Hands-on Lab: Training Generative AI with Incident Data for Better Pattern Recognition

Estimated time: 30 minutes

Introduction

Welcome to the hands-on lab, Training Generative AI with Incident Data For Better Pattern Recognition.

Training generative AI models for cybersecurity with specific requirements helps them focus on detecting and responding to security threats. This approach improves accuracy in identifying malicious patterns, optimizes resource usage, and ensures adaptability to evolving cybersecurity challenges.

Customization for cybersecurity needs to speed up deployment, lower false positives, and enhance overall threat detection efficiency, resulting in a more robust defense against cyber threats.

In this lab, you will train and customize a generative AI model to meet specific requirements.

Objectives

After completing this lab, you will be able to:

- Leverage the benefits of generative AI models for specific requirements and enhanced pattern recognition.

Exercise: Train the Generative AI model for specific pattern analysis

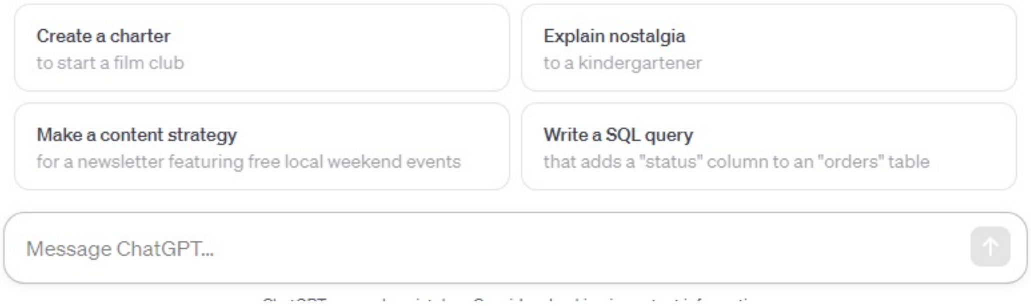
Assume you've received a recent network activity log. After analyzing it with a Generative AI platform like ChatGPT, it incorrectly flagged IP Address 186.20.20.27 as a victim of a DDoS attack. In reality, this IP is hosting an online exam, and multiple requests from various IP addresses are standard and not indicative of a DDoS attack.

Training the ChatGPT model with this information is crucial, clarifying that the IP Address 186.20.20.27 is associated with an online exam and that multiple requests from different sources are typical incidents, not DDoS attacks.

> **Note:** You will need a ChatGPT account for this exercise.

Step 1: Download the Sample Network Log.

Step 2: Open ChatGPT and go to Message Prompt.



Step 3: Type the message in ChatGPT prompt to analyze the network log.

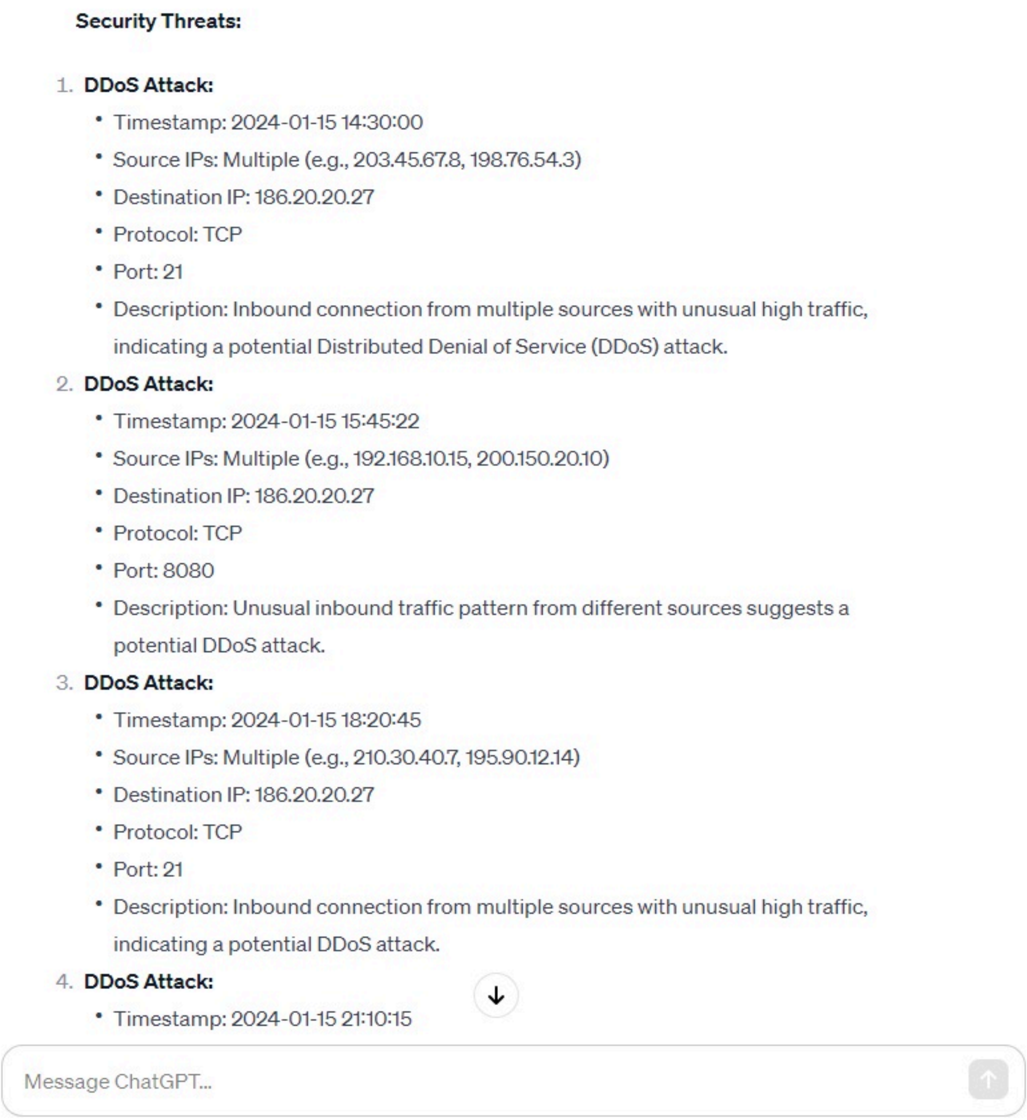
```
Examine the given log entries and provide insights regarding potential security threats or indications of malicious activity.
```

<Copy Paste the Entire Log from M2L2_Network_Log file>

```
Timestamp: 2024-01-16 01:00:00
Source IP: 192.168.45.10
Destination IP: 186.20.20.27
Protocol: TCP
Port: 80
Description: Normal inbound web traffic from a local network source to the server. No signs of unusual activity.
Timestamp: 2024-01-16 02:15:20
Source IP: 10.0.0.5
Destination IP: 186.20.20.27
Protocol: UDP
Port: 53
Description: Routine DNS request from an internal source to the server. Expected behavior with no signs of malicious intent.
.
.
.
.
"
```

Step 4: Observe ChatGPT's response

ChatGPT will generate a response similar to this.



Security Threats:

1. **DDoS Attack:**
   - Timestamp: 2024-01-15 14:30:00
   - Source IPs: Multiple (e.g., 203.45.67.8, 198.76.54.3)
   - Destination IP: 186.20.20.27
   - Protocol: TCP
   - Port: 21
   - Description: Inbound connection from multiple sources with unusual high traffic, indicating a potential Distributed Denial of Service (DDoS) attack.
2. **DDoS Attack:**
   - Timestamp: 2024-01-15 15:45:22
   - Source IPs: Multiple (e.g., 192.168.10.15, 200.150.20.10)
   - Destination IP: 186.20.20.27
   - Protocol: TCP
   - Port: 8080
   - Description: Unusual inbound traffic pattern from different sources suggests a potential DDoS attack.
3. **DDoS Attack:**
   - Timestamp: 2024-01-15 18:20:45
   - Source IPs: Multiple (e.g., 210.30.40.7, 195.90.12.14)
   - Destination IP: 186.20.20.27
   - Protocol: TCP
   - Port: 21
   - Description: Inbound connection from multiple sources with unusual high traffic, indicating a potential DDoS attack.
4. **DDoS Attack:**
   - Timestamp: 2024-01-15 21:10:15

From ChatGPT response, it's evident that it has identified a DDoS attack on IP Address 186.20.20.27. Consequently, there is a necessity to train the ChatGPT model regarding the reality that IP Address 186.20.20.27 is hosting an online exam, and multiple requests from diverse sources are normal incidents, not indicative of DDoS attacks.

> **Note:** ChatGPT's dynamic behavior may result in slightly varied responses during your experimentation. Exploring various prompt instructions may be necessary to observe the expected response.

Step 5: Train the ChatGPT model for specific instruction for better pattern recognition. Type the message in ChatGPT prompt to train the model.

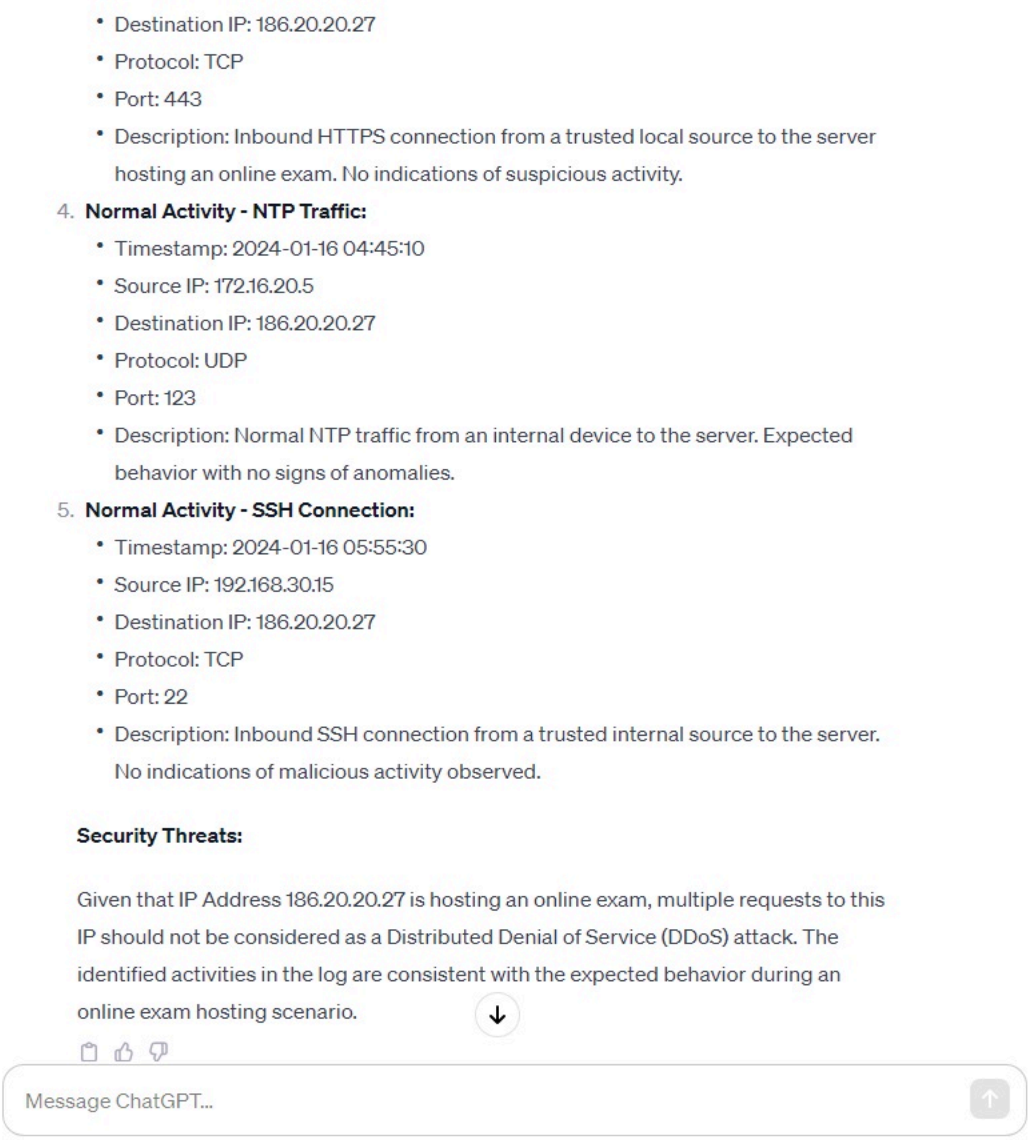
```
"IP Address 186.20.20.27  is hosting a online exam. Do not consider multiple requests to this IP as DDoS attack.
```

> **Note:** ChatGPT will generate a fresh response, just ignore it and move to the next step.

Step 6: Repeat Step 3.

Step 7: Observe ChatGPT's response.

ChatGPT will generate a response similar to this.



- Destination IP: 186.20.20.27
- Protocol: TCP
- Port: 443
- Description: Inbound HTTPS connection from a trusted local source to the server hosting an online exam. No indications of suspicious activity.

4. **Normal Activity - NTP Traffic:**
    - Timestamp: 2024-01-16 04:45:10
    - Source IP: 172.16.20.5
    - Destination IP: 186.20.20.27
    - Protocol: UDP
    - Port: 123
    - Description: Normal NTP traffic from an internal device to the server. Expected behavior with no signs of anomalies.
5. **Normal Activity - SSH Connection:**
    - Timestamp: 2024-01-16 05:55:30
    - Source IP: 192.168.30.15
    - Destination IP: 186.20.20.27
    - Protocol: TCP
    - Port: 22
    - Description: Inbound SSH connection from a trusted internal source to the server. No indications of malicious activity observed.

Security Threats:

Given that IP Address 186.20.20.27 is hosting an online exam, multiple requests to this IP should not be considered as a Distributed Denial of Service (DDoS) attack. The identified activities in the log are consistent with the expected behavior during an online exam hosting scenario.

ChatGPT's response indicates that it has learned the instruction that IP Address 186.20.20.27 is facilitating an online exam, and numerous requests from various sources are normal occurrences, not suggestive of DDoS attacks.

Exercises

1. Get the count for number of Inbound and Outbound Communication for a given log.

▶ Click here for a hint

2. List out the IP address causing unusual traffic. Exempt some IP address from the list.

▶ Click here for a hint

Summary

Congratulations on completing the hands-on lab on training Generative AI with incident data for better pattern recognition!

In this experiment, we explored the effectiveness of training generative AI models with specific requirements in the cybersecurity domain. This strategy focuses on detecting and responding to security threats, leading to improved accuracy in identifying malicious patterns.

The experiment showcased the adaptability of the models to evolving challenges in the cybersecurity landscape through customization for cybersecurity needs. It accelerates deployment and reduces false positives, enhancing the overall efficiency of threat detection. The results contribute to a more robust and specialized defense against cyberthreats.

Author: [Manish Kumar]

Skills Network