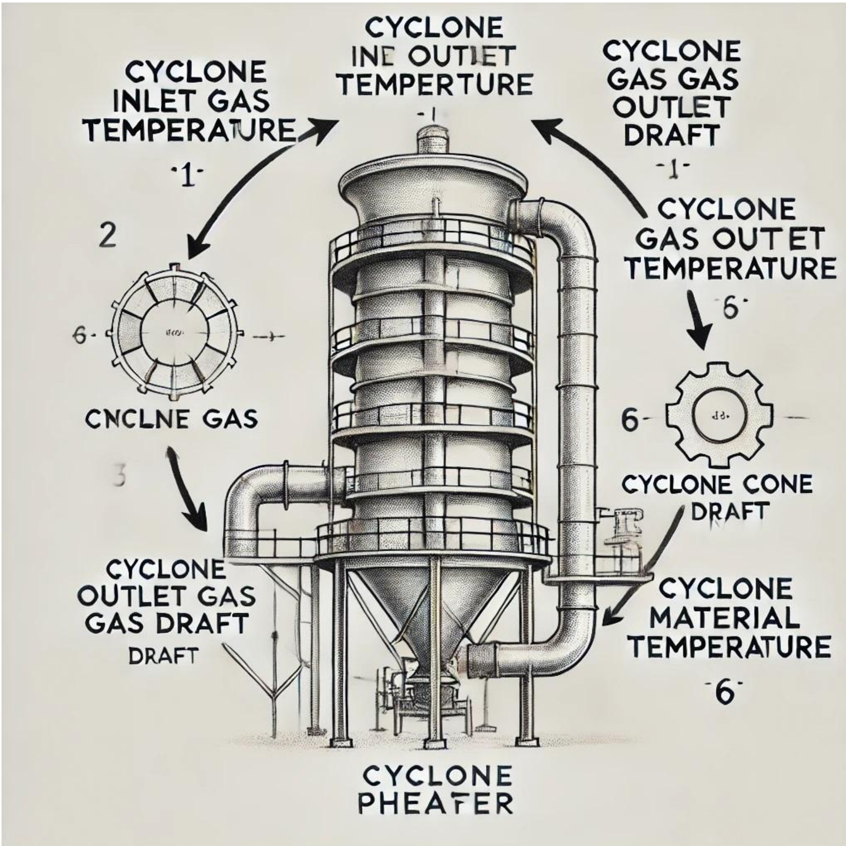
Cyclone Preheater Overview and Dataset Analysis

A cyclone preheater is an essential component in cement manufacturing plants. It is designed to preheat the raw materials before they enter the kiln for clinker production. The preheater uses hot gases from the kiln to improve energy efficiency and reduce fuel consumption. Cyclone preheaters consist of several stages of cyclones that allow the heat exchange between hot gases and raw materials. Monitoring the operational parameters of the cyclone preheater is crucial for ensuring efficiency and safety.



Key Variables Monitored in the Cyclone Preheater

1. **Cyclone Inlet Gas Temperature**: Measures the temperature of gases entering the cyclone. Abnormalities here could indicate issues with heat exchange.
2. **Cyclone Material Temperature**: Reflects the temperature of raw materials entering the cyclone. Variations might suggest inefficiencies in preheating.
3. **Cyclone Outlet Gas Draft**: Represents the pressure of gases exiting the cyclone. High drafts may indicate blockages or process inefficiencies.
4. **Cyclone Cone Draft**: Tracks the pressure within the cyclone cone. Anomalies here might signal mechanical issues or flow disruptions.

5. **Cyclone Gas Outlet Temperature**: Indicates the temperature of gases leaving the system. This helps in assessing heat loss or efficiency.
6. **Cyclone Inlet Draft**: Measures the pressure of gases entering the cyclone. Variations could indicate upstream issues or flow disruptions.

Dataset Overview

Total Records: 377,719 entries

Variables: 6 operational parameters along with a timestamp.

Duration: Data recorded every 5 minutes over a period of 3 years.

Data Type: Initial inspection shows all operational variables are recorded as objects, likely requiring conversion to appropriate numerical formats.

Structure of the data :

Column	Data Type	Non-Null Count	Description
`time`	datetime64[ns]	377,719	Timestamp of each recorded observation.
`Cyclone_Inlet_Gas_Temp`	object	377,719	Temperature of gas entering the cyclone.
`Cyclone_Material_Temp`	object	377,719	Temperature of material in the cyclone.
`Cyclone_Outlet_Gas_draft`	object	377,719	Gas pressure exiting the cyclone.
`Cyclone_cone_draft`	object	377,719	Draft pressure in the cyclone cone.
`Cyclone_Gas_Outlet_Temp`	object	377,719	Temperature of gas exiting the cyclone.
`Cyclone_Inlet_Draft`	object	377,719	Gas pressure entering the cyclone.

Data Processing and Cleaning

Data Cleaning:

- Transformed the datatype of Time column into Datetime.
- Duplicated values are not removed due to each time period is different.
- Non-numeric values were identified in operational variables and converted to numerical data.
- Missing values were imputed .

Exploratory Data Analysis (EDA):

Performed Uni-variant analysis Such as Box plot to check the outliers of data.

Performed Bi-variant analysis Such as scatter plot to check the relationship between the variables.

Performed Heat map to check the correlation among the variable .

Category	Variable Pair	Correlation Coefficient	Insight
Strong Correlation	Cyclone_Inlet_Gas_Temp ↔ Cyclone_Gas_Outlet_Temp	0.99	Inlet gas temperature almost perfectly predicts outlet gas temperature.
Strong Correlation	Cyclone_Inlet_Gas_Temp ↔ Cyclone_Material_Temp	0.96	Material temperature strongly depends on the inlet gas temperature.
Strong Correlation	Cyclone_cone_draft ↔ Cyclone_Outlet_Gas_draft	0.97	Pressure at the cone section and outlet are closely linked.
Weak Correlation	Cyclone_Material_Temp ↔ Cyclone_Outlet_Gas_draft	-0.88	Material temperature has a weaker inverse relationship with outlet draft pressure.
Weak Correlation	Cyclone_cone_draft ↔ Cyclone_Gas_Outlet_Temp	-0.89	Cone pressure and outlet gas temperature show a weaker inverse relationship.

Feature Engineering

Performed Rolling Average for all the variables such as

- Cyclone_Inlet_Gas_Temp_Rolling_Mean
- Cyclone_cone_draft_Temp_Rolling_Mean
- Cyclone_Gas_Outlet_Temp_Rolling_Mean
- Cyclone_Inlet_Draft_Rolling_Mean

- Cyclone_Material_Temp_Rolling_Mean
- Cyclone_Outlet_gas_draft_Rolling_Mean

Anomaly Detection Techniques Applied

1. Isolation Forest

Objective: Identify anomalies by isolating points that are significantly different from the rest of the dataset.

Results: Detected 123 anomaly groups over the 3-year period.

Example: Group 0 occurred between 2017-05-30 20:35:00 and 2017-05-30 23:00:00 with 29 anomalies.

2. LSTM Auto Encoder

Objective: Detect anomalies by learning the normal patterns in time-series data and identifying deviations.

Results:Detected 1,640 anomaly groups.

Example: Group 0 occurred between `2017-01-01 15:50:00` and `2017-01-01 16:20:00` with 5 anomalies.

Insights from the Results

Frequent Anomaly Periods:

Anomalies were often clustered during specific operational hours, possibly due to shifts in process parameters.

Example: Multiple groups between `2017-05-30` and `2017-05-31` suggest recurring operational issues.

Severity of Anomalies:

Some groups, such as Group 1639, had a high number of anomalies (27 anomalies over a 3-hour period).

Correlations and Root Causes:

Strong correlations between `Cyclone_Inlet_Gas_Temp` and `Cyclone_Gas_Outlet_Temp` suggest that issues with inlet temperature directly impact outlet conditions.

Negative correlations with draft variables indicate potential mechanical or flow disruptions.

Key Challenges

Data Quality:

Non-numeric values and missing data needed cleaning.

Object data types slowed analysis.

High Anomaly Volume:

LSTM detected many more anomalies, requiring focus on critical periods.

Root Cause Analysis:

Linking anomalies to specific issues requires expert input and more contextual data.

Conclusion

Isolation Forest: Best for detecting small clusters of anomalies, suited for simpler systems.

LSTM Autoencoders: Identified a wider range of anomalies but needs more computational power.