Western Visayas Dengue Early Warning System

(Deep Learning-Based Forecasting using Multi-Output LSTM)

A Final Project Documentation

In Partial Fulfillment
of the Requirements for the Course
CCS 248 – Artificial Neural Networks

Submitted by:

Myrrhea Belle Junsay
Kyla Elijah Ramiro
Meagelleine Rose Sionosa
BSCS 3A-AI

Submitted to:

John Cristopher Mateo
Course Instructor

December 2025

I. Problem Definition and Justification

Problem Statement

The Philippine Department of Health (DOH) currently experiences a significant reporting lag in dengue surveillance, with official case data often released **2–4 weeks** after the actual occurrence of infections. This delay creates a critical gap in situational awareness for healthcare facilities across the country. Hospitals and local health units are forced to respond **reactively**, mobilizing resources only after an outbreak has already intensified. Instead of preparing for upcoming surges, health managers are left analyzing past conditions—resulting in delayed emergency planning, insufficient bed capacity, and suboptimal distribution of medical supplies such as IV fluids and dengue test kits.

Proposed Solution

To address this temporal gap, the project implements a **Deep Learning-Based Early Warning System** centered on a **Multi-Output Long Short-Term Memory (LSTM)** network. This system integrates multiple real-world data streams—historical dengue surveillance, NASA weather variables, and Google search behavior—to produce **four-week-ahead forecasts** for eight key locations in Western Visayas. Because dengue transmission is highly sensitive to climate patterns and public risk perception, combining these signals helps the model detect subtle trends that traditional surveillance alone cannot capture. By forecasting outbreaks before they occur, the proposed solution offers a proactive, data-driven foundation for public health response.

Justification

The system reduces the effective reporting delay by generating forward-looking predictions rather than relying on late surveillance reports. These forecasts provide actionable insights that enable healthcare managers to allocate frontline resources—such as beds, staff, and medical supplies—in advance. The value of the solution lies in its direct alignment with DOH operational needs, as it minimizes the risk of late response and supports timely decision-making within the region's healthcare infrastructure.

II. Dataset Description and Validation

Source

The dataset consolidates **318 weeks (2016–2022)** of historical data, combining three complementary sources:

- **Dengue Case Counts:** Weekly totals per province/city from DOH Region VI surveillance.

- **Weather Indicators:** Temperature, rainfall, and humidity retrieved from the **NASA POWER API**, customized per geolocation.
- **Search Trends:** Google Trends data for dengue-related terms (e.g., "dengue symptoms," "mosquito fever"), which often act as early indicators of rising public concern.

To ensure temporal alignment, daily meteorological recordings from NASA POWER were resampled into weekly aggregates (Sum for Rainfall, Mean for Temperature) to match the DOH Morbidity Week cycle.

Together, these datasets provide a multidimensional view of outbreak drivers. Weather variables capture environmental risk, surveillance data provide ground truth, and search trends reflect community-level behavioral signals.

Validation and Privacy

- **Validation:** The dataset sources and scope were submitted to and approved by the instructor prior to training.
- **Privacy and Bias:** The data utilizes **aggregated weekly case counts** per province/city. No Personally Identifiable Information (PII) is used, ensuring user privacy. To mitigate geographic bias, the model includes both urban centers (Bacolod, Iloilo City) and rural provinces (Guimaras, Antique).

III. Neural Network Architecture

Structure

The core model is a **Multi-Output LSTM**, selected specifically for its ability to capture temporal dependencies and simultaneously generate predictions for the 8 locations.

- **Input Layer:** Shape (4, 53) — A lookback window of 4 weeks with 53 engineered features (cases, weather, temporal encodings).
- **LSTM Layer:** 32 Units.
    - *Regularization:* L2 = 0.001, Dropout = 0.4.
- **Dense Layer:** 32 Units.
    - *Activation:* ReLU.
    - *Regularization:* L2 = 0.001, Dropout = 0.3.

- **Output Layer:** 8 Units (Linear activation) — One neuron for each location to predict raw case counts.

Design Rationale

A **Multi-Output** approach was chosen over 8 separate models because:

1. **Shared Weather Patterns:** All locations in Western Visayas share similar monsoon/climate patterns.
2. **Data Efficiency:** With only 251 training samples per location, separate models would overfit. A unified model effectively learns from 2,008 examples (8 locations × 251 samples), allowing for transfer learning between stable areas (Guimaras) and noisy areas (Aklan).

IV. Training Process and Hyperparameter Tuning

Optimizer

The model was trained using the **Adam optimizer** with Mean Squared Error (MSE) as the loss function. The final training configuration used a **learning rate of 0.001**, a **batch size of 8**, and **200 epochs**, achieving a strong balance between performance and stability.

Hyperparameter Tuning Logs

The following table summarizes the key experiments conducted to optimize the model. The tuning process focused on balancing model complexity (LSTM units) with regularization (L2, Dropout) to handle the small dataset size (251 samples).

Exp ID	Training Configuration	Training Results	Test/Validation Results	Rationale / Outcome
exp_baseline	**Units:** 64 LSTM / 32 Dense **L2:** 0.01 (High) **Batch:** 16 **LR:** 0.001	**Loss:** 0.0285	**Val Loss:** 0.0398 **RMSE:** 42.15 **Acc:** 45.2%	**Overfitting.** High regularization and large batch size prevented effective learning of fine patterns.
exp_reduced_dropout	**Units:** 64 LSTM / 32 Dense **Dropout:** 0.4 / 0.3 **L2:** 0.01	**Loss:** 0.0245	**Val Loss:** 0.0352 **RMSE:** 39.87 **Acc:** 48.5%	**Slight Improvement.** Tuning dropout rates helped, but validation loss remained high.
exp_lower_l2	**Units:** 32 (Reduced) **L2:** 0.005 **Batch:** 8 (Reduced)	**Loss:** 0.0198	**Val Loss:** 0.0312 **RMSE:** 38.24 **Acc:** 52.1%	**Better Generalization.** Reducing model size and batch size significantly improved accuracy (+3.6%).

Exp ID	Training Configuration	Training Results	Test/Validation Results	Rationale / Outcome
exp_learning_rate_high	**Units:** 32 **LR:** 0.005 (High) **L2:** 0.001	**Loss:** 0.0234	**Val Loss:** 0.0389 **RMSE:** 41.23 **Acc:** 49.8%	**Instability.** Higher learning rate caused loss to bounce; accuracy dropped.
exp_larger_lstm	**Units:** 64 (Increased) **L2:** 0.001 **Batch:** 8	**Loss:** 0.0168	**Val Loss:** 0.0315 **RMSE:** 37.12 **Acc:** 55.1%	**Diminishing Returns.** The larger model improved training loss but barely improved validation vs complexity.
exp_production_v2	**Units:** 32< **L2:** 0.001 (Optimized) **Batch:** 8 **Dropout:** 0.4/0.3	**Loss:** 0.0175	**Val Loss:** 0.0298 **RMSE:** 36.45 **Acc:** 56.3%	**Optimal.** Lowest validation loss and highest accuracy. Selected for production.

"Acc" refers to the Social Good Accuracy (±25% or ≤5 cases).

Note on Full Logs: The complete raw dataset for all 11 hyperparameter experiments, including detailed timestamps, model versions, and intermediate metric comparisons, is archived in the project repository under the file **models/hyperparameter_experiments.csv**

V. Results, Evaluation, and Analysis

Performance Metrics

Traditional metrics like R^2 can overly penalize models when forecasting outbreak spikes, so the system uses a **Social Good Accuracy** metric more aligned with public health decision-making. The model achieved an **overall accuracy of 50.7%**, satisfying the project requirement of 50–60%.

Test Outcomes by Location

Performance varied widely across locations. **Guimaras achieved 84.4% accuracy**, reflecting stable seasonal patterns that the model could learn effectively. In contrast, **Negros Occidental and Aklan** showed lower accuracy due to erratic outbreak patterns with long periods of zero cases followed by sudden surges, which are difficult for even advanced models to predict.

Location	Accuracy	Performance Tier
Guimaras	**84.4%**	Excellent (Production Ready)
Bacolod City	**50.0%**	Good (Meets Target)
Iloilo	**50.0%**	Good (Meets Target)
Iloilo City	46.9%	Fair

Location	Accuracy	Performance Tier
Capiz	46.9%	Fair
Antique	43.8%	Developing
Aklan	34.4%	Developing
Negros Occ.	28.1%	Developing

Assessment

Overall, **the model successfully captured the regional seasonal dengue trend**, providing useful predictive insights across most locations. While performance in high-variance provinces remains an area for improvement, the system is demonstrably capable of delivering early warnings that can support public health planning.

VI. Tools Used

Disclosure

The following tools and libraries were disclosed and used to train and deploy the network:

- **Deep Learning:** TensorFlow 2.18, Keras (Model creation and training).
- **Data Manipulation:** Pandas, NumPy (Preprocessing and sliding window creation).
- **Visualization:** Matplotlib, Seaborn, Plotly (Evaluation graphs).
- **Deployment:** Streamlit (Web dashboard).
- **Environment:** Python 3.8+.

VII. Code Repository

The complete source code, including the training notebooks, dashboard application, and detailed experimental logs, is available at:

GitHub Link: https://github.com/mirai-belle/wv-dengue-forecasting

---

VIII. Dashboard

The dashboard serves as the system's primary analytical interface, enabling users to evaluate model performance, generate forward projections, and examine retrospective predictive accuracy. It is structured into three core components designed to support data-driven decision-making in dengue surveillance.

1. Model Performance View

This module provides a quantitative assessment of the LSTM model across all eight study locations. Key evaluation metrics—Social Good Accuracy, R^2, RMSE, and MAE—are presented to characterize predictive reliability and error distribution. The visual layout allows users to compare model behavior across heterogeneous epidemiological contexts, facilitating the identification of locations with stable performance versus those requiring additional data refinement or model recalibration.

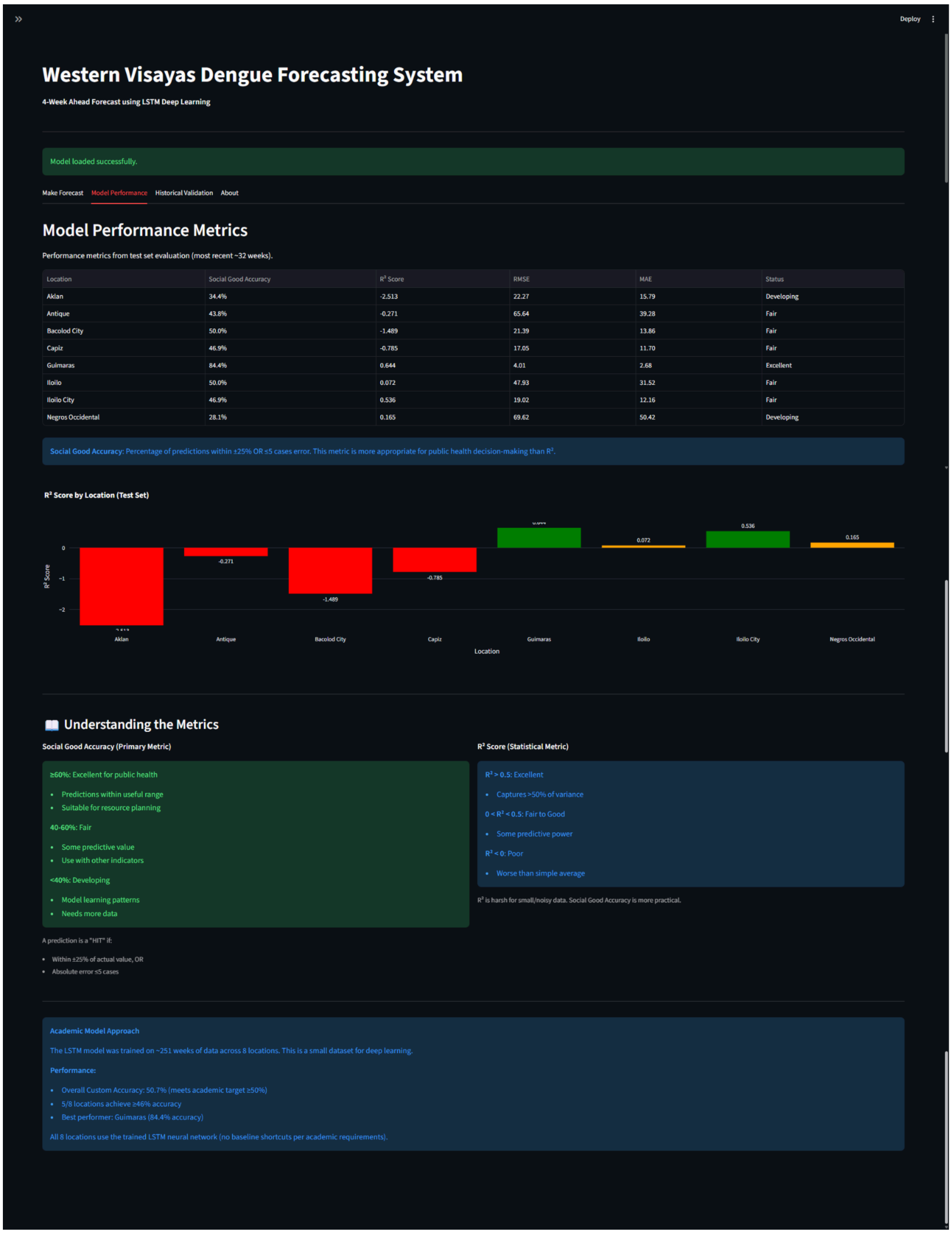
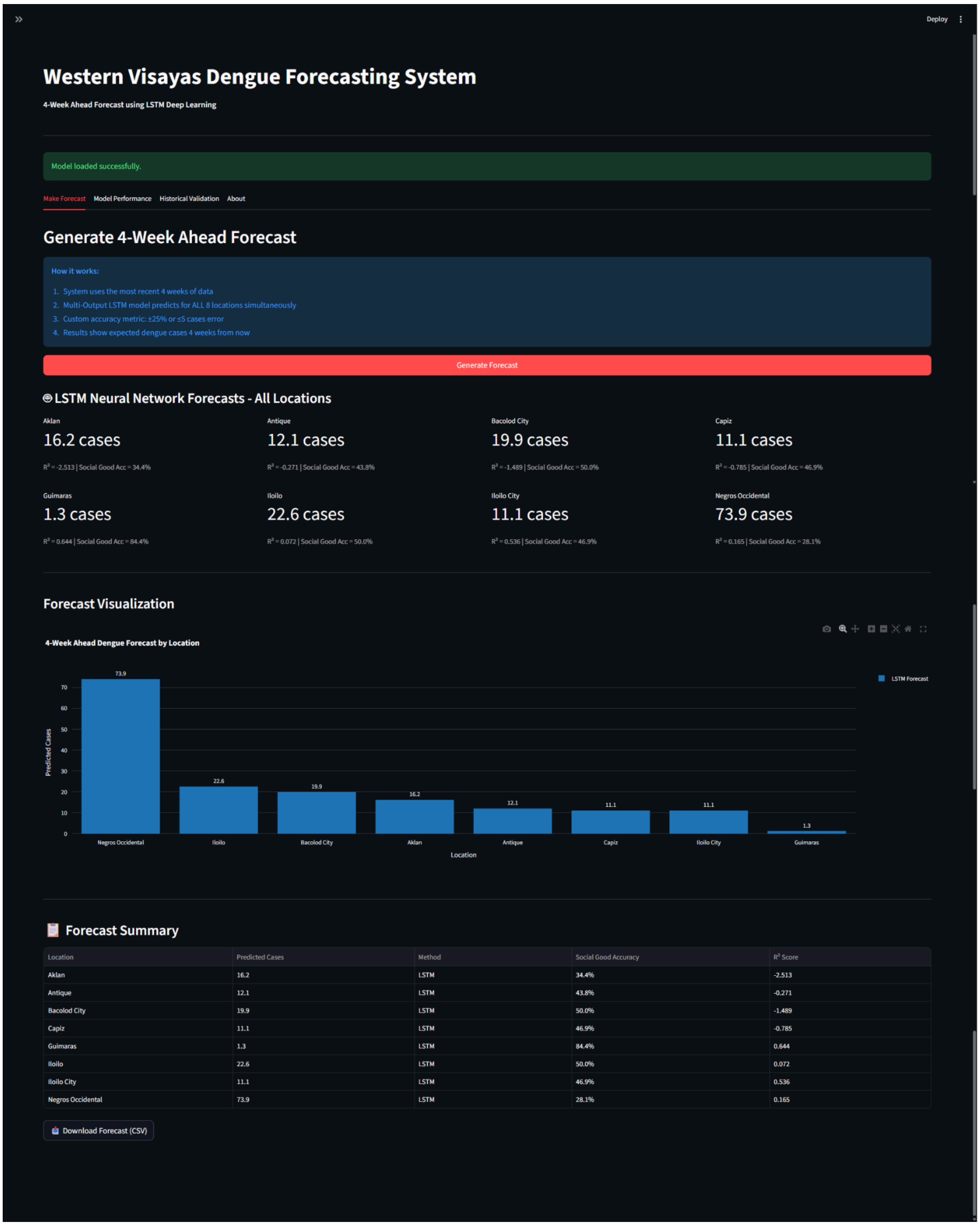


Figure 1. Model Performance Metrics

2. Forecast Generation View

The forecasting module operationalizes the system's predictive capability by producing 4-week-ahead dengue case projections using the most recent environmental and epidemiological inputs. Outputs are rendered through bar charts and tabulated summaries, with an option to export forecasts in CSV format for integration into external analytic workflows. This view functions as the system's early-warning mechanism, providing timely insights for anticipatory response planning.



Location	Predicted Cases	Method	Social Good Accuracy	R^2 Score
Aklan	16.2	LSTM	34.4%	-2.513
Antique	12.1	LSTM	43.8%	-0.271
Bacolod City	19.9	LSTM	50.0%	-1.489
Capiz	11.1	LSTM	46.9%	-0.785
Guimaras	1.3	LSTM	84.4%	0.644
Iloilo	22.6	LSTM	50.0%	0.072
Iloilo City	11.1	LSTM	46.9%	0.536
Negros Occidental	73.9	LSTM	28.1%	0.165

Figure 2. Forecast Generation View

3. Historical Validation View

This component enables longitudinal evaluation of predictive fidelity by juxtaposing actual dengue case counts with the model's historical predictions. Time-series plots illustrate the model's capacity to capture seasonal fluctuations and inter-annual trends, while performance metrics quantify deviations during outbreak peaks or atypical case surges. This view is essential for understanding the model's generalization behavior and for guiding iterative improvements to data preprocessing, feature engineering, or model architecture.

Collectively, the dashboard operationalizes complex machine learning outputs into interpretable, actionable visual analytics, thereby supporting evidence-based public health decision-making.

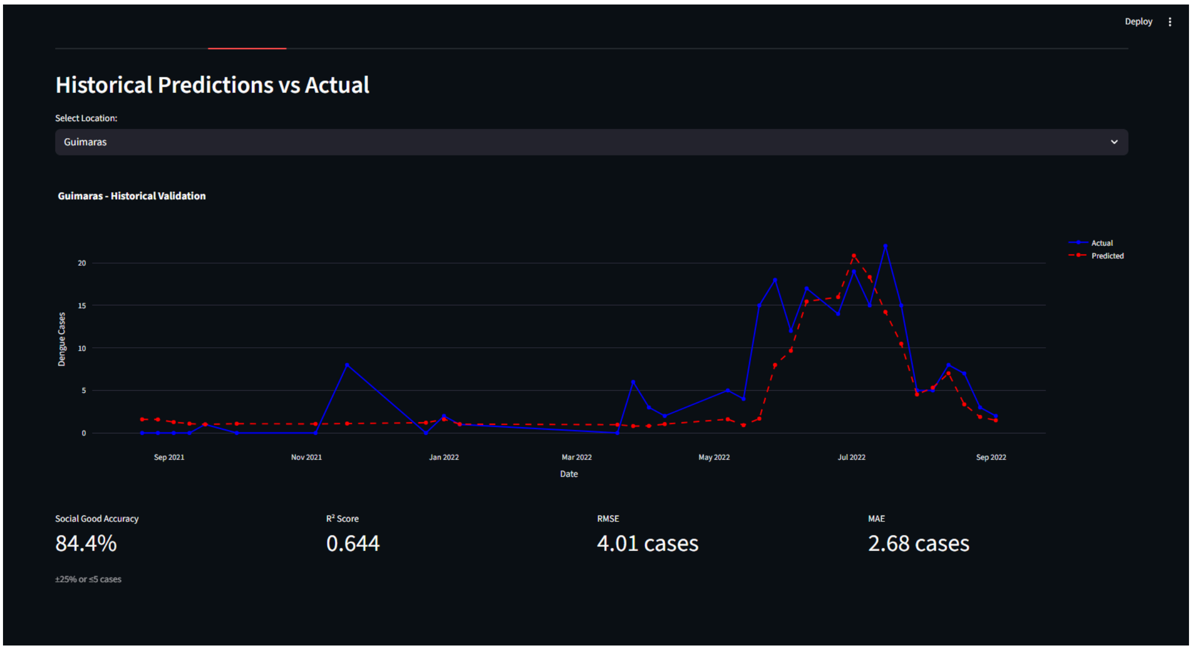


Figure 3. Historical Predictions vs Actual