**Project: Network Attacks Classification Using Deep Learning**

**Abstract:** In this digital era, we know that everything is shifting in online mode it may be confidential meetings, online banking, educational meeting or whatever service we are using through the internet all are under threat and different types of intruders or attackers can attack in our network to get data from our node, so to provide the user a system that in which he/she can identify that which kind of attack can take place in his system, so for this type of problem we are making a attack detection system which will notify the user and provide the information about the websites and other things which can detect their data.

This project is based on the classification of different attack in a network, in this we are developing a model which will detect the type of attack and classify it in various groups like dos attack, intrusion, phishing attack, fuzzing attack, SQL-injection etc. In this model we have used different types of technologies like Deep learning, Machine learning as superset and the different classification algorithms of machine learning, convolutional neural network, and heat map as subset. In this project we got the dataset from the Kaggle, and we will use it as a training data for our model.

The problem statement is very simple that when we try to find out the type of different network attack in network, then there are so many model that are trained for individual attacks not for a set of attacks which may occur in our network so to classify all of them in a single model we are building this model which will provide the information about the attack using CNN.

**Introduction:** This project is basically a combination of machine learning and cyber security. In this we are making a model of machine learning which will be trained for the classification of different types of attacks in a network. In this we are using machine learning and different types of classification algorithm in it to classify the attacks. Our team has used a dataset from Kaggle for the training of our model. The brief introduction of the data set is following.

In this data set we have different things which are following:

1: A pre-processed data set in csv format.

2: Label vector in csv format.

3: the original network in packet capture format.

In this data set each pre-processed dataset contains x rows (packets) and 115 columns which are features without header, the 115 feature were extracted using an extractor by the dataset provider. In brief the 115 features provide a statistical view of the network means host and behaviours with respect to the current packet moving in the network. As we know there are different types of attack in cybersecurity but if we think in large scale, we can divide them into the following four categories:

U2R

R2L

Taxonomy

DOS

Surveillance

In this project we are not dealing with all kinds of attacks in network, we are training the model for 3 to 4 attacks in network, some of the attacks for which we are training the model are following with a brief introduction.

**Phishing Attack:** Phishing attacks are the method of attack in which sending fake communications that seems to come from a trusted source take place. In general, it is done through email and fake links circulated through messages. The aim of this is to capture the confidential data like passwords, usernames stored in our system and banking information, or to install virus on the sufferer’s machine. Many of the time this attack done through the phishing websites.

In general phishing starts with a fraud email or through some communication channel like it may give us a discount coupon in a site and that coupon led us to the phishing website and through several other ways. And when the attack is done through a communication medium then the email or other messages looks so familiar it seems that we know them very well, and they fool the victim, and they ask for the confidential information and sometimes it installs malware in targets computer of system.

It is a very dangerous attack as the data yield from this can be used for the attack in particular company where the victim works, and this information may use to get access to the bank account of the user and many more. Types of phishing attack are Deceptive, Spear, Whaling and Pharming.

**Dos Attack:** It is known as **Denial-of-Service (DoS).** It is an attack used to shut down a machine or network or server, making it unreachable to its contemplated users. DoS attacks complete this task by flooding the target by huge traffic or by sending it information that induce to a crash. In both cases, the DoS attack repels the users (for example employees, or account holders) of the service or resource they expected.

Fortunately, dos attack does not outcome a loss of theft but in general it decreases the performance of the site or server as it sends request in huge amount and slow down the speed of the server and dos cost time and money.

Types of Dos Attack: -

(1) Flooding services: it occurs when there is a heavy traffic in the system for the server to buffer and as a result the system performance decrease and finally stop. The example of this attack is Buffer overflow attacks, ICMP Flood and SYN Flood etc.

(2) Crashing services: it simply manipulates loopholes that causes the target system or service to crash.

**SQL Injection:** SQL injection is a web security vulnerability or openness that gives permission to an attacker to hamper with the queries that an application makes to its datastore or database. It gives access to the attacker to retrieve data modify data and whatever action he wants to do on the data base.

A successful SQL injection attack can able the attacker to capture the sensitive data, such as passwords, bank details, credit card details, or personal user information.

For different situation there are different techniques. Some common examples of SQLi are:

* [Retrieving Hidden data](https://portswigger.net/web-security/sql-injection#retrieving-hidden-data), in this we can change an SQL query to return extra information.
* [Subverting application logic](https://portswigger.net/web-security/sql-injection#subverting-application-logic), in this we can modify a query to hamper with the application's logic.
* [UNION attacks](https://portswigger.net/web-security/sql-injection/union-attacks), in this we can retrieve data from different database tables, and here the duplicity is not allowed.
* [Examining the database](https://portswigger.net/web-security/sql-injection/examining-the-database), in this we can take out information about the editon and structure of the DB.
* [Blind SQL injection](https://portswigger.net/web-security/sql-injection/blind), in this the results of a query we control are not returned in the application's acknowledgement.

The motivation behind this model is simple that we want to develop a model which will be very helpful to classify the different types of attack in a network.

**Literature Survey:** To compare different methods used in the intrusion detection system, we have done survey of some of the research paper and the result from the research paper is in the tabular form, which is following, and for the reference of research paper we have written the reference no of the reference page which we added in the synopsis. In this table we have compared algorithm based on their testing against the KDD-cup 99, Kyoto and DARPA datasets. And the accuracy of the algorithms is also mentioned. Our task in this project to measure the accuracy rate of all the algorithm against the data set which we got from the Kaggle, and we will make a customize data set to check the accuracy again, and we will use CNN with the customized data set and classify the attacks in their group with accuracy rate on which our model is able to detect the attack-type.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Used Algorithm** | **Objective** | **Accuracy** | **Datasets** | **Type of attack** | **Reference** |
| “K-means” + “KNN” | IDS | 93.55 % | KDD-cup 99 | Dos, U2R, R2L, Probing attack | [22] |
| “SVM+ KNN+PSO” | IDS | 88.44 % | KDN-cup 99 | Dos, U2R, R2L, Probing attack | [23] |
| “HC+SVM” | IDS | 95.72 % | KDD-cup 99 | Dos, U2R, R2L, Probing attack | [24-25] |
| “RF+AODE” | IDS | 90.51 % | Kyoto | Attack against honeypots | [26] |
| “FL+ES” | IDS | 91.5 % | DARPA2000 | DDoS, Darpa Attack | [27] |
| “FL+GA” | IDS | 94.6 % | DARPA-KDD99 | Dos, U2R, R2L, Probing attack | [28] |

SVM: support vector machine, KNN: k-nearest algorithm, POS: particle swarm optimization, AODE: average one dependence estimator, HC: hierarchical clustering, RF: random forest, FL: fuzzy logic, ES: expert system, GA: genetic algorithm, IDS: intrusion detection system.

**Methodology:** To solve the problem of classification we are using heat map and CNN.

**Heat map:** A **heat map** is a data visualization technique that used to visualize the data and the matrix which we got from the heat map is used for the input in CNN and it shows magnitude of a abate as colour in multidimension usually 2dimension matrix. The difference in colour may be by tint, giving obvious visual cues to the reader about how the abate is grouped or changes over space.

There are two different categories of heat maps which are (1) the cluster heat map (2) the spatial heat map. In a cluster heat map, enormity is laid out into a matrix of static or constant cell size whose rows and columns are detached phenomena and categories, and the sorting of rows and columns is planned and somewhat capricious, with the goal of suggesting clusters or drawing them as discovered via statistical investigation. The size of the cell is capricious but good enough to be easily visible. By antithesis, the position of an enormity in a spatial heat map is coerced by the position of the enormity in that space, and there is no concept of cells; the occurrence is considered to vary continuously.

To check the attack in our system we contain a security heat map. it is mandatory to find exactly where the problems or challenges are and what we must do fix this and to keep the system fully secured.

In short, the heatmap graphically indicates cyber-risks in a data network. Tables and charts are used to interpret and derive meaning from them. There are individual colours in a matrix that denote the degree of risk that an organisation may be facing. It is decidable by the organization that which colour to choose for higher risk and which colour to choose for lower risk, in general all the organization chooses red as highest risk point and green as lowest risk point.

There are some steps which are involved in heat map designing and the steps are following:

1: Identifying Risk

2: Examining Risk

3: Mapping Risk

4: Monitoring and updating

**CNN:** A Convolutional Neural Network, also known as ConvNet, it is a genre of neural network (which are the part of deep learning) that specialized in processing data that involve a grid-like topology or data in matrix form, such as a matrix of image. A digital image is a 0 or 1 representation of visual data. It has a series of pel arranged in a grid format that holds pel values to denote how intense and what colour each pel will be.

There are three layers in CNN: convolutional layer, pooling layer, and fully connected layer.

****

(1): Convolutional layer: it is primary building block of a Convolutional neural network. It has a set of Kernels, confines of which are to be accomplished in the process of training. Generally, the dimensions of the filters are lesser than the actual image. Each filter combines with the image and form an activation map.

(2): Pooling Layer: The primary goal of pooling layer is to gradually decrease the spatial size of the image provided by the heat map, so that count of computational operations in the network are decreased. Pooling executes down sampling by decreasing the dimensions and sends important data to next layers in CNN.

(3): Fully Connected Layer: In this layer the neurons have full accordance with all neurons in the previous and next layer in FCNN. So, it can be concluded by a matrix multiplication which is amid by a bias effect. The Fully Connected layer supports to map the delineation between the input and the output.

This is the brief about the CNN and there are some classification algorithms in ml which are following:

1. Linear Model: - Logistic Regression, Support vector machine
2. Non-linear Models: - K-nearest neighbours / KNN, Kernel SVM, Naïve Bayes, Decision Tree, Random Forest

This is all about the technologies which we will use to make the model and the data set for the model is available in the Kaggle.

As CNN works on matrices and using certain equation it classifies the images, and the mathematics behind the classification is following:

As we know heat map generates a matrix and that matrix will act as the input image for our CNN. As we know kernel convolutional is a key element of computer vision algorithms, and here kernel convolutional is also used. Kernel convolutional is a process in which we take a input matrix (this will be a number matrix) it is known as kernel or filter, we pass it over image and transform it based on the value from filter.

To calculate the value of subsequent feature map we use the following formula:

In this equation H:

name of the matrix|

a: no of row|

b: no of column|

x: current row index|

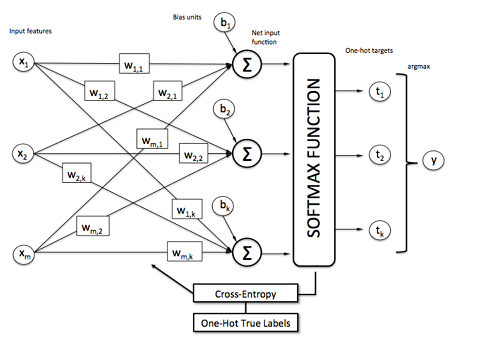
y: current column index|

c: input image|

d: kernel|

Soft Max function: it is a bracing method that convert a vector of R real values to a vector of R real values that accumulate to 1. The inserted values are in the range of real number that it may be negative and positive and may be zero, but the SoftMax transform them into 0 to 1. As a result, the values can be treated as probabilities.

So, in case the input is small then we got small probability and large for the large values.



The formula for the SoftMax function is following:

Here the symbols are:

I: input of SoftMax function in the range of (I0…IK) |

I i: element of the input vector the SoftMax function I in the range of Real number|

It is an exponential function, and it is practised to every element of the input vector, it output a value greater then 0, if input is small value will be small and if large then value will be large|

K: the total count of classes in multi class classifier|

: it is the normalization term. It makes sure that all the resultant values of the method will sum to 1 and each be in range of (0,1) |

As it is synopsis of our project, and we haven’t implemented it completely so the flow in which the solution of the project will take place is mentioned below.

The flow of solution will occur in the sequence of (1) download data set form Kaggle (2) the data set read by google collab (3) label specification and dimension specification (4) pre-processing of data (5) creation of heat map (6) CNN classifier (7) Classification of network.

Download data set from Kaggle

Upload data in google collab

Label , classification, size and dimension

Preprocessing

Heat map

CNN classification

Network which are classified

**References:**

1: <https://www.researchgate.net/figure/Comparison-of-different-attacks_tbl2_301720689>

2: <https://ieeexplore.ieee.org/document/9753934>

3: <https://ieeexplore.ieee.org/document/7351952>

4: <https://ieeexplore.ieee.org/document/8782319>

5: <https://ieeexplore.ieee.org/document/7081776>

6: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9477555>

7: <https://ieeexplore.ieee.org/document/9666581>

8: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9656911>

9: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8359287>

10: https://www.kaggle.com/datasets/ymirsky/network-attack-dataset-kitsune

11: <https://towardsdatascience.com/adversarial-attacks-on-sms-spam-detectors-12b16f1e748e>

12: <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>

13: <https://portswigger.net/web-security/sql-injection>

14: <https://www.paloaltonetworks.com/cyberpedia/what-is-a-denial-of-service-attack-dos>

15: <https://www.cisco.com/c/en_in/products/security/email-security/what-is-phishing.html>

16: <https://jisajournal.springeropen.com/articles/10.1186/s13174-019-0115-x>

17: <https://www.researchgate.net/publication/348604306_A_Survey_on_Network_Intrusion_System_Attacks_Classification_Using_Machine_Learning_Techniques>

18: <https://towardsdatascience.com/>

19: <https://www.annalsofrscb.ro/index.php/journal/article/view/1137/963>

20: Research Article: DMAT: A New Network and Computer Attack Classification

21: Survey Paper: A Survey on Network Intrusion System Attacks Classification Using Machine Learning Techniques by V. Deepa and N. Radha 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1022 012036

22: Tsai, C.-F.; Lin, C.-Y. A triangle area based nearest neighbors approach to intrusion detection. Pattern Recognit. 2010, 43, 222–229.

23: Aburomman, A.A.; Reaz, M.B.I. A novel SVM-kNN-PSO ensemble method for intrusion detection system. Appl. Soft Comput. 2016, 38, 360–372

24: Horng, S.-J.; Su, M.-Y.; Chen, Y.-H.; Kao, T.-W.; Chen, R.-J.; Lai, J.-L.; Perkasa, C.D. A novel intrusion detection system based on hierarchical clustering and support vector machines. Expert Syst. Appl. 2011, 38, 306–313.

25: Khan, L.; Awad, M.; Thuraisingham, B. A new intrusion detection system using support vector machines and hierarchical clustering. VLDB J. 2007, 16, 507–521.

26: Jabbar, M.; Aluvalu, R. RFAODE: A novel ensemble intrusion detection system. Procedia Comput. Sci. 2017, 115, 226–234.

27: Liao, N.; Tian, S.; Wang, T. Network forensics based on fuzzy logic and expert system. Comput. Commun. 2009, 32, 1881–1892.

28: Chadha, K.; Jain, S. Hybrid genetic fuzzy rule-based inference engine to detect intrusion in networks. In Intelligent Distributed Computing; Springer: Cham, Germany, 2015; pp. 185–198.

**Student Details:**

Name Section Roll-number University

RACHIT SHARMA B 1918574 Gehu-Dehradun

RAGHVENDRA SINGH E 1918576 Gehu-Dehradun

PAWAN SINGH B 1918531 Gehu-Dehradun

**Name Of Supervisor:** Dr SATVIK VATS

**Signature of Supervisor:**