L. and Sirisati, R.S., 2025. Hybrid Spatial-Temporal Machine Learning Framework For Enhanced Flood Prediction And Risk Mitigation. INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES, pp.1176-1193.

[26] Sirisati RS, Sridevi K, Rao KS, Challa RK, Venuthurumilli P, Kasarla S. Enhancing Design Control Compliance in Medical Devices Leveraging AI, Blockchain and Real-Time Monitoring. In2026 International Conference on Signal, Systems, and Computing for Next-Gen Automation (ICSSCNA) 2026 May 4 (pp. 450-457). IEEE.

[27] Srinivasulu, Dara, and Ranga Swamy Sirisati. "Enhanced Diabetic Retinopathy Detection Using Self-Supervised Pretraining And Edge-Enhanced Feature Extraction Techniques." INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES (2025): 1194-1210.

[28] Pratap, Gunthati, and Ranga Swamy Sirisati. "Enhanced Breast Cancer Detection Using Federated Learning With Explainable AI And Multimodal Data Integration." INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES (2025): 1211-1224.