

1. Data Preparation



Epidemiological Data

Dengue Cases



Socioeconomic Variables

GDP per capita, social progress, access to information, vaccination coverage, adequate housing, waste collection, sanitation, etc.



Environmental Variables

Monthly temperature, precipitation, elevation, Köppen, etc.



Cleaning & Normalization

Unified by IBGE code, imputed missing values, MinMax Scaling



2. Data Structuring for Modeling

ID	t	$v_1 \dots v_n$
i_1	t_1	0.5 ...
...

Longitudinal Panel Data



$v_1 \dots v_n$
$x(t_{-1}), (1 \dots n)$
...
$x_t, (1 \dots n)$

+

y_i (Target)

Time Series Sequences per Municipality



3. LSTM

Initial LSTM Network

Feature Selection

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

Final LSTM Network

Cross-Validation (k=5)



4. Clustering



Spatial Autocorrelation

Moran's I for risk clusters



Hotspot Analysis (LISA Maps)

Identification of critical areas



K-Means Clustering

Grouping municipalities by risk profile