# Title: Early Disease Outbreak Prediction using Health and Environmental Data

# SDG Goal: SDG 3 (Health)

## SDG Problem

Disease outbreaks (like dengue) often catch public health systems unprepared, especially in climate-sensitive regions. By leveraging historical health and environmental data, this project aims to predict outbreaks early, allowing for timely medical and logistical interventions. This directly supports SDG 3 by improving disease preparedness and reducing mortality.

## Machine Learning Method

- Type: Supervised Learning  
- Model: Random Forest Classifier  
- Input Features: Weekly data on temperature, precipitation, humidity, vegetation indices (NDVI), and time variables (week, year)  
- Target: Binary classification of outbreak vs. no outbreak based on total case thresholds

## Outcome

The trained model achieved an accuracy of ~85% on test data, demonstrating strong predictive power. This outcome suggests high potential for real-world deployment in early warning systems, especially in dengue-prone or resource-limited regions.

## Ethical Considerations

- Data Representation: Ensure the dataset is balanced across regions, seasons, and demographics to avoid biased predictions  
- Transparency: Communicate model limitations and uncertainty to decision-makers  
- Privacy: Use only aggregated, anonymized health data to protect individuals

## Conclusion

Machine learning can power public health tools that detect early signals of disease outbreaks. With careful ethical implementation, this solution could aid governments and health agencies in saving lives through a proactive response.

Prepared by: Faith

Date: June 2025