Research Proposal Outline

Title: Interpretable and Reliable Machine Learning for Epidemic Forecasting

1. Research Problem & Significance
Machine learning has enhanced epidemic forecasting but lacks interpretability, consistent validation, and ethical safeguards, limiting its trustworthiness for public health use. There is a need to balance predictive accuracy with transparency and reliability.

2. Research Question
How can ML models for epidemic forecasting be made more interpretable and reliable while maintaining predictive accuracy?

3. Aim & Objectives
Aim: Develop a hybrid ML-epidemiological forecasting framework with a focus on interpretability and robustness.

Objectives:

-Identify gaps in current ML-based epidemic models

-Build a hybrid ML + SEIR model

-Apply explainability tools (SHAP, LIME)

-Evaluate using benchmark datasets and validation protocols

4. Key Literature
Covers models such as Random Forests, XGBoost, LSTM, and hybrid SEIR-ML systems. Literature highlights key challenges: black-box models, overfitting, limited validation, and ethical concerns (e.g., bias and fairness).

5. Methodology

-Use public datasets

-Data preprocessing, model training

-Compare baseline ML vs hybrid models

-Use SHAP/LIME for interpretability

-Evaluate performance and forecast uncertainty

6. Ethics & Risk

- Address bias and fairness using mitigation strategies

-Ensure data anonymisation and compliance with ethical standards

-Promote reproducibility through transparent documentation and open tools

7. Artefact

Prototype forecasting tool with interpretable outputs, accompanied by documentation and ethical risk checklist.

8. Timeline

9. Summary