Welcome to demo of Asset Health Framework -

The current scenario in the market is complicated due to the significant prevalence of legacy processes and systems.

The main goal of this framework is to identify assets break downs or plantshut downs

The longer the asset downtime or higher the inefficiency in performance

The goal of this frame work is to detect the anomalies in the plant,so to do that we are used machine learning algorithms like Kmeans,GMM and isolation forest

What are the steps to detect anomalies

DATA INGESTION

DATA QUALITY

DATA PRE-PROCESSING

MODEL TRAINING

MODEL EVALUATION

INFERENCE

DATA STORAGE

The first step would be

DATA INGESTION

IN DATA INGESTION WE SELECT THE FILE SOURCE

A meta data table   rotatorytableinfo1 is created with attributes related to each sensor group

BY Ensuring the data format is acceptable so that we can insert into postgres tables

DATA QUALITY

THE NEXT ONE IS DATA QUALITY

HERE The data is read from rotatorytableinfo1 and converted into Spark DataFrames for analysis.

It applies transformations and calculations to the data, such as changing column types and performing specific calculations as mentioned below

The grading metrics such as Completeness , Negative Percentage ,DuplicateValuesPercentage, ZeroValuesPercentage , 1stPercentile , 99thPercentile, ZeroVariance , precision, Mean, Minimum , Maximum , Sum , Distinctness , DatatypeInconsistency , OutlierPercentage are calculated.

The statistics of the data Minimum, 1stPercentile , 3rdPercentile ,5thPercentile,10thPercentile ,25thPercentile, Median , Mean , 90thPercentile, 95thPercentile ,98thPercentile , 99thPercentile , Maximum are calculated

DATA PREPROCESSING

IN DATA PREPROCESSING WE CLEANS THE DATA

HERE IT WILL PERFORM THE FOLLOWING OPERATIONS -

It retrieves base table information and data from the database for further processing.

In preprocess, filtering the good data, removing duplicating values, imputing null values with mean.

Removing Outliers, standardising the data

Model scales are calculated and saved for future use.

THE NEXT ONE IS MODEL TRAINING

HERE WE HAVE THREE MODELS

- K MEANS
- ISOLATION FOREST
- GMM

K MEANS

K-Means is a simple clustering algorithm that partitions data into 'k' clusters, with each cluster represented by its centroid.

Anomaly detection using K-Means typically involves finding data points that are furthest from the centroids. Data points that are distant from all cluster centroids can be considered anomalies.

Gaussian Mixture Model(GMM)

GMM assumes that the data is generated from a mixture of several Gaussian distributions. It tries to fit multiple Gaussian distributions to the data.

While GMM can be used for clustering, it is also used for anomaly detection by considering data points that are unlikely to belong to any of the Gaussian distributions as anomalies.

Isolation Forest

Isolation Forest is an ensemble-based anomaly detection algorithm that works by recursively partitioning the data into subsets, isolating anomalies in the process.

It identifies anomalies by looking at the number of partitions needed to isolate a data point. Anomalies are expected to be isolated with fewer partitions.

Isolation Forest is known for its scalability and efficiency, making it suitable for high-dimensional datasets.

And then

MODEL Evaluation

in this step Rotatorytabletrain is created based on rotatorytable_nfo1- append the new sensorgroup records and update the existing records

It performs various transformations on the loaded data.This includes rounding a column, filtering rows, and extracting specific columns.

AND THE FINAL STEPS ARE

INFERENCE

In this final steps we predict the results from the train models and create final table from which we detect the anomalies

And also we will be creating alert tables and warning tables from the final tables

DATA STORAGE

The final table , alert table and warning table will be stored in the database

So we can extract the data from the data base and find out the alert is real or not

And this is all about AHF and in the way we detect plant breakdowns

THANK YOU

And this is all about AHF and in this way we detect plant breakdowns

THANK YOU