DDOS ATTACK DETECTION (Distributed Denial of Service)

(Cyber Security and Machine Learning)

Team Members :

Rajesh Kumar Jena - IIIT BBSR (CSE 2nd Year)

G.N Venkat Raju - IIIT BBSR (CSE 2nd Year)

Manyata Patra - IIIT BBSR (CSE 2nd Year)

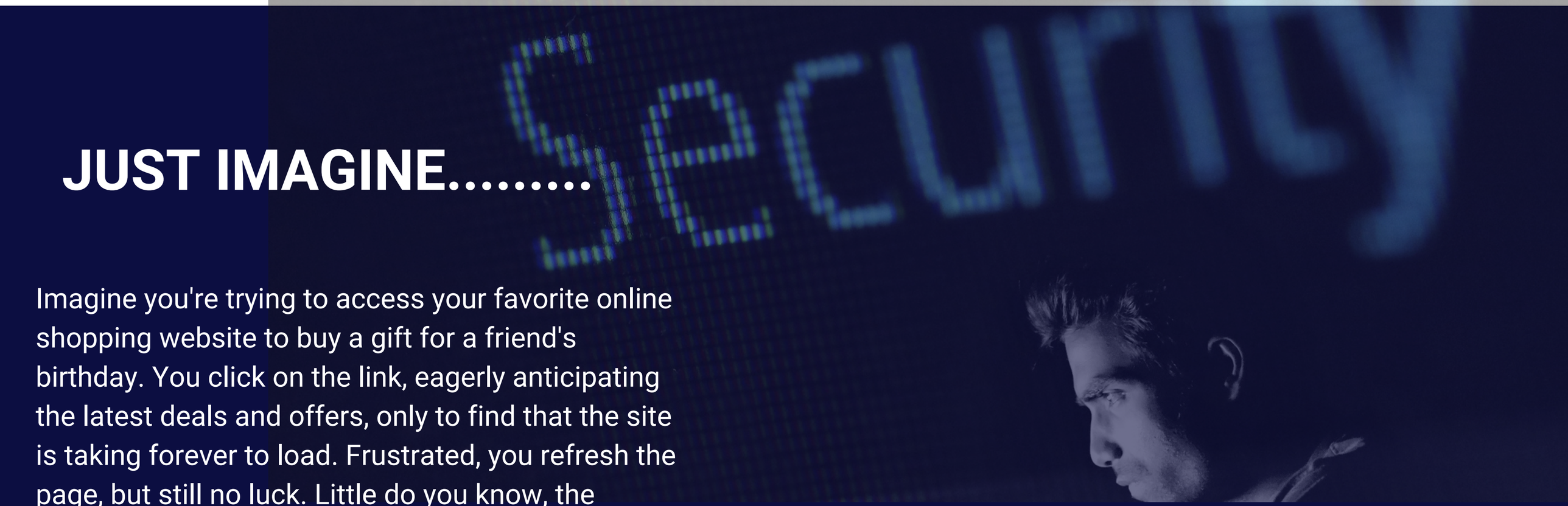
Kaushika Dash - IIIT BBSR (CSE 2nd Year)

Did you know that the largest DDoS attack ever recorded peaked at a staggering 2.3 terabits per second?

That's equivalent to streaming approximately 460 million HD movies simultaneously!

Are you aware of any past instances of your internet-connected devices being compromised or infected with malware or viruses?

JUST IMAGINE.....

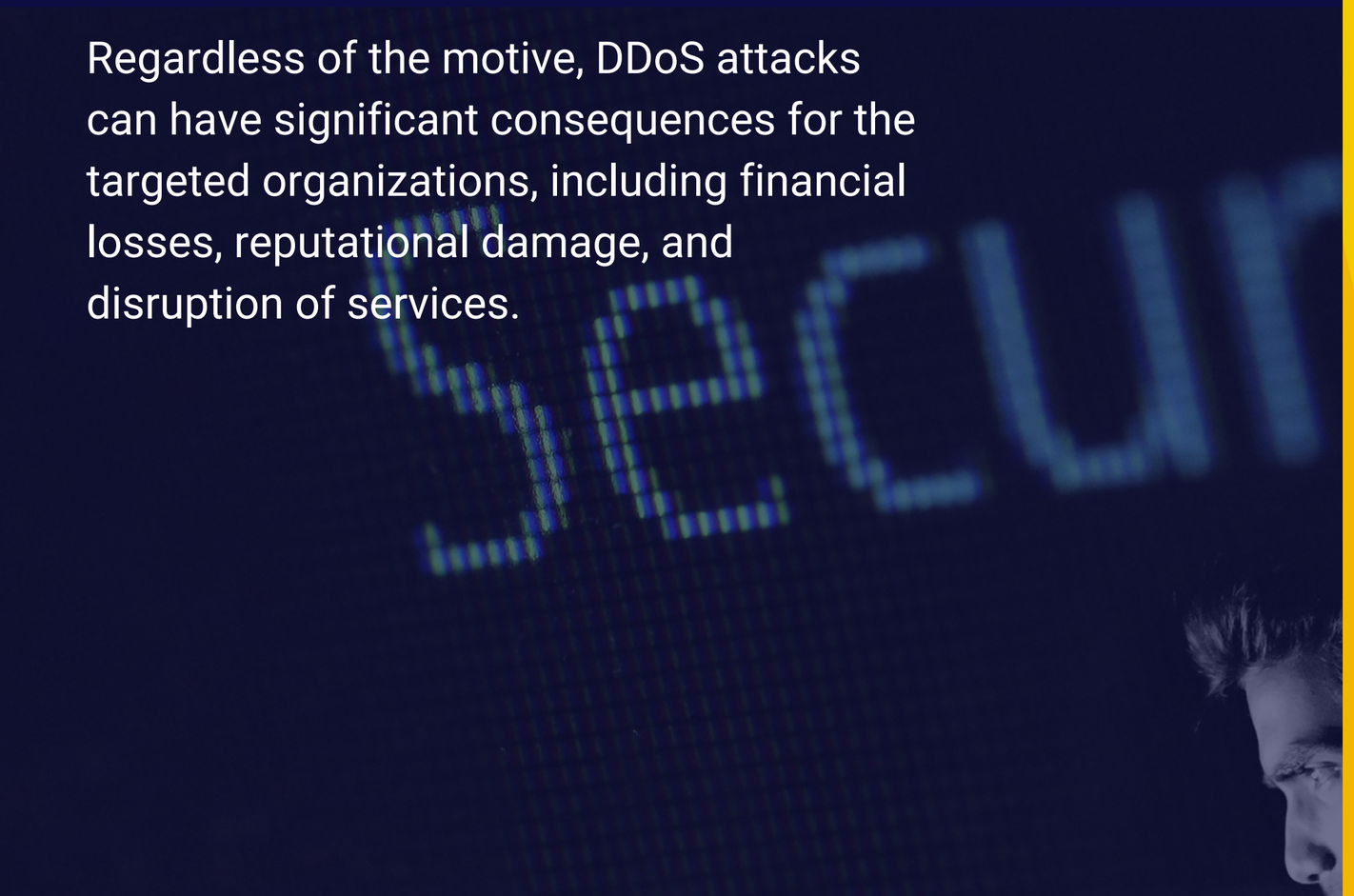


Imagine you're trying to access your favorite online shopping website to buy a gift for a friend's birthday. You click on the link, eagerly anticipating the latest deals and offers, only to find that the site is taking forever to load. Frustrated, you refresh the page, but still no luck. Little do you know, the website is under attack.

This is just one example of a Distributed Denial of Service (DDoS) attack, a malicious attempt to disrupt the normal traffic of a targeted server, service, or network by overwhelming it with a flood of internet traffic. These attacks can come from thousands or even millions of devices connected to the internet, including compromised computers, servers, and Internet of Things (IoT) devices, all working in unison to flood the target with an overwhelming volume of data.

Why hackers do DDOS Attack ?

Regardless of the motive, DDoS attacks can have significant consequences for the targeted organizations, including financial losses, reputational damage, and disruption of services.



- Financial Gain:
Extortion schemes where attackers demand payment to stop the attack.

- Competitive Advantage:
Disrupting rivals' services to divert traffic or cause reputational damage.

- Ideological or Political Reasons:
Promoting ideologies, raising awareness, or protesting against specific entities.

- Revenge or Vendettas:
Retaliating against perceived injustices or personal disputes.

-**Brand Devaluation:** Targeting a brand's online services to undermine its reputation or credibility.

-**Customer Distrust:** Causing service disruptions or data breaches to erode customer trust and confidence.

Project Overview:

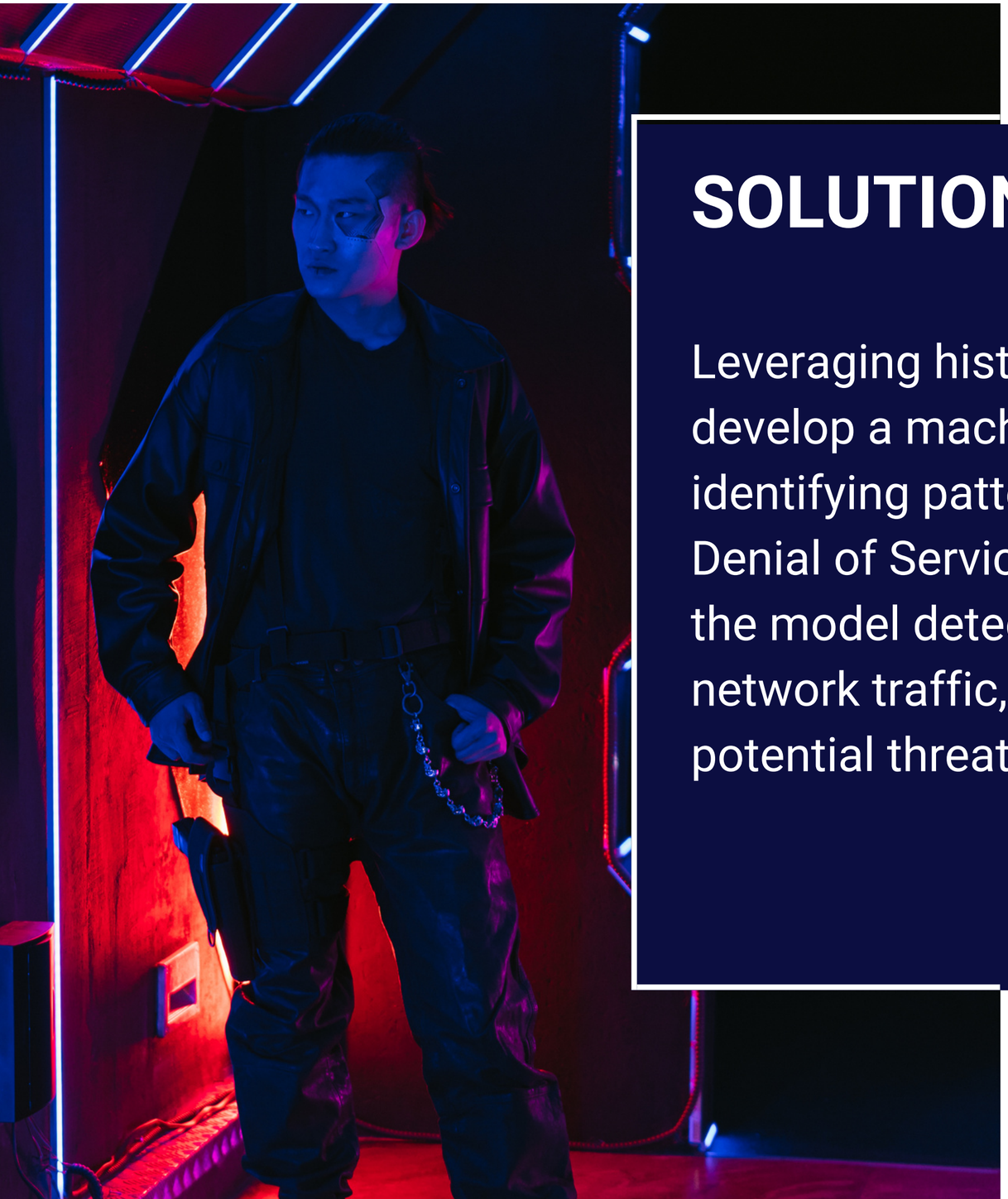
The primary objective of this project is to design and implement a robust machine learning model capable of detecting Distributed Denial of Service (DDoS) attacks within network traffic data. DDoS attacks pose a significant threat to network infrastructure, often resulting in service disruptions and downtime for legitimate users. By developing an effective detection system, we aim to mitigate the impact of such attacks and bolster the resilience of network defenses.

Approach:

Our approach involves leveraging machine learning techniques to analyze and classify network traffic data into two categories: legitimate traffic and DDoS attack traffic. To achieve this, we extract various features from network packets, including packet counts, flow duration, flag counts (e.g., SYN, RST, ACK), and activity metrics (e.g., average packet size, packet length mean).

Feature Selection:

The selection of features is crucial for training an accurate and effective model. We carefully choose features that capture key characteristics of network traffic and are indicative of potential DDoS activity. These features provide valuable insights into the behavior and attributes of network packets, enabling the model to differentiate between normal and malicious traffic patterns.



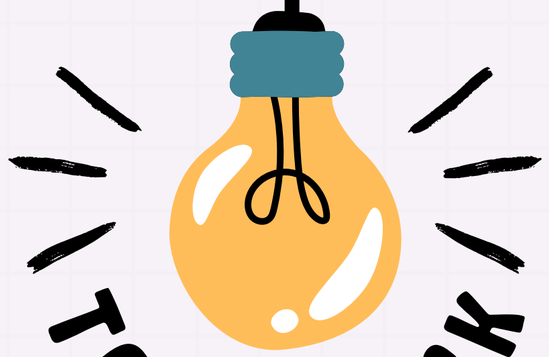
SOLUTION

Leveraging historical data on network traffic to develop a machine learning model capable of identifying patterns associated with Distributed Denial of Service (DDoS) attacks. Additionally, the model detects anomalies in real-time network traffic, enabling proactive responses to potential threats.

HOW ?

PROCESSES

1-Data Collection
2-Data Preprocessing
3-Feature Engineering
4-Model Selection
5-Model Training And Evaluation
6-Web Server Development
7-User Interface Design
8-Integration
9- Deployment



TECH STACK

MACHING LEARNING

PYTHON, FLASK, PANDAS, NUMPY, SKLEARN, REQUEST, JSONIFY, PICKLE, JUPYTER

WEB DEVELOPMENT

HTML , CSS, FLASK , PYTHON

DEPLOYMENT

RENDER,
GITHUB

DATA MODELLING

Data modelling by performing one hot Encoding based internet protocol setting
(UDP, TCP, ICMP)

Splitting the data into training and testing subsets and applying various classification models on them.

GMM

ACCURACY - 0.8651 (86.5%)

KNN

ACCURACY - 0.9629 (96.29%)

K- MEANS CLUSTURING

PURITY SCORE - 0.6226. (62.26%)

RANDOM FOREST CLASSIFIER

ACCURACY - 1.0 (100 %)

SVM

ACCURACY - 0.76. (76 %)