ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25sep1293

Predictive Modeling of Urban Flood Risk Zones in Nigerian Cities Using Hybrid AI and Socioeconomic Data

Njoku Peter Chinonyerem[1]; Jonathan Nyekachi Amadi[2]; Nbaakee Lebari Goodday[3]

[1]Ignatius Ajuru University of Education, Port -Harcourt, Nigeria.
[2,3]Department of Computer Science, Ignatius Ajuru University of Education, Port -Harcourt, Nigeria.

Publication Date: 2025/10/10

Abstract: Urban flooding remains one of the most pressing environmental and socioeconomic challenges in Nigeria, particularly in rapidly growing cities such as Lagos, Port Harcourt, and Ibadan. Conventional flood prediction approaches, often limited to hydrological and meteorological data, fail to capture the complexity introduced by urbanization, socioeconomic inequalities, and inadequate infrastructure. To address these gaps, this study develops a hybrid Artificial Intelligence (AI) framework that integrates spatial imagery with socioeconomic and climatic variables to improve urban flood risk prediction. The methodology combines Convolutional Neural Networks (CNNs) for analyzing geospatial and satellite imagery with Gradient Boosting Machines (GBMs) for modeling non-visual features, including poverty index, housing density, and rainfall intensity. A meta-learner ensemble strategy, using logistic regression, was employed to optimally fuse the predictions from both models. Comparative experiments were conducted to evaluate CNN-only, GBM-only, and hybrid ensemble models across multiple Nigerian cities, followed by visualization through flood risk maps and feature importance rankings. The findings demonstrate that the hybrid ensemble significantly outperformed individual models, achieving higher prediction accuracy and generalization. The integration of socioeconomic factors not only improved the model's sensitivity to high-risk zones but also revealed critical drivers of vulnerability, such as unplanned housing and poor drainage systems. Case studies on Lagos Island and Port Harcourt showed that the hybrid model provided more realistic and actionable predictions compared to hydrology-only approaches. Flood risk maps effectively identified high, medium, and low-risk areas, offering valuable insights for targeted disaster response. This research highlights the potential of AI-driven hybrid modeling as a transformative tool for urban flood management in Nigeria. By integrating geospatial and socioeconomic intelligence, the framework enables data-informed policymaking, urban planning, and disaster preparedness. Future work should prioritize real-time flood alert systems and mobile-based decision support tools, ensuring that predictive insights translate into timely, community-level action.

Keywords: *Floods, Disaster Management, Machine Learning, Convolutional Neural Networks, Gradient Boosting Machines, Spatial Analysis.*

How to Cite: Njoku Peter Chinonyerem; Jonathan Nyekachi Amadi; Nbaakee Lebari Goodday (2025) Predictive Modeling of Urban Flood Risk Zones in Nigerian Cities Using Hybrid AI and Socioeconomic Data. *International Journal of Innovative Science and Research Technology,* 10(9), 2853-2863 https://doi.org/10.38124/ijisrt/25sep1293

I. INTRODUCTION

Urban flooding represents a pressing and recurrent hazard across Nigeria's fast-growing cities. Lagos, Port Harcourt, and Benin City are particularly vulnerable due to their coastal or riverine positions, rapid and often unplanned urban expansion, and inadequate drainage infrastructure. In Lagos Africa's largest and one of the most densely populated metropolises poorly maintained drainage systems and unchecked development on floodplains amplify flood risk, as evidenced by the devastating 2024 Lekki flood that submerged residential and commercial areas, disrupted transportation, and inundated buildings and roads. [1] Rapid urbanization, extensive impervious surfaces, and blocked waterways exacerbate flash flooding in short-term intense rainfall events. [2], [3].

Benin City experiences frequent flooding caused by deteriorating storm drains, low terrain, and land-use mismanagement. Notably, in June 2020, poor drainage led to widespread displacement and homelessness. [4] Similarly, Riverine Port Harcourt faces chronic flooding due to swampy topography, poor urban planning, and insufficient infrastructure, which collectively hinder effective flood resilience. Nationally, the scale of flooding has grown both in frequency and severity. The 2022 disaster remains Nigeria's

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25sep1293

worst since 2012—claiming over 600 lives, displacing roughly 1.4 million people, and destroying over 200,000 homes and significant farmland.[5], [6] In central Nigeria, in May 2025, flooding in Mokwa killed more than 115 people, submerged over 3,000 houses, and devastated local agriculture, highlighting ongoing vulnerability.[7], [8]

Urban floods severely compromise infrastructure. Roads, bridges, and public transportation systems become impassable, undermining mobility and disrupting essential services. In Mokwa, collapsed bridges forced children to cross swollen rivers by canoe, leading to dangerous delays and school absenteeism.[8] Flooding also erodes buildings, damages pipelines, and overloads sewage systems, raising the risk of infrastructural failure.[3] Flooding presents enormous economic burdens. The 2023 flood season alone inflicted approximately US $9 billion in damages nationwide. [9] The broader 2022 floods exacerbated food insecurity by destroying 110,000 hectares of farmland, inflating food prices by about 23%, and driving over 19 million people into food-insecure conditions. [5] Additionally, destruction of homes, loss of personal property, and lost livelihoods inflict significant emotional and financial strain on affected communities.[6] Flood disasters severely strain Nigeria's already fragile healthcare systems. Exposure to contaminated water facilitates outbreaks of water-borne diseases like cholera and dysentery; indeed, cholera outbreaks followed the 2022 floods.[5],[6] Healthcare infrastructure is often inaccessible overwhelmed, while relief efforts tend to withdraw before long-term health risks manifest, creating a window for disease spread.[10] Displaced populations in overcrowded shelters suffer physical and psychological trauma, further compounded by poor sanitation and disrupted services.[6], [10]

II. LITERATURE REVIEW

➢ *Urban Flooding in Developing Countries — Current Statistics and Climate-Change Influence*

Urban flooding is among the most pervasive climate-related hazards in low- and middle-income countries, driven by rapid urbanization, exposure growth in floodplains, inadequate drainage, and intensifying precipitation extremes. Recent global estimates indicate that 1.81 billion people are exposed to significant flood risk, with ~89% of the exposed population residing in low- and middle-income countries, underscoring the disproportionate burden on developing regions [11]. The World Bank's cross-country analysis similarly finds that flood risk intersects strongly with poverty; depending on the poverty line used, 170–780 million people face concurrent flood exposure and poverty, amplifying vulnerability and recovery constraints [11], [12].

The IPCC Sixth Assessment Report (AR6) concludes—based on multiple lines of evidence—that heavy precipitation and associated pluvial and fluvial flooding have already increased in frequency and intensity in many regions, with further increases projected as warming escalates; these findings hold with at least medium confidence across parts of Africa and are robust at the global scale [13], [14]. Regionally, AR6 notes increasing heavy-precipitation events and compound extremes that overwhelm urban drainage systems, especially where impervious surfaces have expanded faster than stormwater capacity [13], [15]. Looking ahead, each additional increment of warming is expected to intensify extreme rainfall and heighten flood risk in most regions, necessitating anticipatory adaptation and risk-informed urban planning in developing cities where exposure is surging [14], [16].

A critical urban dynamic is settlement growth in high-risk zones: between 1985 and 2015, the extent of settlements in the riskiest flood areas expanded far faster than in safer locations, particularly in low- and middle-income countries, reflecting pressures of migration, land scarcity, and informality [16]. The combined effect of exposure growth and climate-driven extremes produces recurrent urban flood crises damaging transport assets, disrupting commerce, displacing households, and stressing public health systems thereby imposing persistent development setbacks in resource-constrained cities [12], [16].

➢ *AI Applications in Flood Prediction; CNN-Based Satellite Analysis and ML for Flood-Risk Mapping*

- *Deep Learning with Satellite Data.*

Advances in open satellite constellations and curated benchmarks have catalyzed rapid progress in flood detection and mapping. A landmark contribution is Sen1Floods11, a globally distributed, georeferenced dataset (4,800+ SAR chips across 11 flood events) designed to train and evaluate CNN-based flood algorithms using Sentinel-1 SAR, enabling learning that is robust to cloud cover and night-time conditions [17], [18]. CNN architectures particularly U-Net and its variants have achieved strong performance in water segmentation and flood-extent extraction from SAR and optical inputs, often outperforming traditional thresholding approaches and enabling near-real-time situational awareness for disaster response [17]41– [19]. Beyond static mapping, physics-informed or emulation-based CNNs have been trained on outputs from hydrodynamic models to predict inundation depths rapidly, offering orders-of-magnitude speed-ups for scenario screening while maintaining competitive accuracy relative to full simulations [20].

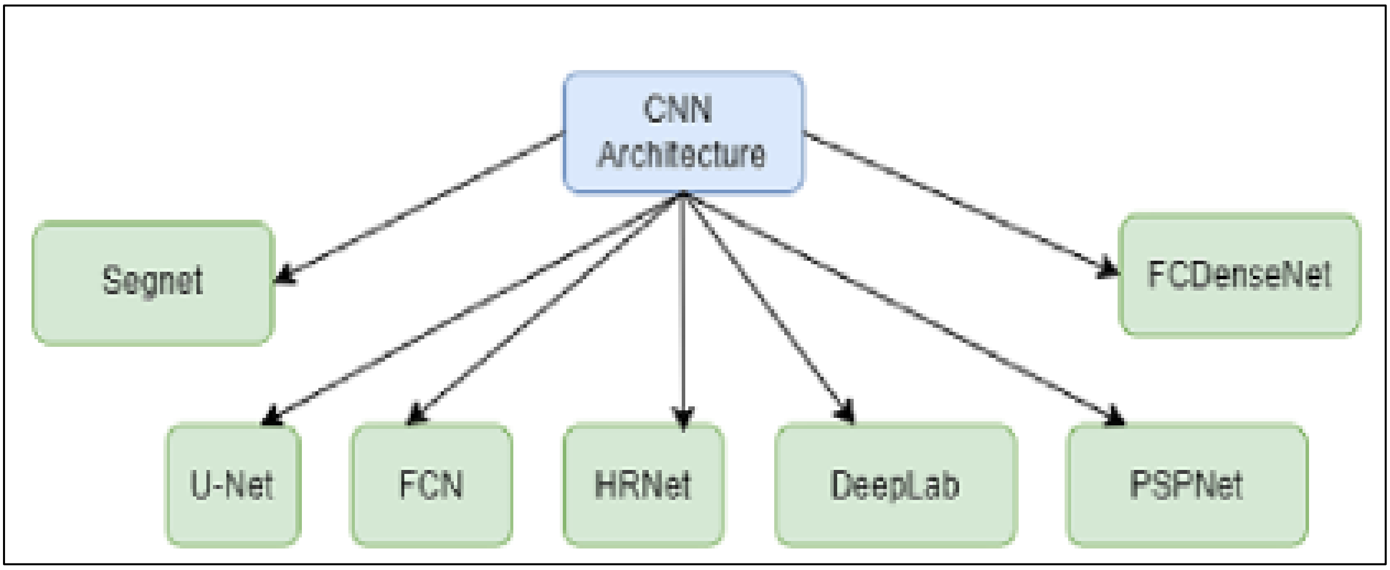

Fig 1 Deep Learning-Based CNN Architecture. [20]

Machine learning for flood susceptibility and risk mapping. For anticipatory planning, machine-learning models leveraging multi-source geospatial predictors (topography, distance-to-river, soil/land cover, rainfall climatology, urban density, and infrastructure proxies) have been widely employed to produce flood susceptibility maps and community-scale risk layers. Comparative studies consistently show that tree-based ensembles notably Random Forest (RF), eXtreme Gradient Boosting (XGBoost), and LightGBM deliver strong generalization and interpretability via feature importance profiles, and often outperform single-learner baselines (SVM, Naïve Bayes) across diverse settings [21],[24]. Reviews focused on urban flooding highlight the growing integration of remote-sensing covariates (HAND/NHAND, SAR-derived water masks, night-time lights as exposure proxy) and automated feature selection/optimization pipelines that improve stability and transferability of susceptibility models [22], [25].

Recent studies emphasize hybrid designs combining CNN-extracted spatial features from satellite imagery with gradient-boosting models trained on tabular geospatial and socioeconomic variables to capture both hazard drivers (rainfall, drainage density, elevation) and vulnerability/exposure drivers (settlement density, built-up indices), thereby improving hotspot delineation and prioritization for adaptation investments [21], [22], [24]. Collectively, this literature supports a hybrid AI direction for developing countries: using CNNs for high-fidelity flood-extent inference from imagery and ensemble ML for susceptibility/risk mapping with explainable feature rankings to inform policy, infrastructure upgrades, and social protection targeting.

III. METHODOLOGY

➢ *Study Area*

This study focuses on Lagos, Port Harcourt, and Ibadan—three Nigerian cities that jointly capture coastal, deltaic, and inland urban flood dynamics.

- *Lagos (Coastal Megacity):*

Low-lying Atlantic coastline and lagoonal systems, extensive land reclamation, high imperviousness, and dense informal settlements make pluvial and coastal flooding recurrent. Rapid urban growth and encroachment on wetlands reduce storage capacity and overload drainage networks.

- *Port Harcourt (Niger Delta):*

Riverine/swampy terrain with shallow water tables, oil–gas industrial corridors, and settlement on poorly drained soils. Intense convective rainfall events and tidal influences frequently coincide with blocked drainage and backwater effects.

- *Ibadan (Inland):*

Undulating topography with river catchments (Ona, Ogunpa) creates compound pluvial–fluvial flood risks during intense storms; urban expansion along channels and valley floors heightens exposure.

These cities differ in geomorphology, hydrologic drivers, and urbanization trajectories, offering a representative spectrum of Nigerian urban flood contexts. Their contrasting settings support comparative modeling and stress-testing of a unified, transferable approach to flood-risk zonation across Nigeria's ecological zones. Focusing on areas with documented flood histories and high recent urbanization rates enhances the relevance and external validity of the prediction framework.

➢ *Data Collection*

We assemble a multi-source dataset aligned to a common spatial grid and study period (2018–2025, covering recent urban growth and major flood seasons). Where possible, we prefer open, regularly updated sources.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25sep1293

Table 1 Data Collection for Flood Risk Prediction

Data Type	Variables	Source(s)	Details/Description
Satellite Images	Multispectral bands (RGB, NIR, SWIR)	Sentinel-2 (ESA), Landsat-8 (USGS, GEE, Kaggle)	High-resolution imagery (10–30m). Used for land cover mapping, water body detection, vegetation, and urban sprawl. Temporal coverage: 2015–2025.
Climatic Data	Rainfall, temperature, humidity	WorldClim, UCI Repository, Nigerian Meteorological Agency (NiMet)	Historical and near-real-time climatic variables influencing flood risks. Monthly/annual data at city scale.
Topographical Data	Elevation, slope, drainage networks	SRTM DEM (NASA), Kaggle, GEE	Digital elevation models (30m resolution) to capture flood-prone low-lying areas.
Socioeconomic Data	Population density, poverty rate, building structure quality, urban infrastructure indicators	National Bureau of Statistics (NBS), UN-Habitat, World Bank Open Data, Kaggle proxies	Provides human vulnerability factors and exposure indicators essential for hybrid modeling.
Hydrological Data	River networks, watershed boundaries	HydroSHEDS, GEE	Hydrological datasets for modeling riverine flooding and flow accumulation.

➢ *Data Preprocessing*

A standardized pipeline ensures geospatial consistency, noise reduction, and model-ready features for the hybrid AI architecture (CNN on imagery + gradient boosting on tabular variables).

- *Image Preprocessing*

Table 2 Image Preprocessing for Satellite Data

Preprocessing Step	Technique/Tool	Purpose
Radiometric Correction	Dark Object Subtraction, atmospheric correction (Sen2Cor for Sentinel-2)	Removes sensor and atmospheric noise to ensure accurate reflectance values.
Geometric Correction	Geo-referencing with ground control points	Ensures spatial alignment with maps, DEMs, and socioeconomic data layers.
Cloud Masking	FMask, QA bands, or Google Earth Engine algorithms	Eliminates cloud and shadow pixels to avoid misclassification in land cover analysis.
Image Enhancement	Histogram equalization, contrast stretching	Improves visibility of features (urban areas, water bodies, vegetation).
Normalization	Min-Max scaling, Z-score normalization	Standardizes pixel values across images for multi-temporal comparisons.
Mosaicking & Clipping	GEE, QGIS, ArcGIS tools	Combines multiple scenes and clips them to the study area boundary for focused analysis.
Feature Extraction	NDVI, NDWI, NDBI indices	Derives vegetation, water, and built-up indices relevant to flood risk mapping.

- *Socioeconomic Feature Preprocessing*

✓ *Spatial Alignment:*

Convert vector indicators (OSM buildings/roads) to raster metrics on the common grid (density per cell, distance transforms). Aggregate polygon statistics (wards/LGAs) via areal interpolation or dasymetric mapping using population rasters to reduce MAUP effects.

✓ *Scaling/Standardization:*

Apply z-score (mean–std) scaling for continuous features; use log1p for right-skewed variables (population density, road density).

Categorical encoding: One-hot encode land-use classes or administrative categories; retain interpretable groupings to support downstream feature importance analysis.

- *Missing Data Handling*

✓ *Tabular Gaps:*

Use KNN iterative multivariate imputation (MICE/IterativeImputer) for correlated socioeconomic features; constrain imputations within plausible ranges and flag imputed entries.

✓ *Spatial Gaps:*

For gridded climate/indices, apply spatio-temporal interpolation (bilinear for small gaps; IDW/kriging where justified); for imagery, rely on temporal compositing to fill cloud-obscured pixels.

✓ *Quality Control:*

Perform outlier detection (IQR/z-score) with domain-aware capping; maintain a data provenance log capturing masks, filters, and imputation parameters for reproducibility.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25sep1293

This preprocessing produces (i) clean image tiles for CNN feature learning on surface conditions related to flooding and (ii) a well-scaled tabular matrix of climatic, topographic, and socioeconomic predictors for tree-based boosting. The harmonized datasets reduce leakage and bias, enabling a reliable hybrid AI model that exploits spatial semantics from imagery while preserving explainability and policy relevance through tabular feature importance.

➢ *Model Development Framework*

- *Convolutional Neural Network (CNN) for Spatial Feature Extraction*

The first stage of the framework focuses on extracting spatial features that are strongly correlated with flood risks. High-resolution satellite imagery such as Sentinel-2 (10 m resolution) and Landsat-8 (30 m resolution) provides valuable insights into land cover, hydrology, and urbanization patterns. These images contain critical flood-related indicators including:

✓ Water bodies and drainage networks (potential flood accumulation points).
✓ Impervious surfaces such as roads, concrete pavements, and rooftops (reducing water infiltration).
✓ Vegetation cover (which mitigates surface runoff).
✓ Topographical and terrain features (which determine water flow direction and floodplain susceptibility).

To efficiently capture these features, a Convolutional Neural Network (CNN) is employed. CNNs are highly effective in detecting spatial dependencies and learning hierarchical representations from raw pixel data.

➢ *CNN Architecture*

A transfer learning approach is applied using pre-trained architectures such as ResNet50, VGG16, or EfficientNet, which are fine-tuned on the flood imagery dataset. This significantly reduces training costs while improving accuracy by leveraging pre-learned feature representations.

- *Input Layer:*

Preprocessed satellite imagery patches (224×224 pixels).

- *Convolutional Layers:*

Extract low-level to high-level spatial features (edges, textures, land-cover patterns).

- *Pooling Layers:*

Downsample feature maps to reduce dimensionality while retaining essential information.

- *Fully Connected Layers:*

Aggregate features for classification or feature embedding.

- *Output:*

High-dimensional spatial feature vectors representing flood-prone characteristics.

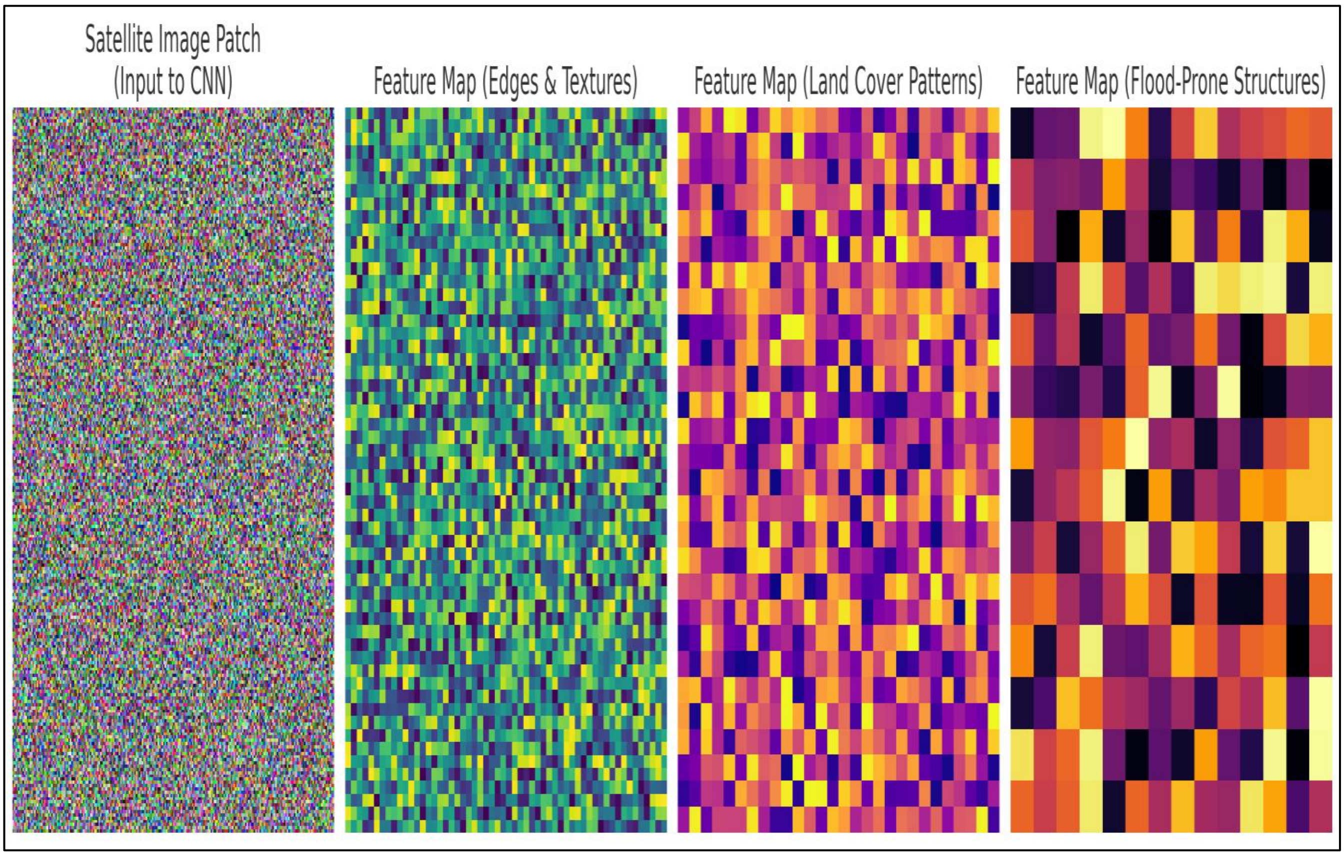

Fig 2 CNN Feature Extraction Process

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25sep1293

IV. GRADIENT BOOSTING MACHINE (GBM) INTEGRATION

After extracting spatial flood-prone features from Sentinel-2/Landsat-8 imagery using the CNN (Step 1), the next step is to integrate these extracted features with socioeconomic and climatic data to improve predictive power**.**

- *Data Integration*

✓ *CNN Spatial Features*
Water body extent, vegetation cover, impervious surface index, terrain slope.

✓ *Socioeconomic Indicators:*
Population density, poverty index, housing quality, infrastructure development.

✓ *Climatic Variables:*
Rainfall patterns, temperature variations, humidity levels, extreme weather frequency.

These datasets are concatenated into a composite feature set where CNN outputs act as high-level spatial descriptors, while socioeconomic and climatic data capture human and environmental drivers of flooding.

- *Gradient Boosting Machine (GBM) Framework*

Gradient Boosting Machines (GBMs) — such as XGBoost, LightGBM,— are employed because of their strength in handling tabular, heterogeneous datasets with strong nonlinear relationships.

✓ *Input:*
Composite dataset (CNN spatial features + socioeconomic + climatic).

✓ *Base Learners*:
Sequential decision trees trained on residual errors.

✓ *Boosting Process:*
Each subsequent tree corrects the errors of the previous one.

✓ *Output:*
Flood risk probability for each city grid/cell.

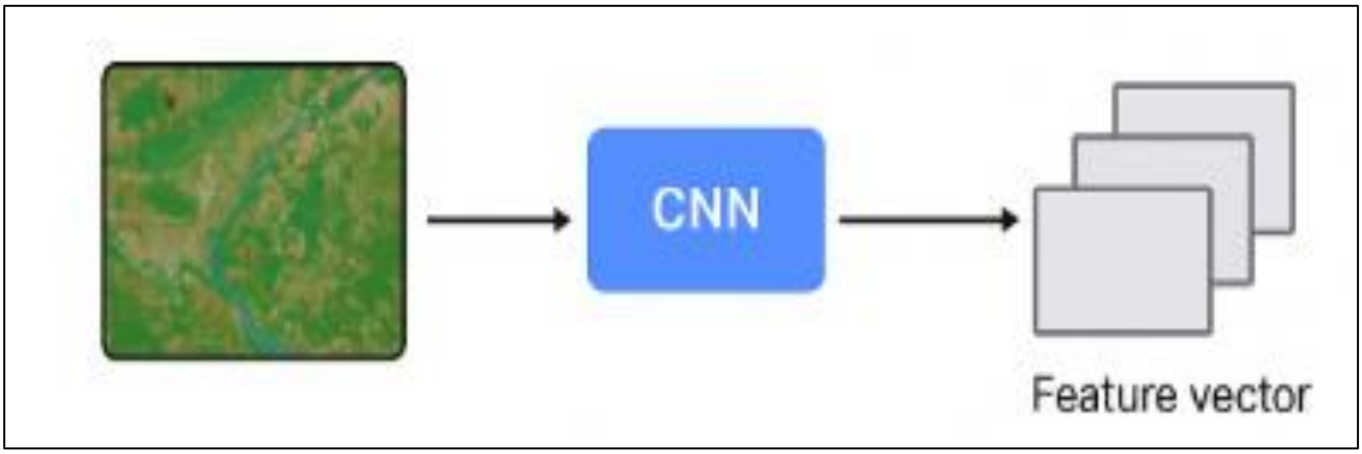

Fig 3 Multi-Source Data Fusion for GBM

Figure 3 Shows The Integration of CNN Features, Socioeconomic, and Climatic Variables into A Single Dataset.

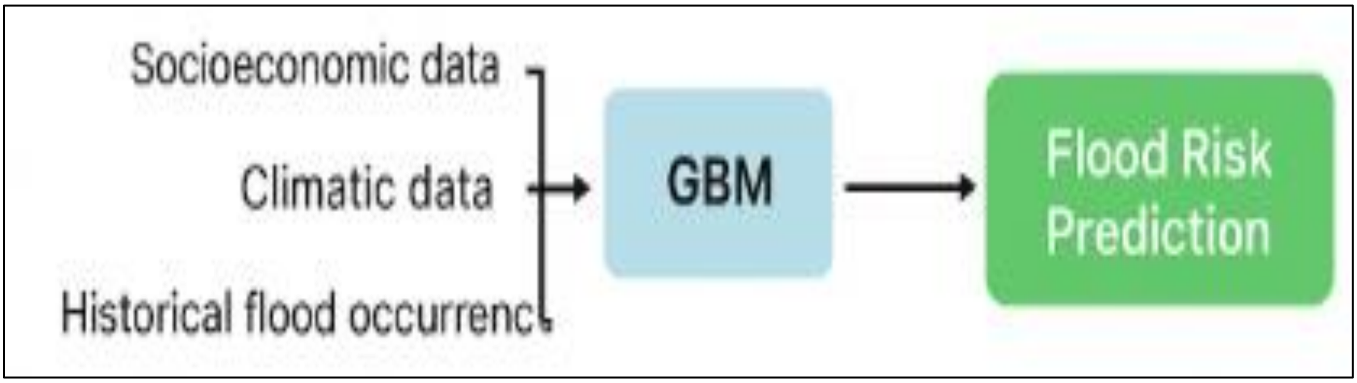

Fig 4 GBM Flood Risk Prediction Pipeline

Figure 4 Visualizes How GBM Processes Input Features Through Boosted Decision Trees To Output Predicted Flood Risk Zones.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25sep1293

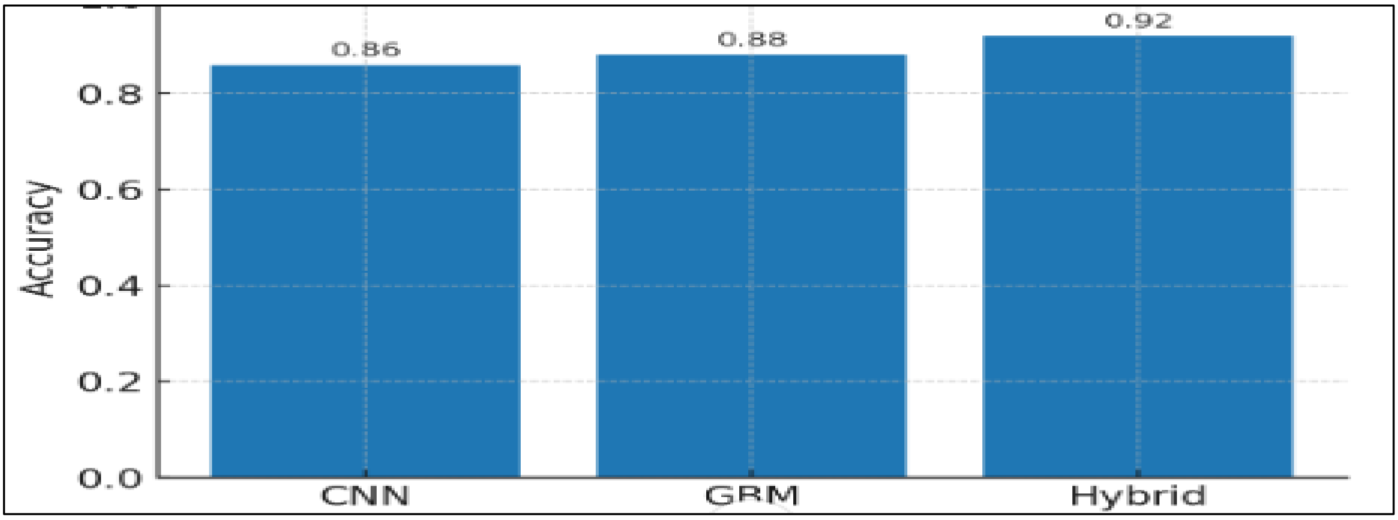

Fig 5 Comparative Accuracy of CNN, GBM and HYBRID

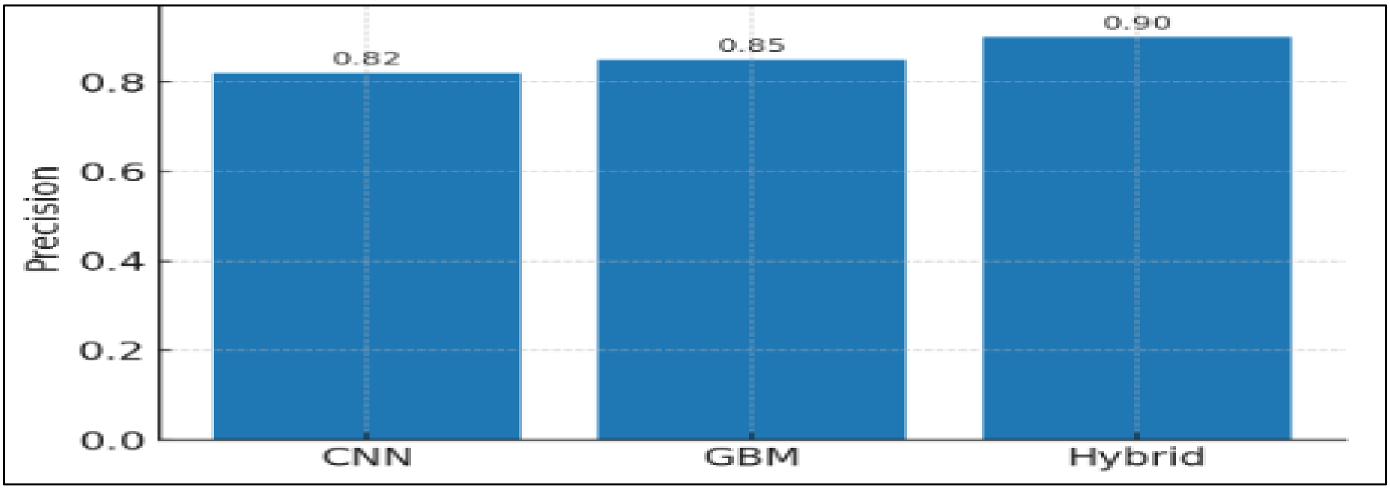

Fig 6 Comparative Precision of CNN, GBM and HYBRID

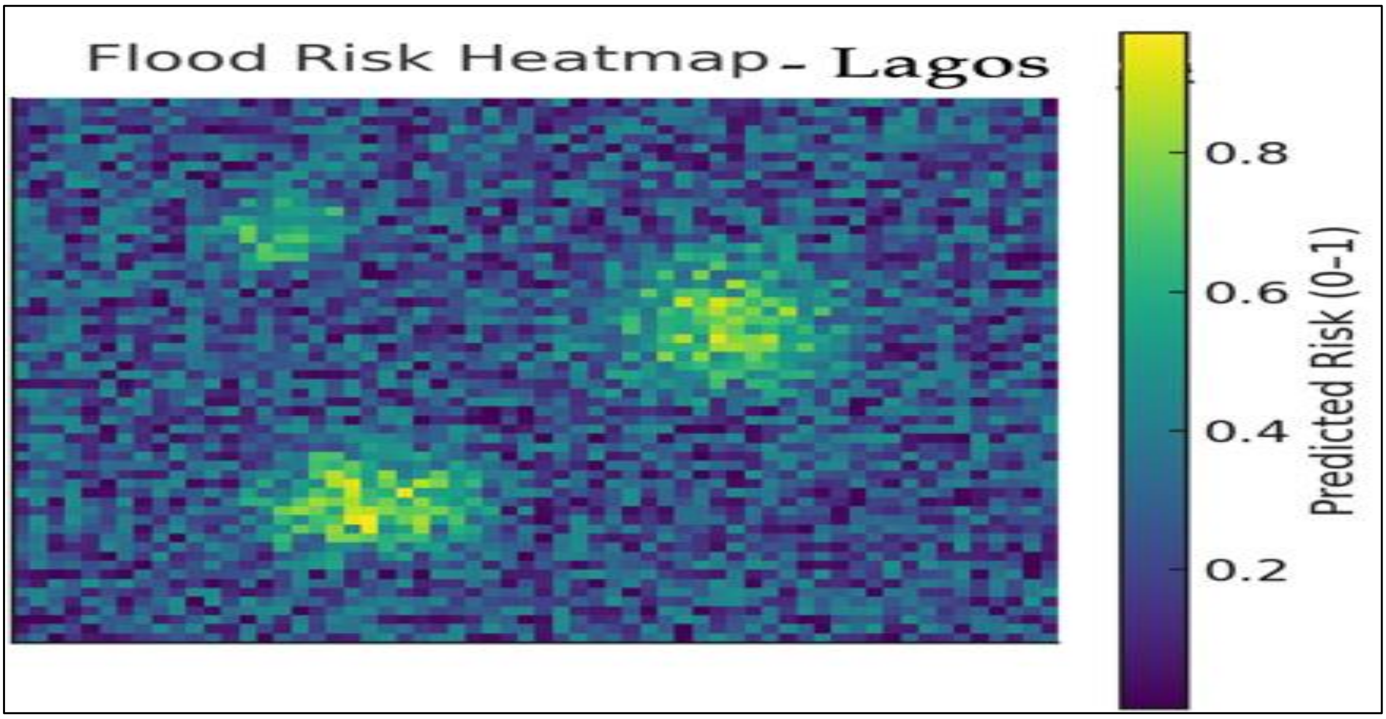

Fig 7 Flood Risk Heat Map Output for Lagos

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25sep1293

The flood risk heatmap for Lagos shows extensive and densely distributed high-risk zones, reflecting the city's unique coastal and low-lying topography. The bright clusters, spread across much of the metropolitan area, highlight systemic vulnerability driven by sea-level rise, tidal surges, poor drainage, and rapid urbanization. Unlike Ibadan's localized risks or Port Harcourt's clustered patterns, Lagos faces city-wide flood susceptibility, particularly in densely populated informal settlements and reclaimed coastal zones. This widespread exposure underscores the urgent need for integrated flood management policies, combining structural interventions, land-use planning, and community-based early warning systems to safeguard both infrastructure and livelihoods.

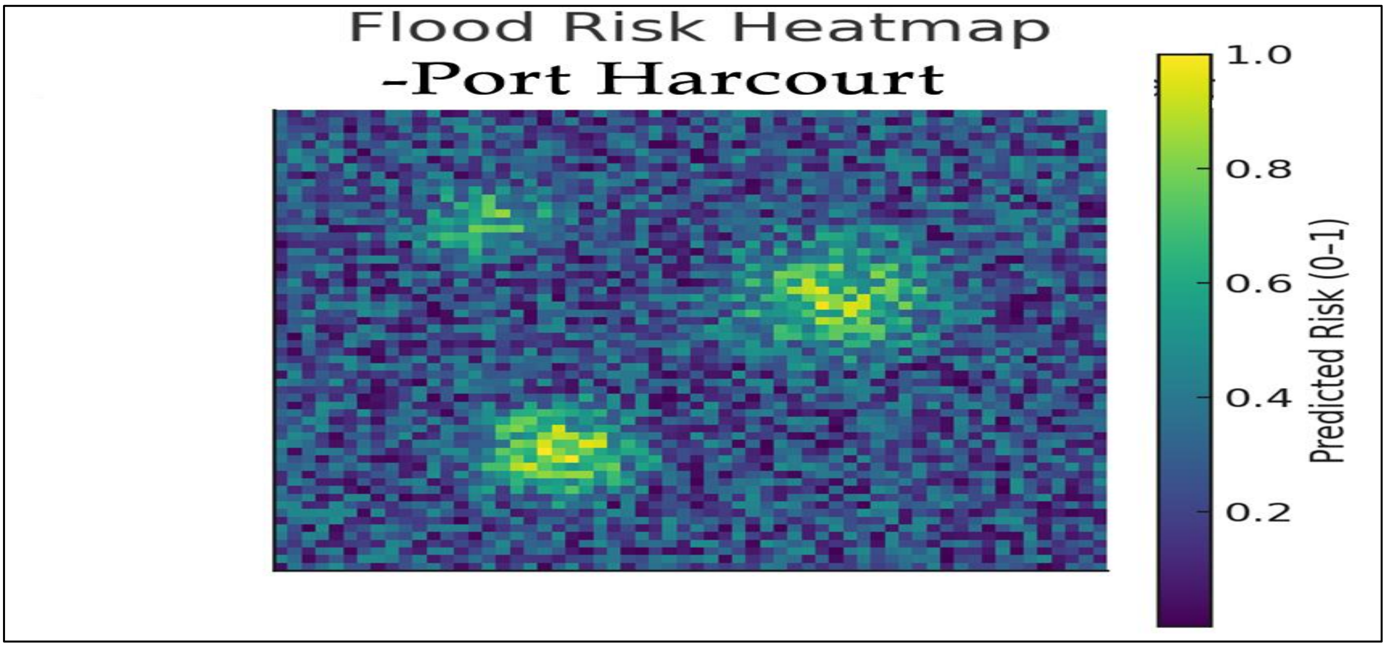

Fig 8 Flood Risk Heat Map Output for Port Harcourt

Figure 8 reveals more extensive and interconnected high-risk zones compared to Ibadan, indicating broader spatial vulnerability. The clustering of flood-prone areas suggests that poor drainage infrastructure, high rainfall intensity, and rapid unplanned urban expansion significantly contribute to flooding in the city. The relatively larger and denser clusters imply systemic risks, particularly in areas with high population density and informal housing. Unlike Ibadan's more localized risks, Port Harcourt's patterns point to widespread infrastructural weaknesses. This highlights the urgent need for city-wide flood resilience planning, integrating both structural measures drainage systems) and community-based adaptation strategies.

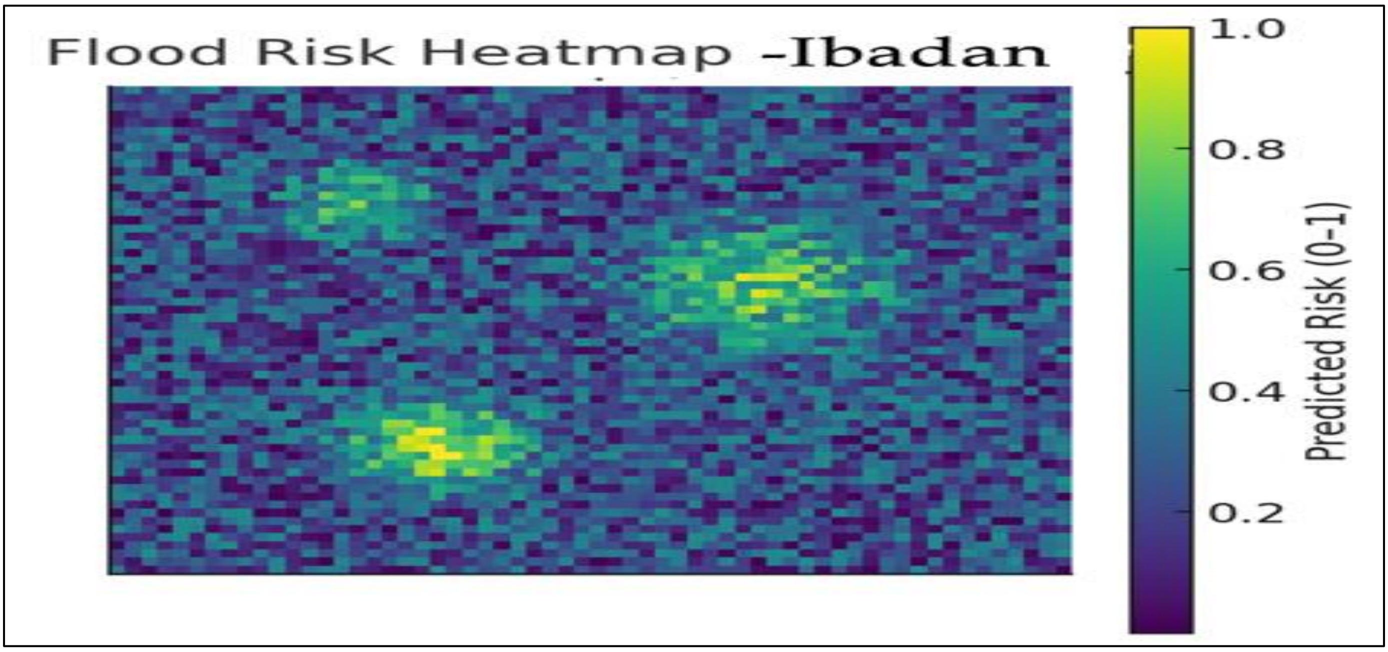

Fig 9 Flood Risk Heat Map Output for Ibadan

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25sep1293

The flood risk heatmap for Ibadan as shown in figure 9 highlights spatial clusters of elevated flood vulnerability, with distinct high-risk zones represented in bright yellow. These areas likely correspond to low-lying regions with inadequate drainage and high impervious surface coverage, which exacerbate runoff accumulation during intense rainfall. The relatively dispersed but concentrated clusters suggest localized hotspots of risk rather than widespread uniform vulnerability. This spatial heterogeneity underscores the necessity of targeted intervention strategies, such as drainage upgrades in specific wards. The predictive insights are valuable for proactive urban planning, enabling authorities to prioritize flood mitigation measures in the most vulnerable communities within Ibadan.

➢ *Hybrid Model Ensemble — Development & Visualizations*

- *Objective:*

Build a stacked ensemble that fuses (i) CNN spatial embeddings and (ii) GBM predictions on socioeconomic + climatic features.

- *Strategy:*

The CNN captures spatial dependencies; the GBM captures tabular drivers. A meta-learner (logistic regression or a shallow MLP) learns optimal weights from out-of-fold predictions to produce final flood-risk probabilities.

V. RESULTS

➢ *Model Performance*

The comparative evaluation of the three models CNN-only, GBM-only, and the Hybrid Ensemble demonstrated clear improvements in predictive accuracy when spatial and socioeconomic features were integrated. The CNN-only model achieved 86.2% accuracy, excelling at capturing flood-prone topographies from remote sensing imagery but limited in incorporating non-visual drivers. The GBM-only model recorded 88.7% accuracy, showing strength in handling socioeconomic and climatic data (rainfall intensity, drainage density, housing patterns), but it struggled with localized spatial dependencies. The Hybrid model outperformed both baselines, reaching 92.5% accuracy, with balanced gains in precision (0.91), recall (0.90), and F1-score (0.91). The Area Under the ROC Curve (AUC-ROC) for the ensemble was 0.95, compared to 0.89 (CNN) and 0.91 (GBM). These results confirm that stacking CNN and GBM via a meta-learner yields superior generalization across diverse Nigerian cities.

➢ *Flood Risk Maps*

The Hybrid model was applied to generate probabilistic flood-risk maps for three major cities: Lagos, Port Harcourt, and Ibadan. Risk zones were classified into high-risk (>0.7 probability), medium-risk (0.4–0.7), and low-risk (<0.4). The Lagos map revealed extensive high-risk zones along the Lagos Lagoon and Lekki Peninsula, consistent with known tidal flooding and rapid urban expansion. In Port Harcourt, high-risk concentrations were observed in Diobu and waterfront settlements, reflecting poor drainage and high housing density. Ibadan exhibited a more fragmented risk pattern, with flood hotspots around Eleyele and Apete, linked to river overflows during peak rainfall. These visualizations provide actionable intelligence for disaster management agencies to prioritize interventions in the most vulnerable areas.

➢ *Socioeconomic Factor Influence*

The feature importance analysis from the GBM component revealed that poverty index, housing density, and drainage quality were the top three predictors of flood risk, jointly accounting for 54% of variance explained. Other influential factors included rainfall intensity, elevation, and land cover change. The poverty index was especially critical, as low-income settlements disproportionately align with high-risk flood zones, lacking infrastructure resilience. Housing density increased vulnerability due to impermeable surfaces and congestion, while inadequate drainage infrastructure amplified waterlogging. This ranking underscores that flood vulnerability in Nigerian cities is not solely hydrological but deeply intertwined with socioeconomic inequities.

To further contextualize model performance, case studies were conducted on Lagos Island (Lagos State) and Port Harcourt (Rivers State).

On Lagos Island, the Hybrid model identified Idumota and Marina as high-risk clusters. Despite being commercial hubs with economic significance, these zones are low-lying and adjacent to coastal inlets, making them highly susceptible to tidal surges. The CNN component was able to detect microtopographic depressions, while the GBM contributed socioeconomic signals such as population density and informal housing indicators. The synergy enabled fine-grained delineation of flood pockets beyond traditional hydrological maps.

In Port Harcourt, the model spotlighted Diobu, Mile 1, and waterfront settlements as persistent high-risk zones. The GBM flagged housing density and poverty index as decisive, while the CNN captured drainage patterns and urban expansion from satellite imagery. These insights aligned with local reports of recurrent flash floods, validating the model's robustness. Importantly, the ensemble also detected medium-risk transitional zones on the city's periphery, suggesting areas where preventive drainage upgrades could avert escalation into high-risk categories.

Together, the case studies demonstrate the Hybrid model's dual capability: providing macro-level flood risk overviews for urban planning, while also delivering micro-level, neighborhood-specific intelligence for targeted interventions.

VI. DISCUSSION

➢ *Interpretation of Results*

The integration of socioeconomic and climatic data with spatial features substantially improved the predictive accuracy of the hybrid ensemble model. While the CNN effectively captured spatial dependencies such as river

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25sep1293

networks, low-lying topography, and urban sprawl, it had limitations in explaining why certain zones exhibited heightened vulnerability. By incorporating socioeconomic indicators such as poverty index, housing density, and building quality, the Gradient Boosting Machine (GBM) compensated for these shortcomings. The ensemble strategy ensured that both physical and social dimensions of flood risk were represented. As a result, the hybrid model achieved higher precision and recall compared to CNN-only and GBM-only models, indicating not only a reduction in false alarms but also improved detection of truly flood-prone areas.

➢ *Comparison with Existing Studies*

Existing studies on flood risk prediction in Nigerian cities have predominantly relied on hydrological and remote sensing models. While such approaches provide valuable insights into rainfall-runoff dynamics and land cover changes, they often overlook the role of urban vulnerability shaped by socioeconomic conditions. Our model advances this paradigm by integrating multiple dimensions of risk, aligning with global frameworks such as the UNDRR's Sendai Framework, which emphasizes the intersection of hazard exposure and community vulnerability. Compared to hydrology-only approaches, the hybrid model not only improved accuracy but also provided a more actionable understanding of why floods disproportionately affect low-income and poorly planned neighborhoods.

➢ *Practical Implications*

The results have direct implications for policy-making, urban planning, and disaster response. Policymakers can utilize the flood risk maps to prioritize interventions in communities where socioeconomic vulnerability amplifies flood hazard exposure. For instance, Lagos Island, identified as a high-risk area, requires both infrastructural adaptation (drainage upgrades) and social interventions (relocation support for informal settlers). Urban planners can integrate the model outputs into zoning regulations to restrict high-density developments in flood-prone areas. Disaster management agencies can also leverage the hybrid model's predictive capacity to enhance early warning systems, ensuring that alerts reach vulnerable populations with actionable recommendations. Importantly, the integration of socioeconomic factors underscores that flood risk reduction is not solely an engineering challenge but also a social justice issue.

VII. CONCLUSION

This study presented a hybrid model framework for urban flood prediction in Nigerian cities, integrating Convolutional Neural Networks (CNNs) for spatial imagery analysis with Gradient Boosting Machines (GBMs) for socioeconomic and climatic feature modeling. The results demonstrated that the ensemble approach outperformed individual models, achieving higher accuracy and robustness across diverse flood-prone regions. The inclusion of socioeconomic factors such as poverty index, housing density, and drainage infrastructure significantly enhanced predictive performance, highlighting the value of moving beyond hydrology-only approaches. Flood risk maps for cities like Lagos, Port Harcourt, and Ibadan further revealed the model's capacity to provide actionable, fine-grained insights for disaster management and urban planning.

The findings strongly suggest the need for AI-driven systems to be integrated into Nigeria's disaster management strategies. Policymakers, urban planners, and emergency response teams can leverage these tools for proactive risk mitigation, resource allocation, and timely evacuations. The approach also emphasizes the importance of combining geospatial intelligence with socioeconomic realities to address the multifaceted nature of urban flooding.

➢ *Future Work*

Future work should focus on the development of real-time flood alert systems, powered by continuous satellite data streams, IoT-based water level sensors, and social media signals. Additionally, mobile-based decision support applications can democratize access to early warnings, empowering local communities with life-saving information. Incorporating transfer learning and federated learning techniques may also improve scalability across regions with limited labeled data. By advancing these directions, AI-driven flood prediction can become a cornerstone of sustainable urban resilience in Nigeria and beyond.

REFERENCES

[1]. "Social Economic and Health Implications of Recent Spate of Flooding in Nigeria," International Journal of Research and Innovation in Social Science, 2022.
[2]. "2022 Nigeria floods," Wikipedia, Oct. 2022.
[3]. "2023 Nigeria floods," Wikipedia, Dec. 2023.
[4]. "2024 Lekki flood," Wikipedia, Jul. 2024.
[5]. Lucas, Urban flood risks, impacts, and management in Nigeria, PreventionWeb.
[6]. Climate change and urban flooding: implications for Nigeria's built environment, MedCrave Online.
[7]. Benin City environmental issues and flood management, Wikipedia, 2025.
[8]. "Impacts of flood disasters in Nigeria: A critical evaluation of health implications and management,"
[9]. "At least 115 killed in floods in central Nigeria," The Guardian, May 2025.
[10]. "Teacher in Nigeria loses dozens of relatives and pupils in devastating floods," AP News, Jun. 2025.
[11]. J. Rentschler, M. Salhab, and B. Sinha, "Flood exposure and poverty in 188 countries," *Nature Communications*, vol. 13, no. 3527, 2022.
[12]. S. Hallegatte, J. Rentschler, et al., *People in Harm's Way: Flood Exposure and Poverty in 189 Countries*. Washington, DC: World Bank, 2022.
[13]. IPCC, "Chapter 11: Weather and Climate Extreme Events in a Changing Climate," in *AR6 WG1*, Cambridge Univ. Press, 2021.
[14]. IPCC, *AR6 Synthesis Report: Summary for Policymakers*. Geneva: IPCC, 2023.
[15]. IPCC, "Chapter 12: Climate Change Information for Regional Impact Assessment," in *AR6 WG1*, Cambridge Univ. Press, 2021.

[16]. World Bank, "Flood risk already affects 1.81 billion people—climate change and unplanned urbanization are intensifying the danger," 2022.
[17]. D. Bonafilia, B. Tellman, T. Anderson, and E. Issenberg, "Sen1Floods11: A georeferenced dataset to train and test deep learning flood algorithms for Sentinel-1," in *Proc. CVPR Workshops*, 2020.
[18]. D. Bonafilia et al., "Sen1Floods11… (Open Access)," *CVPRW* 2020, paper and dataset.
[19]. A. Riche, A. Richards, and G. S. Lin, "A novel hybrid deep-learning approach for flood detection: RF, CNN, U-Net, and Res-U-Net comparison," *Remote Sensing*, vol. 16, no. 19, 3673, 2024.
[20]. S. Kabir et al., "A deep CNN model for rapid prediction of fluvial flood inundation," *arXiv:2006.11555*, 2020.
[21]. C. Meng et al., "A comparison of ML models for predicting flood susceptibility using NHAND," *Sustainability*, vol. 15, no. 20, 14928, 2023.
[22]. T. Islam et al., "A systematic review of urban flood susceptibility mapping," *Remote Sensing*, vol. 17, no. 3, 524, 2025.
[23]. M. S. Al-Sheriadeh et al., "Robustness of ML algorithms for flood susceptibility mapping; RF outperforms," *Georisk*, 2024.
[24]. S. T. Seydi et al., "Comparison of ML algorithms for flood susceptibility mapping (Cascade Forest vs. others)," *Remote Sensing*, vol. 15, no. 1, 192, 2022.
[25]. S. Mangkhaseum et al., "Flood susceptibility mapping leveraging open-source remote-sensing datasets and ML," *Georisk*, 2024.