# Anomaly Detection in Ion Beam Etching Processes

**Objective:**
To develop an effective and scalable anomaly detection framework leveraging LSTM models, aimed at identifying rare faults in time-series data generated by industrial ion beam etching systems.

**Key Methods:**

- Conducted exploratory data analysis (EDA) to understand patterns, correlations, and feature importance in the dataset.
- Applied Clustering (MiniBatchKMeans) to group similar behaviors in the dataset, aiding initial insights.
- Designed and trained an LSTM-based neural network to analyze temporal sequences for detecting anomalies.
- Tuned probabilistic thresholds to balance sensitivity and specificity in anomaly identification.

**Evaluation:**

- The **Training** dataset was evaluated by metrics such as Accuracy, Precision, Recall, F1-Score and Confusion Matrix.
- The **Test** dataset was evaluated by calculating the MSE of both dataset and prioritizing low MSE value and small gap.

**Results:**

- Achieved a low Mean Squared Error (MSE) of approximately $2.37\times10^{-6}$ on the training dataset and approximately $7.72\times10^{-3}$ on the test dataset, small gap and low test data MSE value, showcasing high model generalization.
- Successfully scaled from a sample dataset of 600k rows to the full Training dataset of 3.7M rows, maintaining high accuracy, precision and F1-Score.

This project highlights the efficiency of the LSTM model for anomaly detection in a time-series dataset.