Take-Home Assignment: DDoS Detection

February 20, 2025

Contents

1. Introduction

In this assignment, you are provided with a dataset of network traffic flows collected using CICFlowMeter-V3. The dataset contains both benign traffic and traffic from DDoS attacks. Your task is to develop a solution that distinguishes between these two types of traffic. This document provides a high-level overview of the approach, methodology, and deliverables expected in your submission.

2. Assignment Breakdown

The project is divided into several key sections. Please address each area in your submission using the following guidelines:

2.1. Data Exploration & Analysis

Begin by exploring the dataset to gain an understanding of its structure and features. In your analysis, describe any quality issues or notable patterns you encounter. Use descriptive statistics and visualizations to support your findings. Your summary should highlight insights that may inform your modeling choices.

2.2. Modeling Strategy

Outline the method you choose for detecting DDoS attacks. Briefly explain your reasoning. Describe the advantages and potential limitations of your selected strategy without revealing excessive implementation details.

2.3. Implementation

Provide your code in a clear and well-organized manner. Your submission should include:

- A well-commented codebase using any preferred libraries or frameworks (e.g., `scikit-learn`, `TensorFlow`, `PyTorch`).
- A data pipeline that preprocesses the data prior to model training.
- Optionally, any extra features such as model serving or automation.

Focus on clarity and reproducibility in your implementation.

2.4. Results & Evaluation

Explain how you will measure the performance of your solution. Discuss the metrics you deem appropriate (such as accuracy, precision, recall, F1-score, ROC-AUC, etc.) and present your results. Support your evaluation with visual aids (for example, confusion matrices or ROC curves) where possible.

2.5. Submission

Your final submission should include:

- Your code, organized either in a Git repository or as a compressed file.
- A concise explanation of your approach and key findings.
- Clear instructions on how to run your solution.

Keep your explanations succinct and focused on demonstrating your technical approach.

3. Conclusion

This assignment challenges you to address a real-world network security problem: detecting DDoS attacks through data analysis. Emphasize comprehensive data exploration, a justified modeling strategy, robust implementation, and clear presentation of your results. Your submission should showcase both technical ability and effective communication of your methodology.