Network Intrusion Detection using Machine Learning

This project focuses on detecting network intrusions by applying both supervised classification and unsupervised anomaly detection techniques on network traffic data. The goal is to distinguish normal traffic from potentially malicious or suspicious activity using machine learning models.

Project Structure:

- Network_Intrusion_detection_using_machine_learning.ipynb
- Midterm_53_group.csv
- README.pdf

Objectives:

- Preprocess and clean network traffic data.
- Perform feature engineering and transformation.
- Apply anomaly detection models (e.g., Isolation Forest, One-Class SVM).
- Train classification models (e.g., SVM, Random Forest).
- Visualize data distributions, anomalies, and model performance.
- Evaluate model effectiveness using confusion matrices and classification reports.

Models Used:

Anomaly Detection (Unsupervised):

- Isolation Forest
- One-Class SVM
- Local Outlier Factor (LOF)
- KMeans Clustering

Classification (Supervised):

- Support Vector Machine (SVM)

- Random Forest Classifier

Visualizations:

- Packet Length Distribution (Histogram + KDE)

- Boxplot by Anomaly Status

- Correlation Heatmap

- 3D PCA Plot for Anomaly Clusters

- Confusion Matrices

Data Preprocessing:

- Removed duplicates

- Forward-filled missing values using ffill

- One-hot encoded categorical variables like Protocol

- Dropped irrelevant columns like Info, Source, and Destination

Dataset:

- Midterm_53_group.csv

Contains packet-level network traffic logs captured in a CSV format with time, protocol, length, and other metadata.

Evaluation Metrics:

- Accuracy

- Precision

- Recall

- F1-Score

- Confusion Matrix

How to Run:

1. Clone the repository or upload the notebook and dataset to your environment.

2. Open the .ipynb notebook.

3. Run all cells sequentially.

4. Modify or switch models to compare performance.

Requirements:

- Python 3.8+

- Jupyter Notebook

- pandas, numpy

- scikit-learn

- seaborn, matplotlib

Authors:

- Group 53

License:

This project is for academic use only.