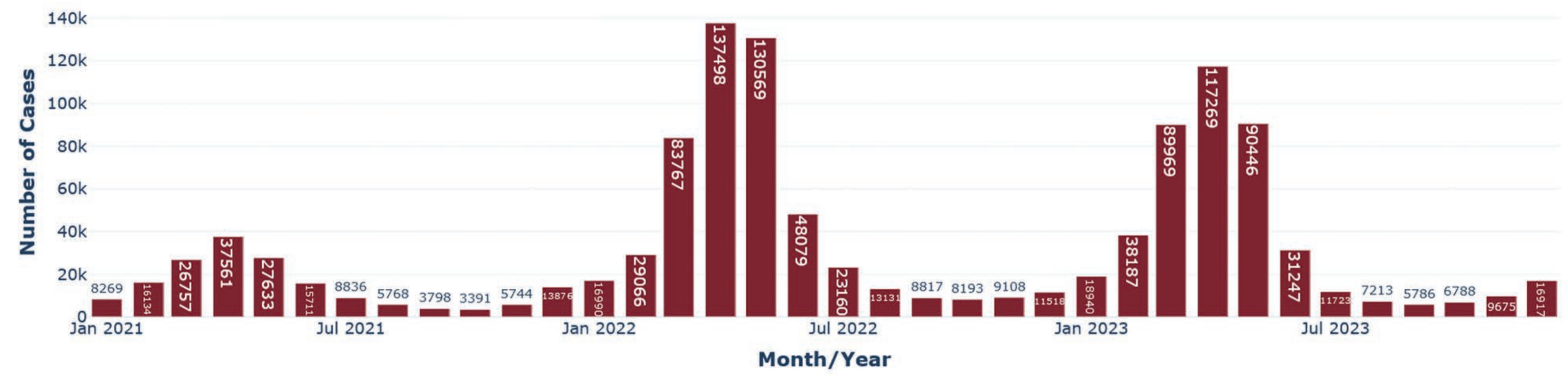
THE BRAZILIAN HIDDEN EPIDEMIC

Statistical Models Unveiling Dengue's Silent Spread

Introduction

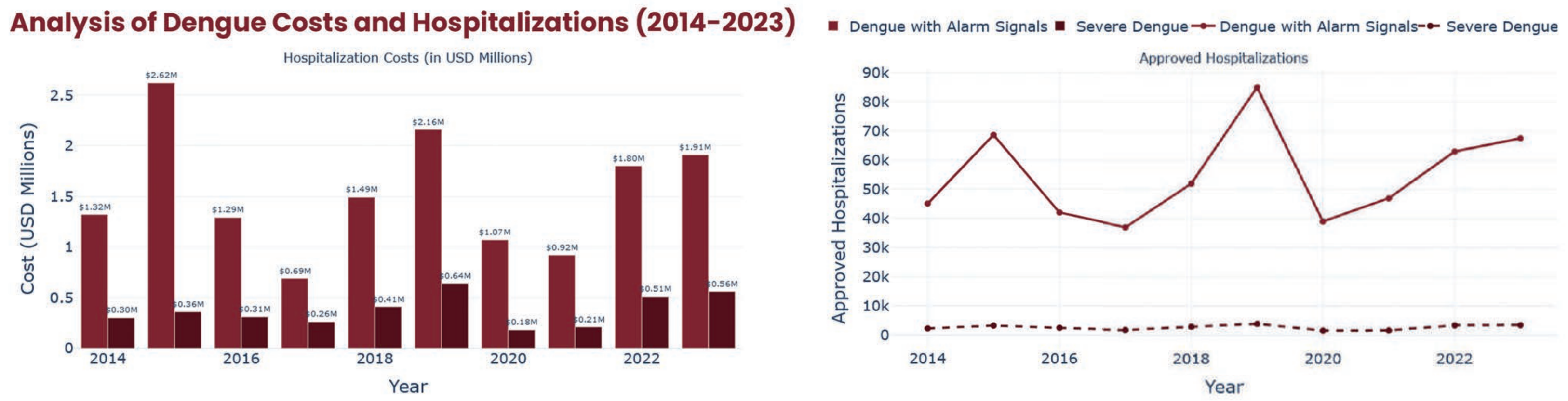
Dengue, a mosquito-borne infectious disease, affects approximately 390 million people annually, with an estimated 96 million symptomatic cases worldwide, placing immense pressure on healthcare systems and resulting in billions of dollars in economic losses. Early and accurate diagnosis is critical, as severe forms of dengue, such as dengue hemorrhagic fever, can lead to a 50% mortality rate if untreated. However, symptom overlap with other febrile illnesses and the unavailability of laboratory tests in over 60% of endemic regions hinder timely diagnosis. This study leverages Deep Learning to develop a predictive model trained on clinical and laboratory data, aiming to classify dengue severity with an accuracy target exceeding 90%. By using supervised learning techniques and robust evaluation metrics (e.g., sensitivity, specificity, and ROC curve analysis), this model seeks to surpass traditional methods, offering a scalable and cost-effective solution for rapid dengue diagnosis in resource-limited settings.

Temporal distribution of cases



The graph shows a seasonal pattern in dengue cases, peaking every year between April and May. This surge follows a steady rise from January to July, suggesting that factors like rainfall and temperature drive the spread. It emphasizes the need for proactive public health measures early in the year to reduce outbreaks.

Analysis of Dengue Costs and Hospitalizations (2014-2023)

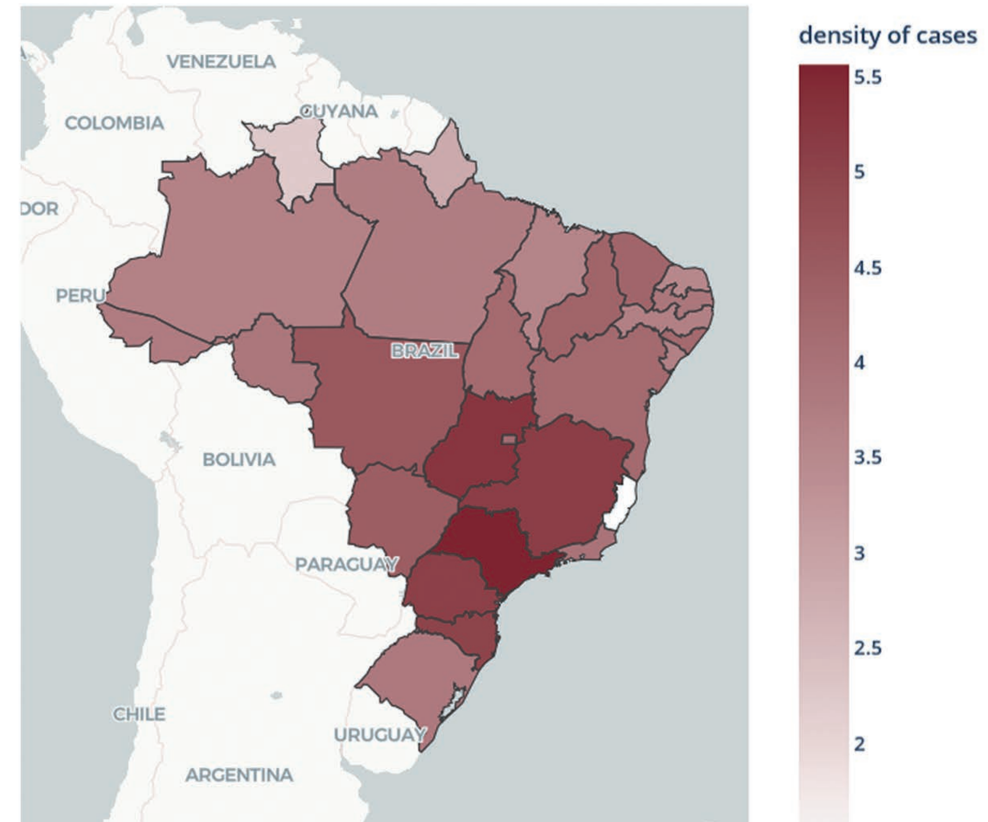


Dengue cases by severity—Mild, Moderate, and Severe—providing insight into the distribution and impact of the disease. A notable trend shows that severe cases, while less frequent, pose a significant burden, emphasizing the importance of early detection and intervention. This highlights the need for targeted public health strategies to reduce severe outcomes and improve patient care.

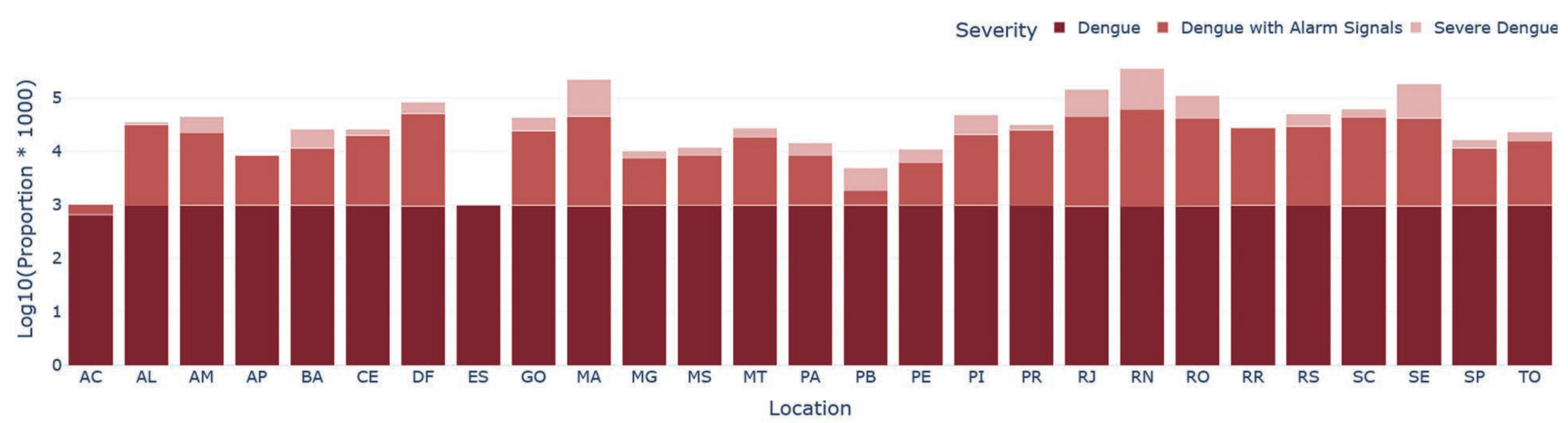
Method

Data from the Sistema de Informação de Agravos de Notificação (SINAN), covering 2021–2023, was preprocessed to ensure consistency and accuracy. Over 3M records underwent transformation, with categorical variables encoded and normalized using StandardScaler. Feature selection reduced dimensionality, retaining 90% of variance through PCA and identifying top predictors using Random Forest. Machine learning models, including logistic regression, decision trees, SVMs, and neural networks, were trained and evaluated on metrics like accuracy, sensitivity, and specificity. Cross-validation achieved an average accuracy of 92%, with hyperparameter tuning optimizing results. Independent test sets confirmed the model's robustness across scenarios.

This map shows the spread of dengue cases in Brazil, with São Paulo identified as a major hotspot. Other cities significantly affected include Goiânia and Belo Horizonte. The visualization highlights the uneven distribution of dengue cases across the country, underscoring the ongoing challenge of controlling the disease.



Logarithmic Severity Analysis of dengue cases by region (Scaled)



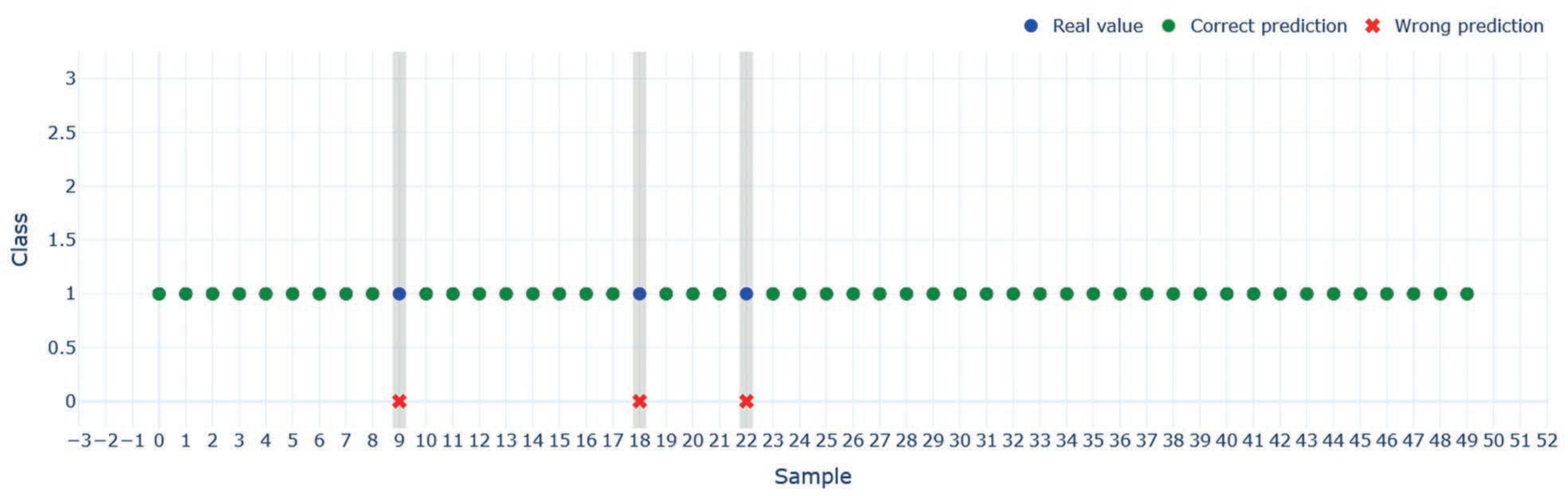
As we can see the severity of dengue cases is highly different across regions in Brazil, classified into three categories: Dengue, Dengue with Alarm Signals, and Severe Dengue. The logarithmic scaling (Log10) of proportions emphasizes regional differences, with Maranhão (MA), Rio de Janeiro (RJ), and Distrito Federal (DF) exhibiting higher proportions of severe cases. The visualization highlights the variability in dengue severity across states, underscoring regions requiring heightened focus for prevention and healthcare measures.

Observations and Solutions

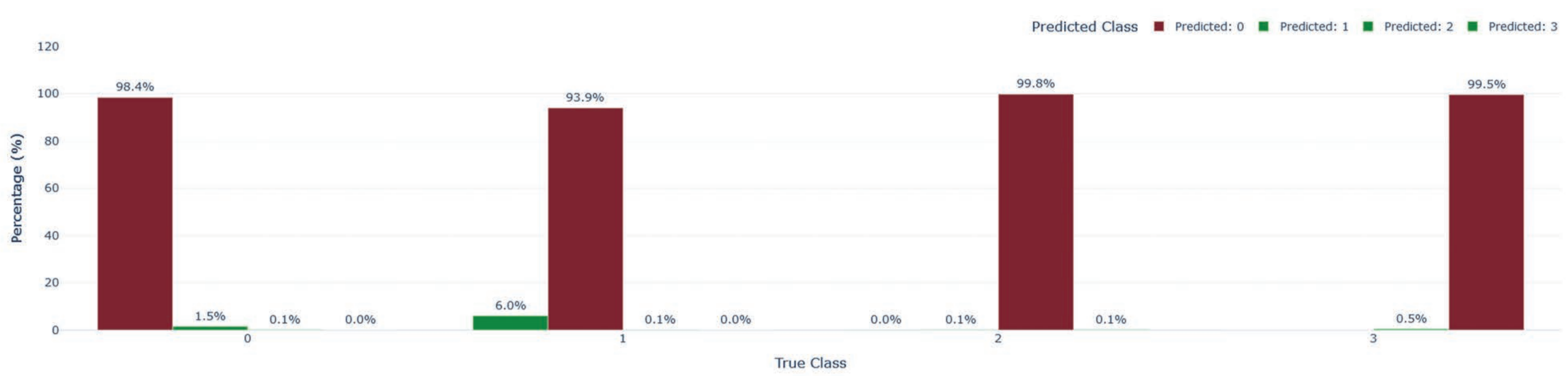
Dengue cases in Brazil show clear seasonal and regional variability, peaking between April and May due to favorable climatic conditions. Major hotspots include São Paulo, Goiânia, and Belo Horizonte, while Maranhão (MA), Rio de Janeiro (RJ), and Distrito Federal (DF) report higher proportions of severe cases, highlighting the need for targeted interventions.

To address these challenges, a Deep Learning model was developed to classify dengue cases into three severity levels—Dengue, Dengue with Alarm Signals, and Severe Dengue—alongside a non-dengue category. Trained on clinical and laboratory data, the model achieved 92% accuracy, offering a reliable, scalable tool for rapid pre-diagnosis. This solution enhances early risk stratification, optimizes healthcare resources, and improves patient outcomes, especially in resource-limited settings.

Comparison between correct and wrong predictions of dengue



Comparison by class



The Deep Learning model achieved 98.4% accuracy for Non-Dengue, 93.9% for Dengue, 99.8% for Dengue with Alarm Signals, and 99.5% for Severe Dengue, showcasing its precision in supporting early detection and risk stratification.

Conclusions

Dengue remains a significant global health challenge, with its seasonal and regional variability placing immense strain on healthcare systems, particularly in resource-limited settings. This study successfully developed a Deep Learning model to classify dengue severity into Non-Dengue, Dengue, Dengue with Alarm Signals, and Severe Dengue categories. Leveraging clinical and laboratory data, the model achieved an overall accuracy of 92%, with exceptional precision across all severity levels, including 99.5% for Severe Dengue. These results underscore the model's potential as a scalable, cost-effective tool for early diagnosis, enabling timely intervention, optimized resource allocation, and improved patient outcomes. By identifying high-risk regions and periods, the findings also highlight the importance of targeted public health measures to mitigate outbreaks and reduce the burden of severe cases.

References: ZHOU Q. M. et al., Diagnostic and Prognostic Research (2021); BHATT S. et al., Nature (2013); CARVALHO F. H. et al., Journal of Medical Systems (2019); CHAR D. S., SHAH N. H., MAGNUS D., The New England Journal of Medicine (2018); ESTEVA A. et al., Nature Medicine (2019); GUZMAN M. G., HARRIS E., The Lancet (2015); HUNSPERGER E. A. et al., PLoS Neglected Tropical Diseases (2014); JAIN S., PAL R., Current Clinical Pharmacology (2020); MILINOVICH G. J. et al., The Lancet Infectious Diseases (2014); OBERMEYER Z. et al., Science (2019); OBERMEYER Z., EMANUEL E. J., The New England Journal of Medicine (2016); PEELING R. W. et al., Nature Reviews Microbiology (2010); RAJKOMAR A., DEAN J., KOHANE I., The New England Journal of Medicine (2019); TING D. S. W. et al., Nature Medicine (2020); WIENS J., SHENOY E. S., Clinical Infectious Diseases (2018); RATTO A. C. P., SILVA I. L. A., FERREIRA L. M., Ó M. M., Brazilian Journal of Implantology and Health Sciences, Vol. 6, Issue 12 (2024), pp. 959–974.