**Celebal AnaVerse\_N**

**The problem is to predict anomaly based on the readings captured by sensors.**

## **Overview**

The problem is to predict anomaly based on the readings captured by sensors.

• Objective: Evaluate candidates on their ability to work with tabular data, perform advanced analysis, and build robust predictive models.  
• Format: Individual participation  
• Deliverables: Jupyter Notebook with code

Evaluation Criteria  
• Data Exploration and Preprocessing (20%):  
o Handling missing values, outliers  
o Feature engineering and correlation analysis

• Modeling (60%):  
o Application of classical models (Logistic Regression, Support Vector Machine, KNN, Decision Trees) and advanced models (e.g. Random Forest, XGBoost, LightGBM, CatBoost, Neural Network etc.)  
o Justification for model selection and tuning strategies

• Model Evaluation (20%):  
o Use of appropriate metrics (e.g. Accuracy , Precision, Recall, F1 Score)  
o Robustness checks, e.g., backtesting and residual analysis

### Description

Key Expectations from Participants  
• Creativity in feature engineering  
• Use of both classical and deep learning models to show expertise  
• Hyperparameter tuning, cross-validation, and interpretation of results  
• Effective communication of insights through visualizations  
Tools and Resources Allowed  
• Open-book format with access to online resources  
• Tools: Python (NumPy, Pandas, Scikit-Learn, TensorFlow/PyTorch  
Submission Guidelines  
• Notebook: Clean and well-commented Jupyter Notebook

### Evaluation

Evaluation would be done based on F1 score and accuracy corresponding to each class.

For reference please check "sample\_submission" file in data section

## Dataset Description

Data is captured from various sensors from a energy manufacturing plant on specific time intervals.  
The problem is to predict anomaly based on the readings captured by sensors.  
Columns : {'Date': datetime64[ns],  
'X1' : reading from sensor one,  
'X2' : reading from sensor two,  
'X3' : reading from sensor three,  
'X4' : reading from sensor four,  
'X5' : reading from sensor five,  
'target' : Anomaly(0 or 1) }