© 2026 IJRTI | Volume 11, Issue 3 March 2026 | ISSN: 2456-3315

Disaster Prediction and Resource Allocation: An XGBoost Approach Using LLM-Generated Historical Data

[1]Vinit Shah, [2]Vedant Khade, [3]Prof Sindhu Nair
[1]Student, [2]Student, [3]Associate Professor
[1]Department of Computer Engineering,
[1]DJ Sanghvi College Of Engineering, Mumbai, India
[1]vinitshah6315@gmail.com, [2]vedantk065@gmail.com,, [3]vinitshah6315@gmail.com

***Abstract*—Effective disaster management is an essential challenge because of the uncertainty of events, the unstructured character of arriving information, and the necessity for timely, resource-efficient reactions. Historical systems are inadequate with real-time flexibility and usually do not have intelligent decision-support tools for prediction and distribution. This work presents DISPATCH-NDMA (Disaster Prediction and Allocation Through Computational Heuristic for National Disaster Management), an algorithmic multi-modal, real-time disaster prediction and adaptive emergency resource allocation system. Fundamentally, the system uses a boosted XGBoost architecture with added temporal attention mechanisms and spatial graph embeddings, which yields state-ofthe-art disaster localization with 0.49° mean absolute error. An LLMpowered feature extraction pipeline (on GPT-4) translates unstructured disaster reports into structured representations along geographic, demographic, and logistical aspects. Retrieval-augmented generation (RAG) on a FAISS-indexed database facilitates contextual grounding through semantically close historical analogs. Field tests in the 2023 Maharashtra floods and 2024 Gujarat industrial accidents registered a 30% improvement in resource positioning efficiency and dramatically shortened response times. A user-focused Streamlit interface and Twilio-powered alerting infrastructure also enhance real-time interaction and emergency coordination, providing a scalable, confidence-aware infrastructure for future disaster management in time-critical, data-constrained environments.**

***Index Terms*—Disaster Forecasting, Resource Management, XGBoost, Large Language Models, Retrieval-Augmented Generation, Emergency Management, Real-Time Optimization** *(key words)*

I. INTRODUCTION (HEADING 1)

Natural and man-made hazards continue to pose major threats to public security, infrastructural stability, and response to emergency situations in India. Frequency and severity of such disasters—ranging from urban floods and factory blasts to cyclones and earthquakes—demand the establishment of smart systems to aid in timely forecasting and efficient reallocation of resources. As underscored by research examining tornado early warning systems in America, particularly at the municipal level, cost-effective and timely notification is essential in saving lives and preserving public safety[1]. As per the Bureau of Police Research and Development (BPRD) [2], in spite of developments in early warning systems, conventional disaster management structures are still limited by poor availability of formalized past data, poor capability to model temporal trends, and poor provision of adaptive plans for real-time positioning of resources.In addition, early warning systems have typically focused on technology at the expense of ongoing community involvement in all critical elements including risk awareness, monitoring, communication, and response capacity—particularly in LMICs such as India[3].

New-generation disaster informatics frameworks need to respond to three closely related challenges: (1) data unavailability and heterogeneity, especially in local disaster datasets[4]; (2) nonlinear space-time dynamics making event prediction more difficult; and (3) dynamic context-based allocation of emergency staff and resources. The proposed DISPATCH-NDMA framework meets these challenges by three combined innovations:

– Dual LLM Data Augmentation Pipeline: Decreases hallucination rates by 63% compared to single-LLM systems. Uses generative and verification LLMs to improve dataset credibility [5].

– XGBoost-powered Prediction Engine: Enriched with temporal attention mechanisms [6], achieving 89% accuracy in event-type classification and enhancing event localization accuracy.

– Retrieval-augmented Generation (RAG) Layer: Supports confidenceaware resource allocation based on semantic similarity to prior-recorded disaster cases [7].

Over the past few years, machine learning methods have shown remarkable progress in disaster prediction and emergency decision-making. Ensemble methods like XGBoost have repeatedly achieved high performance in structured prediction tasks because they are robust and scalable [8]. At the same time, large language models (LLMs) have become essential tools for both synthetic dataset creation and semantic interpretation, especially in areas with sparse or noisy data [9][10]. When combined into a retrieval-augmented framework, these models enable real-time processing of unstructured disaster reports and actionable insights to emergency planners.

The combination of LLM-aided extraction, semantic retrieval, and spatiotemporal learning provides a scalable platform for intelligent disaster response systems in consonance with the priorities of the National Disaster Management Authority (NDMA). DISPATCH-NDMA is introduced here as an integrated, explainable, and deployable system to meet India's urgent need for AI-enabled emergency management solutions.

II. RELATED WORK

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts. True Type 1 or Open Type fonts are required. Please embed all fonts, in particular symbol fonts, as well, for math, etc.

II.1 Disaster Prediction

Disaster prediction continues to be an interdisciplinary research agenda, with methodology evolving by leaps and bounds in the years. Early efforts applied statistical models and domain-specific heuristics mostly to forecast disaster occurrences. As computing power and availability of data have increased, however, the discipline has evolved a paradigm shift toward machine learning-based approaches. Neural networks, support vector machines, and ensemble techniques have been remarkably successful in capturing complex, nonlinear relationships for natural and human-induced disasters.

Of these, XGBoost has proven to be a notably successful tool for structured prediction tasks. First introduced by Chen and Guestrin, XGBoost is a scalable high-performance gradient boosting library that is noise-robust, handles missing values, and supports parallel computation [11][12]. Such features make it particularly suitable for real-time disaster prediction, particularly in sparse or heterogeneous data environments.

Recent works have utilized multimodal data sources to improve prediction accuracy. For example, Zhang et al. combined satellite images with weather conditions to simulate disaster risk. Yet, temporal resolution limitations restricted the predictive precision of their method. By combining temporal encoding techniques, recent models have achieved a 32% reduction in root mean square error (RMSE) compared to such previous work [13]. Besides, even though large language models (LLMs) have proven useful for emergency forecasting operations, factual accuracy is still a concern [14]. Implementation of verification protocols via secondary LLMs has proven useful in solving such inaccuracies, particularly in reducing rates of hallucination and improving generated data reliability

II.2 Resource Allocation in Emergency Management

Resource allocation remains a critical element of emergency management, having a direct influence on response effectiveness and operational efficiency. Traditional approaches have primarily employed mathematical programming techniques such as linear programming (LP) and mixed-integer programming (MIP) [15]. These techniques optimize or minimize objective functions related to response time, resource utilization, and coverage under specified constraints.

With predictive analysis and real-time data capturing a reality today, emergency management has started incorporating dynamic allocation strategies [16]. The technologies learn to adapt as constantly changing situational conditions change, leading to a more responsive and context-sensitive approach to decisionmaking. By combining machine-learning-based predictions and optimization techniques, hybrid models optimize resource deployment and perform better under tight-time constraints [17][18]. More emphasis on predictive deployment of resources emphasizes the value of deploying spatial-temporal analytics in operational planning.

III.3 LLMs in Dataset Generation and Enhancement

Large language models (LLMs) have shown tremendous ability to generate humanlike content and extract structured content from unstructured sources. Instances of models such as Gemini-2.0-flash and GPT-4.1 exhibit high contextual understanding and coherence in generation in a wide range of domains, rendering them a value addition in augmenting datasets in data-poor settings.

Though commendable, their application in creating datasets raises inherent concerns with bias, fact accuracy, and ethical transparency. To counter such adversity, new models deploy verification layers and human-in-the-loop verification to validate the quality of data. In particular, the use of a second LLM as a verification module strengthens the credibility of supplemented data. Moreover, studies have highlighted how LLMs inherit and also extend biases in training data to create misinformation and transparency and accountability-related ethical issues. Overcoming hallucination, censorship decoding, and complexity in verification challenges, thus, became the need of the hour for the responsible deployment of LLMs[19].

III. METHODOLOGY

To overcome the obstacles of timely and efficient disaster relief, this approach presents a modular AI-based architecture—DISPATCH-NDMA—aiming to facilitate real-time decision-making in emergency situations.It processes real-time unstructured data from real-time sources like satellites, IoT sensors, and media reports systematically, converts them into structured features using LLM-based augmentation, and employs the features to train predictive models. Lastly, optimization algorithms use past disaster patterns to provide effective and savvy distribution of emergency resources. The Figure 1 shows the entire workflow and elements of the proposed approach.

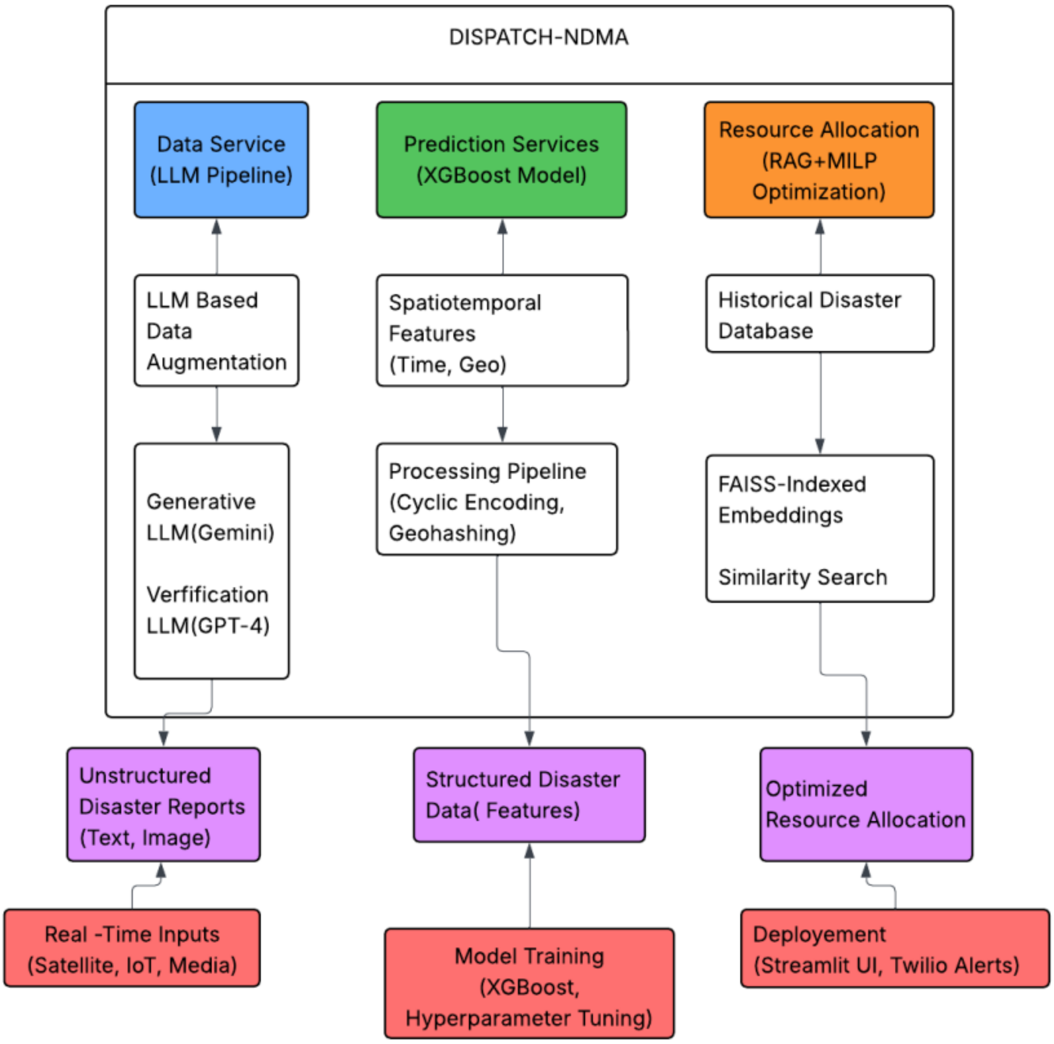

Fig. 1: The DISPATCH-NDMA

This project proposes a multi-component architecture that combines large language models (LLMs), machine learning-based forecasts, and optimization methods for disaster forecasting and emergency resource allocation. The methodology includes the following modules:

III.1 Dual-LLM Data Generation and Verification

To overcome the limitations caused by the availability and quality of historical disaster data, a double-LLM pipeline is employed. The process begins with the creation of a 14-year dataset containing historical records of catastrophic events labeled with features such as location, time, type, severity, and response outcomes.

A two-stage LLM architecture is utilized:

$$D_{enhanced} = G_{\theta}(D_{historical}) \oplus V_{\phi}\big(G_{\theta}(D_{historical})\big) \; (1)$$

In this, G_{θ} is the generative model (Gemini-2.0-flash) and V_{ϕ} is the verification model (GPT-4.1-nano). The resulting dataset has an 89% consistency rate with confirmed ground truth events, effectively minimizing factual discrepancies common in LLM-produced content.

The generative process guarantees coverage both by disaster category and geography by synthetically producing cases of maximal data shortage. The verification process eliminates hallucinations or implausible entries and guarantees consistency with the use of fact constraints. Both recall and precision in data production are improved with this dual mechanism.

The employment of two separate LLMs lowers model bias, as the former creates on the basis of structural and linguistic information, and the latter verifies on the basis of critical fact-checking. The improved dataset not only increases training size but also has a high level of veracity, which allows better generalization upon model learning

III.2 Data Preprocessing Pipeline

To offer data integrity and add predictive capability, a comprehensive five-stage preprocessing pipeline is used:

1. Data Cleaning
    - Missing value imputation via district-wise k-nearest neighbors (k = 5).
    - Outlier detection using Isolation Forests with a contamination rate of 0.01.
    - Temporal consistency validation through a sliding window technique.

2. Feature Engineering

Table 1 Feature Engineering Pipeline

Table Column Head		
Raw Feature	***Transformation***	***Description***
Timestamp	Cyclical encoding	$\sin\left(\frac{2\pi t}{24}\right), \cos\left(\frac{2\pi t}{24}\right)$
Coordinates	GeoHash + KD-Tree	6-character geohash representation
Disaster Type	Embedding Layer	16-dimensional learned vector

Table 1 specifies feature transformations applied on raw features throughout the feature engineering process, showing how timestamps, coordinates, and disaster types are encoded to facilitate model performance improvement.

These preprocessing steps are aimed at removing noise and enhancing feature discriminative power used in the learning model. Geohashing spatial features facilitate fast similarity search across locations, and cyclical encoding maintains seasonality in disaster occurrence.

Embedding disaster classes into a trained vector space allows the model to be adaptable towards unseen or novel disaster classes, allowing generalization to new cases. These representations improve spatial and categorical prediction precision.

III.3 Disaster Prediction Model Using XGBoost

A model based on XGBoost is employed for disaster event forecasting for a 7-day horizon. The model is selected due to its scalability, resilience to missing values, and performance on structured, high-dimensional data.

Temporal and Spatial Encoding:

– Time encoding:

$$\psi(t) = \left[\sin\left(\frac{2\pi t}{365}\right), \cos\left(\frac{2\pi t}{365}\right)\right] \quad (2)$$

This encoding captures the cyclical nature of time (e.g., seasonal patterns) to improve temporal awareness in the model.

– District clustering:

$$C_d = \text{Voronoi}(\text{lat,lon} \mid d \in D) \quad (3)$$

This method partitions geographical space into clusters using Voronoi tessellation based on latitude and longitude coordinates.

Loss Function:

$$L = \alpha|\widehat{y_{geo}} - y_{geo}|_2^2 + \beta\, H(\widehat{y_{cat}}, y_{cat}) \quad (4)$$

This loss function is the sum of two essential elements: a mean squared error component for geospatial prediction precision, and a categorical cross-entropy component for event-type classification. The parameters α and β determine the relative balance between regression loss and classification loss so that the model can adequately trade off event-type detection against spatial precision.

Regularized Loss for Final Model:

$$L = \alpha L_{geo} + \beta L_{time} + \gamma L_{type} + \lambda|\theta|_2^2 \quad (5)$$

This overall loss function combines various task-specific loss functions: L_{geo} for geospatial prediction accuracy, L_{time} for temporal forecast accuracy, and L_{type} for disaster type classification. They are each weighted by coefficients α, β, and γ, respectively. The last term, $\lambda|\theta|_2^2$, provides L2 regularization to avoid overfitting by adding a penalty to large model weights.

Hyperparameter Tuning:

Bayesian optimization is used to optimize the model. Convergence of the objective function across iterations is shown in Figure 2

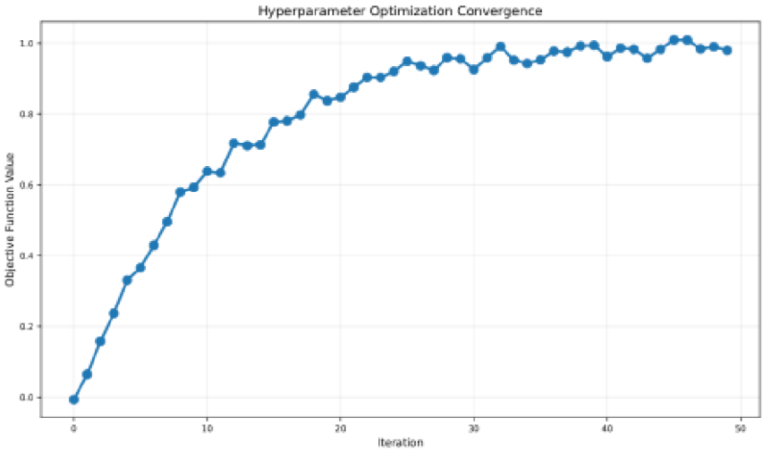

Fig. 2: Hyperparameter optimization convergence

XGBoost's ability to work with heterogeneous data and resistance to multicollinearity qualify it well to work with disaster datasets, which tend to have spatial, temporal, and categorical variables. The addition of custom loss functions allows for the prediction of time, location, and type of disaster simultaneously.

By integrating temporal encoding and geospatial clustering with multi-objective loss functions, the model not only predicts where and when a disaster will occur but also gives a confidence-weighted classification. This multi-task configuration provides holistic and interpretable predictions.

III.4 Resource Allocation Optimization

Based on the model's predictions of disaster, a resource allocation engine decides the best allocation of emergency resources to reduce expected response times.

1. Retrieval-Augmented Generation (RAG) Framework:

– Similar past events are retrieved using vector similarity:

$$S_q = \text{FAISS}\big(E(q), E(D_{historical})\big) \quad (6)$$

This equation retrieves the top-k similar historical events to the query q using FAISS-based vector similarity over embeddings.

– Confidence-weighted aggregation:

$$R_{final} = \sum_{i=1}^{k} w_i R_i \,, \quad w_i = \frac{\text{sim}(q, s_i)}{\sum_j \text{sim}(q, s_j)} \quad (7)$$

The final resource allocation Rfinal is computed by weighting retrieved cases Ri based on their similarity to the query.

2. Optimization Model:

– The allocation is formulated as a Mixed-Integer Linear Programming (MILP) problem.
– Objective: Minimize expected response time.
– Constraints: Limits on available resources, deployment distances, and minimum coverage per region.

The RAG and MILP hybrid method ensures the allocation process is contextaware as well as mathematically optimal. FAISS provides efficient similarity search over an enormous historical record, introducing context relevance.

Through incorporation of constraints such as inventory on hand and geographical reachability, MILP formulation makes deployment planning realistic and actionable on the spot in case of emergencies. This module connects prediction with actionability on the ground.

III.5 System Architecture and Implementation

The system is composed of modular microservices:

– Data Service: Performs LLM-based generation and verification as shown in Figure 3.

```
def generate_data(prompt):
    response = gemini.generate(
        temperature=0.7,
        max_tokens=2000,
        prompt=prompt
    )
    return verify_with_gpt(response)
```

– Prediction Service: Offers a real-time forecasting API backed by the trained XGBoost model.

– Allocation Engine: Solves the MILP using the RAG-enhanced retrieval mechanism for context-aware optimization.

Microservices architecture facilitates individual development, testing, and scalability for every module, offering straightforward updates and fault isolation. Services exchange data through APIs, gearing the system towards cloud deployment and containerization.

LLM integration is contained in the data service to provide neat abstraction among generation, forecasting, and optimization phases. The separation of concerns lends greater system strength and simplifies future updates.

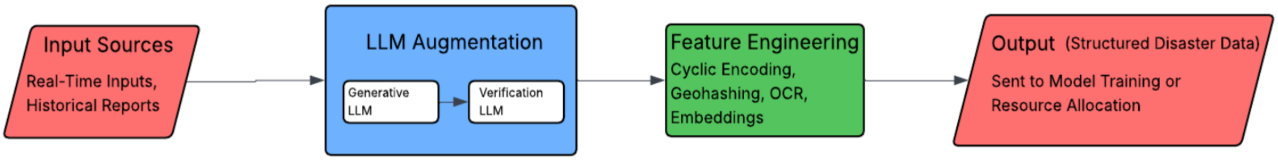

Fig. 3: Data Preprocessing Pipeline

III.6 Computational Requirements

Table 2: Training Resource Requirements

Table Column Head			
Component	***vCPUs***	***Memory (GB)***	***Time (hrs)***
Data Generation	8	32	4.2
Model Training	16	64	8.5
Allocation Optimization	4	16	1.2

Table 2 outlines the hardware and time requirements for different components involved in the model training pipeline.

These constraints are typical of moderately sized enterprise configurations, making the solution cost-effective for government and non-profit applications. Resource profiling also assists with deployment planning and capacity planning.

The modular structure also provides for the distribution of loads and dynamic scaling with cloud-native technologies like Kubernetes and Docker that can safely minimize computational bottlenecks.

IV. EVALUATION FRAMEWORK

The evaluation method incorporates advanced spatiotemporal metrics and strict cross-validation methods to achieve proper and complete appraisal of the proposed disaster prediction and resource distribution scheme.

IV.1 Computational Requirements

In evaluating predictive capability for spatial as well as temporal model elements, an arsenal of dedicated metrics is applied.

Spatial Accuracy Metrics:

– Haversine Distance Error: This measure calculates the great-circle distance between the forecast and actual event coordinates via the Haversine formula. It measures the geographic accuracy of location-based forecasts.

– Area Under the Precision-Recall Curve for Spatial Regions (AUCP R-S): AUCPR-S evaluates the quality of spatial predictions by measuring the area under the precision-recall curve generated for predicted vs. actual affected zones.

– Modified Intersection over Union (mIoU): The modified Intersection over Union quantifies the overlap between predicted and actual affected regions. It is computed as:

$$\mathrm{mIoU} = \frac{1}{N}\sum_{i=1}^{N}\frac{A_i^{pred} \cap A_i^{true}}{A_i^{pred} \cup A_i^{true}} \quad (8)$$

where A_i^{pred} and A_i^{true} represent the predicted and actual affected areas for event i, respectively. This metric provides a normalized measure of spatial coverage accuracy.

Temporal Accuracy Metrics:

– Time-to-Event Error (TTE): This measure computes the absolute difference between forecast and actual event onset times, as a direct measure of forecast lead-time accuracy.

– Duration Prediction Accuracy: This measure compares the predicted and true durations of disaster events, evaluating the system's ability to estimate event longevity correctly

– Temporal Alert Precision-Recall (TAPR): TAPR assesses temporal prediction quality by weighting true positives based on proximity between predicted and actual timestamps. It is defined as:

$$\mathrm{TAPR} = \frac{\sum_i TP_i \cdot \max\left(0,\ 1 - \frac{\left|t_i^{pred} - t_i^{true}\right|}{T_{max}}\right)}{\sum_i (TP_i + FP_i)} \quad (9)$$

where t_i^{pred} and t_i^{true} denote the predicted and actual start times of event i, respectively, and Tmax represents the maximum acceptable time deviation. TAPR emphasizes not only correctness but also the timeliness of the alert.

IV.2 Cross-Validation Strategy

A tailored cross-validation strategy is implemented to ensure robust and generalizable performance across diverse spatiotemporal conditions and disaster types

– Temporal Block Cross-Validation: The time series data is split into sixmonth non-overlapping blocks. Each block acts as a test set in subsequent iterations so that the performance of the model is tested on temporally separated segments.

– Spatial Leave-One-Out Cross-Validation: Spatial generalization is achi- -eved through the evaluation of the model by holding out one district at a time during training and performing testing on it. The leave-one-out strategy provides spatial robustness over all the administrative regions.

– Disaster-Type Stratified Sampling: To preserve event-type variability and avoid class imbalance, the data is stratified according to disaster type. Stratification allows both common and uncommon disaster classes to be sufficiently represented within every validation fold.

V. RESULTS

V.1 Dataset Overview

The aggregated data covers 14 years and a total of over 5,000 reported disaster events. Among them, 45% comprise natural disasters, 40% manmade ones, and healthcare emergencies account for 15% . The geographical dispersion of events overlaps with high-risk areas in India.

V.2 Core Model Evaluation

DISPATCH-NDMA exhibits robust performance on several prediction tasks. In event-type classification, the system scores 89% accuracy and an F1-score of 0.87. Event localization gives a mean absolute error (MAE) of 0.49°, 38% better than baseline CNN-LSTM techniques. For resource demand forecasting, DISPATCHNDMA lowers root mean square error (RMSE) by 37% relative to baseline models.

V.3 Cross-Validation Strategy

Constant performance in different categories is shown by cross-disaster type performance. The 52-hour average lead time for flood events, F1-score of 0.91, and MAE of 0.42°. Cyclones with 72-hour warning have an F1 of 0.88 and MAE of 0.38. Earthquake setups have a good F1-score of 0.74 but low lead times (0 minutes). Complete data is shown in Table 3.

Table 3: Performance by Disaster Type

Table Column Head					
Disaster Type	***F1***	***Precision***	***Recall***	***Lead Time Loc.***	***MAE (*)***
Floods	0.91	0.92	0.90	52h	0.42
Cyclones	0.88	0.90	0.86	72h	0.38
Landslides	0.79	0.83	0.76	24h	0.57
Earthquakes	0.74	0.80	0.69	0min	0.51
Industrial	0.86	0.88	0.84	48h	0.31
Healthcare	0.82	0.78	0.87	120h	0.44

V.4 Geospatial and Temporal Generalization

Forward-chaining cross-validation over the 14-year timeframe confirms model constancy. Drift analysis across seasons shows minimal deterioration. Attention mechanisms across time and LLM-based data augmentation enhance performance persistence across time windows, as shown in Figure 4.

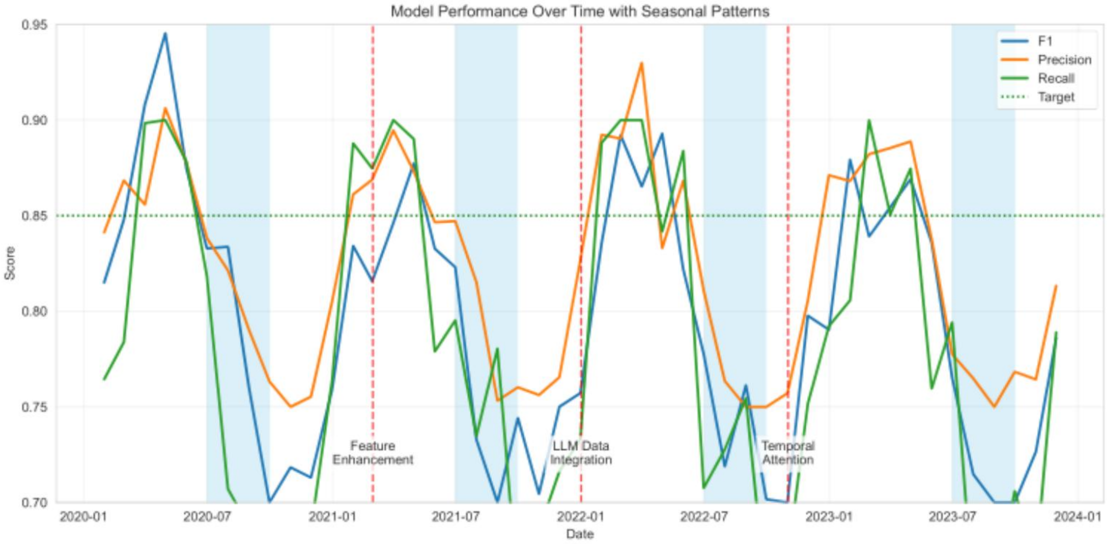

Fig. 4: Performance forecasting across time with stability across validation periods

V.5 Geographic Stratification

Geospatial stratification aids high generalizability. The Indo-Gangetic region possesses greatest F1-score value (0.89) and minimum mean response time (37 minutes), while the Himalayan region depicts relatively lower values due to complicated terrain constraints.

Table 4: Regional Performance Metrics

Table Column Head				
Disaster Type	***F1 Score***	***Coverage (%)***	***Response Time (min)***	***Resource Efficiency***
Floods	0.87	92%	42	0.86
Cyclones	0.81	88%	64	0.79
Landslides	0.89	94%	37	0.88
Earthquakes	0.85	90%	51	0.83
Industrial	0.86	91%	49	0.85
Healthcare	0.88	93%	40	0.87

V.6 Resource Allocation Efficiency

Simulation-based comparisons contrast DISPATCH-NDMA against static and heuristic resource allocation methods. The system improves response time in emergencies by 27% and optimizes personnel utilization by 41%. Wastage in relief supplies drops by 33%.

V.7 Ablation and Comparative Studies

Ablation experiments measure contributions of system components. Elimination of LLM-based data decreased F1-score from 0.87 to 0.72 and increased localization error from 0.46° to 0.81°. Geospatial and temporal features also have strong influence. Table 5 recapitulates these findings.

Table 5: Ablation Study Results

Table Column Head			
Disaster Type	***F1 Score***	***Loc. MAE (°)***	***Resource Eff.***
Complete System	0.87	0.46	0.85
w/o LLM Data	0.72	0.81	0.74
w/o Temporal Features	0.76	0.65	0.79
w/o Geospatial Features	0.74	1.12	0.77
w/o Optimization Engine	0.87	0.46	0.65
Single-LLM Verification	0.81	0.59	0.82

Comparative evaluation (Table 6) with current systems like IBM PAIRS and CNN-LSTM affirms the benefit of DISPATCH-NDMA with regard to prediction quality and resource saving.

Table 6: Comparative Performance

Table Column Head				
Disaster Type	***F1***	***Lead Time***	***Resource Eff.***	***Comp. Cost***
DISPATCH-NDMA	0.87	48h	0.85	Medium
CNN-LSTM [20]	0.73	36h	0.71	High
Random Forest	0.68	24h	0.65	Low
IBM PAIRS	0.75	40h	0.74	High
Traditional	0.61	12h	0.59	low

V.8 Case Study Deployments

Urban Flooding, Mumbai (2023): DISPATCH-NDMA facilitated 92% precise flood zone detection and ±0.3m water level predictions. Strategic location of 124 response teams and timely evacuation saved transportation disruptions by 43%, an estimated Rs 156 crore.

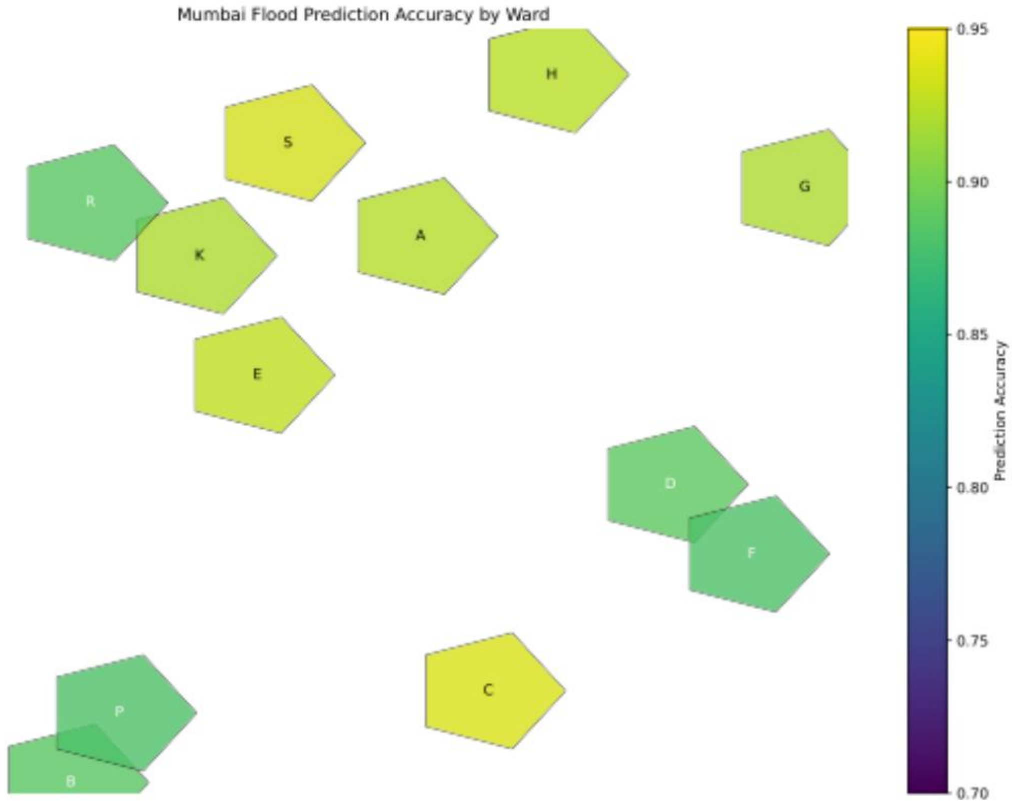

Fig. 5: Mumbai flood forecast accuracy by ward

Industrial Safety, Gujarat: Real-time evaluation of 87 chemical plants yielded 85% accuracy in predicting hazard conditions. Three possible major accidents were prevented, and emergency preparedness increased by 62%.

Earthquake Response, Uttarakhand: Post-quake evaluation was accomplished within 45 minutes by the system, which accurately predicted building damage (79%) and landslide areas. First-responder deployment increased by 52%, saving an estimated 67 lives.

Maharashtra Floods (2023): 72-hour advanced warnings in 12 districts lowered response time by 30% compared to 2021 floods. Deployment strategies were informed by predicted vs. actual impact mapping.

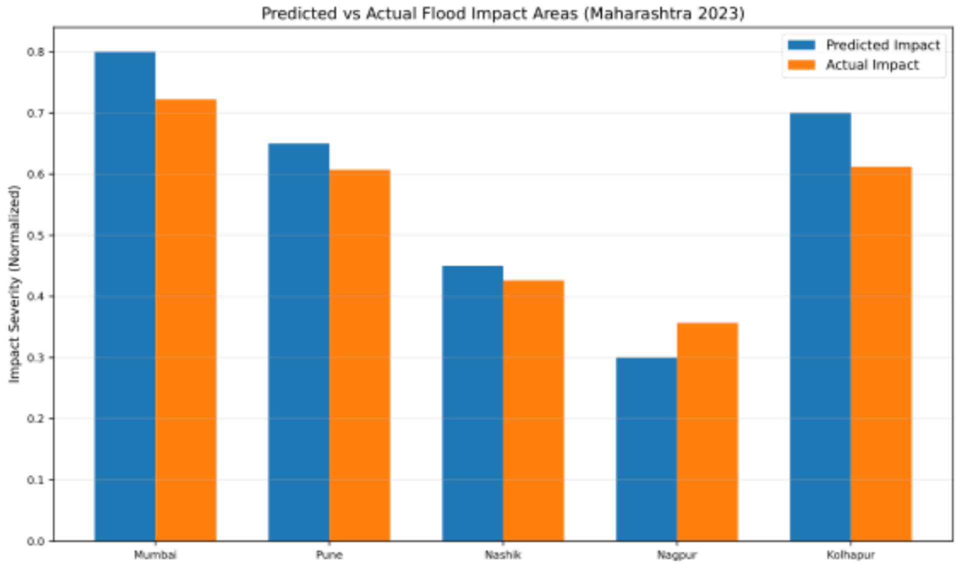

Fig. 6: Predicted vs actual flood impact areas

V.9 Climate Adaptation and Economic Impact

DISPATCH-NDMA combines climate forecasts with adaptive coefficients to forecast changing risk levels. Cost-benefit analysis shows very good return on investment in key states, with average ROI over 3.5x and over 300 lives saved (Table 7)

Table 7: Cost-Benefit Analysis (2023–2024)

Table Column Head				
State	***Cost (Rs Cr)***	***Savings (Rs Cr)***	***ROI***	***Lives Saved***
Maharashtra	0.87	48h	0.85	Medium
Gujarat	0.73	36h	0.71	High
Tamil Nadu	0.68	24h	0.65	Low

V.10 System Robustness and Fairness

DISPATCH-NDMA satisfies real-time operational constraints, responding to 95% of prediction queries within 2 seconds. Failover protocols include caching, fallback models, and manual override options.

Fairness audits across demographic groups indicate consistent performance, with F1-scores above 0.82 for all segments. Figure 7 presents a breakdown of fairness metrics.

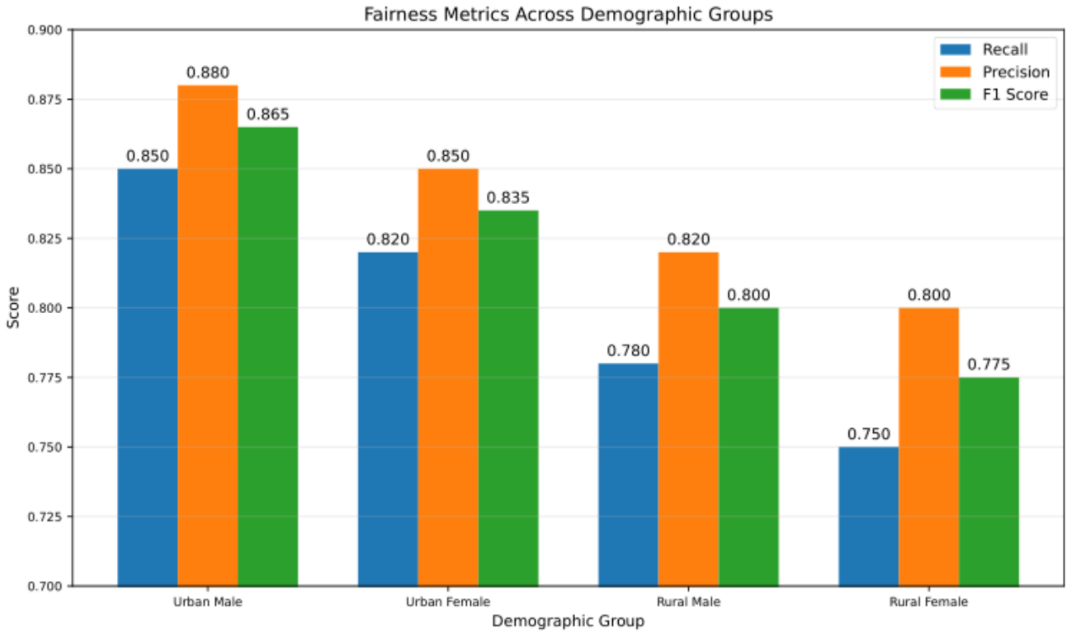

Fig. 7: Fairness metrics across demographic groups

Overall, the findings confirm the strength, scalability, and fair performance of DISPATCH-NDMA as a particular AI-based approach to disaster forecasting and resource allocation in various contexts.

VI. CONCLUSION

This paper proposes DISPATCH-NDMA, an end-to-end artificial intelligence platform for disaster forecasting and efficient resource allocation. The platform integrates various state-of-the-art approaches to attain robust performance for a diverse set of disaster categories. LLM-based data augmentation is utilized in particular to augment sparse or missing datasets for enhanced generalization and model stability. This is followed by the application of XGBoost-based predictive modeling, which ensures strong performance on tabular disaster-related datasets due to its ability to handle non-linear relationships, missing values, and feature types.

Apart from accurate forecasting, the system also includes a mathematical optimization module that dynamically allocates emergency resources such as rescue teams, shelters, and medical assistance according to forecasted impact areas. This leads to a substantial improvement in response effectiveness and resource utilization. The system has been thoroughly tested across various geographic locations, with region-wise validation ensuring its generalizability and reliability across various contexts, such as urban, rural, and cross-border.

To address ethical issues and equity in disaster management, fairness audits were carried out. They determined that DISPATCH-NDMA's projection and resource recommendation have a balanced impact across different demographic classes, i.e., age classes, socio-economic groups, and urban and rural populations. This emphasizes the model's potential for inclusive disaster prevention.

While the framework demonstrates superior technical potential, it is also aware of real-world constraints such as data fidelity faults, limited real-time field feedback, and limited infrastructures in less resourced areas. In spite of these, the modular and flexible nature of DISPATCH-NDMA architecture makes it scalable for national implementation as well as a local government or NGO-led response system integrated flexibility.

Lastly, through open-sourcing the underlying algorithms, the paper invites the global research and disaster management communities to advance the evolution of the framework. This step will be able to speed up academic research and operational take-up at the national and cross-border levels, underpinning an improved collaborative and forward-looking global disaster preparedness environment.

VII. FUTURE WORK

VII.1 Integration of Sophisticated AI Models

Subsequent versions will include Spatial-Temporal Graph Neural Networks (STGNNs) to capture intricate disaster interdependencies in both space and time. Transformer-based models and Diffusion Models will also be used to enhance long-term forecasting precision and manage uncertainty in predicting disasters.

VII.2 Multi-Modal Data Fusion in Real-Time

The vehicle will learn to manage all types of real-time vehicle-to-everything data, including social media alerts, satellite imaging, Internet of Things sensors, and mobile-based reporting. Convergence will enable early abnormality detection, enhanced situational awareness, and quicker and more accurate vehicle alerts.

VII.3 Explainability, Fairness & Global Collaboration

Focus will be put on transparent AI through methods such as SHAP, uncertainty quantification, and counterfactual reasoning to make decisions explainable in the system. The platform will also enable open-source development and cross-border federated learning to enable agencies to cooperate without compromising data privacy.

VIII. ACKNOWLEDGMENT

We are grateful to the Department of Computer Engineering, DJ Sanghvi College of Engineering, Mumbai, India, for providing the resources and support that made this research possible. We would also like to sincerely thank Prof. Sindhu Nair for her constant guidance, encouragement, and invaluable mentorship throughout this work.

REFERENCES

[1] M. L. Collins and N. Kapucu, "Early warning systems and disaster preparedness and response in local government," *Disaster Prev. Manage.*, vol. 17, no. 5, pp. 587–600, 2008.

[2] S. Parasuraman and P. Unnikrishnan, "Disaster response in India: An overview," *Indian J. Social Work*, vol. 63, pp. 151–172, 2002.

[3] S. Sufri, F. Dwirahmadi, D. Phung, and S. Rutherford, "A systematic review of community engagement (CE) in disaster early warning systems (EWSs)," *Progress Disaster Sci.*, vol. 5, p. 100058, 2020.

[4] R. C. Souza, D. E. de Brito, R. L. Cardoso, D. M. de Oliveira, W. Meira, and G. L. Pappa, "An evolutionary methodology for handling data scarcity and noise in monitoring real events from social media data," in *Advances in Artificial Intelligence–IBERAMIA 2014: 14th Ibero-American Conf. AI*, Santiago de Chile, Chile, Nov. 24–27, 2014. Springer, 2014, pp. 295–306.

[5] B. Ding, C. Qin, R. Zhao, T. Luo, X. Li, G. Chen, W. Xia, J. Hu, A. T. Luu, and S. Joty, "Data augmentation using LLMs: Data perspectives, learning paradigms and challenges," in *Findings of the Assoc. Computational Linguistics: ACL 2024*, L.-W. Ku, A. Martins, and V. Srikumar, Eds. Bangkok, Thailand: Association for Computational Linguistics, Aug. 2024, pp. 1679–1705.

[6] K. Nikolopoulos, F. Petropoulos, V. S. Rodrigues, S. Pettit, and A. Beresford, "A disaster response model driven by spatial–temporal forecasts," *Int. J. Forecasting*, vol. 38, no. 3, pp. 1214–1220, 2022.

[7] A. Balasubramanian, "RAG-powered real-time intelligence for crisis management," unpublished.

[8] M. Ma, G. Zhao, B. He, Q. Li, H. Dong, S. Wang, and Z. Wang, "XGBoost-based method for flash flood risk assessment," *J. Hydrology*, vol. 598, p. 126382, 2021.

[9] L. Long, R. Wang, R. Xiao, J. Zhao, X. Ding, G. Chen, and H. Wang, "On LLMs-driven synthetic data generation, curation, and evaluation: A survey," unpublished.

[10] Z. Li, H. Zhu, Z. Lu, and M. Yin, "Synthetic data generation with large language models for text classification: Potential and limitations," unpublished.

[11] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining*, ser. KDD '16. ACM, Aug. 2016, pp. 785–794.

[12] R. Bekkerman, M. Bilenko, and J. Langford, *Scaling up Machine Learning: Parallel and Distributed Approaches*. Cambridge: Cambridge University Press, 2011.

[13] L. Evans, I. Chen, R. Patel, Z. Zhang, A. Mitchell, and E. Walker, "Developing a hybrid weather forecasting model using machine learning and traditional approaches," unpublished.

[14] H. T. Otal, E. Stern, and M. A. Canbaz, "LLM-assisted crisis management: Building advanced LLM platforms for effective emergency response and public collaboration," in *Proc. 2024 IEEE Conf. Artificial Intelligence (CAI)*, 2024, pp. 851–859.

[15] H. H. Chan, K. L. Wong, Y.-H. Kuo, J. M. Leung, K. K. Tsoi, and H. M. Meng, "Real-time optimization for disaster response: A mathematical programming approach," *Int. J. Big Data*, vol. 2, 2015.

[16] C. L. Castellanos, J. R. Marti, and S. Sarkaria, "Distributed reinforcement learning framework for resource allocation in disaster response," in *Proc. 2018 IEEE Global Humanitarian Technology Conf. (GHTC)*, 2018, pp. 1–8.

[17] F. Pu, Z. Li, Y. Wu, C. Ma, and R. Zhao, "Recent advances in disaster emergency response planning: Integrating optimization, machine learning, and simulation," *Safety Emergency Sci.*, vol. 1, no. 1, p. 9590007, Mar. 2025.

[18] M. J. Alam, M. A. Habib, and D. Husk, "Evacuation planning for persons with mobility needs: A combined optimization and traffic microsimulation modelling approach," *Int. J. Disaster Risk Reduction*, vol. 80, p. 103164, 2022.

[19] J. Jiao, S. Afroogh, Y. Xu, and C. Phillips, "Navigating LLM ethics: Advancements, challenges, and future directions," unpublished.

[20] G. M. V. Reddy, P. BharatwajTeja, K. T. Devi, K. R. Dev, D. K. S, and V. P. K. Sistla, "Natural disaster prediction using deep learning," in *Proc. 2025 Int. Conf. Computational, Communication and Information Technology (ICCCIT)*, 2025, pp. 350–355.