**Microsoft: Classifying Cybersecurity Incidents with Machine Learning**

1. **Data Acquisition**

- Collect and understand the structure of the provided datasets: `train.csv` and `test.csv`.

2. **Data Preprocessing**

- Handle missing values and clean the data to remove noise.

- Normalize or standardize features for consistent scaling.

- Encode categorical variables, if any, to make the data machine-readable.

3. **Feature Engineering**

- Analyse and select the most relevant features for classification.

- Generate new features or transform existing ones to enhance model performance.

4. **Model Development**

- Choose appropriate machine learning algorithms for classification.

- Train the model using the `train.csv` [dataset](https://drive.google.com/drive/folders/18vt2lkf69MggXitrTSn9qnZ8s-ToeKcH).

- Perform hyperparameter tuning to optimize the model.

5. **Evaluation on Test Data**

- Use the `test.csv` dataset to evaluate the model's performance.

- Compute key metrics: **Macro-F1 Score, Precision,** and **Recall**.

6. **Model Generalization**

- Assess how well the model generalizes to unseen data to ensure reliability for real-world applications.

7. **Insights and Documentation**

- Record insights and observations from the evaluation results.

- Document the process, methodology, and findings to make the project easily reproducible.