

Feature Engineering Approach

To forecast pollutant concentrations, we will first feature engineer variables based on the time-series data.

1. Time-Based Features:
 - a. Hour: Extraction of hours of the day from the timestamp to account for daily cycles.
 - b. Day of the week: Extract the day of the week to account for weekly cycles.
 - c. Month: Extract the month to account for seasonality.
2. Lagged features:
 - a. Lag 1: Pollutant concentration from the previous time period (1 Hour ago)
 - b. Lag 2: Pollutant concentration from two time period ago (2 Hours ago)
3. Rolling Statistics:
 - a. Rolling Mean(Window = 3): Average concentration over the last 3 hours.
 - b. Rolling Standard Deviation(Window=3): standard Deviation over the last 3 hours.

These features will help the model understand the temporal dependencies and trends in the pollutant concentrations.