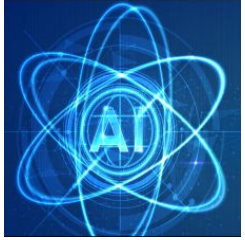
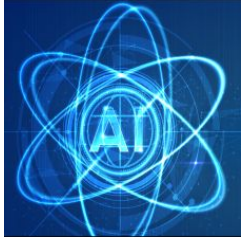

AI for Good
AI for Fusion Energy Challenge
In partnership with:
ITU
IAEA
Ning Jia
40 UN PARTNERS
ITU



Methodology: Strategy

Seeking Easy Solution:

- Consider the sample sizes
- Generalize well for unseen data
- Easy to implement
- Quick training and testing
- Less input signals, less features, less hyper parameter tuning to avoid over-fitting



Methodology: Model

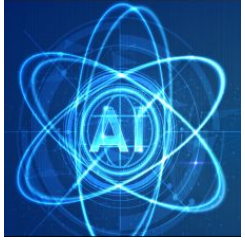
- Algorithm:

  Tree-based **lightgbm** classifier
- Signal Extraction and Preprocessing

  Read with **jddb** package and normalize raw signals
- Observation Window:

  300 (60 ms for 5000 HZ)
- Sample Weights:

  15 for C-Mod samples and **1** for J-Text samples



Methodology: 25 Input Signals

MIR signals:

'poloidal Mirnov probes_01',

'poloidal Mirnov probes_02',

'poloidal Mirnov probes_03',

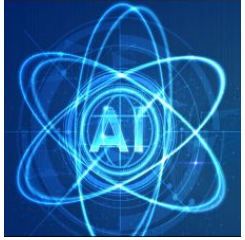
'poloidal Mirnov probes_05',

'poloidal Mirnov probes_06',

SXR signals:

soft-X-ray from 1 to 20

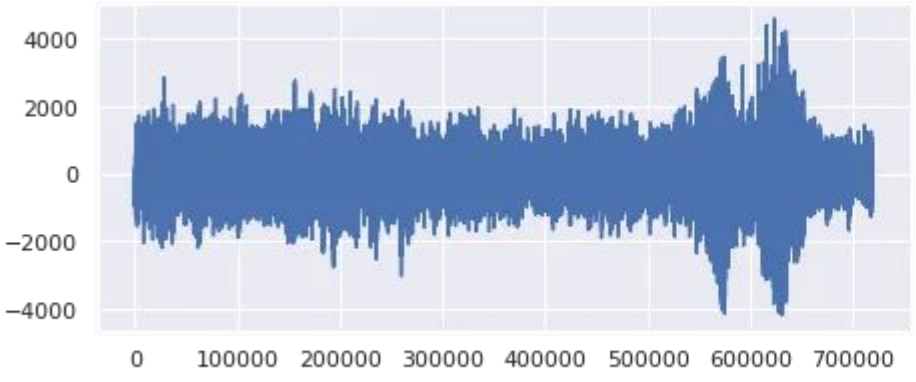
Other non-sensor signals may leak target

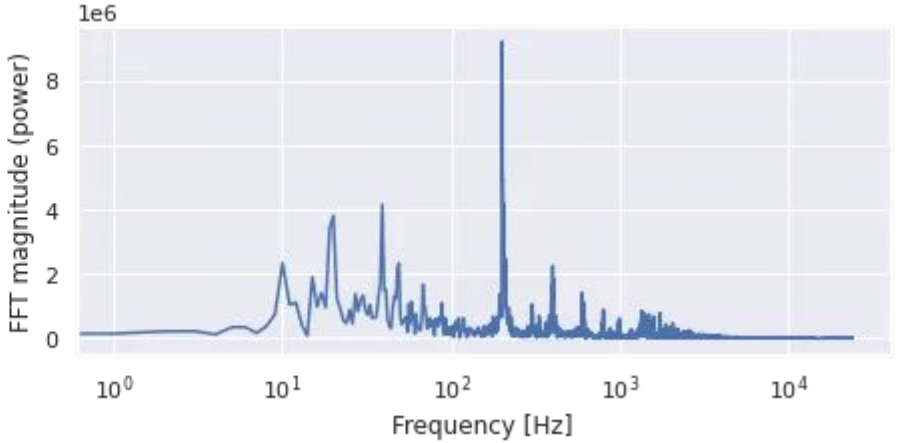


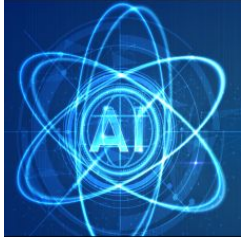
Methodology: Feature Engineering

- Frequency domain: spectral entropy
- Time domain: variance

```
row[f'{col}_std'] = np.nanstd(x)
row[f'{col}_entropy'] = spectral_entropy(x, sample_rate)
row[f'{col}_roll_std_mean']= np.mean(pd.Series(x).rolling(roll_win).std())
row[f'{col}_roll_std_std']= np.std(pd.Series(x).rolling(roll_win).std())
```



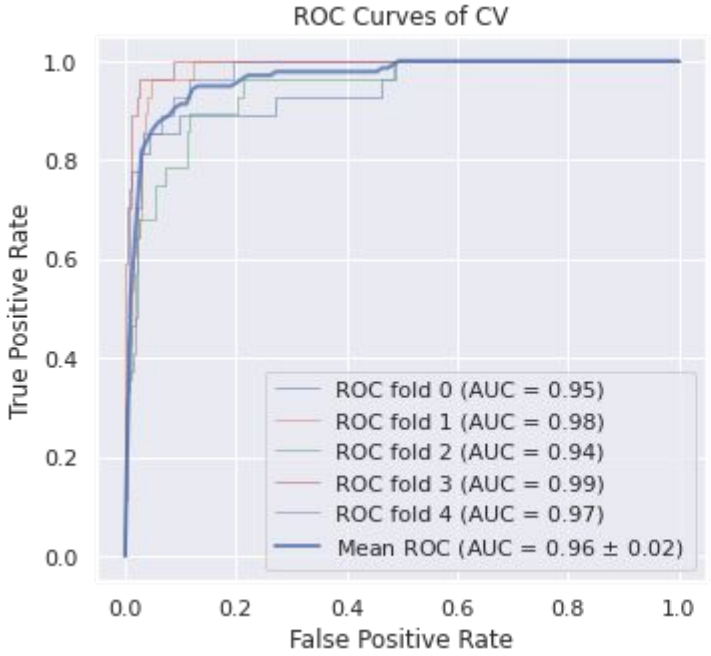


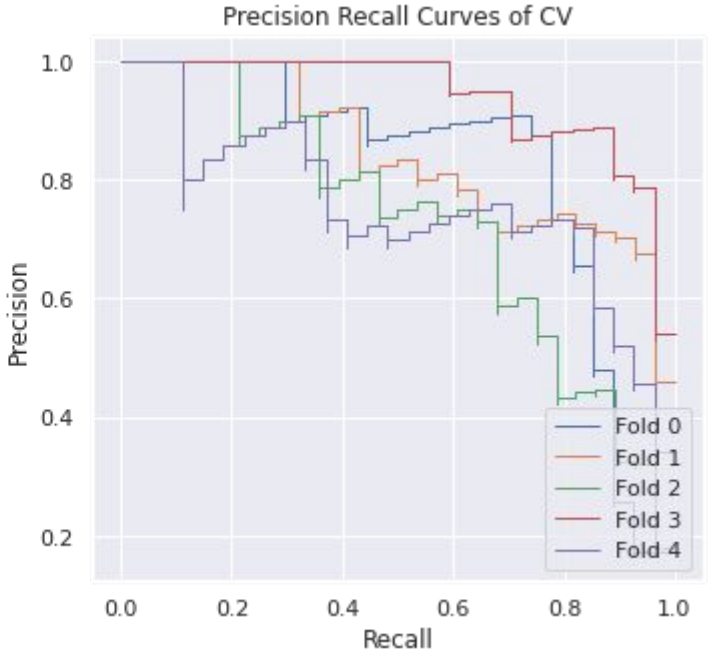


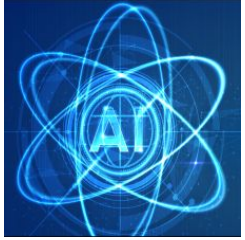
Solution Results (1)

- Mean AUC (area of under ROC curve) for 5 folds: **0.97**
- True positive and false positive rates at 60 ms before disruption: **0.91** and **0.1**.
- Public F1 score: **0.857**, Private F1 score: **0.898**
- May consider using **average precision of detected events**.

Solution Results (2)







Q & A

Thank you!