Detection of Cyber Attack in Network using Machine Learning Techniques

Ms.BANDARUPRAHARSHA (18N81A05C4)

Mr.CH.SHIVAPRASAD (18N81A05C7)

Ms. SINGAM AKHILA (18N81A05C8)

Ms.BUDHAVARAM SHAAMBHAVI (18N81A05D6)

Abstract

contrasted with the past, improvements in PC and correspondence innovations have given broad and propelled changes. The use of new innovations give incredible advantages to people, organizations, and governments, be that as it may, messes some up against them. For instance, the protection of significant data, security of put away information stages, accessibility of information and so forth. Contingent upon these issues, digital fear based oppression is one of the most significant issues in this day and age. Digital fear, which made a great deal of issues people and establishments, has arrived at a level that could undermine open and nation security by different gatherings, for example, criminal association, proficient people and digital activists. Along these lines, Intrusion Detection Systems (IDS) has been created to maintain a strategic distance from digital assaults.

Introduction

- Introduction PC wrong doings keep on expanding throughout the years. They are not just confined to inconsequential acts, for example, assessing the login accreditations of a framework yet in addition they are significantly more dangerous. Data security is the way toward shielding data from unapproved get to, use, exposure, decimation, change or harm. The expressions "Data security", "PC security" and "data protection" are regularly utilized reciprocally. These territories are identified with one another and have shared objectives to give accessibility, secrecy, and honesty of data. Studies show that the initial step of an assault is disclosure [1]. Surveillance is made so as to get data about the framework right now.

Cont

- Finding an once-over of open ports in a structure gives incredibly essential information to an attacker. Consequently, there are lots of gadgets to recognize open ports [2], for instance, ant viruses and IDS. At this moment, learning and SVM AI computations were been applied to make IDS models to perceive port yield tries the models were presented with the explanation of used material and techniques

EXISTING SYSTEM

- Blameless Bayes and Principal Component Analysis (PCA) were been used with the KDD99 dataset by Almansob and Lomte [9].Similarly, PCA, SVM, and KDD99 were used Chithik and Rabbani for IDS [10]. In Aljawarneh et al's. Paper, their assessment and examinations were conveyed reliant on the NSL-KDD dataset for their IDS model [11] Composing inspects show that KDD99 dataset is continually used for IDS [6]–[10].There are 41 highlights in KDD99 and it was created in 1999. **Consequently, KDD99 is old and doesn't give any data about cutting edge new assault types**, example, multi day misuses and so forth. In this manner we utilized a cutting-edge and new CICIDS2017 dataset [12] in our investigation.

PROPOSED SYSTEM

- Important steps of the algorithm are given in below.
  1) Normalization of every dataset.
  2) Convert that dataset into the testing and training.
  3) Form IDS models with the help of using RF, ANN, CNN and SVM algorithms.
  4) Evaluate every model's performances

DISADVANTAGE

- Strict Regulations
- Difficult to work with for non-technical users
- Restrictive to resources
- Constantly needs Patching
- Constantly being attacked

ADVANTAGE

- Protection from malicious attacks on your network.
- Deletion and/or guaranteeing malicious elements within a preexisting network.
- Prevents users from unauthorized access to the network.
- Deny's programs from certain resources that could be infected.
- Securing confidential information

Literature survey

Cyber criminals often use port scanning to identify weak points in computer systems before launching an attack. When using a modem or a local area network, you'll often find services listening on both well-known and obscure ports. By port scanning, an attacker can discover what services are running on a targeted system, who owns those services and whether anonymous logins are allowed. Port scanning is accomplished by sending a message to each port one at a time. The type of response you receive lets you know whether or not the port is in use and whether or not it can be further probed for security flaws and vulnerabilities. Port scanners are used by network security professionals because they can detect security flaws on the targeted system.

REQUIRMENTS:

Hardware requirements:

- Operating system: windows, linux
- Processor : minimum intel i3
- Ram: minimum 4 gb
- Harddisk : minimum 250gb

REQUIRMENTS

Software requirements

- Python idel 3.7 version (or)
- Anaconda 3.7 (or)

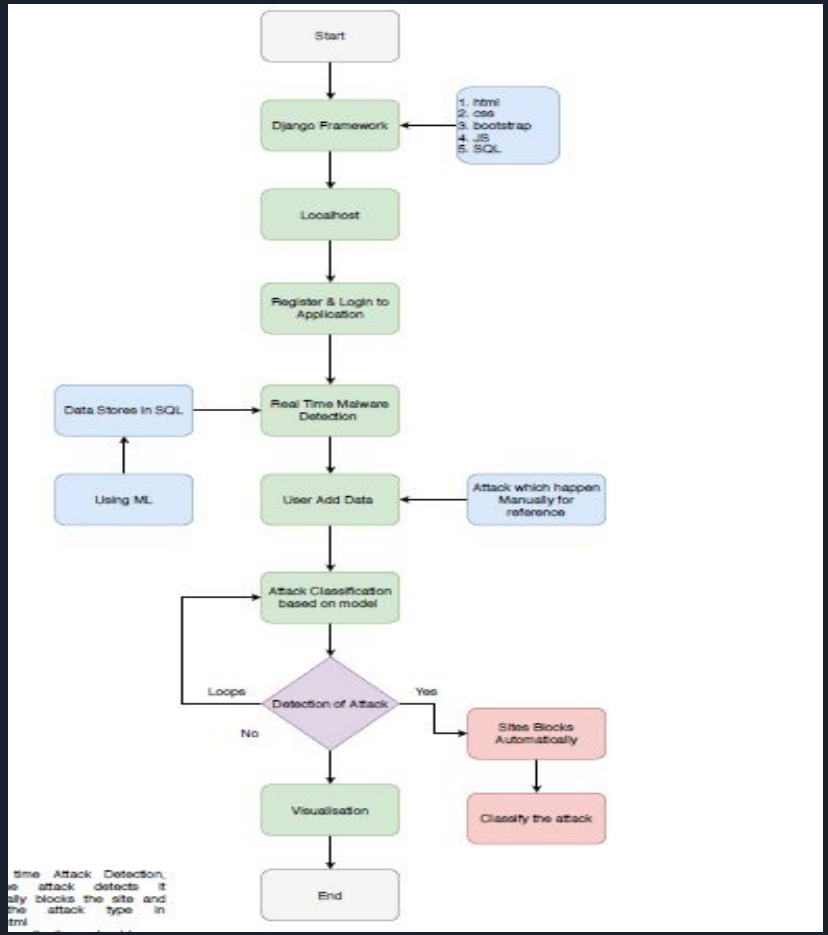
  Jupiter (or)

  Google colab

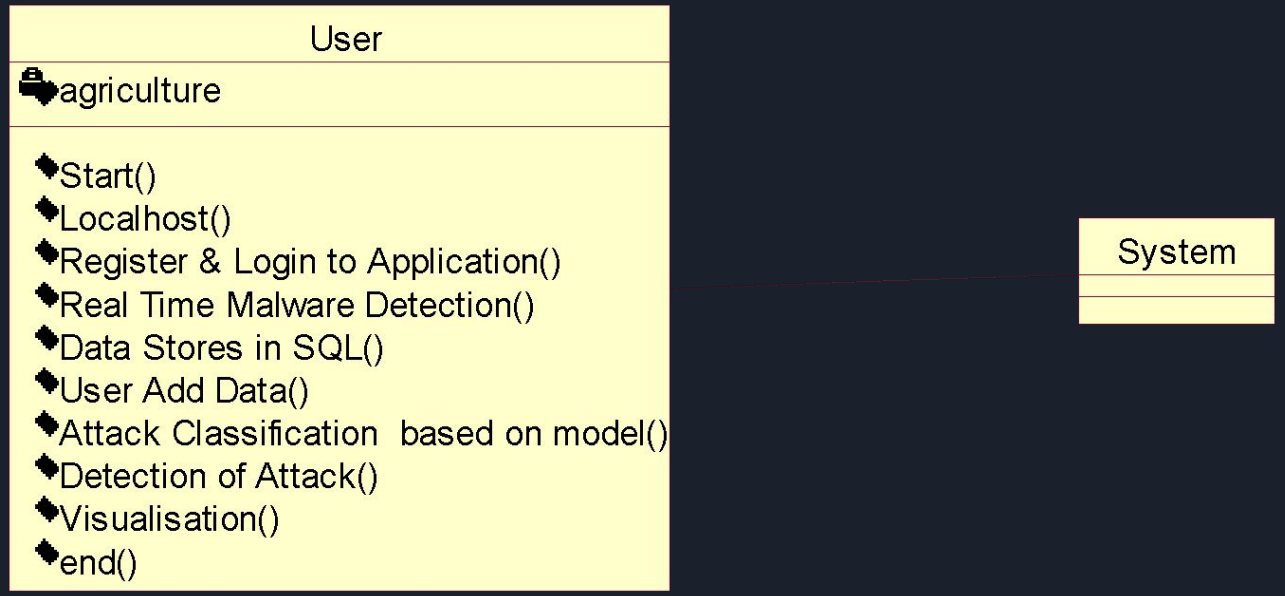
Programming language and packages

- Python
- Numpy, pandas,matplotlib.
- Tensorflow, opencv, nlp, nltk etc

FLOW CHART


Start
Django Framework
1. html
2. css
3. bootstrap
4. JS
5. SQL
Localhost
Register & Login to Application
Data Stores in SQL
Real Time Malware Detection
Using ML
User Add Data
Attack which happen Manually for reference
Attack Classification based on model
Loops
Detection of Attack
Yes
No
Sites Blocks Automatically
Visualisation
Classify the attack
End

UML DIAGRAM


User
agriculture
Start()
Localhost()
Register & Login to Application()
Real Time Malware Detection()
Data Stores in SQL()
User Add Data()
Attack Classification based on model()
Detection of Attack()
Visualisation()
end()
System

Machine learning Algorithm models

Random forest

Artificial neural network

Convolutional neural network

Support vector machine

Conclusion

- Right now, estimations of help vector machine, ANN, CNN, Random Forest and profound learning calculations dependent on modern CICIDS2017 dataset were introduced relatively. Results show that the profound learning calculation performed fundamentally preferable outcomes over SVM, ANN, RF and CNN. We are going to utilize port sweep endeavors as well as other assault types with AI and profound learning calculations, apache Hadoop and sparkle innovations together dependent on this dataset later on. All these calculation helps us to detect the cyber attack in network. It happens in the way that when we consider long back years there may be so many attacks happened so when these attacks are recognized then the features at which values these attacks are happening will be stored in some datasets. So by using these datasets we are going to predict whether cyber attack is done or not. These predictions can be done by four algorithms like SVM, ANN, RF, CNN this paper helps to identify which algorithm predicts the best accuracy rates which helps to predict best results to identify the cyber attacks happened or not.

Thank you!