Datathon 2025 - Deliverables Overview

Overview

Participants will submit a comprehensive analysis of their findings, structured into key deliverables. These include predictive model outputs, feature importance analysis, and visualizations that support their conclusions. The deliverables should clearly communicate how weather conditions impact emissions, backed by data-driven insights.

Predictive Model Output

Participants will use a predictive model to estimate emissions levels based on weather conditions. The model's performance should be evaluated using appropriate accuracy metrics. The goal is to demonstrate the ability to make reasonable forecasts and assess model reliability.

Key Points:

1. Training and validating the model using historical data.
2. Making predictions for future emissions levels based on available weather conditions.
3. Evaluating the accuracy of predictions using appropriate performance metrics.
4. Comparing predicted emissions with real values to assess effectiveness.
5. Discussing challenges and potential improvements to the model.

Feature Importance Analysis

This section highlights which weather and operational parameters most influence emissions. Understanding these relationships helps determine which factors contribute most to pollution levels and can inform decision-making in emissions management.

Key Points:

1. Identifying which variables have the strongest impact on emissions levels.
2. Using statistical methods or model-based techniques to determine feature importance.
3. Explaining why certain variables matter more than others in influencing emissions.
4. Discussing real-world implications of identified key features.
5. Suggesting additional factors that could improve analysis in future research.

Visualization

Participants should create visualizations to effectively communicate their findings. Graphs, charts, and other data visualizations should highlight key trends, relationships, and insights from the analysis.

Key Points:

1. Creating scatter plots to show relationships between weather and emissions variables.
2. Using correlation heatmaps to identify key trends and dependencies.
3. Presenting time-series graphs to illustrate emissions changes over time.
4. Highlighting feature importance using bar charts or ranking tables.
5. Ensuring clarity in visualizations with proper labels, legends, and concise explanations.