**Air Quality Prediction Using Time Series Forecasting Techniques**

**INTRODUCTION**

**Brief Background:**

Air pollution is a major environmental issue impacting public health, climate change, and quality of life. The Air Quality Index (AQI) is applied to measure and communicate pollution levels to the public. Predicting air quality through time series analysis is important for early warnings, pollution control strategies, and policy-making.

**Motivation:**

* Growing pollution will be a major health risk, causing respiratory diseases and cardiovascular problems.
* It will aid in policy formulation and prevent various measures.
* Integrating time series models with machine learning techniques will improve the accuracy.
* This project will provide a comparative analysis of various forecasting models that will prove the best.

**OBJECTIVES**

* To analyze the historical AQI and identify trends, seasonality, and anomalies.
* Development of predictive models based on statistical and machine learning techniques.
* Comparing the accuracy of different forecasting techniques.
* Actionable insights for government agencies and urban planners to mitigate pollution.

**DATASET DESCRIPTION**

**Source of Data:** <https://www.kaggle.com/code/rohanrao/calculating-aqi-air-quality-index-tutorial/input>

**Nature of Data:**

* The dataset contains **daily air quality measurements** for multiple cities in India.
* It includes concentrations of various air pollutants.
* The dataset is structured as **time series data**, where each row represents a specific city’s air quality readings on a given date.

**Features Included:**

* Pollutants: PM2.5, PM10, NO₂, SO₂, CO, O₃, NH₃,
* Meteorological Variables: Temperature, Humidity, Wind Speed, Pressure
* AQI Calculation: AQI & AQI\_Bucket

**Granularity:**  Daily air quality readings recorded for multiple cities.

**Time Period Covered & Data Points:**

* **Time Period:** The dataset spans from **January 1, 2015, to 2020.**
* **Number of Data Points:** The dataset contains **29,531** records across various cities.

**EXPECTED OUTCOMES**

* Identify seasonal patterns, trends, and anomalies in AQI data.
* Determine the most suitable model for accurate air quality forecasting.
* Provide a comparative study of traditional vs. machine learning-based forecasting approaches.
* Develop insights that can assist policymakers in reducing air pollution levels.

**TIMELINE:**

|  |  |
| --- | --- |
| Data Collection | Week 1 |
| Data Cleaning & Preprocessing | Week 2 |
| Exploratory Data Analysis (EDA) | Week 3 |
| Model Selection & Implementation | Weeks 4-5 |
| Model Evaluation & Optimization | Week 6 |
| Final Presentation & Report | Week 7 |

**REFERENCES**

* " Application of Deep Learning Models and Network Method for Comprehensive Air-Quality Index Prediction " (MDPI,2022)
* " An air quality index prediction model based on CNN‑ILSTM " (IEEE, 2022)
* " Prediction of Air Quality Index Using Machine Learning Techniques " (Hindawi,2022)

**CONCLUSION**

This project aims to leverage time series forecasting techniques for air quality prediction, helping in proactive pollution control measures. The study will evaluate different models to find the best-suited forecasting approach.