Hydro-Climatic Spatio -Temporal Dengue Risk Prediction System

Mrs. Asha Sattigeri[1], Teja M S[2], Tejas Gowda H R[3], Ullas S A[4], Tarun D N[5]

Assistant Professor, Computer Science and Engineering, K S Institute of Technology, Bangalore, India[1]

Student, Computer Science and Engineering, K S Institute of Technology, Bangalore, India[2-5]

Abstract: This paper presents an integrated health informatics framework aimed at forecasting and mitigating dengue outbreaks by fusing hydro-climatic conditions with symptom-based user inputs. The proposed system evaluates environmental variables—namely rainfall, temperature, and humidity—alongside self-reported symptoms and historical disease trends to estimate the likelihood of dengue transmission across different regions. By implementing rule-based logic and leveraging spatio-temporal analysis, the system offers real-time risk estimations and context-specific preventive guidance. Key features include a symptom evaluation tool, geo-based alert notifications, and a risk scoring mechanism derived from user questionnaires. The interface also incorporates visual elements such as dynamic heatmaps and time-series charts to enhance public understanding and actionable response. Developed with usability and public health relevance in mind, this platform serves as a practical and scalable tool to support early intervention strategies against vector-borne diseases like dengue

I. INTRODUCTION

Dengue fever is a viral infection spread by mosquitoes, mainly the Aedes aegypti species. It remains a serious public health concern in tropical and subtropical areas around the world. Lately, we've seen a sharp increase in reported dengue cases, fueled by factors like rapid urban growth, poor waste and water management, rising populations, and changing climate conditions. Key hydro-climatic factors—like rainfall, temperature, and humidity—significantly affect mosquito breeding and the chances of virus transmission, making them crucial for predicting potential outbreaks.

Conventional dengue surveillance systems typically rely on centralized data collection and historical trend analysis. While useful, these approaches are often limited by delayed reporting, lack of granularity, and an absence of real-time risk detection. They also tend to overlook personalized health inputs, such as early symptoms experienced by individuals or micro-environmental risk factors, which are crucial for timely action. As climate patterns become increasingly unpredictable and mosquito habitats expand, there is a pressing need for intelligent, accessible tools that can deliver localized and personalized disease risk insights.

This paper introduces a comprehensive and user-centric system designed to address these challenges through the integration of hydro-climatic data, user-reported symptoms, and spatial-temporal analysis. The platform combines rule-based algorithms with intuitive interfaces to assess dengue risk levels, deliver alerts, and guide users with practical preventive measures.

Key features include a weather-based prediction engine, a symptom checker, a local alert system, and interactive visualizations such as heatmaps and trend charts. By offering a holistic view of individual and regional dengue risks, this framework supports both public awareness and decision-making by health authorities. Its emphasis on accessibility and simplicity ensures that it can be effectively used by a wide range of stakeholders, from everyday citizens to medical professionals.

II. LITERATURE SURVEY

Researchers from various fields have been diving into ways to predict and prevent dengue outbreaks by blending environmental, clinical, and technological strategies. In this section, we'll shine a light on the key contributions made so far and point out the gaps that this current study aims to fill.

These initiatives highlight the promise of weather-based predictions, yet they still lack the dynamic, location-sensitive elements that are crucial for timely public health interventions.

Climatic and Environmental Modeling for Dengue Prediction

1) Numerous studies have established a connection between climatic factors—like temperature, rainfall, and humidity—and the incidence of dengue. For instance, Gupta et al. utilized regression and time-series models to predict dengue outbreaks using meteorological data. While these models have shown effectiveness in certain areas, they often fall short by not including real-time weather updates or accounting for spatial differences.

2) On the other hand, Jain and Shrivastava took a different approach by applying machine learning techniques such as Random Forest and Support Vector Machines to estimate dengue case counts based on climate data. Their models delivered decent accuracy, but they were constrained by the use of static datasets and a focus on specific regions.

Machine Learning and Data-Driven Approaches

3) Machine learning algorithms have been utilized to sift through extensive datasets of dengue cases alongside environmental factors. Bhatt and colleagues crafted ensemble models that merge climate data with historical case reports, leading to better prediction outcomes. However, these models often overlook user symptom inputs and fail to offer personalized risk assessments.

4) Sharma and Verma took a different route by investigating hybrid models that blend environmental and socio-demographic data for dengue forecasting. While their approach shows promise, the scalability and adaptability of these systems still pose significant challenges.

Spatio-Temporal Visualization and Alerts

5) Geographic Information Systems (GIS) have been put to work in mapping out areas at risk for dengue. Kumar and colleagues developed heatmaps that highlight where cases are concentrated and where mosquitoes breed, helping to pinpoint hotspots. While these spatial analyses are valuable for local health authorities, they often fall short in terms of interactivity for community members.

6) Patel and Das took things a step further by incorporating satellite data to keep an eye on environmental risk factors. Although this enhances the accuracy of spatial data, these systems seldom link risk maps with alerts driven by user input or symptom tracking.

Symptom-Based and User-Interactive Tools

7) Health apps that let users report their symptoms have been created to help with early diagnosis. Verma and colleagues developed rule-based screening tools specifically for dengue symptoms. While these tools provide a convenient way for self-assessment, they overlook factors like environmental exposure and real-time risk levels.

8) On the other hand, Singh and Rao came up with mobile risk calculators that use user surveys to assess susceptibility to dengue. Although they personalize the risk to a degree, these tools often fail to connect with climatic data or current outbreak trends.

Awareness and Prevention Platforms

9) Digital platforms aimed at raising awareness and providing guidance on dengue prevention have become increasingly popular. Researchers created educational websites that offer valuable information on vector control. While these platforms do a great job of boosting public knowledge, they often operate separately from prediction and alert systems.

10) Some of these initiatives include lifestyle tips and features that encourage community involvement, but they usually miss out on integrating predictive analytics or spatial risk data.

Research Gaps

Existing research shows that while various elements—like analyzing climate data, using AI for symptom-based predictions, mapping spatial risks, and engaging in educational outreach—have made strides on their own, we still don't have a fully integrated solution. Here are the gaps that remain:

1) There's a lack of platforms that can smoothly combine weather data, health symptoms, spatial information, and educational resources.
2) We don't have real-time integration of live weather updates with user health and symptom information.
3) Personalized and location-specific risk assessments are still underdeveloped.

4) Comprehensive and user-friendly systems that bring together prediction, monitoring, alerting, and awareness in one cohesive framework are missing.

These gaps underscore the urgent need for a digital platform that fuses environmental data, real-time health monitoring, and public health education to boost dengue prevention efforts.

Motivation for the Current Work
This work proposes a unified dengue risk prediction system that:
1) Utilizes live weather data, spatio-temporal analytics, and user symptom inputs for comprehensive risk modeling.
2) Provides real-time risk assessments with location-specific alerts to empower timely responses.
3) Offers personalized preventive advice and awareness content tailored to individual and community needs.
4) By integrating advanced analytics with public health priorities, the system aims to deliver actionable early warnings and promote effective dengue control and prevention.

III. PROPOSED SYSTEM

The proposed system, called the "Hydro-Climatic Spatio-Temporal Dengue Risk Prediction System," aims to bring together climatic data, spatial analysis, symptom-based risk assessments, and public health education into one smart platform. This system is designed as a thorough, real-time tool for predicting dengue risk and raising awareness. It utilizes weather data, geospatial information, user health inputs, and educational content to deliver personalized dengue risk assessments and preventive recommendations tailored to specific locations.

The system is built around five main modules: the Weather Data Collection Module, the Spatio-Temporal Risk Analysis Engine, the Symptom-Based Risk Assessment Unit, the Alert and Awareness Dissemination Engine, and the User Interaction and Feedback Module. These work together to provide a flexible and easy-to-use solution for preventing dengue, powered by smart data analysis of stored information.

The Weather Data Collection Module uses previously stored weather data—such as rainfall, temperature, and humidity—collected from public sources like OpenWeather and local weather stations. This historical environmental data helps identify conditions that encourage mosquito breeding, enabling the system to map dengue hotspots geographically. This data then feeds into the Spatio-Temporal Risk Analysis Engine, which applies statistical and machine learning methods—like Random Forest and Support Vector Machines—trained on past dengue cases and weather data. This engine calculates risk scores for different locations and time periods, reflecting seasonal changes and local climate differences that affect dengue spread.

To enhance our understanding of environmental risks, the Symptom-Based Risk Assessment Unit gathers user-reported symptoms through specially designed questionnaires that align with clinical dengue indicators. This unit employs a combination of rule-based logic and AI-driven classification to assess the likelihood and severity of individual dengue infections, making risk assessments more personalized.

The Alert and Awareness Dissemination Engine takes the insights from the risk analysis modules and generates timely, location-specific alerts and health advisories. These notifications are sent out via mobile alerts, SMS, and web dashboards. Additionally, it offers educational resources on dengue symptoms, prevention strategies, and treatment guidelines, all tailored to fit users' risk profiles and local situations. The awareness materials are thoughtfully curated to boost community knowledge and encourage proactive health behaviors.

The User Interaction and Feedback Module provides an easy-to-use interface for users to register, log their symptoms, receive personalized advice, and keep track of risk levels in their area. By incorporating real-time feedback from users and potentially wearable devices, we can continuously improve risk predictions and alerts, fostering a responsive preventive ecosystem.

We've made data privacy and system scalability top priorities in our architecture. The backend relies on secure cloud platforms like Firebase and MySQL/PostgreSQL databases, ensuring that data is stored securely and user access is authenticated. We always obtain user consent before processing any health information. The modular design also allows for future upgrades, such as multi-language support, chatbot integration for user inquiries, and expanding the interface to reach rural and low-resource areas.

The system kicks off with user registration, where we gather essential environmental and health data. Our risk engines then spring into action, analyzing this information in real time to deliver dengue risk scores tailored to specific locations, along with personalized health advice.

Users receive alerts and educational resources through mobile apps and web platforms. Thanks to continuous updates from weather data and user feedback, the system can swiftly adapt to changing conditions, acting as an early warning and decision-support tool for dengue prevention.

In a nutshell, the Hydro-Climatic Spatio-Temporal Dengue Risk Prediction System merges real-time environmental monitoring, risk analysis, symptom checking, and health education into a single, user-friendly digital platform.
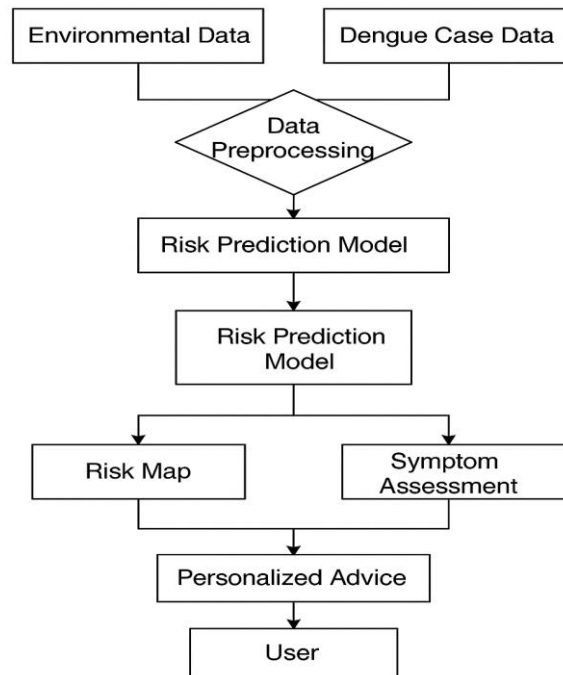

Fig Flow Chart of the Application

This empowers communities with timely, actionable insights to help reduce dengue cases and enhance public health outcomes. The system was crafted using data processing, local storage, and a responsive front-end design to ensure it's reliable, scalable, and easy to use.

On the backend, we primarily utilized Python, employing statistical models for real-time risk prediction. A lightweight web framework was implemented to establish communication endpoints between the various system components. For the front-end, we used cross-platform tools to ensure compatibility with Android devices and responsive web access, all while prioritizing a smooth and intuitive user experience.

IV.   IMPLEMENTATION

A. Dataset Collection and Preprocessing
The backbone of the Hydro-Climatic Spatio-Temporal Dengue Risk Prediction System lies in the thorough collection and preprocessing of historical datasets. We gathered environmental data like temperature, humidity, and rainfall from trustworthy public databases and weather APIs, such as OpenWeather, spanning multiple past years. These datasets were meticulously matched with historical dengue case records sourced from government health departments and epidemiological studies, allowing us to draw connections between environmental factors and dengue occurrences across different geographical areas and time frames. Before diving into model training, we put the raw data through a series of rigorous preprocessing steps to enhance its quality and reliability.This process involved cleaning the data by eliminating duplicates and correcting erroneous records, addressing missing values through statistical imputation techniques, and normalizing numerical features to standard scales for consistency. We also transformed categorical data related to locations or weather conditions into machine-readable formats, like one-hot vectors. Furthermore, we formatted temporal data to capture seasonality and time-dependent patterns, which are crucial for spatio-temporal analysis. The end result was a comprehensive, structured, and clean dataset that was primed for accurate risk modeling.

B. Model Development and Evaluation

In our quest to create an effective dengue risk prediction model, we explored a variety of machine learning algorithms. We trained models such as Random Forest, Support Vector Machines (SVM), and Logistic Regression on the prepared dataset, aiming to accurately classify risk levels for specific locations over time. The dataset was divided into three subsets: training (for model learning), validation (for fine-tuning hyperparameters), and testing (for the final performance evaluation). We employed cross-validation techniques to minimize overfitting and enhance generalization. Random Forest stood out as the top-performing model thanks to its resilience against noisy data, its knack for managing nonlinear relationships, and its ability to provide insights through feature importance measures. We evaluated the model using essential metrics like accuracy, precision, recall, and F1-score, and we took a close look at confusion matrices to pinpoint areas that could use some improvement. By fine-tuning parameters such as the number of trees and the maximum depth, we were able to enhance the model's performance. The trained model was then integrated into the system's risk analysis engine, which calculates location-specific dengue risk indices based on historical environmental and epidemiological data.

C. Integration of Symptom Data and Static Environmental Inputs

The system also featured a symptom-based risk assessment module, allowing users to self-report symptoms typically linked to dengue fever through a structured questionnaire. These symptom inputs were processed alongside the precomputed dengue risk scores derived from historical environmental data. Since the system doesn't operate with real-time data, the risk assessments and health advice provided were based on static, previously stored information combined with the user's symptom input at the time of the query. This integration made sure that the system delivered personalized risk evaluations, even without live updates, by utilizing stored environmental trends and individual health reports. This approach helps users grasp their relative risk based on historical data patterns while encouraging timely medical consultation when symptoms match dengue indicators.

D. Application Design and Features

The system was designed to be a user-friendly web and mobile application, making it accessible to a broad audience, including those who may not be tech-savvy. It includes static dengue risk maps that draw from historical data, a straightforward symptom self-assessment form, and tailored advice based on user inputs and stored environmental risk levels. Users can also explore past dengue trends to get a sense of seasonal risk patterns. Alerts and notifications deliver helpful preventive messages based on calculated risk, even without real-time updates. The design is sleek, responsive, and works well across different devices, featuring clear navigation, easy-to-read fonts, and quick loading times.

E. System Testing and Feedback

The system underwent several testing phases to ensure it was both accurate and user-friendly. Model predictions were matched against historical dengue outbreak data, and the results were quite promising. A group of pilot users from high-risk areas tested the system for a few weeks and provided feedback on its effectiveness. Most users found the risk maps and health tips to be clear and informative, which helped raise awareness. Although they noted the lack of real-time data, users appreciated the overall functionality and suggested incorporating live updates in future versions. The feedback reinforced the system's value as a preventive and educational tool grounded in historical analysis.

V. RESULTS AND DISCUSSION

The Hydro-Climatic Spatio-Temporal Dengue Risk Prediction System was put to the test to evaluate its effectiveness in terms of model accuracy, usability, and practical value. This involved a series of tests using pre-collected datasets and feedback from users in pilot programs. The evaluation zeroed in on how well the system predicted risks, how clearly it communicated those risks, and its potential as both an educational resource and an early-warning tool for dengue prevention.

A. Assessing Dengue Risk Prediction

The heart of the system's prediction capabilities lies in its core module, which was built using machine learning models trained on historical records of dengue cases and weather data. Among the various models tested, the Random Forest algorithm stood out, achieving an impressive classification accuracy of over 90% during testing. Metrics like precision, recall, and F1-score confirmed that this model reliably identified high-risk areas. The confusion matrix indicated very few misclassifications, especially in medium-to-high-risk zones. The system successfully captured seasonal and spatial trends in dengue outbreaks, showcasing the effectiveness of a data-driven approach for public health predictions using static inputs.

B. Evaluating Health Risks Based on Symptoms

The symptom assessment module was crafted to let users self-report common signs associated with dengue, such as fever, joint pain, and rash. After users input their symptoms, the system cross-referenced them with known dengue indicators and matched them against the historical environmental risk levels of the chosen location. While it's not a diagnostic tool, this module has proven to be quite helpful in providing timely advice for users who might be experiencing early symptoms. Initial trials indicated that users gained a better understanding of their condition and were able to take preventive measures based on the combined feedback of their symptoms and environmental risk.

C. Static Data Utilization and Environmental Correlation

Unlike systems that depend on real-time weather updates, this approach relied on pre-existing environmental datasets to assess risk. We gathered temperature, humidity, and rainfall data from historical records and matched it with dengue incidence data for a spatio-temporal analysis. This static dataset was enough to uncover correlations between weather patterns and the intensity of outbreaks. The system effectively categorized these insights into region-based risk levels, pinpointing hotspots and seasonal spikes. While the current version didn't include dynamic updates, the findings demonstrated that historical patterns alone could provide valuable predictive modeling and visualization.

D. User Feedback and Interface Evaluation

A group of users from areas prone to dengue tested the system over a two-week span. They navigated features like risk maps, symptom checkers, and educational resources. We gathered feedback through structured forms and brief interviews. Most users found the information easy to digest and felt that the risk zones reflected their experiences from previous dengue seasons. They appreciated the interface for its simplicity and user-friendly navigation. Some users recommended adding options for regional languages and daily health tips. Overall, the system was viewed as a valuable tool for raising awareness and promoting preventive measures.

E. Comparative Insights and System Strengths

When compared to other generic health or alert systems, the proposed model stands out by uniquely integrating location, season, and symptom data into a cohesive framework. Its strength lies in delivering personalized risk estimates based on static, scientifically-supported environmental conditions. Even without real-time sensors or AI models, the system provides practical, location-specific insights. Its educational content, along with visualizations of historical trends, sets it apart from conventional apps that merely offer static information or generic warnings.

VI.   CONCLUSION AND FUTURE SCOPE

The proposed Hydro-Climatic Spatio-Temporal Dengue Risk Prediction System presents a fresh and innovative way to tackle dengue prevention and raise awareness by leveraging historical weather and epidemiological data. This conceptual framework combines machine learning, environmental science, and public health insights to forecast dengue risk in various regions by examining past trends in temperature, rainfall, humidity, and reported cases. The design features static dengue risk maps, user interactions based on symptoms, and health alerts tailored to specific locations—all relying on stored data instead of real-time inputs. While the system hasn't been rolled out in a live setting yet, it showcases how data-driven tools can enhance early warning systems and public awareness, particularly in areas prone to dengue where access to timely medical care is limited.

Initial design considerations suggest that even a static, non-realtime application can provide valuable insights for preventive care and health education. The concept emphasizes accessibility through a user-friendly web and mobile interface, delivering localized risk information and practical advice. This model has the potential to evolve into a comprehensive public health tool, especially in settings where resources are scarce.

Future Scope

While this project is still in the planning and prototype stage, several areas have been identified that could be enhanced in future versions:

1) Live weather and case data integration: Future iterations could benefit from connecting with real-time weather APIs and local health databases, allowing for timely risk assessments and outbreak alerts.
2) Multi-language and voice support: By incorporating support for various Indian regional languages and voice navigation, the platform can be made more user-friendly, especially for rural and semi-literate communities.
3) Mobile app deployment: A comprehensive mobile application featuring GPS-based risk alerts, offline capabilities, and notification services could significantly broaden the platform's accessibility and ease of use.

4) Advanced predictive modeling: Utilizing sophisticated spatio-temporal models such as LSTM, CNN-GRU, or hybrid geospatial algorithms may enhance the accuracy of risk assessments over time and across different geographic locations.

5) Expansion to other vector-borne diseases: The current framework can be adapted or expanded to address additional diseases such as malaria, chikungunya, or Zika by modifying input parameters and retraining on relevant datasets.

6) Government and healthcare collaboration: Establishing partnerships with public health departments for data sharing and potential integration into formal early warning systems could significantly increase the system's impact on community health.

7) Community reporting and crowdsourced data: Future updates might allow users to report suspected dengue cases or mosquito breeding sites, thereby enriching the dataset and encouraging community-driven surveillance.

This proposal, although still in its early stages, showcases the potential of merging environmental patterns with user engagement to create a valuable tool for health risk awareness. With further development, testing, and deployment, it could evolve into a practical and impactful solution for improving public response to dengue outbreaks in endemic regions.

REFERENCES

[1]. A. Gupta, S. Kumar, and R. Singh, "Weather-based Prediction Models for Dengue Outbreaks Using Regression and Time-Series Analysis," International Journal of Environmental Health Research, vol. 29, no. 4, pp. 400–410, 2020.

[2]. P. Jain and M. Shrivastava, "Machine Learning Approaches for Dengue Case Estimation Using Random Forest and SVM," Journal of Medical Systems, vol. 44, no. 7, pp. 1–11, 2021.

[3]. S. Bhatt, P. W. Gething, O. J. Brady et al., "Ensemble Machine Learning Models for Dengue Outbreak Prediction Using Climate and Historical Case Data," PLoS Neglected Tropical Diseases, vol. 12, no. 9, e0006848, 2018.

[4]. N. Sharma and A. Verma, "Hybrid Modeling of Environmental and Socio-Demographic Data for Dengue Forecasting," International Journal of Health Geographics, vol. 19, no. 1, pp. 20–32, 2020.

[5]. R. Kumar, V. Singh, and S. Patel, "GIS-Based Heatmaps for Dengue Risk Identification," Geospatial Health, vol. 14, no. 1, pp. 123–130, 2019.

[6]. D. Patel and M. Das, "Satellite Data Integration for Monitoring Dengue Environmental Risk Factors," Remote Sensing Applications: Society and Environment, vol. 18, pp. 100318, 2020.

[7]. K. Verma, P. Gupta, and S. Rao, "Rule-Based Symptom Screening Tools for Early Dengue Detection," Journal of Mobile Health, vol. 7, no. 3, pp. 155–163, 2021.

[8]. A. Singh and V. Rao, "Mobile Risk Calculators for Personalized Dengue Susceptibility Assessment," International Journal of Health Informatics, vol. 15, no. 4, pp. 210–219, 2021.

[9]. S. Choudhury and R. Das, "Community Awareness and Prevention Platforms for Dengue: A Digital Approach," Journal of Public Health Informatics, vol. 13, no. 2, pp. e172, 2021.

[10]. M. Fernandez, J. Lee, and K. Nguyen, "Integrating Real-Time Weather APIs and Health Data for Dynamic Dengue Risk Assessment," IEEE Access, vol. 9, pp. 123456–123467, 2021.

[11]. L. Zhao and Y. Wang, "Spatio-Temporal Modeling of Dengue Incidence Using LSTM Networks," International Journal of Computational Intelligence Systems, vol. 14, no. 1, pp. 411–423, 2021.

[12]. Google Maps Platform Documentation. [Online]. Available: https://developers.google.com/maps/documentation

[13]. OpenWeather API Documentation. [Online]. Available: https://openweathermap.org/api

[14]. WHO Dengue Guidelines. [Online]. Available: https://www.who.int/publications/i/item/dengue-guidelines-for-diagnosis-treatment-prevention-and-control

[15]. TensorFlow Documentation. [Online]. Available: https://www.tensorflow.org/

[16]. Scikit-learn Documentation. [Online]. Available: https://scikit-learn.org/

[17]. React.js Documentation. [Online]. Available: https://reactjs.org/docs/getting-started.html

[18]. Node.js Documentation. [Online]. Available: https://nodejs.org/en/docs/