MICROSOFT

CLASSIFYING CYBERSECURITY INCIDENTS – PROJECT REPORT

BY

KARAN KUMAR M

1. Objective:

The goal of this project is to classify cybersecurity incidents by predicting their triage grade using machine learning techniques. This classification is designed to assist Security Operation Centers (SOCs) in efficiently managing and prioritizing incidents, enabling faster responses to critical threats.

2. Methodology:

2.1 Data Exploration & Preprocessing:

- **Data Source:**
  The training and test datasets consist of millions of rows with features such as AlertTitle, Category, EntityType, EvidenceRole, among others. The target variable for prediction is IncidentGrade.
- **Handling Missing Values:**
  Significant missing values were observed in columns such as MitreTechniques, IncidentGrade, and others. These missing values were handled by imputing values where possible or by dropping irrelevant columns.
- **Feature Extraction:**
  Features such as Year, Month, Day, and Hour were extracted from timestamp data, while irrelevant timestamp columns were removed.
- **Feature Encoding:**
  Categorical columns were label encoded to make them interpretable by machine learning models, ensuring more efficient training.

2.2 Model Selection & Training:

Several models were evaluated using a down sampled version of the dataset to speed up processing, including:

- Logistic Regression
- Decision Tree
- Random Forest
- Gradient Boosting
- Support Vector Machine (SVM)
- K-Nearest Neighbors (KNN)

3. Model Evaluation:

Model performance was assessed using the following metrics:

- **Accuracy:** Measures the overall correctness of the model.
- **Precision:** Assesses the accuracy of positive predictions.
- **Recall:** Evaluates the model's ability to capture all positive instances.
- **F1 Score:** The harmonic mean of precision and recall.
- **Macro-F1 Score:** Averages the F1 scores across all classes, treating each class equally, regardless of size.

4. Findings:

4.1 Best Performing Model:

- The **SVM (Support Vector Machine)** model demonstrated the best performance, achieving a high accuracy score of **0.69**. This indicates that the model effectively balanced precision and recall, providing reliable classification across all incident categories.
- The model also showed robustness against overfitting, thanks to the regularization parameter, making it well-suited for classifying incident grades in cybersecurity.
- Compared to other models Decision Tree, Logistic Regression, Random Forest, Gradient Boosting and KNN, SVM consistently outperformed them in terms of both precision and recall, providing more accurate predictions for critical cybersecurity incidents.

5. Rationale for Model Selection:

- The **SVM (Support Vector Machine)** model was selected as the final model due to its strong performance across key metrics, including accuracy, F1 score, precision, and recall.
- While models like Decision Tree and Random Forest achieved higher precision scores, their lower recall for minority classes made them less optimal for this task

6. Model Improvement:

6.1 Hyperparameter Tuning:

- Further improvements can be achieved by using **RandomizedSearchCV** to optimize key hyperparameters for the **SVM model**, including C, kernel, gamma, and class_weight. In this case, the best parameters identified were:
    - kernel: 'rbf'
    - gamma: 'scale'
    - class_weight: 'balanced'
    - C: 1
- This tuning process helps improve the **Macro-F1 score**, which reached **0.644** with the optimal parameters, enhancing the model's performance in balancing precision and recall.

6.2 Cross-Validation:

- To ensure the robustness of the model, **k-fold cross-validation** can be applied. This method evaluates the model on different data splits, reducing the risk of overfitting and improving the model's ability to generalize to unseen data.

6.3 Feature Engineering:

- Further exploration of feature interactions or the creation of new features, such as adding contextual information related to entities (e.g., 'EntityType', 'Category'), could enhance the model's understanding of the incidents and improve its predictive power.

7. Conclusion:

This project demonstrated the successful application of machine learning in classifying cybersecurity incidents using historical data. The **SVM (Support Vector Machine)** model provided the best performance, achieving a strong balance between accuracy, precision, and recall across all incident categories. By utilizing techniques such as hyperparameter tuning (via RandomizedSearchCV) and exploring feature engineering, the model achieved an improved Macro-F1 score, making it highly effective in handling the complexities of cybersecurity data.

Further refinements, including additional hyperparameter optimization and feature engineering, have the potential to enhance the model's performance even further. This approach provides a valuable tool for Security Operation Centers (SOCs) to prioritize and address cybersecurity threats more efficiently, ultimately improving response times and threat mitigation.