ISSN No. 0976-5697

VOLUME 17 SPECIAL ISSUE 1, JULY 2026
International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Sanket: A Smart Community Health Monitoring and Early Warning System for Water-Borne Diseases in Rural India

Harshit Borana
Dept. of Computer Science and Engineering
Geetanjali Institute of Technical Studies
Udaipur, Rajasthan, India
harshitborana2@gmail.com

Parth Sharma
Dept. of Computer Science and Engineering
Geetanjali Institute of Technical Studies
Udaipur, Rajasthan, India
sharmaparth2004@gmail.com

Shiv Kumar
Dept. of Computer Science and Engineering
Geetanjali Institute of Technical Studies
Udaipur, Rajasthan, India
shivbadyal123@gmail.com

Dr. Mayank Patel
Dept. of Computer Science and Engineering
Geetanjali Institute of Technical Studies
Udaipur, Rajasthan, India
mayank999_udaipur@yahoo.com

Abstract: Project SANKET, a comprehensive Smart Community Health Monitoring and Early Warning System designed to combat the recurrent and fatal outbreaks of water-borne diseases in rural Northeast India. Every monsoon, flooded rivers contaminate vulnerable groundwater sources, triggering severe epidemics of cholera, diarrhea, dysentery, and hepatitis. Current public health responses are hindered by slow, manual water testing and fragmented, reactive disease surveillance. This results in a critical time lag, allowing outbreaks to spread unchecked and causing preventable loss of life and heavy economic burdens. To bridge this technological gap, SANKET introduces a multi-layered ecosystem that synergizes Internet of Things (IoT) hardware, Artificial Intelligence (AI), and community participation. Low-cost IoT sensor nodes, termed "Jal Suraksha," continuously monitor water quality parameters such as pH, turbidity, and dissolved oxygen, transmitting data via LoRaWAN or GSM networks. Concurrently, a multilingual, offline-first mobile application empowers villagers and Accredited Social Health Activists (ASHAs) to report health symptoms and manual water test results directly to a centralized cloud. An integrated AI/ML predictive engine analyzes this fused environmental and health data to identify anomalies, forecast outbreak probabilities, and autonomously generate localized alerts for both citizens and health officials. By shifting the public health paradigm from post-outbreak reaction to real-time, predictive intervention, SANKET safeguards vulnerable populations, empowers local women's Self-Help Groups (SHGs), and provides a scalable, cost-effective framework ready for integration into national digital health infrastructures.

KeyWords— Water Quality Monitoring, **Internet of Things (IoT)** sensor nodes and **Artificial Intelligence (AI)**, specifically utilizing **Machine Learning** algorithms for **Predictive Analytics** to forecast health crises.

I. INTRODUCTION

1.1. Context and Background Water is universally essential for life, but in rural India, particularly during the monsoon season, it frequently becomes a silent killer. Every year, swollen rivers carry not just silt and debris, but invisible pathogens that infect thousands across vulnerable regions. States such as Assam, Tripura, and neighboring areas are especially susceptible, facing recurrent waves of cholera, diarrhea, hepatitis, and dysentery. The geographic landscape of the Northeast, dominated by the Brahmaputra and Barak river basins, experiences annual flooding that submerges latrines, contaminates wells, and creates ideal breeding grounds for pathogens. The threat is magnified by the fact that 85% of rural drinking water in India comes from groundwater sources—like wells, hand pumps, and shallow tube wells—which are highly vulnerable to contamination during these floods.

1.2. The Scale of the Crisis The urgency of this issue is starkly highlighted by recent, devastating outbreaks. In Tripura, between December 2022 and January 2023, over 130 cases of acute diarrhea were reported in a single village within days. Investigations traced the source to a contaminated public water tap, with laboratory tests confirming the presence of *E. coli* and *Shigella*. Similarly, in May 2024, a suspected water contamination epidemic in a Tinsukia tea estate in Assam claimed 11 lives in just 10 days. The crisis extends beyond the Northeast; in Indore, officially India's "cleanest city," contaminated drinking water recently claimed 16 lives and left over 1,400 residents ill. According to national health estimates, 1 in every 5 hospital admissions is linked to waterborne diseases. These outbreaks also trigger staggering economic losses, with families spending up to 20% of their monthly income on medical treatment, while missed workdays further cripple fragile rural livelihoods.

1.3. The Technological Gap At the heart of this recurring crisis lies a critical technological gap: the absence of early detection and coordinated response mechanisms. Currently, water testing in rural areas remains manual, slow, and sample-based. By the time samples are transported and results reach district laboratories, outbreaks have already spread through communities. Government monitoring efforts, such as those by Pollution Control Boards, rely on periodic sampling (monthly or quarterly), resulting in massive time lags where data is only available weeks after contamination occurs. Furthermore, health tracking systems like the National Health Mission's Integrated Disease Surveillance Program (IDSP) are entirely reactive, tracking outbreaks only after they are established and suffering from an inherent 7-10 day delay in confirmation. There is currently no real-time monitoring or predictive warning system for most rural communities, and villagers lack reliable, localized channels to report symptoms.

1.4. Objective of the Study The cycle of floods, subsequent contamination, delayed detection, and resulting fatalities repeats yearly. Without intervention, these recurring epidemics will continue to claim lives, worsen child malnutrition, and erode trust in public health systems. To address this systemic failure, this paper introduces Project SANKET (Smart Alert and Notification for Knowledge-Enabled Tracking). SANKET is proposed as an integrated, multi-layered early warning system tailored for rural, flood-prone India. By uniquely uniting real-time IoT water monitoring, AI-powered outbreak prediction, and community-driven symptom reporting, this research aims to demonstrate a proactive framework capable of preventing waterborne disease outbreaks before they escalate into epidemics.

II. Literature Survey

A. Environmental and Socio-Economic Context The vulnerability of Northeast India to waterborne diseases is driven by a complex intersection of environmental and socio-economic factors. Research and government data, including the Census of India 2011 and NFHS-5 surveys, highlight significant sanitation and health challenges in these regions. Geographically, the Brahmaputra and Barak river basins dominate Assam and Tripura, experiencing annual floods that submerge latrines, contaminate wells, and create breeding grounds for pathogens. Additionally, academic studies by institutions like IIT Guwahati have documented severe water pollution in the Brahmaputra basin. Wetlands such as Deepor Beel, once biodiversity hotspots, now face frequent ecological collapse due to toxic industrial discharge and city sewage.

From a socio-economic perspective, the burden of water collection disproportionately affects rural women, who spend 30-40 minutes daily fetching water, translating to 27 lost workdays annually. Furthermore, studies indicate that children in households tasked with water collection face a 22% higher risk of school dropout, a risk that exacerbates after flood events. Joint reports by UNICEF and WHO, alongside peer-reviewed studies, also confirm that flood exposure is directly linked to increased stunting and malnutrition among children in rural India.

B. Historical Outbreaks and Epidemic Trends A review of recent media reports and health data reveals a predictable, yearly cycle of floods leading to contamination, delayed detection, and subsequent fatal outbreaks. Key incidents documented over the past few years include:

- **2019:** A widespread diarrhea outbreak occurred in Dibrugarh following post-monsoon flooding.
- **2020:** Parts of Assam experienced a hepatitis outbreak linked to contaminated, leaking pipelines.
- **2022:** In Tripura, an acute diarrhea outbreak resulted in over 130 cases, which were later traced to a contaminated public water supply.
- **2024:** An epidemic in a Tinsukia tea estate in Assam claimed 11 lives within 10 days due to suspected contaminated water.

C. Review of Existing Monitoring Systems Current approaches to mitigating these crises are largely fragmented. Government initiatives, while heavily funded, face critical operational limitations:

- **Pollution Control Boards (PCBA, CPCB):** These bodies conduct water quality sampling, but the process is sporadic (monthly or quarterly) and not real-time. This creates a significant time lag, with results often arriving weeks after contamination has occurred.
- **National Health Mission (IDSP):** The Integrated Disease Surveillance Program tracks disease cases manually at the district level. This approach is inherently reactive, tracking outbreaks only after they are established, and suffers from a 7-10 day delay in confirmation.
- **Jal Jeevan Mission (JJM):** While this mission focuses on large-scale piped water infrastructure to provide access to "improved" sources, it does not guarantee continuous safety monitoring; unmonitored infrastructure can inadvertently become a vector for disease during floods.

D. Gap Analysis in Current Approaches Beyond government frameworks, manual field test kits and community awareness campaigns are periodically deployed. However, manual testing requires trained personnel, cannot be easily scaled to every household, and the resulting data is rarely digitized for historical tracking. Awareness campaigns provide general hygiene education but fail to deliver targeted, real-time alerts about imminent threats. While academic institutions and NGOs have piloted small-scale sensor deployments and filter distributions, these projects struggle to scale and lack integration with mainstream government health data systems.

Ultimately, the existing literature and system analyses expose three critical gaps: a lack of integration between siloed health and water data, a lack of speed due to reliance on manual testing, and a lack of empowerment, as villagers have no reliable channels to report symptoms or receive early warnings.

III. Proposed Methodology

A. System Overview Project SANKET (Smart Alert and Notification for Knowledge-Enabled Tracking) is proposed as an integrated, multi-layered system designed for the real-time monitoring of water quality, early disease outbreak detection, and community-driven reporting. The solution shifts away from purely reactive measures by resting on three foundational pillars: Technology, Community, and Governance. By uniting real-time Internet of Things (IoT) water surveillance, Artificial Intelligence (AI) predictive analytics, and localized community participation, SANKET establishes a seamless early warning ecosystem tailored for rural India.

B. Hardware Architecture The physical data collection infrastructure is built upon a dual-tiered approach to balance advanced monitoring with cost-effective scalability:

- **IoT Sensor Nodes ("Jal Suraksha" kits):** Mounted at critical water sources such as handpumps, community taps, and river inlets, these nodes continuously measure key water quality parameters including pH, turbidity, dissolved oxygen (DO), conductivity, chlorine levels, and temperature.
- **Microcontrollers and Transmission:** The sensors interface with ESP32 microcontrollers, which serve as the core processing units. Data is transmitted primarily via LoRaWAN protocols, which offer long-range (10-15 km) and low-power communication ideal for clustered villages. In areas where LoRa coverage is unfeasible, GSM modules act as a fallback.
- **Edge Gateways and Power:** A Raspberry Pi or similar low-power edge device functions as a local gateway to aggregate readings from multiple sensor nodes before forwarding them to the cloud. To counter unreliable electricity grids in remote villages, the hardware is supported by solar panels and battery backup units for uninterrupted operation.
- **Manual Test Kits ("Jal Parikshan"):** For broader household coverage, affordable strip-based test packs are distributed. While low-tech, the results are digitized by simply clicking a photo of the test strip and uploading it via the SANKET mobile app.

C. Software Ecosystem and Cloud Architecture The software stack is designed for reliability, offline capabilities, and cross-platform accessibility:

- **Cloud Infrastructure:** Data processing and storage are hosted on scalable platforms like AWS or Azure. The database architecture utilizes MongoDB to handle flexible, high-volume time-series data (health and water metrics) and PostgreSQL for structured relational data.
- **Mobile Application:** Developed using the Flutter framework for cross-platform (Android/iOS) compatibility, the app serves two main user groups. For villagers, it provides a multilingual interface (Assamese, Bengali, Hindi, Khasi, Mizo) to report symptoms, receive hygiene advisories, and check water safety status. It is designed with an offline-first architecture that syncs data via Firebase when a connection is established, alongside
- SMS fallback. For ASHA workers, it provides a dedicated dashboard for case reporting, patient referrals, and test uploads.

Fig1 : Asha Worker Screen Fig2: Report Case Screen

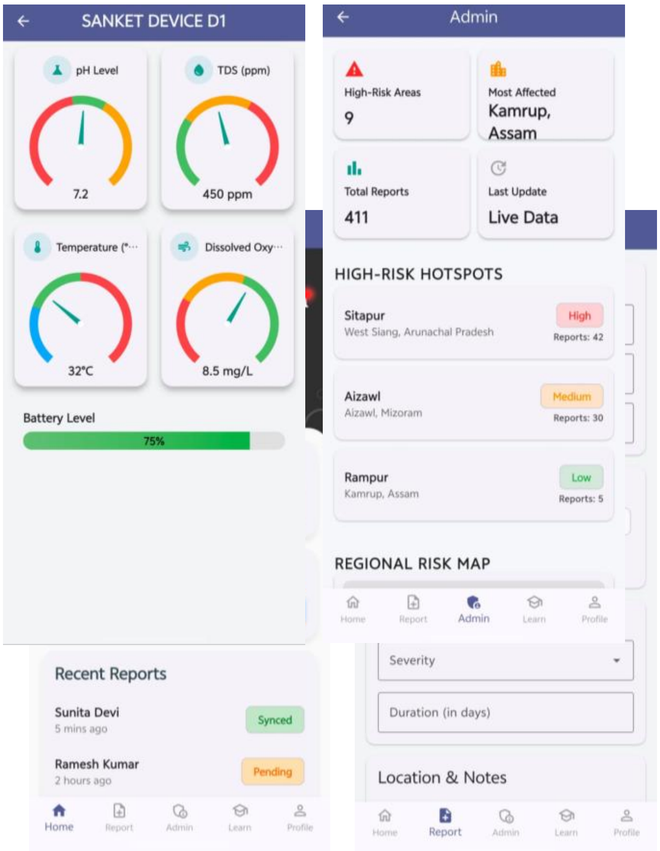

Fig 3: IOT Data Screen Fig 4: Admin Screen

Web Dashboard: Built with React.js (frontend) and Node.js (backend), the web dashboard is designed for health officials at Primary Health Centres (PHCs) and state authorities.

Fig 5: Web Dashboard

D. AI-Powered Predictive Engine SANKET employs an AI/ML layer developed in Python, utilizing libraries such as TensorFlow for deep learning and Scikit-learn for traditional machine learning models.

- **Data Fusion:** The ML pipeline merges multiple data inputs—including real-time water quality metrics, community symptom clusters, and environmental factors like rainfall and temperature—into a single integrated dataset.
- **Outbreak Forecasting:** The models are trained on historical outbreak data to recognize patterns and generate outbreak risk scores. For instance, the system logic cross-checks anomalies to flag events such as a "sudden rise in turbidity + spike in diarrhea reports = likely outbreak," identifying high-risk hotspots before an epidemic escalates.

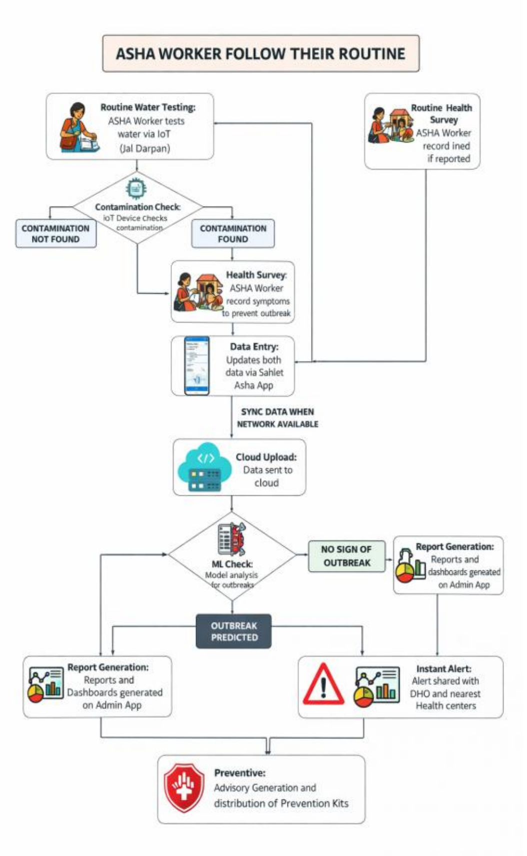

Fig 6: ASHA Worker Flowchart

E. System Workflow and Data Flow The real-world implementation of the SANKET methodology follows a coordinated, automated workflow

- **Without SANKET:** When a flood hits a vulnerable area like Barpeta, villagers unknowingly consume contaminated water. Due to the lack of real-time monitoring, within a week, cases escalate to over 100, overwhelming local Primary Health Centres (PHCs), forcing children out of school, and draining family finances on medical treatments.
- **With SANKET:** Following a flood event, IoT sensors flag water contamination within 24 hours. The AI predictive engine calculates the outbreak risk, allowing health officials to issue immediate boil-water advisories and dispatch ASHA workers with chlorine tablets. As a result of this proactive intervention, the simulated outbreak is averted, with total cases limited to fewer than 10.

IV. Results and Expected Outcomes

A. System Integration and Functional Capabilities The primary result of the SANKET project design is the successful conceptualization of a unified, multi-layered early warning system. Unlike existing fragmented approaches, SANKET successfully integrates real-time IoT water quality monitoring with community-driven symptom reporting and AI-powered outbreak prediction. The system design proves technically feasible by utilizing mature, low-cost technologies such as ESP32 microcontrollers, LoRaWAN for long-range communication in low-connectivity rural areas, and solar-powered nodes to bypass unreliable electricity grids.

B. Simulated Comparative Analysis (Before vs. After SANKET)

A simulated operational scenario demonstrates the dramatic reduction in response time and disease propagation achieved by the system.

C. Projected Economic Impact and Cost-Benefit Analysis The financial analysis of the SANKET framework reveals a highly favorable return on investment for public health sectors. The estimated cost to deploy a comprehensive pilot program across 50 villages for one year is approximately ₹16.24 lakhs. In contrast, preventing just a single major waterborne disease outbreak—similar to the 2022 Tripura acute diarrhea crisis—can save an estimated ₹20 to 25 lakhs in medical treatments, lost wages, and emergency relief costs. Therefore, the economic benefits of deploying SANKET far outweigh the operational costs, with economies of scale expected to drive per-village expenses even lower during statewide expansions.

D. Social and Environmental Outcomes

Beyond immediate health interventions, the implementation of SANKET is projected to yield extensive secondary benefits:

- **Social Empowerment:** The system actively shifts communities from passive victims to active participants. By training women's Self-Help Groups (SHGs) to manage the "Jal Suraksha" kits, the project fosters local leadership in community health monitoring. Additionally, by preventing prolonged illnesses, the system helps lower school dropout rates among children and indirectly tackles flood-linked child malnutrition.
- **Environmental Restoration:** Continuous, real-time monitoring establishes strict accountability for water quality. By instantly flagging toxic industrial discharge and sewage dumping, SANKET aids in preventing the ecological collapse of vital water bodies and wetlands (such as Deepor Beel), thereby preserving local biodiversity and supporting the livelihoods of local fisheries)
-

V. Acknowledgement

The guidance of our faculty mentor, Dr. Mayank Patel, was instrumental in shaping the technical direction of this project, particularly in system architecture design and practical implementation strategies. His insights helped maintain a balance between theoretical concepts and real-world applicability. The Department of Computer Science and Engineering at Geetanjali Institute of Technical Studies, Udaipur, provided the necessary infrastructure, resources, and academic support required to successfully develop and evaluate the Sanket. We also acknowledge the Smart India Hackathon (SIH) for providing the problem statement, which served as the foundation for this work and guided the development of a solution focused on Internet Of Things and Mobile Based Real Time Application.

VI. Reference

[1] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease prediction by machine learning over big data from healthcare communities," IEEE Access, vol. 5, pp. 8869–8879, 2017.

[2] Sen, S., Patel, M., Sharma, A.K. (2021). Software Development Life Cycle Performance Analysis. In: Mathur, R., Gupta, C.P., Katewa, V., Jat, D.S., Yadav, N. (eds) Emerging Trends in Data Driven Computing and Communications. Studies in Autonomic, Data-driven and Industrial Computing. Springer, Singapore. https://doi.org/10.1007/978-981-16-3915-9_27

[3] A.K. Tripathi and S. K. Dubey, "Early warning systems for waterborne diseases using data analytics," Procedia Computer Science, vol. 167, pp. 1892–1901, 2020.

[4] P. Rajalakshmi and S. Balamurugan, "IoT-based smart health monitoring system for rural healthcare," International Journal of Advanced Computer Science and Applications, vol. 10, no. 6, pp. 145–152, 2019.

[5] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," Future Generation Computer Systems, vol. 29, no. 7, pp. 1645–1660, 2013.

[6] H. Alemdar and C. Ersoy, "Wireless sensor networks for healthcare: A survey," Computer Networks, vol. 54, no. 15, pp. 2688–2710, 2010.

[7] Ameta, Upasana, Mayank Patel, and Ajay Kumar Sharma. "Scaled Agile Framework Implementation in Organizations', its Shortcomings and an AI Based Solution to Track Team's Performance." 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT). IEEE, 2022.

[8] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," Computer Networks, vol. 54, no. 15, pp. 2787–2805, 2010.

[9] Ameta, Upasana and Patel, Mayank and Rathore, Narendra Singh, Fusing Artificial Intelligence with Scrum Framework (April 25, 2023). Available at SSRN: https://ssrn.com/abstract=4428286 or http://dx.doi.org/10.2139/ssrn.4428286

[10] A.Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," IEEE Internet of Things Journal, vol. 1, no. 1, pp. 22–32, 2014.

