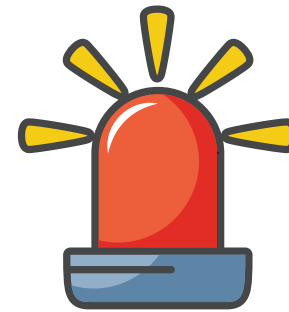
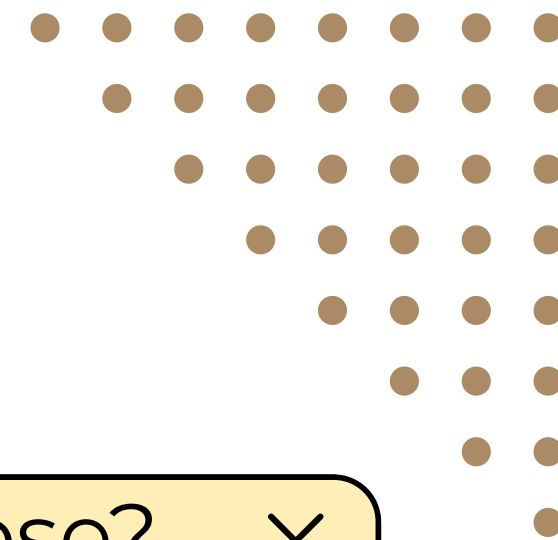


# Early Detection of Respiratory Diseases with Random Forest



Mónica Lozada - Carolina Saavedra





○

What is?

×


This project implements a Random Forest Regressor for the early detection of acute respiratory diseases (IRA) in Peru. The model predicts the weekly incidence rate by region, using climate and historical case data, while excluding the pandemic years (2020–2021).

○

What is the purpose?

×

We process raw official health data and, using a Random Forest model, built an early detection system for acute respiratory diseases across the departments of Peru.



○

Data (2017-2025)

×

MINSA

INEI

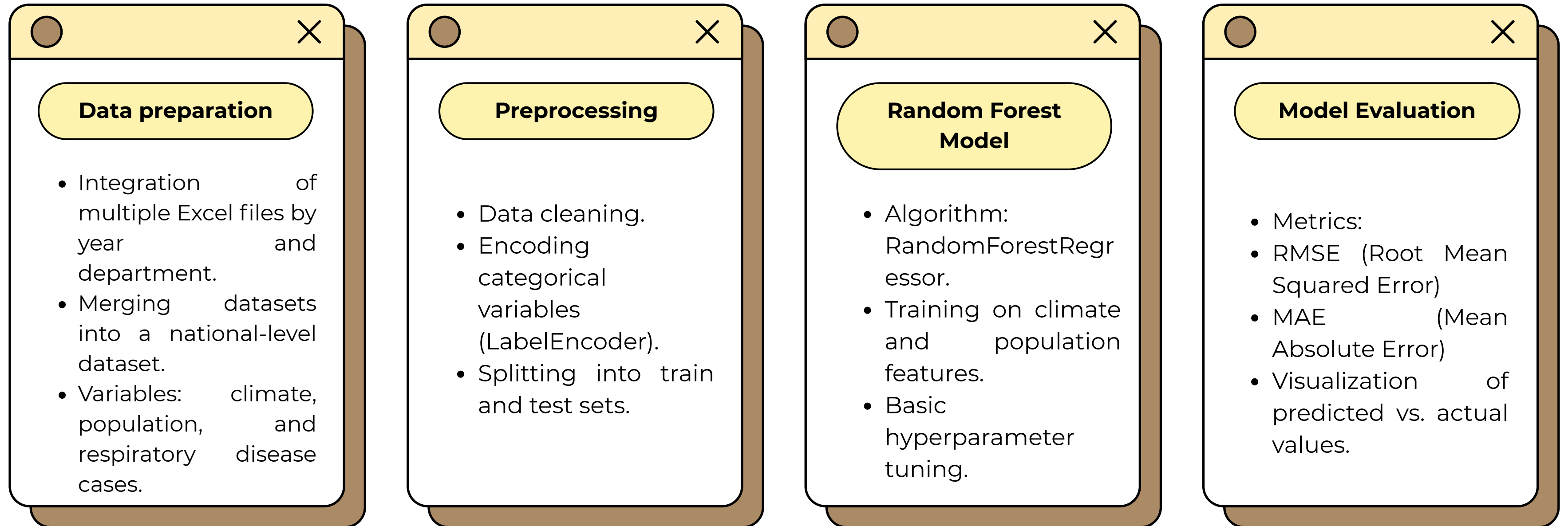
Clima and Health

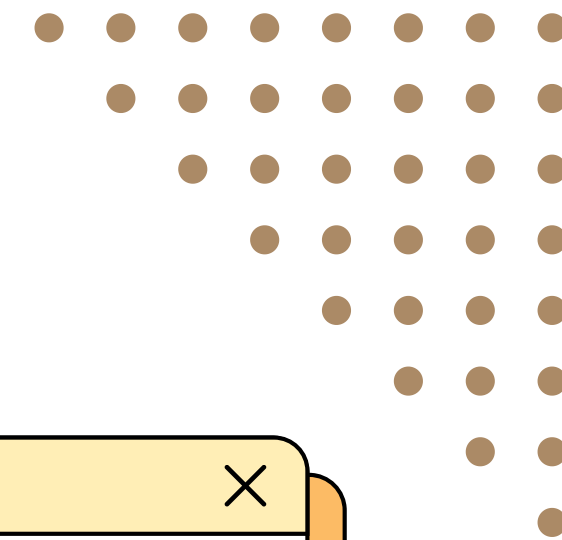
Population





# What was the process?





# Key results



The model shows good accuracy predicting disease incidence by department.


The approach can be extended with new datasets (e.g., air pollution, mobility).

Climate and population features are highly relevant.






# Next steps




Define and validate alerts

- Establish thresholds (absolute, relative, or probabilistic).
- Test on historical data to balance sensitivity vs. false alarms.



Build prediction & Alert Pipeline

- Automate data ingestion and preprocessing.
- Deploy Random Forest as an API/service.
- Apply alert rules and send notifications (dashboard, email, SMS).



Monitor and improve

- Track model performance and data drift.
- Retrain regularly and refine thresholds.
- Collect feedback from health agencies for system adjustments.



