Microsoft : Classifying Cybersecurity Incidents with Machine Learning

# Domain : Cybersecurity and Machine Learning

# Submitted by ,

R.Gayathri

Project Overview :

The task is to enhance the efficiency of Security Operation Centers (SOCs) by developing a machine learning model that can accurately predict the triage grade of cybersecurity incidents. Utilizing the comprehensive GUIDE dataset. The goal is to create a classification model that categorizes incidents as true positive(TP), benign positive (BP), or false positive (FP) based on historical evidence and customer responses. The model should be robust enough to support guided response systems in providing SOC analysts with precise context-rich recommendations, ultimately improving the overall security posture of enterprise environments.

Objectives :

1. A machine learning model capable of accurately predicting the triage grade of cybersecurity incidents(TP,BP,FP) with high macro F1 score, precision and recall.
2. Implement a machine learning model for the classification and detection of attack patterns using MITRE ATT&CK data.
3. A comprehensive analysis of model performance, including insights into which features are most influential in the prediction process.

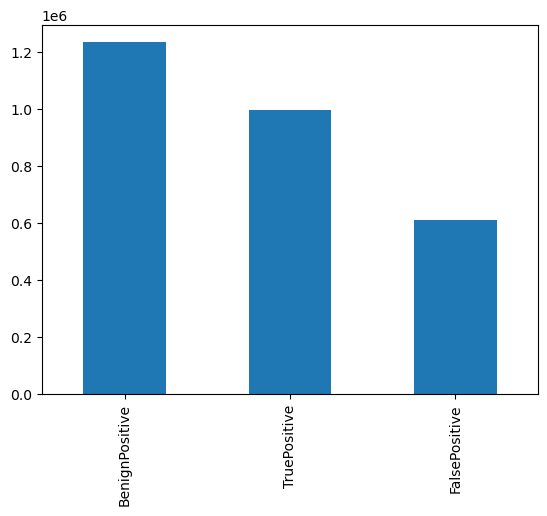
Data Preprocessed :

The dataset used in this project includes high dimensional features from various MITRE ATTA&CK techniques and tactics, resulting in large volumes of data.

1. Handling missing data
2. Feature engineering
3. Encoding Categorical variables

Exploratory Data Analysis :

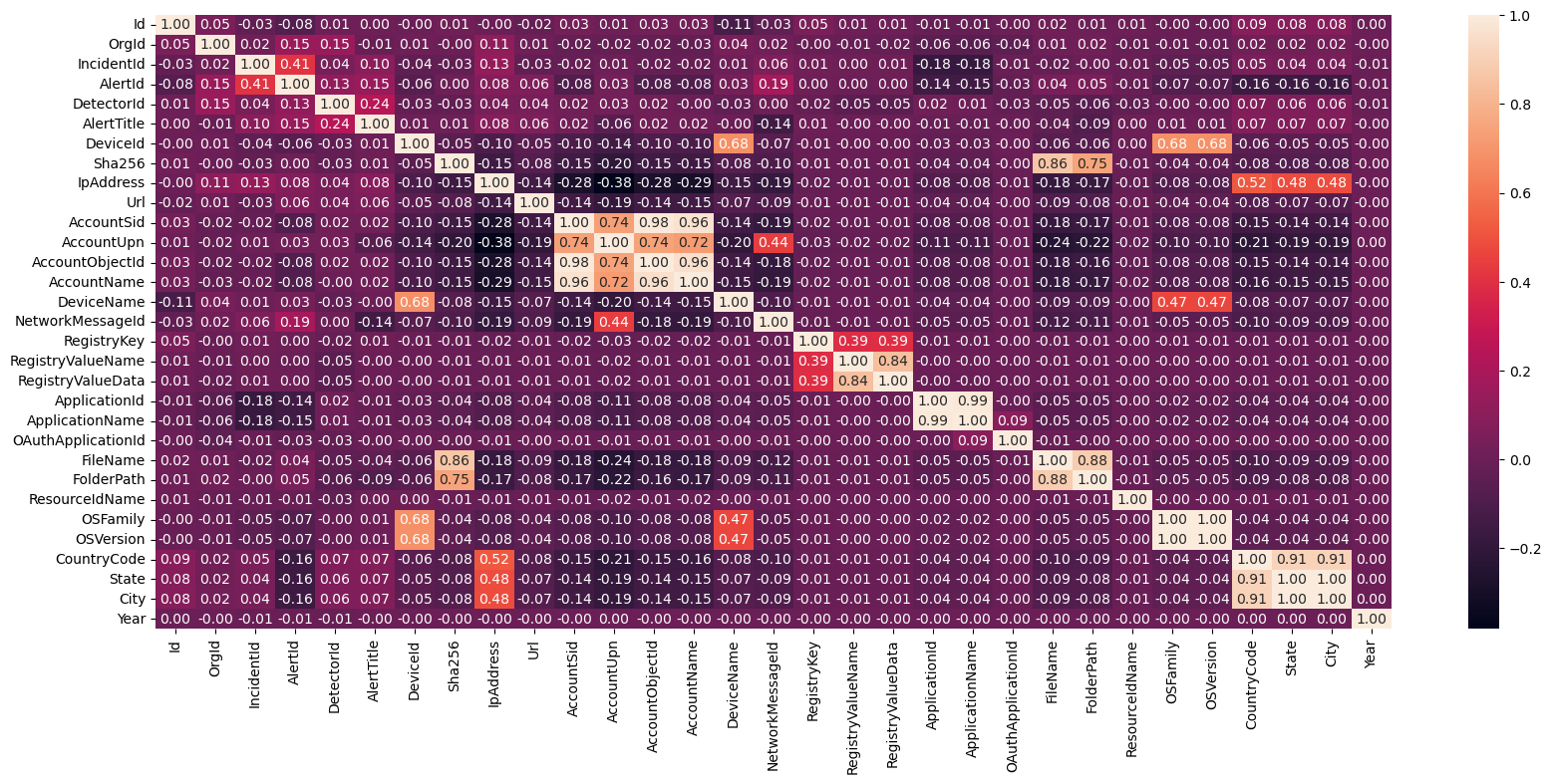
1. Count of categories in target



From the dataset , the category Benignpositive has the high rate , next Is Truepositive and last Falsepositive.

The category benignpositive has high frequency of above 1.2.

1. Correlation Matrix :



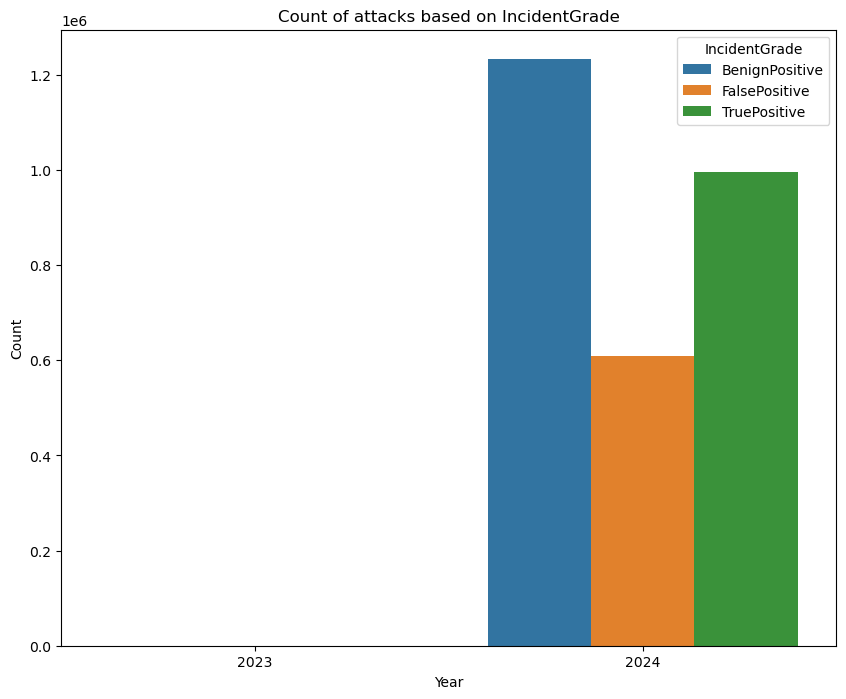
From the correlation matrix , the features related to accounts have high correlation between eachother.

The feature RegistryvalueName is highly correlated with Registry value data.

There is a high correlation between state , country and city . But these features are not necessary for this model .

The feature Filename is correlated with Sha256 with a correlation value of 0.80.

1. Count of attacks in years :



Most of the cyber attacks are happened in the year of 2024 .

Most of its are Benign positive .

Model Building :

The target feature has multiple classes , so Random forest classifier algorithm is used.

Randomforestclassifier model is build for prediction but it gave an accuracy of 67.48% with

Class : 0 , Precision : 0.61 , recall : 0.95 , f1-score : 0.75

Class : 1 , Precision : 0.83 , recall : 0.22 , f1-score : 0.35

Class : 2 , Precision : 0.79 , recall : 0.61 , f1-score : 0.69

Because of low accuracy XGB classifier algorithm is used. It gave an accuracy of 74.09%

Class : 0 , Precision : 0.68 , recall : 0.92 , f1-score : 0.79

Class : 1 , Precision : 0.79 , recall : 0.45 , f1-score : 0.57

Class : 2 , Precision : 0.83 , recall : 0.69 , f1-score : 0.76

To improve accuracy , the class weight is increased in random forest classifier model.

Now the accuracy of the model is 75.90%

Class : 0 , Precision : 0.85 , recall : 0.73 , f1-score : 0.79

Class : 1 , Precision : 0.66 , recall : 0.66 , f1-score : 0.66

Class : 2 , Precision : 0.73 , recall : 0.85 , f1-score : 0.78

The test data is predicted in this model , got 55.75% accuracy.

Class : 0 , Precision : 0.60 , recall : 0.61 , f1-score : 0.60

Class : 1 , Precision : 0.36 , recall : 0.27 , f1-score : 0.31

Class : 2 , Precision : 0.59 , recall : 0.67 , f1-score : 0.63

Issued faced and resolution :

Memory Error :

Because of the size of the dataset it led to memory error during initial data processing due to the computational weight of handling extensive data fields.

To overcome the memory issues encountered, two key preprocessing steps were applied :

Downsampling technique :

Due to the large volume of data, downsampling was performed to reduce the overall size of the dataset. This involved,

1. Random sampling of the dataset to create a smaller representative subset of the data.
2. Selective downsampling of classes with disproportionately large data points, ensuring a balanced class representation and manageable data size.

Datatype Optimisation :

The dataset’s data types were modified to minimize memory usage. Specifically,

1. Reducing integer and float datatypes : Floating point variables were converted from float64 to float32 and integer fields from int64 to int32 .
2. Optimizing categorical fields by converting string based categories into numerical codes, thereby reducing the memory footprint .

The combination of downsampling and datatype optimization allowed the model to efficiently process the data without exceeding memory limitations.

Model Development :

Once the data was prepared, machine learning models were trained to classify attack techniques. Key stages in model development included.

1. Feature Engineering : Features were engineered based on MITRE ATT&CK tactics and techniques, focusing on distinguishing factors between different types of attacks .
2. Model Selection : Various machine learning model were evaluated, including tree based algorithms and neural networks, with an emphasis on balancing accuracy and computational efficiency .
3. Training and Evaluation : The models were trained on the optimized dataset and performance was evaluated based on accuracy , precision , recall and F1 score.

Result and analysis :

The best performing model was able to classify cybersecurity threats effectively within the constraints of reduced memory usage. The memory optimization techniques ensured that the model could handle the data without crashes, making it suitable for deployment on machines with limited memory resources.

Conclusion :

The project demonstrates the feasibility of using machine learning with the MITRE ATT&CK framework for cybersecurity threat detection. By employing downsampling and datatype optimization, memory constraints were successfully mitigated, enabling efficient processing of high – dimensional cybersecurity data.