Ensemble Deep Learning for Forecasting Lassa Fever Outbreaks in Nigeria: Integrating Incidence and Weather Data for Early Warning

Introduction

Lassa fever is a deadly viral hemorrhagic fever endemic to some parts of West Africa, including Sierra Leone, Liberia, Guinea, and Nigeria that affects between 300,000 and 500,000 annually [1]. In the absence of an effective vaccine against the deadly Lassa virus, predicting its outbreak becomes very crucial for early public health care intervention. Previous studies applied traditional models and machine learning for forecasting Lassa fever, but deep learning, ensemble techniques and external variables (weather) remain underexplored. To address these identified gaps, this study develops an ensemble deep learning model that combines Lassa fever incidence data and weather data for forecasting Lassa fever outbreaks in Ondo and Edo State, Nigeria. The three deep learning techniques explored include Long-Short Term Memory (LSTM), Bidirectional LSTM, and Gated Recurrent Unit (GRU).

Methodology

Weekly confirmed Lassa fever case data were obtained from the Nigeria Centre for Disease Control (NCDC), while weather variables (temperature, precipitation, humidity) were collected from the ECMWF ERA5 reanalysis dataset. Data were normalized to the [0,1] range using min–max scaling. Three deep learning models including (LSTM), Bidirectional LSTM (BiLSTM), and Gated Recurrent Unit (GRU) were trained, tuned, and evaluated. Their outputs were combined through a weighted ensemble to generate 8-week ahead forecasts. The overall forecasting framework and ensemble architecture are illustrated in Figure 1a–b.

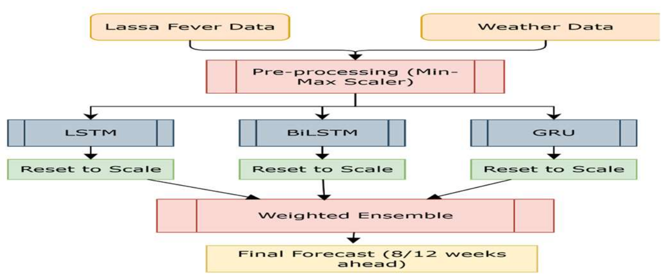

Figure 1a: Lassa Fever Forecasting Framework

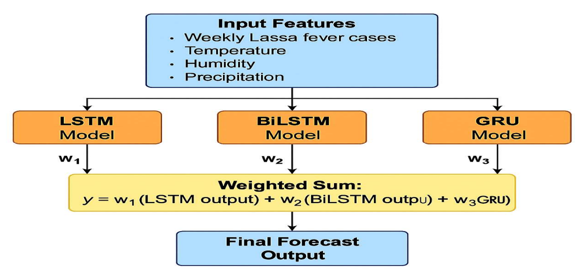

Figure 1b: Weighted Ensemble Architecture

Results and Discussion

The ensemble deep learning model outperformed individual models and baseline approaches across both Ondo and Edo State forecasts. For Ondo State, the ensemble achieved **RMSE = 7.49, MAE = 6.57, R^2 = 0.52**. For Edo State, the ensemble also achieved the lowest error (**RMSE = 6.40, MAE = 5.56**) and higher stability (R^2 = 0.10) as summarized in Table 1. These results demonstrate the effectiveness of ensemble deep learning over traditional and standalone models for improving outbreak prediction and strengthening early warning systems.

Table 1: Comparative Results of Ensemble vs. Baseline models

State	Model	RMSE	MAE	R^2
Ondo	Ensemble	**7.49**	**6.57**	**0.52**
	Best baseline (LSTM)	7.69	6.76	0.49
Edo	Ensemble	**6.40**	**5.56**	**0.10**
	Best baseline (BiLSTM)	6.44	5.42	0.08

References

[1] Al-Mustapha, A. I., Adesiyan, I. M., Orum, T. G., Ogundijo, O. A., Lawal, A. N., Nzedibe, O. E & Abubakar, A. T. (2024). Lassa fever in Nigeria: epidemiology and risk perception. *Scientific Reports*, *14*(1), 27669.