AI Integration of BlockShield Project

Initial ML-Based Intrusion Detection Using the 5G-NIDD Dataset

A Comparative Approach Using PCA and Genetic Algorithm-Based Feature Selection

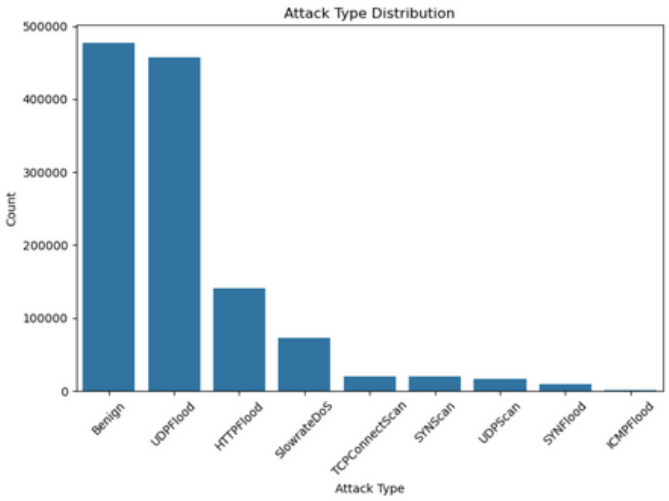
Nashat Alfarajat – BlockShield AI Development Team

Introduction:

This report presents the development of a baseline intrusion detection model using the **5G-NIDD** dataset. Two advanced feature reduction strategies were evaluated: Principal Component Analysis (PCA) and Genetic Algorithm (GA)-based feature selection. While other basic feature importance techniques (e.g., information gain, model-based importance) were tested, they failed to yield significant performance gains or meaningful feature subsets. The GA approach, in contrast, selected a compact and highly predictive feature set. Both PCA and GA-based models were trained using XGBoostc and Random forest, and evaluated through accuracy scores, confusion matrix, and k-fold cross-validation.

Dataset Description

The 5G-NIDD dataset contains 1,215,890 network flow records with 52 features. Each record is labeled with an attack type, which serves as the target variable for classification. These include various forms of Denial-of-Service (DoS), scanning, and other network-based attacks, alongside benign traffic.
The dataset provides a multi-class classification challenge, as the model must distinguish between several different attack types as well as benign activity. This makes it suitable for evaluating the effectiveness of feature selection and dimensionality reduction techniques in intrusion detection.
The dataset also contains the Attack Tool used.



Data Preprocessing

- Standardization of Column Names:
  All column names were cleaned by removing white spaces and ensuring consistent formatting.
- Duplicate Removal:
  Duplicate records were identified and removed to ensure data quality.
- Handling Missing Values:
  - Features with a high percentage of missing values and low relevance were dropped.
  - For numerical columns, missing values were filled using the median.
  - Columns with very few missing values were imputed using the mode or a constant value.
- Encoding Categorical Variables:
  Categorical features, including the target variable (Attack Type), were label encoded to prepare them for model training.

Feature Reduction Approaches

Two approaches were used for feature reduction before training the model: **Principal Component Analysis (PCA)** and **Genetic Algorithm-based feature selection**.

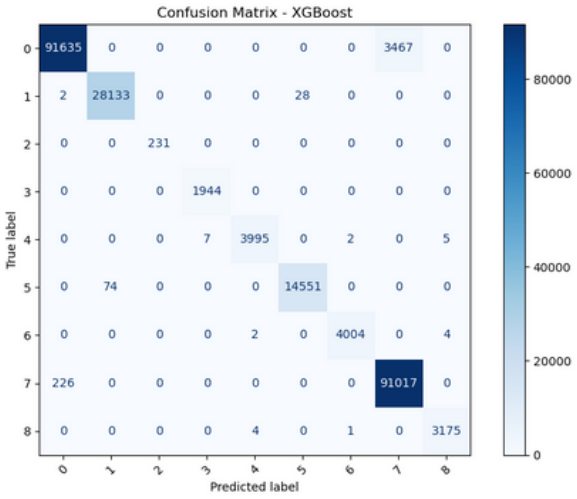
Genetic Algorithm for Feature Selection and XGBoost Modeling

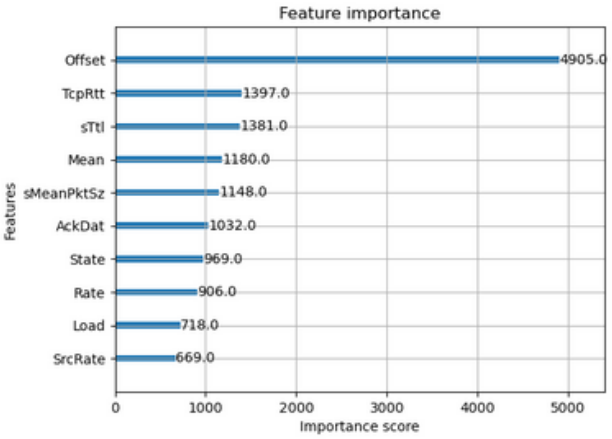
A genetic algorithm (GA) is a stochastic, population-based optimization technique inspired by the process of natural selection. It iteratively evolves a population of candidate solutions—in this case, subsets of features—through selection, crossover, and mutation to find the most effective feature combination for the model.

Using GA on a sample from the dataset, the top 15 most relevant features were selected from the dataset. These features were then used to split the data into training and testing sets. An XGBoost classifier was trained on the selected features to build the intrusion detection model.

Model Evaluation

The XGBoost model trained on the features selected by the genetic algorithm was evaluated using a confusion matrix and classification metrics. The model achieved an overall accuracy of **98.42%**, demonstrating strong performance across all attack classes. Precision, recall, and F1-scores were consistently high, with most classes reaching perfect or near-perfect scores. This indicates that the model reliably detects different types of attacks with minimal misclassification.





Cross-Validation

To ensure the model's robustness and generalizability, training and evaluation were also performed using 5-fold stratified cross-validation. The cross-validation accuracy scores were consistently high across all folds, ranging from approximately 98.34% to 98.44%. The **mean accuracy was 98.39%** with a very low standard deviation of 0.03%, indicating stable and reliable model performance without significant variance between folds.

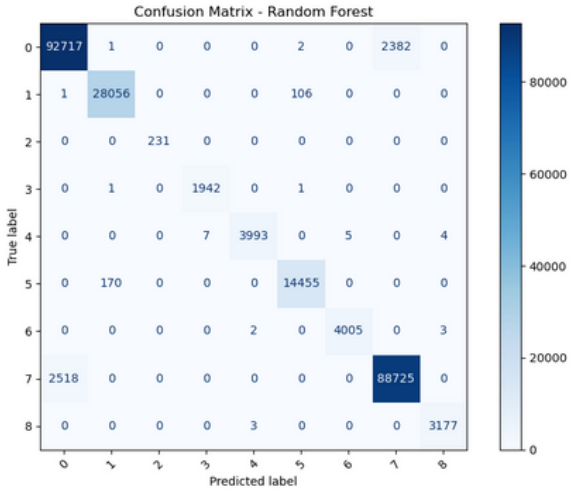
Principal Component Analysis for Dimensionality Reduction and Modeling with XGBoost and Random Forest

Principal Component Analysis (PCA) is a statistical technique used to reduce the dimensionality of data while preserving as much variability as possible. It transforms the original features into a smaller set of uncorrelated components (principal components) ranked by the amount of variance they explain.

In this approach, the dataset was first split into training and testing sets. Both sets were then standardized to ensure consistent scaling. PCA was applied to reduce the number of features before training two classifiers: XGBoost and Random Forest.

Random forest

The Random Forest classifier was trained on the PCA-transformed data and achieved an overall accuracy of **97.85%** on the test set. The model showed strong performance across all attack types, with precision, recall, and F1-scores all consistently high. Most classes had scores close to or equal to 1.00, indicating reliable classification even for less frequent attack types.



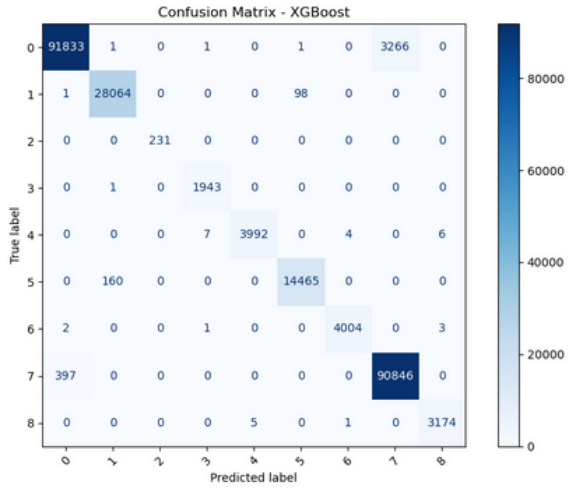

```
Accuracy: 0.9785325784410347
              precision    recall  f1-score   support

           0       0.97      0.97      0.97     95102
           1       0.99      1.00      1.00     28163
           2       1.00      1.00      1.00       231
           3       1.00      1.00      1.00      1944
           4       1.00      1.00      1.00      4009
           5       0.99      0.99      0.99     14625
           6       1.00      1.00      1.00      4010
           7       0.97      0.97      0.97     91243
           8       1.00      1.00      1.00      3180

    accuracy                           0.98    242507
   macro avg       0.99      0.99      0.99    242507
weighted avg       0.98      0.98      0.98    242507
```

XGBoost

The XGBoost classifier, trained on the PCA-transformed data, achieved a test accuracy of **98.37%**. It demonstrated strong classification capabilities across all attack types, with particularly high recall and precision scores. Most classes had near-perfect metrics, including rare attack types, indicating that the model generalized well and handled class imbalance effectively.




```
Accuracy: 0.9836911924191879
              precision    recall  f1-score   support

           0       1.00      0.97      0.98     95102
           1       0.99      1.00      1.00     28163
           2       1.00      1.00      1.00       231
           3       1.00      1.00      1.00      1944
           4       1.00      1.00      1.00      4009
           5       0.99      0.99      0.99     14625
           6       1.00      1.00      1.00      4010
           7       0.97      1.00      0.98     91243
           8       1.00      1.00      1.00      3180

    accuracy                           0.98    242507
   macro avg       0.99      0.99      0.99    242507
weighted avg       0.98      0.98      0.98    242507
```

Cross-Validation

K-fold cross-validation was also applied to evaluate the stability and generalization of the XGBoost classifier. The cross-validation accuracy scores across 5 folds were: [0.9834, 0.9832, 0.9840, 0.9838, 0.9837], with a **mean accuracy of 98.36%** and a standard deviation of only 0.0003. These consistent results across folds further confirm the model's reliability and low variance in performance.
The close match between test accuracy and cross-validation accuracy, along with the low variance across folds, suggests that the model is not overfitting and generalizes well to unseen data.

Conclusion and Future Work

The trained model is designed to automatically detect and classify malicious activity within a 5G-NIDD (Next Generation Intrusion Detection Dataset) environment. By analyzing features extracted from network communication flows, the model successfully distinguishes between benign traffic and multiple types of cyberattacks — including DoS, scanning, and protocol misuse.

This enables **real-time or near-real-time intrusion detection**, helping network operators identify and respond to threats targeting both **mobile and non-mobile devices** across the 5G infrastructure. The system supports proactive defense in complex, high-speed 5G environments, where traditional signature-based or rule-based detection methods may fall short.

Future improvements to the system may include:

- Hyperparameter tuning for models such as XGBoost and Random Forest, which was not applied in this study.
- Exploring deep learning approaches, such as LSTM or transformer-based models, which might better capture temporal or sequential patterns in attack behavior.
- Developing a second-stage model to predict the attack tool used (not just the type), which could enhance forensic capabilities and improve response strategies.

These enhancements would help push the system towards a more intelligent, adaptive, and robust intrusion detection framework suitable for next-generation networks.