Project Name

- DDoSNetGuard: A Machine Learning Approach to Identifying DDoS Attacks

Brief Description/Title

- Implementing an ML System for DDoS Attack Detection using CICIDS2017 Dataset

Team Members

- Sean Sica (just me!)

Problem Statement

- Despite advancements in network security, Distributed Denial of Service (DDoS) attacks remain a critical threat to internet infrastructure, causing significant service disruption. Traditional detection systems struggle to adapt to evolving DDoS tactics, necessitating more dynamic and intelligent solutions.

Objective

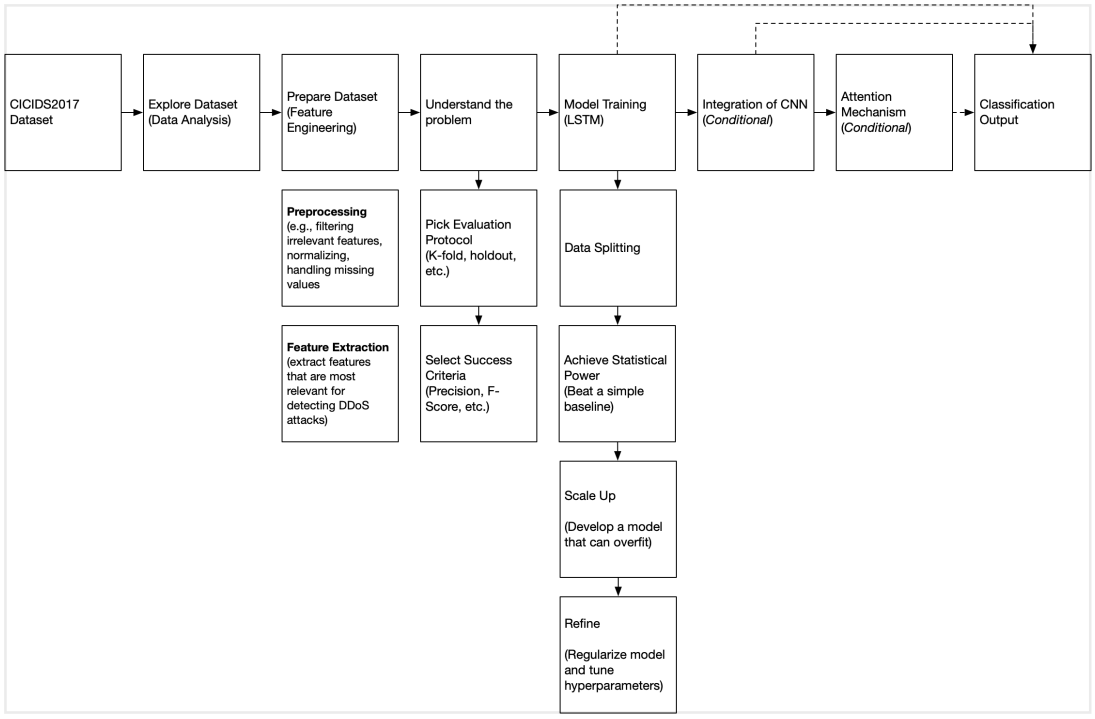
- To develop a machine learning-based system capable of accurately identifying DDoS attacks in network traffic, with potential scalability towards distinguishing various DDoS sub-types.

Approach/Methodology

1. **Data Preparation**: Utilize the CICIDS2017 dataset for training and testing, focusing on labeled network flows that include benign and attack vectors.

2. **Initial Model Development**: Start with an LSTM network to capture temporal dependencies in network flow data indicative of DDoS patterns.
3. **Iterative Enhancement**:
    - Upon successful LSTM implementation, incorporate a **CNN layer** to extract spatial features from the data, improving detection accuracy.
    - Integrate an **attention mechanism** to prioritize features most indicative of DDoS attacks, enhancing the model's focus and efficiency.
4. **Evaluation and Optimization**: Continuously evaluate the model's performance using the CICIDS2017 dataset and refine the architecture and hyperparameters for optimal detection accuracy.

Block Diagram



Datasets (Potential)

- CICIDS2017 dataset, which includes detailed network flows labeled as benign or various attack types, including DDoS.
  - The dataset features labeled network flows, including benign and various DDoS attack vectors, captured from July 3 to July 7, 2017.
  - It encompasses over 80 network flow features derived from full packet payloads in pcap format and CSV files, designed to simulate real-world data including HTTP, HTTPS, FTP, SSH, and email protocols.
  - The dataset includes a range of attack types such as Brute Force FTP/SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet, and DDoS, alongside naturalistic benign traffic.

What is Considered Success/Failure?

- **Success**: Achieving a high accuracy rate in identifying DDoS attacks, with low false positives and negatives, and potential scalability towards classifying DDoS sub-types.
- **Failure**: Inability to significantly outperform traditional detection methods or to adapt to new DDoS patterns not covered in the training dataset.

Evaluation Parameters (Potential)

- Accuracy: Percentage of total predictions that were correct.
- Precision: Of the identified DDoS attacks, how many were actually DDoS attacks.
- Recall: Of all actual DDoS attacks, how many were identified by the model.
- F1 Score: Harmonic mean of precision and recall, providing a balance between the two in cases of class imbalance.