Statistical Time Series Modelling for Predictive Epidemic Surveillance (STMES)

GROUP NO- 8

Submitted by-

1. Mrudula Deshmukh (50605669)
2. Aishvarya Samuel Salvi (50604254)
3. Siddhi Sunil Nalawade (50604254)
4. Ragavi Muthukrishnan (50604772)

Executive Summary:

This project developed a statistical framework for early detection and prediction of infectious disease outbreaks using advanced time series methodologies. By implementing and comparing multiple statistical approaches—particularly Change-Point ARIMA modeling—the framework provides reliable early warning signals for public health authorities. The model demonstrated effectiveness in predicting outbreaks such as COVID-19 and offers potential applications for other infectious diseases.

Problem Statement:

Traditional surveillance systems detect outbreaks only after they begin, limiting opportunities for timely interventions. While machine learning models exist, they often lack interpretability and require significant computational resources. Statistical time series methods offer a promising alternative but face challenges such as noisy data, seasonal patterns, and structural changes in epidemic dynamics.

Objectives:

1. Develop a robust time series modeling framework for outbreak prediction.
2. Incorporate change-point detection to identify structural shifts in disease incidence.
3. Evaluate model performance using historical outbreak data.
4. Provide interpretable outputs to support public health decision-making.

Methodology:

Data Sources

- COVID-19 case data from Johns Hopkins University.
- Influenza-like illness data from CDC FluView.
- Historical outbreak reports from WHO.

Statistical Framework:

1. Change-Point Detection: Using the PELT algorithm to identify significant shifts in disease trends.
2. Change-Point ARIMA Modeling: Segmenting time series data based on detected change points and fitting ARIMA models to each segment.
3. Alternative Models: Implementing Prophet and traditional ARIMA models for comparison.

Evaluation Metrics:

- Forecast accuracy: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE).
- Outbreak detection: Sensitivity, specificity, lead time.
- Interpretability: Visualizations and actionable insights.

Key Findings:

- Improved Accuracy: Change-Point ARIMA outperformed traditional ARIMA models during transition phases between epidemic waves.
- Early Warning: The framework detected outbreaks 5–7 days earlier than traditional threshold-based systems.
- Adaptability: The model successfully adapted to different diseases and geographical contexts.

Conclusion:

The STMES framework provides a powerful tool for predictive epidemic surveillance by combining statistical rigor with practical interpretability. Its ability to detect early warning signals can significantly enhance public health preparedness and response strategies.