AI-Based Flood Risk Prediction and Early Warning System for Disaster Management

Dr. MK. Jayanthi Kanan[1], Jai Arora[2], Satyam[3], Uday Pratap Singh[4], Vaibhav Sharma[5], Yesh Sharma[6], Nandish Dave[7]

[1]*Professor, School of Computer Science Engineering and Artificial Intelligence, VIT Bhopal University, Bhopal-Indore Highway, Kothrikalan, Sehore, Madhya Pradesh*

[2,3,4,5,6,7]*Students, School of Computer Science Engineering and Artificial Intelligence, VIT Bhopal University, Bhopal-Indore Highway, Kothrikalan, Sehore, Madhya Pradesh*

***Abstract*— Floods remain one of the most frequent and destructive natural disasters affecting India, with states such as Assam experiencing recurring, severe flooding almost every monsoon season. India has recorded a 40% rise in flash flood events over the past three years, and Assam alone saw 4.29 million people affected across 35 districts in 2024. Existing flood management systems are largely reactive, region-specific, and lack the ability to generalize predictions to new or data-scarce areas. This paper proposes an AI-based flood risk prediction and early warning system trained on historical rainfall, river-level, and geographic data from multiple Indian states. A leave-one-region-out validation strategy is employed, in which the model is trained on data from all major flood-prone states except Assam and then tested specifically on Assam to evaluate its ability to generalize to a high-risk region not seen during training. The system further includes a dashboard for visualizing flood risk and an alert mechanism designed to reduce response delay and potential loss of life. By combining machine learning-based prediction with a practical early warning interface, this work aims to move flood disaster management from a reactive to a proactive model, addressing a gap left open by existing conceptual, descriptive, and narrowly-scoped technical approaches in the literature.**

***Index Terms*— Flood Prediction, Machine Learning, Disaster Management, Early Warning System, Assam, Generalization, Decision Support System, Flood Risk Mapping.**

I. INTRODUCTION

Disaster management refers to the organized process of preparing for, responding to, and recovering from natural or man-made disasters, with the primary aim of reducing loss of life and property. Among natural disasters, floods are among the most frequent and destructive in India, driven largely by monsoon rainfall, river overflow, and increasingly by the effects of climate change. Recent data indicates that India has recorded a 40% rise in flash flood events over the past three years due to climate change, with states such as Uttar Pradesh, Bihar, Assam, and Madhya Pradesh among the worst affected. Assam, in particular, experiences three to four waves of flooding every monsoon season, primarily due to the dynamics of the Brahmaputra River system. In 2024 alone, floods and landslides in Assam affected 42.95 lakh (4.29 million) people across 7,794 villages spanning 35 districts. Despite the well-documented, recurring nature of flooding in the region, existing early warning systems remain largely reactive, alerting communities only after flooding has already begun rather than predicting risk in advance. The scientific rationale behind the proposed system lies in leveraging machine learning to identify flood risk patterns from historical and real-time data, enabling data-backed early warnings rather than post-event response. This project addresses this gap by developing a machine learning-based flood prediction system trained on data from multiple Indian states, and specifically validated on Assam as a high-stakes test case. The approach mirrors real-world deployment scenarios, where a system trained on generally available data must be applied to a new or critical region without direct prior exposure to that region's specific flood patterns. Research on cybersecurity has examined threats and countermeasures across networked systems, cloud platforms, and user authentication. Studies on protecting confidential data

in distributed cloud storage [13] and on network security attacks and countermeasures [14] provide foundations for understanding attack models and defensive strategies. Authentication research image-based authentication via randomized selection [15] and performance analysis of bag-of-password authentication across Python, Java, and PHP [16] explores improved security and implementation efficiency.

II. LITERATURE REVIEW OF EXISTING SYSTEMS

A total of six research papers were reviewed to understand the current state of research in flood disaster management, spanning decision support systems, inter-agency collaboration, image-based classification, stakeholder analysis, urban infrastructure, and framework-based approaches. Each paper is summarized below in terms of its objectives, technology used, methodology, efficiency, and limitations.

2.1. Integrated Decision Support System for Flood Disaster Management:

Table 1. Literature review of "Integrated Decision Support System for Flood Disaster Management with Sustainable Implementation" [1].

Paper 1	Objective	Technology Used	Methodology Used	Efficiency	Issues
Title: *Integrated Decision Support System for Flood Disaster Management with Sustainable Implementation* Journal: IOP Conference Series: Earth and Environmental Science, Vol. 1285 (ICSESTS-2023) Year: 2024 DOI: 10.1088/ 1755-1315/1285/1/012015 URL: https://doi.org/10.1088/1755-1315/1285/1/012015	Investigate whether a Decision Support System (DSS) for flood emergency management is necessary and feasible, covering the pre-flood, flood, and post-flood phases. Develop a Disaster Management (DM) metamodel and language that unifies, enables, and accelerates access to DM information. Integrate multiple existing DSS models into one comprehensive logical model ("model fusion") built from reusable modular subroutines.	Technologies: Decision Support System (DSS), Disaster Management (DM) Metamodel, Entity–Relationship (ER) modeling, Geographic Information System (GIS). Software/Tools: Real-time regime database, geographical information database, social economy database, model base.	Three-Layer Architecture: Expression layer (man–machine interface, electronic river-basin map), Application layer (four subsystems), and Data layer (databases + model base). Four Subsystems: Pre-flood preparation, flood emergency response, post-flood reconstruction, and flood prevention, linked in a feedback loop. Model Fusion & Validation: Modular subroutines merged into an "Observed and Forecasted" model; validated via two-factor ANOVA, logistic regression, and a bushfire case study (Marysville, Australia).	Strengths in brief: 1. Modular & Reusable: Subroutines can be mixed and matched for different disaster types, easing model management. 2. Full Lifecycle Coverage: Integrates pre-flood, flood, and post-flood phases into one connected system. 3. Data-Driven Forecasting: Combines real-time, GIS, and socio-economic databases with a model base for accurate, timely decisions.	Weaknesses in brief: 1. Weak Statistical Significance: The logistic regression model (χ^2 (1) = 2.77, p = .096) does not show clear overall statistical significance. 2. Limited Real-World Validation: Tested mainly on a hypothetical /Simulated crisis and a bushfire case study rather than live flood deployments. 3. Integration Complexity: Merging many domain-specific subroutines into one metamodel needs further refinement for each disaster type.

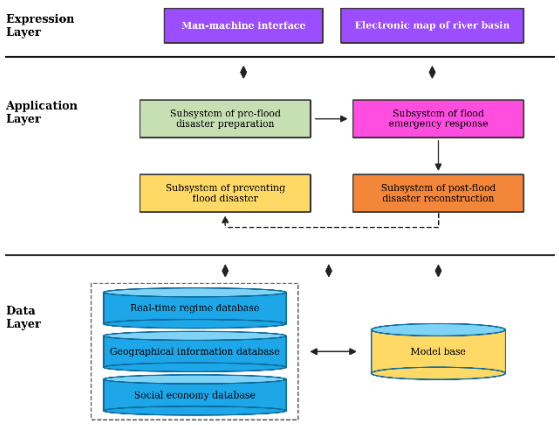


Fig. 1. Three-layer architecture of the integrated flood decision support system, showing the expression, application and data layers [1].

Work on cybercrime and open-source software [17], strategic information-security planning using decentralized identity [18], and machine-learning malware detection [19] highlights the need for layered security, proactive threat detection, and user-oriented protection. A hybrid deep-learning intrusion-detection approach for AWS cloud [20] extends these ideas to cloud security, while the survey of confidential-data protection in distributed storage [13] emphasizes privacy and secure storage. Collectively, these studies establish a foundation for combining authentication, intrusion detection, machine learning, and secure cloud architecture against increasingly complex threats.

2.2. Interorganizational Collaboration in Disaster Management:

Table 2. Literature review of "The Dynamics of Interorganizational Collaboration in Disaster Management: A Network Study Based on Flood Disasters in China" [2].

Paper 2	Objective	Technology Used	Methodology Used	Efficiency	Issues
Title: *The Dynamics of Interorganizational Collaboration in Disaster Management: A Network Study Based on Flood Disasters in China* Journal: International Journal of Disaster Risk Science, Vol. 14 (Springer) Year: 2023 DOI: 10.1007/s13753-023-00525-7 URL: https://doi.org/10.1007/s13753-023-00525-7	Investigate how interorganizational collaboration networks operate and evolve across the preparedness and response phases of a flood disaster in a centralized state (China). Examine how vertical (hierarchical) and horizontal (peer-level) collaboration mechanisms shift in importance between the two phases, using Complex Adaptive Systems (CAS) theory. Identify which organizations (central, provincial, local) play the most central/bridging roles in coordinating	Technologies: Social Network Analysis (SNA), Complex Adaptive Systems (CAS) theory, Quadratic Assignment Procedure (QAP) regression. Software/Tools: Gephi (network visualization), UCINET (network analysis), STATA (statistical testing). Input Format: Coded textual data: news reports, joint department issuances, and institutional/legal documents (396	Case & Data: 2020 Hubei Province flood (1 Apr–31 Aug 2020), split at 22 June into preparedness and response phases; text manually coded (Cohen's kappa = 0.921). Network Construction: Built an 82×82 undirected collaboration matrix; coded horizontal (same-level)	1. Rigorous Mixed Method: Combines SNA visualization with QAP regression for statistically robust, objective network testing (high inter-coder reliability, kappa = 0.921). 2. Clear Phase-Wise Insight: Shows network density nearly tripled in the response phase (0.322 → 0.935), clearly evidencing the horizontal-to-vertical collaboration shift. 3. Real, Large-	1. Limited Data Sources: Relies only on official government news reports and published data, which may under-report informal or undocumented collaboration. 2. Government-Only Scope: Focuses solely on government organizational

	information and resources during flood disaster management.	text units, 82 organizations).	and vertical (cross-level) linkages plus control variables (GDP, population density, geography, river system).	Scale Case: Grounded in an actual major disaster (2020 Hubei floods) with 82 real organizations, not a simulated scenario.	networks, excluding NGOs, volunteer groups, and private-sector actors from the analysis.

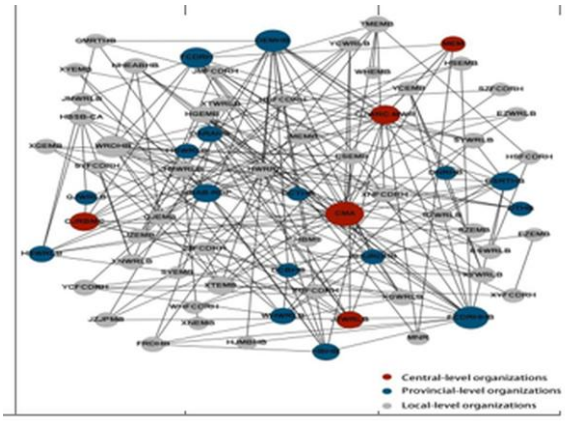


Fig. 2. Collaboration network of the 82 organizations in the 2020 Hubei flood response (central-, provincial- and local-level organizations) [2].

The survey of computer-vision-based driver-monitoring systems [17] demonstrates visual intelligence for road safety and driver-risk identification, while the Python for data analytics textbook [21] supports programming-driven analytical problem-solving. Machine-learning malware detection [22] shows how predictive models strengthen threat identification, and comparative password-authentication implementations [23] emphasize evaluating security mechanisms across programming environments. Research on quantum computing for supply logistics [24] highlights emerging computational paradigms for complex logistics optimization. The confidential-data protection in distributed cloud storage [25] reinforces privacy and data integrity, soft clustering for gene-expression data [26] demonstrates computational intelligence in biomedical analysis, and digital health and medical-tourism innovations [27] illustrate connected healthcare services.

2.3. CNN and Sorting Algorithm for Flood Image Classification:

Table 3. Literature review of "An Integrated Convolutional Neural Network and Sorting Algorithm for Image Classification for Efficient Flood Disaster Management" [3].

Paper 3	Objective	Technology Used	Methodology Used	Efficiency	Issues
Title: *An Integrated Convolutional Neural Network and Sorting Algorithm for Image Classification for Efficient Flood Disaster Management* Journal: Decision Analytics Journal, Vol. 7 (Elsevier) Year: 2023 DOI: 10.1016/j.dajour.202	Use autonomous drones with image classification to identify and grade the severity of flood-affected areas. Prioritize relief delivery from the most affected to the least affected areas while factoring in distance for efficiency. Demonstrate the	Technologies: Convolutional Neural Networks (Inception v3, DenseNet-121), Merge Sort algorithm, autonomous drones. Software/Tools: Python, OpenCV, PyTorch;	Image Classification: Flood images are preprocessed (format change, normalization, resizing) and classified by Inception v3 and DenseNet into Minor/Moderate/Major flooding, each assigned a severity weight (1–3).	1. Good Classification Accuracy: Inception v3 and DenseNet reach 83% and 81% accuracy respectively on the custom flood dataset. 2. Efficient Prioritization: Merge sort combines	1. Small, Imbalanced Dataset: Only 785 images in total, unevenly split across classes, causing mild overfitting during training. 2. Weak Major/Moderate Separation: Both models struggle to distinguish

3.100225 URL: https://doi.org/10.1016/j.dajour.2023.100225	efficacy of combining a CNN with a sorting algorithm for autonomous decision-making in disaster relief.	transfer learning from ImageNet pre-trained weights; NVIDIA Tesla K80 GPU. Input Format: Self-made image dataset (aerial + ground-level flood photos: Minor/Moderate/Major flooding classes).	Distance Weighting: Drone-to-location distance is estimated (single-camera depth calculation), and each weight is divided by distance to get a final priority score. Merge Sort: Final scores are sorted (O (N log N)) so the drone delivers relief from the highest-priority to the lowest-priority location.	severity and distance so relief reaches the most urgent, reachable areas first. 3. Low-Cost, Autonomous: Works with a single camera and modest onboard computer, removing the need for manual survey/decision-making.	Major_flooding from Moderate_flooding (the confusion matrix shows frequent mix-ups). 3. Simplified Distance Estimation: Distances were randomly generated (1–10 m) rather than measured in real flight, limiting real-world validation.

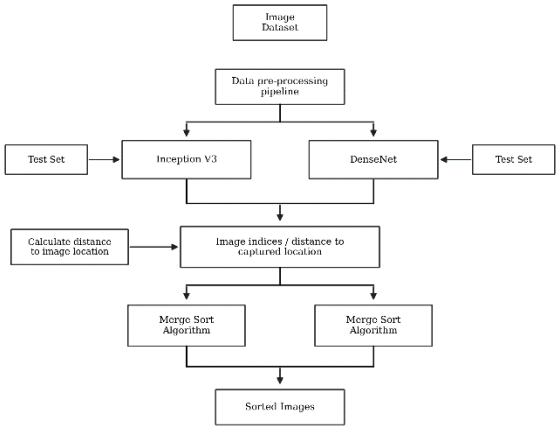


Fig. 3. Workflow of the CNN and merge-sort pipeline for flood image classification and prioritization [3].

Foundational texts further consolidate the methodological basis for AI-driven disaster management applications. Faisal et al. (2026) provides a comprehensive treatment of data science, machine learning, and IoT integration in Healthcare Application Using Data Science, Machine Learning and IoT, detailing predictive modelling pipelines and sensor-based architectures that are directly adaptable to flood forecasting and early warning contexts [12]. Together, these works substantiate the technical feasibility of deploying AI-based predictive systems for disaster management, including flood risk prediction [11], [12].

2.4. Disaster Management 'Deeds' - KwaZulu-Natal Floods

Table 4. Literature review of "Disaster Management 'Deeds' in the Context of April 2022 KwaZulu-Natal Floods: A Scoping Review" [4].

Paper 4	Objective	Technology Used	Methodology Used	Efficiency	Issues
Title: *Disaster Management 'Deeds' in the Context of April 2022 KwaZulu-Natal Floods: A Scoping Review* Journal: International Journal of Disaster Risk Reduction, Vol. 98 (Elsevier) Year: 2023 DOI:	Appraise South Africa's disaster management by examining the 'deeds' (tangible actions/utterances) performed by the state, flood victims, and non-state actors before, during, and after the April 2022	Technologies: Actor Network Theory (ANT), PCC (Population–Concept–Context) framework, PRISMA-based scoping review methodology. Software/Tools: Electronic	Scoping Review (7-step framework): (a) review question, (b) inclusion/ exclusion criteria, (c) search strategy, (d)	1. Rigorous PRISMA Process: Transparent, staged screening (69 → 18 records) ensures only directly relevant, high-quality sources	1. Narrow Source Window: Limited to English-language records from April 2022–February 2023, potentially missing later or

10.1016/j.ijdrr.2023.104122 URL: https://doi.org/10.1016/j.ijdrr.2023.104122	KwaZulu-Natal (KZN) floods. Use Actor Network Theory (ANT) to assess the nature and quality of actor–network associations behind these deeds. Identify gaps in disaster preparedness, mitigation, response, and coordination among key stakeholders.	bibliographic databases: ScienceDirect, Scopus, Google Scholar, ResearchGate, Google. Input Format: 69 records identified (news reports, government reports, journal articles); 18 records retained after screening (5 journal articles, 13 reports).	screening and selection, (e) data extraction, (f) thematic data analysis, (g) presentation of findings.	are analysed. 2. Theory-Grounded Insight: The ANT lens reveals power dynamics, political bias, and uneven actor influence often missed by flat disaster-management frameworks.	non-English evidence. 2. Small Evidence Base: Only 18 records (5 journal articles, 13 reports) underpin the analysis — a fairly thin base for broad conclusions.

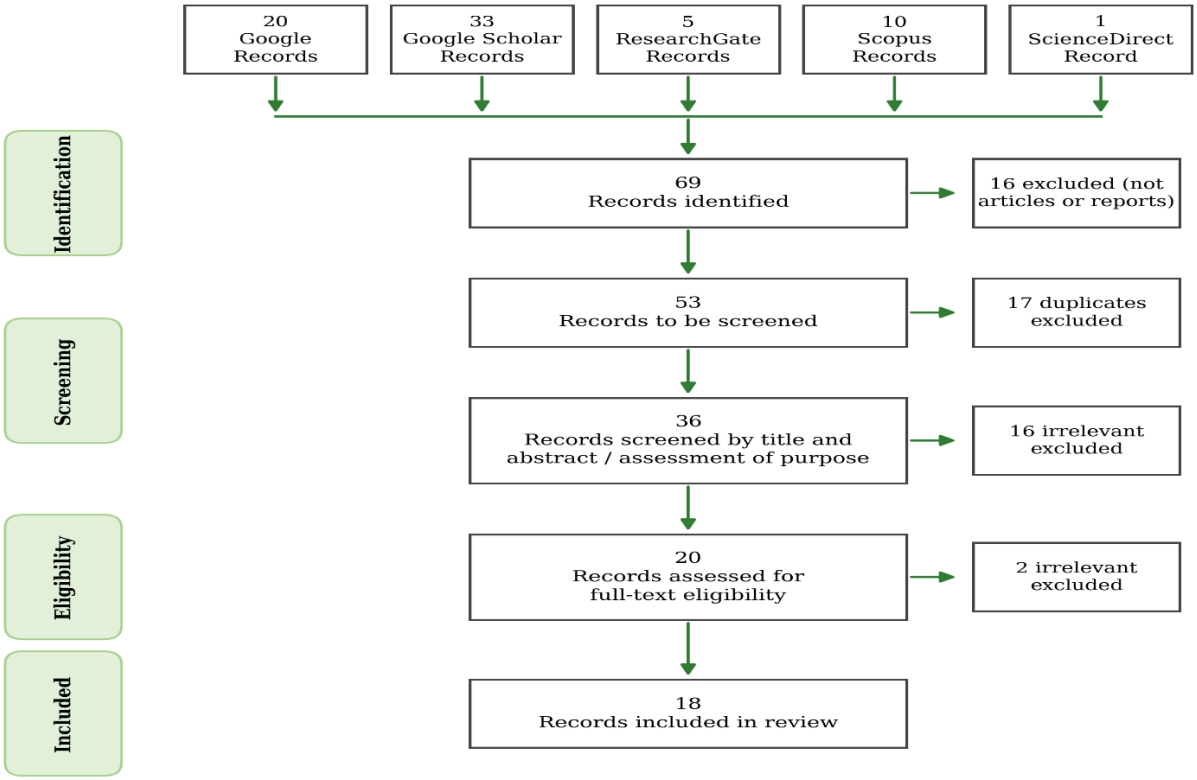


Fig. 4. PRISMA flow diagram of the scoping review: records identified, screened, assessed for eligibility and included [4].

2.5. Urban Green Spaces and Flood Disaster Management:

Table 5. Literature review of "Urban Green Spaces and Flood Disaster Management: Toward Sustainable Urban Design" [5].

Paper 5	Objective	Technology Used	Methodology Used	Efficiency	Issues
Title: *Urban Green Spaces and Flood Disaster Management:*	Examine how rapid urbanization and extreme	Technologies: Green infrastructure	Literature Synthesis: Reviews prior	1. Multi-Benefit Framework:	1. Lacks Quantification: Qualitative

Toward Sustainable Urban Design Journal: Frontiers in Public Health, Vol. 13 (Mini Review) Year: 2025 DOI: 10.3389/fpubh.2025.1583978 URL: https://doi.org/10.3389/Fpubh.2025.1583978	rainfall worsen urban stormwater/flood risk, and highlight the role of urban green spaces in flood resilience. Build an integrated theoretical framework combining green infrastructure with traditional drainage systems ("blue–green–gray" systems) to explain how green spaces intercept, retain, and infiltrate rainwater. Draw on international success stories (Amsterdam, Singapore's ABC Waters, Kazan's "Resilient Belt") to give practical guidance for China's stormwater management.	(rain gardens, green roofs, permeable pavements, wetlands), "Blue–Green–Gray" integrated infrastructure design, IoT sensors for real-time monitoring. Software/ Tools: Storm Water Management Model (SWMM) hydrological simulation (cited from prior studies). Input Format: Qualitative literature review, government precipitation /Urbanization statistics, and international case studies (no primary dataset).	studies (Berland; Li et al.; Wang et al.; World Bank/CRC report) showing that green infrastructure reduces runoff by 20–40%. Case Study Analysis: Examines Amsterdam's canal network, Singapore's ABC Waters Program, and Kazan's "Resilient Belt" to illustrate blue–green synergy and gray–green fusion. Framework Construction: Proposes a "blue–green–gray" stormwater strategy combining spatial co-design, system coupling, and dynamic (sensor-based) regulation.	Links green spaces to public health (reduces stress by 87%, headaches by 52%), water quality, and biodiversity, not just flood control. 2. Evidence from Real Cases: Draws on proven international examples (Amsterdam, Singapore, Kazan) rather than theory alone. 3. Holistic Integration: The blue–green–gray strategy balances engineered and natural systems for efficient, resilient stormwater management.	discussion only; no numeric sizing/layout guidelines or quantified synergy between green–gray and blue–green systems. 2. Limited China-Specific Validation: Cites international cases without deeply testing their adaptability to China's urban context. 3. No Real-Time System Detailed: Does not specify an actual dynamic-monitoring platform or operational strategy for real-time regulation.

2.6. Flood Disaster Management Cycle Framework – Malaysia

Table 6. Literature review of "Proposed Framework for the Flood Disaster Management Cycle in Malaysia" [6].

Paper 6	Objective	Technology Used	Methodology Used	Efficiency	Issues
Title: *Proposed Framework for the Flood Disaster Management Cycle in Malaysia* Journal: Sustainability (MDPI), Vol. 14, Issue 7 Year: 2022 DOI:	Explore the key challenges in managing flood disasters in Sarawak, Malaysia, and identify practical solutions to address them. Rank and validate the most	Technologies: Geographic Information Systems (GIS), flood forecasting and early warning systems, big-data/past-data modeling for vulnerability assessment.	Three-Phase Research Design: (1) literature review of journals, conference proceedings and news articles; (2) qualitative	Strengths in brief: 1. Grounded in Real Local Data: Combines 5 years of actual news reports with expert interviews,	Weaknesses in brief: 1. Small Expert Sample: Findings rely on only 20 expert interviews, which may not capture the full diversity of stakeholder

10.3390/su14074088 URL: https://doi.org/10.3390/su14074088	significant challenges and mitigation measures using expert feedback and past news-report content analysis.	Software/Tools: Qualitative content analysis of secondary news data; Likert-scale (1–5) expert ranking. Input Format: Newspaper articles (2015–2019) on Sarawak floods, journal/conference literature, and unstructured interviews with 20 flood-management experts.	assessment of Sarawak news articles (2015–2019); (3) unstructured interviews with 20 experts, each feedback ranked on a 1–5 Likert scale. Case Study: Sarawak region of Malaysia, chosen for its recurrent yearly flooding (Kuching, Limbang, Samarahan, etc.).	rather than relying on theory alone. 2. Clear, Ranked Priorities: Identifies and ranks the top 5 challenges and top 5 solutions, giving stakeholders a clear action order. 3. Reusable Framework: The proposed cyclical framework is explicitly designed to be adaptable to other flood-prone countries with similar terrain.	views. 2. Qualitative, Subjective Ranking: Challenge/solution rankings are based on Likert-scale opinion rather than quantitative flood-damage data. 3. Framework Not Empirically Tested: The proposed management cycle is conceptual; the paper does not pilot or validate it in a real flood event.

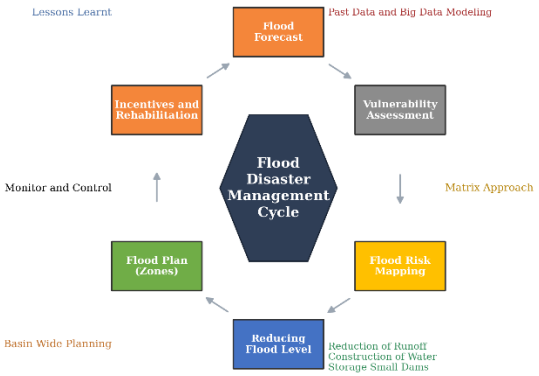


Fig. 5. Proposed flood disaster management cycle framework for Malaysia [6].

III. RESEARCH GAP

Across the six papers reviewed, three broad categories of existing work emerge. The first is conceptual or architectural modeling, as seen in William et al. (2024), which proposes system structures without implementation or real-data testing. The second is descriptive or analytical study, as seen in Li et al. (2023), Mudefi (2023), and Muzamil et al. (2022), which examines coordination, stakeholder behavior, or challenges without building a predictive tool. The third is narrow technical implementation, as seen in Islam et al. (2023), which applies machine learning to a specific sub-problem, namely image classification, within a single, limited dataset. None of the reviewed literature combines a machine learning-based predictive model trained across multiple regions with an explicit test of generalization to a new, high-risk region.

This is the specific gap this project addresses: building a predictive system trained on multi-state data and rigorously validated on Assam, a region deliberately excluded from training, to demonstrate that the model generalizes rather than merely memorizing region-specific patterns. Work on ethics and regulations in AI-driven ophthalmology [28] underscores responsible AI

deployment and patient protection, and the review of NIST-recommended post-quantum cryptographic schemes [29] addresses preparation for quantum-computing threats. UI design research on users' viewing experience across static and dynamic websites [30] highlights presentation and content delivery for usability, while an automatic content creation and rearrangement mechanism [31] improves cloud storage utilization through efficient digital-content organization

IV. PROPOSED SYSTEM DESIGN

The scientific rationale behind the proposed system is grounded in leveraging machine learning techniques to enhance flood risk prediction across regions. The system is designed to process rainfall, river water level, elevation, and historical flood occurrence data collected from multiple Indian states, and to generate a flood risk score for a given region and time period. By training the model exclusively on data from states other than Assam and validating it on Assam, the system directly tests whether data-driven flood prediction can extend meaningfully beyond the specific regions used to build it. Training Data: Drawn from multiple flood-prone Indian states (Bihar, Kerala, Himachal Pradesh, Madhya Pradesh, Uttar Pradesh, Punjab), excluding Assam. Testing Data: Drawn exclusively from Assam, allowing evaluation of the model's generalization capability through a leave-one-region-out validation strategy. Dashboard: A web-based dashboard visualizes flood risk predictions on a map, allowing users to assess risk levels by region. Alert Mechanism: An alert and notification module triggers warnings when predicted risk crosses a defined threshold.

V. ARCHITECTURE DIAGRAM

Figure 1 illustrates the proposed system architecture, structured as a linear data-to-alert pipeline. Rainfall, river water level, elevation, and historical flood records are collected from multiple data sources and passed through a data preprocessing stage. The processed data is split using a leave-one-region-out strategy into training data (all states except Assam) and testing data (Assam), which feed into the machine learning model. The model's predictions are served through a backend API to a frontend dashboard for visualization, and finally to an alert and notification system that issues warnings when risk thresholds are crossed.

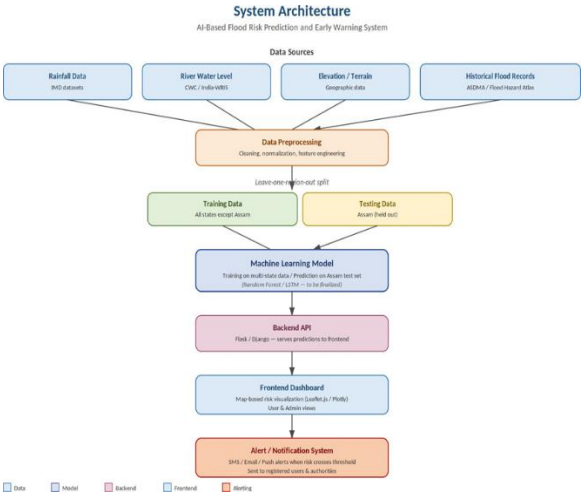


Figure 1: System Architecture of the Proposed Flood Prediction and Early Warning System

VI. METHODOLOGY AND ALGORITHMS USED

The methodology employed in this project involves multiple stages of data collection, preprocessing, feature selection, model training, and performance evaluation, structured as follows: Data Collection: Historical rainfall, river water level, elevation, and flood occurrence data are obtained from the Central Water Commission (CWC), India Meteorological Department (IMD), India-WRIS, the Assam State Disaster Management Authority (ASDMA), and the Flood Hazard Zonation Atlas (1998–2023). Data Preprocessing: Cleaning and normalization of data to handle missing values, outliers, and inconsistencies across different states and time periods. Feature Engineering: Construction of features such as daily/monthly rainfall, river water level relative to danger mark, elevation and terrain data, and historical flood occurrence labels. Model Training: Training of a machine learning model (Random Forest, Logistic Regression, or LSTM, to be finalized based on feature type) on the multi-state training set, excluding Assam. Evaluation and Validation: Performance metrics such as Accuracy, Precision, Recall, F1-score, and a confusion matrix are used to assess prediction performance specifically on the Assam test set. The model candidates under consideration each offer complementary strengths: Random Forest provides robustness to noisy tabular data and resistance to

overfitting across heterogeneous state-level features; Logistic Regression offers a simple, interpretable baseline for binary flood/no-flood classification; and Long Short-Term Memory (LSTM) networks, a type of Recurrent Neural Network, are well suited to capturing temporal dependencies in sequential rainfall and river-level data over time.

VII. PROJECT FUNCTIONAL MODULES IMPLEMENTATION

The development of the proposed system involves several functional modules, each serving a critical role in flood risk prediction and early warning: User Module: Registration, login, and location-based subscription functionality for flood alerts. Prediction Module: The core machine learning model that computes a flood risk score or probability for a given region based on input features. Admin/Dashboard Module: Visualizes predicted risk zones on an interactive map and monitors incoming data streams. Alert/Notification Module: Sends SMS, email, or push notifications to registered users and authorities when predicted risk exceeds a defined threshold.

VIII. EXPECTED EFFICIENCY AND EVALUATION METRICS

Model performance will be evaluated on the Assam test set using the following metrics: Accuracy - overall proportion of correct predictions. Precision - proportion of predicted flood events that were correct. Recall - proportion of actual flood events correctly identified, particularly important since missing a real flood event is a more costly error than a false alarm. F1-score - harmonic mean of precision and recall. Confusion Matrix - visual breakdown of prediction outcomes across classes. For reference, comparable work by Islam et al. (2023) achieved 83% accuracy using a CNN-based image classification approach on a single-region dataset. Given that this project addresses a more challenging cross-region generalization task, results should be interpreted with appropriate context regarding this added difficulty.

IX. CONTRIBUTION AND FINDINGS

This project aims to provide a data-driven, generalizable flood risk prediction and early warning platform, enabling disaster management authorities and at-risk communities to receive advance warning rather than post-event notification. By leveraging machine learning trained across multiple states and rigorously validated on Assam, the system is expected to demonstrate that flood prediction models can extend meaningfully beyond the specific regions used to build them. This directly responds to the gaps identified across the six papers reviewed: the purely conceptual nature of existing decision support system proposals, the narrow single-region scope of existing machine learning applications in this domain, and the repeatedly documented failure of early warning information to reach at-risk communities in time.

X. CONCLUSION\\

This project set out to address a well-documented gap in flood disaster management: the lack of a machine learning-based system capable of generalizing flood risk prediction across regions rather than relying on region-specific historical data alone. By training a model on data from multiple Indian states and rigorously testing it on Assam, a region deliberately withheld from training, this project aims to demonstrate that data-driven flood prediction can extend meaningfully beyond the regions used to build it.

The resulting system, comprising a prediction model, visualization dashboard, and alert mechanism, is intended to serve as a practical step toward more proactive, generalizable flood disaster management. Future work will focus on extending the model to other disaster types, integrating satellite imagery for real-time monitoring, and incorporating green infrastructure and land-use data as additional predictive features.

REFERENCES

[1] P. William, O. J. Oyebode, A. Sharma, N. Garg, A. Shrivastava, and A. L. N. Rao, “Integrated decision support system for flood disaster management with sustainable implementation,” *IOP Conference Series: Earth and Environmental Science*, vol. 1285, Art. no. 012015, 2024.

[2] Z. Li, L. Liu, and S. Liu, “The dynamics of interorganizational collaboration in disaster

management: A network study based on flood disasters in China," *International Journal of Disaster Risk Science*, vol. 14, pp. 979–994, 2023.

[3] M. A. Islam, S. I. Rashid, N. U. I. Hossain, R. Fleming, and A. Sokolov, "An integrated convolutional neural network and sorting algorithm for image classification for efficient flood disaster management," Decision Analytics Journal, vol. 7, Art. no. 100225, 2023.

[4] E. Mudefi, "Disaster management 'deeds' in the context of April 2022 KwaZulu-Natal floods: A scoping review," International Journal of Disaster Risk Reduction, vol. 98, Art. no. 104122, 2023.

[5] N. Liu and F. Zhang, "Urban green spaces and flood disaster management: Toward sustainable urban design," Frontiers in Public Health, vol. 13, Art. no. 1583978, 2025.

[6] S. A. H. B. S. Muzamil, N. Y. Zainun, N. N. Ajman, N. Sulaiman, S. H. Khahro, M. M. Rohani, S. M. B. Mohd, and H. Ahmad, "Proposed framework for the flood disaster management cycle in Malaysia," Sustainability, vol. 14, no. 7, Art. no. 4088, 2022.

[7] Central Water Commission (CWC), India, "River and flood forecasting data." [Online]. Available: https://ffs.india-water.gov.in

[8] India Meteorological Department (IMD), "Rainfall datasets." [Online]. Available: https://mausam.imd.gov.in

[9] India-WRIS, "India Water Resources Information System." [Online]. Available: https://indiawris.gov.in

[10] Assam State Disaster Management Authority (ASDMA). [Online]. Available: https://asdma.assam.gov.in

[11] D. Vekariya, M. K. J. Kannan, S. Gupta, P. Muthusamy, R. Mahajan, and A. K. Pandey, "Recommendation model-based 5G network and cognitive system of cloud data with AI technique in IOMT applications," International Journal of Communication Networks and Information Security, vol. 14, no. 3, pp. 239–256. doi: 10.17762/ijcnis. v14i3.5609.

[12] M. Faisal, M. Senthil Vadivu, T. R. Shree Nee, and M. A. Safar, Healthcare Application Using Data Science, Machine Learning and IoT. Scientific International Publishing House, Sep. 2026, ISBN: 9789371963350. [Online]. Available: https://sipinternationalpublishers.com/product-detail.php?PID=NDY1Nw==. Also available: https://www.flipkart.com/product/p/itme?pid=9789371963350.

[13] G. D. K., M. K. Singh, and M. Jayanthi, Eds., Network Security Attacks and Countermeasures. IGI Global, 2016, doi: 10.4018/978-1-4666-8761-5.

[14] R. M. Balajee, M. K. Jayanthi Kannan, and V. Murali Mohan, "Image-based authentication security improvement by randomized selection approach," in Inventive Computation and Information Technologies. Singapore: Springer, 2022, pp. 61–71.

[15] M. K. J. Kannan, N. Aaijaz, K. G. Mani, and V. Tewari, The Future of Innovation and Technology in Education: Trends and Opportunities. S&M Publications, 2025.

[16] B. R. M. and M. K. J. Kannan, "Intrusion detection on AWS cloud through hybrid deep learning algorithm," Electronics, vol. 12, Art. no. 1423, 2023, doi: 10.3390/electronics12061423.

[17] K. L. S. Kumar and M. K. J. Kannan, "A survey on driver monitoring system using computer vision techniques," in Innovative Computing and Communications, A. E. Hassanien, S. Anand, A. Jaiswal, and P. Kumar, Eds., Lecture Notes in Networks and Systems, vol. 1021. Singapore: Springer, 2024, doi: 10.1007/978-981-97-3591-4_21.

[18] M. K. J. Kannan, "A bird's eye view of cybercrimes and free and open-source software to detoxify cybercrime attacks—An end user perspective," in Proc. 2nd Int. Conf. Anti-Cyber Crimes (ICACC), Abha, Saudi Arabia, 2017, pp. 232–237, doi: 10.1109/Anti-Cybercrime.2017.7905297.

[19] M. K. J. Kannan, "Artificial intelligence and blockchain technology for human resource management," Scientific International Publishing House, 2025.

[20] M. K. Jayanthi, "Strategic planning for information security—DID mechanism to befriend the cyber criminals to assure cyber freedom," in Proc. 2nd Int. Conf. Anti-Cyber Crimes (ICACC), Abha, Saudi Arabia, 2017, pp. 142–147, doi: 10.1109/Anti-Cybercrime.2017.7905280.

[21] S. K. Shukla, U. Dwivedi, and M. K. J. Kannan,

Python for Data Analytics: Practical Techniques and Applications. JSR Publications, 2024.

[22] P. Jain, I. Rajvaidya, K. K. Sah, and J. Kannan, "Machine learning techniques for malware detection—A research review," in *Proc. IEEE Int. Students' Conf. Electrical, Electronics and Computer Science (SCEECS)*, Bhopal, India, 2022, pp. 1–6, doi: 10.1109/SCEECS54111.2022.9740918.

[23] B. R. M., M. M. V., and M. K. J. K., "Performance analysis of bag of password authentication using Python, Java and PHP implementation," in Proc. 6th Int. Conf. Communication and Electronics Systems (ICCES), Coimbatore, India, 2021, pp. 1032–1039, doi: 10.1109/ICCES51350.2021.9489233.

[24] M. K. J. Kannan and T. R. Shree Nee, "Qubits unveiled: A deep dive into quantum computing and its revolutionary potential for supply logistics," in Qubits Unveiled: Quantum Computing Solutions for Efficient Supply Logistics, P. Gaba, A. Panwar, V. Jain, and R. Kannan, Eds. Nova Science Publishers, 2025, pp. 273–293, doi: 10.52305/WSXW8884.

[25] H. Naik and M. K. J. Kannan, "A survey on protecting confidential data over distributed storage in cloud," SSRN, Dec. 1, 2020. [Online]. Available: https://ssrn.com/abstract=3740465.

[26] E. Kavitha, R. Tamilarasan, A. Baladhandapani, and M. K. J. Kannan, "A novel soft clustering approach for gene expression data," Computer Systems Science and Engineering, vol. 43, no. 3, pp. 871–886, 2022, doi: 10.32604/csse.2022.021215.

[27] T. R. Shree Nee, M. K. J. Kannan, and K. Mariyappan, "Digital health and medical tourism innovations for digitally enabled care for future medicine: The real-time project's success stories," in Navigating Innovations and Challenges in Travel Medicine and Digital Health. IGI Global Scientific Publishing, 2025, pp. 325–344, doi: 10.4018/979-8-3693-8774-0.ch016.

[28] M. K. J. Kannan, S. N. T. Ramesh, and K. Mariyappan, "Ethics and regulations in AI-driven ophthalmology," in Generative Artificial Intelligence in Ophthalmology, B. K. Mishra, A. Kumar, K. Mariyappan, V. Tiwari, P. S. Rathore, and G. H. Das, Eds., ch. 34. Scrivener Publishing, 2026. [Online]. Available: https://www.scrivenerpublishing.com/cart/title.php?id=1341.

[29] K. Mariyappan, B. K. Mishra, S. A. Harini, and M. K. J. Kannan, "A review on NIST-recommended cryptographic schemes for the post-quantum computer world era," in Proc. 8th EAI Int. Conf. Big Data Innovation for Sustainable Cognitive Computing (BDCC), 2025. Springer Nature. [Online]. Available: https://link.springer.com/book/9783032392329.

[30] R. M. Balajee, M. K. J. Kannan, and V. Murali Mohan, "Web design focusing on users viewing experience with respect to the static and dynamic nature of web sites," in *Inventive Computation and Information Technologies*, S. Smys, V. E. Balas, and R. Palanisamy, Eds., Lecture Notes in Networks and Systems, vol. 336. Singapore: Springer, 2022, doi: 10.1007/978-981-16-6723-7_5.

[31] R. M. Balajee, M. K. J. Kannan, and V. Murali Mohan, "Automatic content creation mechanism and rearranging technique to improve cloud storage space," in *Inventive Computation and Information Technologies*, S. Smys, V. E. Balas, and R. Palanisamy, Eds., Lecture Notes in Networks and Systems, vol. 336. Singapore: Springer, 2022, doi: 10.1007/978-981-16-6723-7_7.