Project Documentation Report

Project Title:

Microsoft: Classifying Cybersecurity Incidents with Machine Learning

---

Gunvant M Ingale

Project Documentation Report

Project Title:

Microsoft: Classifying Cybersecurity Incidents with Machine Learning

---

1. Introduction

Cybersecurity has become an essential aspect of organizational security, with cyber threats and incidents rapidly evolving. This project aims to classify cybersecurity incidents using machine learning (ML) techniques to improve the accuracy and efficiency of incident detection. The key objectives of the project are to preprocess cybersecurity data, apply machine learning classification models, and evaluate their effectiveness using metrics like Macro-F1 Score, Precision, and Recall. The project also covers concepts from cybersecurity frameworks like MITRE ATT&CK and addresses the challenge of handling imbalanced datasets, a common issue in cybersecurity.

---

2. Objectives

- To classify different types of cybersecurity incidents using machine learning algorithms.
- To learn and apply data preprocessing and feature engineering techniques specific to cybersecurity.
- To handle imbalanced datasets effectively to improve model performance.
- To apply appropriate evaluation metrics like Macro-F1 Score, Precision, and Recall for model performance.
- To gain an understanding of the MITRE ATT&CK framework and its role in cybersecurity classification.
- To optimize machine learning models for better performance using techniques like model benchmarking and hyperparameter tuning.

Project Workflow

Step 1: Understanding the Dataset

The project dataset contains cybersecurity incident data with labeled classes representing different types of incidents (e.g., malware, phishing, insider threat). The dataset could include the following fields:

- Incident type
- Source IP
- Destination IP
- Timestamps
- Log data (user actions, system flags)
- Attack vector (e.g., phishing, brute-force)
- Event outcomes (e.g., compromised system, failed attack)

Step 2: Data Preprocessing and Feature Engineering

- **Data Cleaning**: Remove or impute missing values and handle outliers.
- **Feature Engineering**: Transform raw data into relevant features. For instance, convert categorical data into numeric format using encoding techniques (One-hot Encoding, Label Encoding).
- **Handling Temporal Data**: Convert timestamps into usable features, such as calculating time intervals between events to detect patterns.
- **Feature Scaling**: Normalize or standardize features to prepare them for machine learning algorithms.

Step 3: Exploratory Data Analysis (EDA)

- Identify data distribution and correlations between features.
- Understand class distribution, especially the imbalance between different incident types.
- Visualize attack patterns using techniques like heatmaps and histograms.

Step 4: Handling Imbalanced Data

Cybersecurity data is often imbalanced, where certain incident types (e.g., malware) may dominate others (e.g., insider threat). To counter this:

- **Resampling Techniques**: Use techniques like **undersampling** the majority class or **oversampling** the minority class.
- **SMOTE**: Generate synthetic samples for minority classes to balance the dataset.
- **Class Weighting**: Modify the model's loss function to account for imbalanced class distribution.

Step 5: Model Selection

- Apply and compare various **machine learning classification algorithms**:
    - **Logistic Regression**
    - **Decision Trees**
    - **Random Forest**

Step 6: Model Evaluation

- Use performance metrics that are particularly suited for imbalanced datasets:
    - **Precision**: The proportion of true positive predictions among all positive predictions.
    - **Recall**: The proportion of true positives among all actual positive instances.
    - **Macro-F1 Score**: A balanced evaluation metric that averages F1 scores across all classes, giving equal weight to minority classes.

Step 7: Cybersecurity Concepts (MITRE ATT&CK Framework)

- Learn to map machine learning predictions to the **MITRE ATT&CK** framework, which is a comprehensive matrix of tactics and techniques used by adversaries during cyberattacks. This provides a practical application of ML in the context of threat detection.

Step 8: Model Benchmarking and Optimization

- Use techniques such as **Grid Search** or **Randomized Search** for **hyperparameter tuning**.
- Benchmark models based on their Macro-F1 Score, Precision, Recall, and other relevant metrics.
- Compare the performance of different models and select the best-performing one.