PREDICTION OF CYBER ATTACK USING DATA SCIENCE TECHNIQUE IN MACHINE LEARNING

Akshayashri G, Dhivya M, Harshini Priya A, Nithisha J

Student, Information Technology, Jeppiaar Engineering College, Tamilnadu,India.
Student, Information Technology, Jeppiaar Engineering College, Tamilnadu,India.
Student, Information Technology, Jeppiaar Engineering College, Tamilnadu,India
Assistant Professor,Information Technology, Jeppiaar Engineering College, Tamilnadu,India

ABSTRACT

Cyber-attack, via cyberspace, targeting an enterprise's use of cyberspace for the purpose of disrupting, disabling, destroying, or maliciously controlling a computing environment/infrastructure; or destroying the integrity of the data or stealing controlled information. The state of the cyberspace portends uncertainty for the future Internet and its accelerated number of users. New paradigms add more concerns with big data collected through device sensors divulging large amounts of information, which can be used for targeted attacks. Though a plethora of extant approaches, models and algorithms have provided the basis for cyber-attack predictions, there is the need to consider new models and algorithms, which are based on data representations other than task-specific techniques. However, its non-linear information processing architecture can be adapted towards learning the different data representations of network traffic to classify type of network attack. In this paper, we model cyber-attack prediction as a classification problem, Networking sectors have to predict the type of Network attack from given dataset using machine learning techniques. The analysis of dataset by supervised machine learning technique (SMLT) to capture several information's like, variable identification, uni-variate analysis, bi-variate and multi-variate analysis, missing value treatments etc. A comparative study between machine learning algorithms had been carried out in order to determine which algorithm is the most accurate in predicting the type cyber-Attacks. We classify four types of attacks are DOS Socio-technical attack is an organized approach which is defined by the interaction among people through maltreatment of technology with some of the malicious intent to attack the social structure based on trust and faith. Awful advertisement over internet and mobile phones may defame a person, organization, group and brand value in society which may be proved to be fatal. People are always very sensitive towards their religion therefore mass spread of manipulated information against their religious belief may create pandemonium in the society and can be one of the reasons for social riots, political misbalance etc.

Cyber-attack on water, electricity, finance, healthcare, food and transportation system may create chaos in society within few minutes and may prove even more destructive than that of a bomb as it does not attack physically but it attacks on the faith and trust which is the basic pillar of our social structure. Attack, R2L Attack, U2R Attack, Probe attack. The results show that the effectiveness of the proposed machine learning algorithm technique can be compared with best accuracy with entropy calculation, precision, Recall, F1 Score, Sensitivity, Specificity and Entropy.

Keyword:— *Datasets, Prediction, Protocol type, GUI ,Logistic regression, Decision tree, Random forest, Support vector machine*

1.INTRODUCTION

There are supervised learning, unsupervised learning and reinforcement learning. Supervised learning program is both given the input data and the corresponding labeling to learn data has to be labeled by a human being beforehand. Unsupervised learning is no labels. It provided to the learning algorithm. This algorithm has to figure out the clustering of the input data. Finally, Reinforcement learning dynamically interacts with its environment and it receives positive or negative feedback to improve its performance.

Data scientists use many different kinds of machine learning algorithms to discover patterns in python that lead to actionable insights. At a high level, these different algorithms can be classified into two groups based on the way they "learn" about data to make predictions: supervised and unsupervised learning. Classification is the process of predicting the class of given data points. Classes are sometimes called as targets/ labels or categories. Classification predictive modeling is the task of approximating a mapping function from input variables(X) to discrete output variables(y). In machine learning and statistics, classification is a supervised learning approach in which the computer program learns from the data input given to it and then uses this learning to classify new observation. This data set may simply be bi-class (like identifying whether the person is male or female or that the mail is spam or non-spam) or it may be multi-class too. Some examples of classification problems are: speech recognition, handwriting recognition, bio metric identification, document classification etc.

2.LITERATURE SURVEY

Related work

2.1.Apriori Viterbi Model for Prior Detection of Socio-Technical Attacks in a Social Network[1]

Socio-technical attack is an organized approach which is defined by the interaction among people through maltreatment of technology with some of the malicious intent to attack the social structure based on trust and faith. Awful advertisement over internet and mobile phones may defame a person, organization, group and brand value in society which may be proved to be fatal. People are always very sensitive towards their religion therefore mass spread of manipulated information against their religious belief may create pandemonium in the society and can be one of the reasons for social riots, political misbalance etc. Cyber-attack on water, electricity, finance, healthcare, food and transportation system are may create chaos in society within few minutes and may prove even more destructive than that of a bomb as it does not attack physically but it attacks on the faith and trust which is the basic pillar of our social structure.

2.2. Cyber-Attack Features for Detecting Cyber Threat Incidents from Online News[2]

There is large volume of data from the online news sources that are freely available which might contain valuable information. Data such as cyber-attacks news keep growing bigger and can be analyzed to gather informative insights of current situation. However, news is reported in many styles, added with the emerging of new cyber-attack and the ambiguous terms used have made the detection of the related news become more difficult. Thus, to handle these situations, the aim of this paper is to propose a scheme on detecting the related news about cyber-attacks.
The scheme starts with identifying the cyber-attack features which will be used to classify the cyber-attack news. The scheme also includes a machine learning approach using Conditional Random Field (CRF) classifier and Latent Semantic Analysis (LSA) for further analysis. The results from this research should help people by showing the actual picture of cyber-attack occurrences in our surrounding and give valuable information to public thus raising social awareness about cyber-attack activities

2.3. Adversarial Examples: Attacks and Defenses for Deep Learning[3]

It reviewed the recent findings of adversarial examples in DNNs. It investigated the existing methods for generating adversarial examples. A taxonomy of adversarial examples was proposed. It also explored the applications and countermeasures for adversarial examples. It attempted to cover the state-of-the-art studies for adversarial examples in the DL domain. Compared with recent work on adversarial examples, analyzed and discussed the current challenges and potential solutions in adversarial examples. However, deep neural networks (DNNs) have been recently found vulnerable to well-designed input samples called adversarial examples.
Adversarial perturbations are imperceptible to human but can easily fool DNNs in the testing/deploying stage. The vulnerability to adversarial examples becomes one of the major risks for applying DNNs in safety-critical environments.

Therefore, attacks and defenses on adversarial examples draw great attention review recent findings on adversarial examples for DNNs, summarize the methods for generating adversarial examples, and propose taxonomy of these methods. Under the taxonomy, applications for adversarial examples are investigated and further elaborate on counter measures for adversarial examples. In addition, three major challenges in adversarial examples and the potential solutions are discussed.

2.4. Distributed Secure Cooperative Control Under Denial-of-Service Attacks from Multiple Adversaries[4]

It investigated the distributed secure control of multiagent systems under DoS attacks. This focus on the investigation of a jointly adverse impact of distributed DoS attacks from multiple adversaries. In this scenario, two kinds of communication schemes, that is, sample-data and event-triggered communication schemes, have been discussed and, then, a fully distributed control protocol has been developed to guarantee satisfactory asymptotic consensus. Note that this protocol has strong robustness and high scalability. Its design does not involve any global information, and its efficiency has been proved. For the event-triggered case, two effective dynamical event conditions have been designed and implemented in a fully distributed way, and both of them have excluded Zeno behavior. Finally, a simulation example has been provided to verify the effectiveness of theoretical analysis. Our future research topics focus on fully distributed event/self-triggered control for linear/nonlinear multiagent systems to gain a better understanding of fully distributed control.

2.5. Modeling of Security Risk for Industrial Cyber-Physics System under Cyber-attacks[5]

Due to the insufficient awareness of decision makers on the security risks of industrial cyber-physical systems(ICPS) under cyber-attacks, it is difficult to take effective defensive measures according to the characteristics of different cyber-attacks in advance. To solve the above problem, it gives a qualitative analysis method of ICPS security risk from the perspective of defenders. The ICPS being attacked is modeled as a dynamic closed-loop fusion model where the mathematical models of the physical plant and the feedback controller are established. According to the disruption resources and detectability, a general security risk level model is further established to evaluate the security risk level of the system under attacks. The simulation experiments are conducted by using MATLAB to analyze the destructiveness and detectability of attacks, where the results show that the proposed qualitative analysis method can effectively describe the security risk under the cyber-attacks.

III.EXISTING SYSTEM

The Existing system is to create a contrastive self-supervised learning to the anomaly detection problem of attributed networks. CoLa ,is mainly consists of three components: contrastive instance pair sampling, GNN-based contrastive learning model, and multiround sampling-based anomaly score computation. Their model captures the relationship between each node and its neighbouring structure and uses an anomaly-related objective to train the contrastive learning model. We believe that the proposed framework opens a new opportunity to expand self-supervised learning and contrastive learning to increasingly graph anomaly detection applications. The multiround predicted scores by the contrastive learning model are further used to evaluate the abnormality of each node with statistical estimation. The training phase and the inference phase. In the training phase, the contrastive learning model is trained with sampled instance pairs in an unsupervised fashion. After that the anomaly score for each node is obtained in the inference phase.

3.1 Disadvantage

1. The performance is not good and its get complicated for other networks.
2. The performance metrics like recall F1 score and comparison of machine learning algorithm is not done.

IV.PROPOSED SYSTEM

The proposed model is to build a machine learning model for anomaly detection. Anomaly detection is an important technique for recognizing fraud activities, suspicious activities, network intrusion, and other abnormal events that may have great significance but are difficult to detect. The machine learning model is built by applying proper data science techniques like variable identification that is the dependent and independent variables. Then the visualization of the data is done to insights of the data .The model is build based on the previous dataset where the algorithm learn data and get trained different algorithms are used for better comparisons. The performance metrics are calculated and compared.

4.1. Advantages

1. The anomaly detection can be automated process using the machine learning.
2. Performance metric are compared in order to get better model.

V.SYSTEM ARCHITECTURE

This diagram represents how the attacks is Predicted, Where the past data is pre-processed and the accuracy is displayed

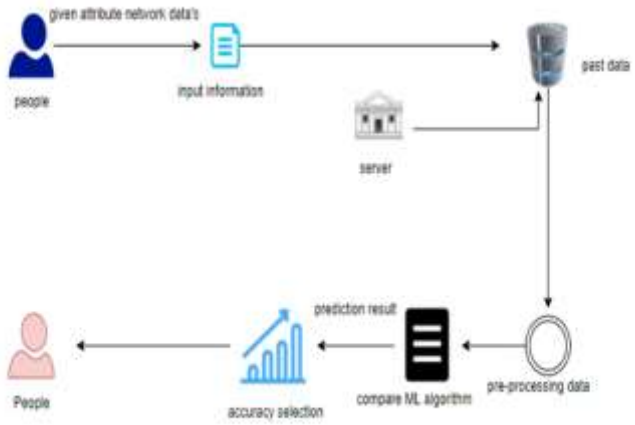


Fig.1 Architecture diagram

VI.MODULE DESCRIPTION

6.1. Variable Identification Process / data validation process

Validation techniques in machine learning are used to get the error rate of the Machine Learning (ML) model, which can be considered as close to the true error rate of the dataset. If the data volume is large enough to be representative of the population, you may not need the validation techniques. However, in real-world scenarios, to work with samples of data that may not be a true representative of the population of given dataset. To finding the missing value, duplicate value and description of data type whether it is float variable or integer. The sample of data used to provide an unbiased evaluation of a model fit on the training dataset while tuning model hyper parameters. The evaluation becomes more biased as skill on the validation dataset is incorporated into the model configuration. The validation set is used to evaluate a given model, but this is for frequent evaluation. It as machine learning engineers uses this data to fine-tune the model hyper parameters. Data collection, data analysis, and the process of addressing data content, quality, and structure can add up to a time -consuming to-do list. During the process of data identification, it helps to understand your data and its properties; this knowledge will help you choose which algorithm to use to build your model. For example, time series data can be analyzed by regression algorithms; classification algorithms can be used to analyze discrete data. (For example to show the data type format of given dataset)

Importing the library packages with loading given dataset. To analyzing the variable identification by data shape, data type and evaluating the missing values, duplicate values. A validation dataset is a sample of data held back from training your model that is used to give an estimate of model skill while tuning model's and procedures that you can use to make the best use of validation and test datasets when evaluating your models. Data visualization is an important skill in applied statistics and machine learning. Statistics does indeed focus on quantitative descriptions and estimations of data. Data visualization provides an important suite of tools for gaining a qualitative understanding. This can be helpful when exploring and getting to know a dataset and can help with identifying patterns, corrupt data, outliers, and much more.

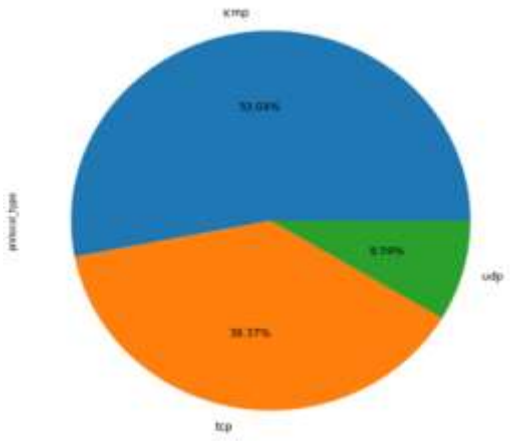


Fig.2 .Percentage level of protocol type

Sometimes data does not make sense until it can look at in a visual form, such as with charts and plots. Being able to quickly visualize of data samples and others is an important skill both in applied statistics and in applied machine learning. It will discover the many types of plots that you will need to know when visualizing data in Python and how to use them to better understand your own data.

(i)How to chart time series data with line plots and categorical quantities with bar charts.

(ii)How to summarize data distributions with histograms and box plots.

(iii)How to summarize the relationship between variables with scatter plots.

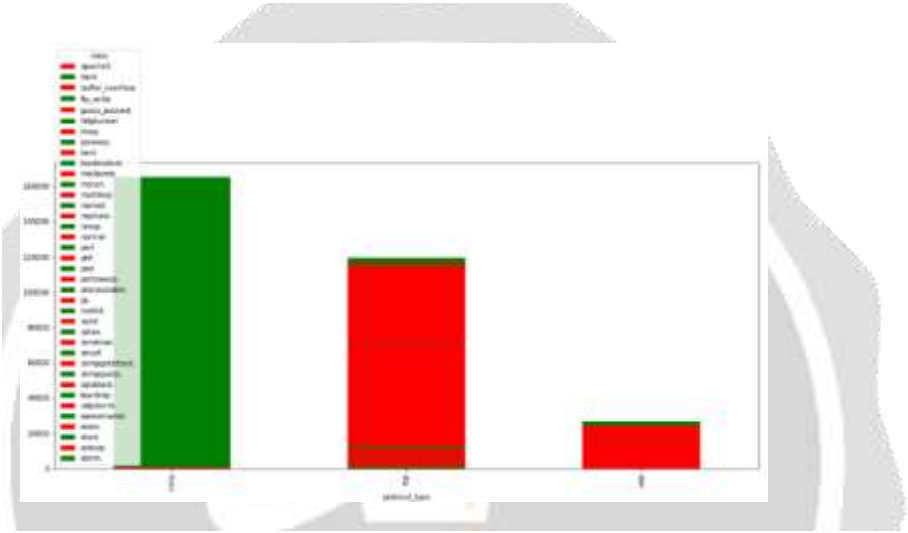


Fig.3.Comparison of service type and protocol type

The three steps involved in cross-validation are as follows:
1.Reserve some portion of sample data-set.
2.Using the rest data-set train the model.
3.Test the model using the reserve portion of the data-set.

6.2. Performance measurements of DoS attacks

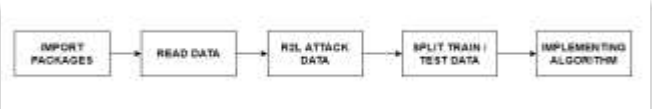
In computing, a denial-of-service attack (DoS attack) is a cyber-attack in which the perpetrator seeks to make a machine or network resource unavailable to its intended users by temporarily or indefinitely disrupting services of a host connected to the internet. Denial of service is typically accomplished by flooding the targeted machine or resource with superfluous requests in an attempt to overload systems and prevent some or all legitimate requests from being fulfilled. In a distributed denial-of-service attack (DDoS attack), the incoming traffic flooding the victim originates from many different sources. This effectively makes it impossible to stop the attack simply by blocking a single source. A DoS or DDoS attack is analogous to a group of people crowding the entry door of a shop, making it hard for legitimate customers to enter, disrupting trade.

A distributed denial-of-service (DDoS) is a large-scale DoS attack where the perpetrator uses more than one unique IP Address, often thousands of them. A distributed denial of service attack typically involves more than around 3–5 nodes on different networks; fewer nodes may qualify as a DoS attack but is not a DDoS attack. Since the incoming traffic flooding the victim originates from different sources, it may be impossible to stop the attack simply by using ingress filtering. It also makes

it difficult to distinguish legitimate user traffic from attack traffic when spread across multiple points of origin. As an alternative or augmentation of a DDoS, attacks may involve forging of IP sender addresses further complicating identifying and defeating the attack. An application layer DDoS attack (sometimes referred to as layer 7 DDoS attack) is a form of DDoS attack where attackers target application layer processes. The attack over-exercises specific functions or features of a website with the intention to disable those functions or features. This application-layer attack is different from an entire network attack, and is often used against financial institutions to distract IT and security personnel from security breaches.

6.3. Performance measurements of R2L attacks

Now-a-days, it is very important to maintain a high-level security to ensure safe and trusted communication of information between various organizations. But secured data communication over internet and any other network is always under threat of intrusions and misuses. To control these threats, recognition of attacks is critical matter. Probing, Denial of Service (DoS), Remote to User (R2L) attacks is some of the attacks which affect large number of computers in the world daily. Detection of these attacks and prevention of computers from it is a major research topic for researchers throughout the world.



6.4. Performance measurements of U2R attacks

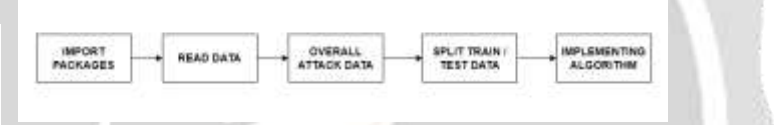
Remote to local attack (r2l) has been widely known to be launched by an attacker to gain unauthorized access to a victim machine in the entire network. Similarly, user to root attack (u2r) is usually launched for illegally obtaining the root's privileges when legally accessing a local machine. Buffer overflow is the most common of U2R attacks. This class begins by gaining access to a normal user while sniffing around for passwords to gain access as a root user to a computer resource. Detection of these attacks and prevention of computers from it is a major research topic for researchers throughout the world.

6.5. Performance measurements of Probe attacks

Probing attacks are an invasive method for bypassing security measures by observing the physical silicon implementation of a chip. As an invasive attack, one directly accesses the internal wires and connections of a targeted device and extracts sensitive information. A probe is an attack which is deliberately crafted so that its target detects and reports it with a recognizable “fingerprint” in the report. The attacker then uses the collaborative infrastructure to learn the detector's location and defensive capabilities from this report. This is an attack where the attacker attempts to gather information about the target machine or the network, to map out the network. Information about target may reveal useful information such as open ports, its IP address, hostname, and operating system. Network Probe is the ultimate network monitor and protocol analyzer to monitor network traffic in real-time, and will help you find the sources of any network slow-downs in a matter of seconds.

6.6 Performance measurements of overall network attacks

Increasingly, attacks are executed in multiple steps, making them harder to detect. Such complex attacks require that defenders recognize the separate stages of an attack, possibly carried out over a longer period, as belonging to the same attack. Complex attacks can be divided into exploration and exploitation phases. Exploration involves identifying vulnerabilities and scanning and testing a system. It is how an attacker gathers information about the system. Exploitation involves gaining and maintaining access. At this stage, the attacker applies the know-how gathered during the exploration stage. An example of a complex attack that combines exploration and exploitation is a sequence of a phishing attack, followed by an exfiltration attack. First, attackers will attempt to collect information on the organization they intend to attack, e.g., names of key employees. Then, they will craft a targeted phishing attack. The phishing attack allows the attackers to gain access to the user's system and install malware. The purpose of the malware could be to extract files from the user's machine or to use the user's machine as an attack vector to attack other machines in the organization's network. A phishing attack is usually carried out by sending an email purporting to come from a trusted source and tricking its receiver to click on a URL that results in installing malware on the user's system. This malware then creates a backdoor into the user's system for staging a more complex attack. Phishing attacks can be recognized both by the types of keywords used in the email (as with a spam email), as well as by the characteristics of URLs included in the message. Features that have been used successfully to detect phishing attacks include URLs that include IP addresses, the age of a linked-to domain, and a mismatch between anchor and text of a link.



6.7. GUI based prediction results of Network attacks

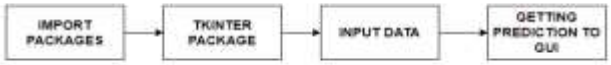
GUI means Graphical User Interface. It is the common user Interface that includes Graphical representation like buttons and icons, and communication can be performed by interacting with these icons rather than the usual text-based or command-based communication. A common example of a GUI is Microsoft operating systems.

The graphical user interface (GUI) is a form of user interface that allows users to interact with electronic devices through graphical icons and audio indicator such as primary notation, instead of text-based user interfaces, typed command labels or text navigation. GUIs were introduced in reaction to the perceived steep learning curve of command-line interfaces (CLIs) which require commands to be typed on a computer keyboard.

The actions in a GUI are usually performed through direct manipulation of the graphical elements. Beyond computers, GUIs are used in many handheld mobile devices such as MP3 players, portable media players, gaming devices, smartphones and smaller household, office and industrial controls. The term GUI tends not to be applied to other lower-display resolution types of interfaces, such as video games (where head-up display (HUD) is preferred), or not including flat screens, like volumetric displays because the term is restricted to the scope of two-dimensional display screens able to describe generic information, in the tradition of the computer science research at the Xerox Palo Alto Research Centre.

Graphical user interface (GUI) wrappers find a way around the command-line interface versions (CLI) of (typically) Linux and Unix-like software applications and their text-based user interfaces or typed command labels. While command-line or text-based applications allow users to run a program non-interactively, GUI wrappers atop them avoid the steep learning curve of the command-line, which requires commands to be typed on the keyboard. By starting a GUI

wrapper, users can intuitively interact with, start, stop, and change its working parameters, through graphical icons and visual indicators of a desktop environment, for example. Applications may also provide both interfaces, and when they do the GUI is usually a WIMP wrapper around the command-line version. This is especially common with applications designed for Unix-like operating systems. The latter used to be implemented first because it allowed the developers to focus exclusively on their product's functionality without bothering about interface details such as designing icons and placing buttons. Designing programs this way also allows users to run the program in a shell script.



ALGORITHMS

Logistic Regression:

It is a statistical method for analysing a data set in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes). The goal of logistic regression is to find the best fitting model to describe the relationship between the dichotomous characteristic of interest (dependent variable = response or outcome variable) and a set of independent (predictor or explanatory) variables. Logistic regression is a Machine Learning classification algorithm that is used to predict the probability of a categorical dependent variable. In logistic regression, the dependent variable is a binary variable that contains data coded as 1 (yes, success, etc.) or 0 (no, failure, etc.).

In other words, the logistic regression model predicts P(Y=1) as a function of X. Logistic regression Assumptions:

- Binary logistic regression requires the dependent variable to be binary.
- For a binary regression, the factor level 1 of the dependent variable should represent the desired outcome.
- Only the meaningful variables should be included.
- The independent variables should be independent of each other. That is, the model should have little.
- The independent variables are linearly related to the log odds.
- Logistic regression requires quite large sample sizes.

Decision tree:

It is one of the most powerful and popular algorithms. Decision-tree algorithm falls under the category of supervised learning algorithms. It works for both continuous as well as categorical output variables. Assumptions of Decision tree:

- At the beginning, we consider the whole training set as the root.
- Attributes are assumed to be categorical for information gain, attributes are assumed to be continuous.
- On the basis of attribute values records are distributed recursively.
- We use statistical methods for ordering attributes as root or internal node.

Decision tree builds classification or regression models in the form of a tree structure. It breaks down a data set into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. A decision node has two or more branches and a leaf node represents a classification or decision. The topmost decision node in a tree which corresponds to the best predictor called root node. Decision trees can handle both categorical and numerical data. Decision tree builds classification or regression models in the form of a tree structure. It utilizes an if-then rule set which is mutually exclusive and exhaustive for classification. The rules are learned sequentially using the training data one at a time. Each time a rule is learned, the tuples covered by the rules are removed.

This process is continued on the training set until meeting a termination condition. It is constructed in a top-down recursive divide-and-conquer manner. All the attributes should be categorical. Otherwise, they should be discretized in advance. Attributes in the top of the tree have more impact towards in the classification and they are identified using the information gain concept. A decision tree can be easily over-fitted generating too many branches and may reflect anomalies due to noise or outliers.

Random Forest:

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of over fitting to their training set. Random forest is a type of supervised machine learning algorithm based on ensemble learning. Ensemble learning is a type of learning where you join different types of algorithms or same algorithm multiple times to form a more powerful prediction model. The random forst algorithm combines multiple algorithms of the same type i.e., multiple decision *trees*, resulting in a *forest of trees*, hence the name "Random Forest". The random forest algorithm can be used for both regression and classification tasks.

The following are the basic steps involved in performing the random forest algorithm:

- Pick N random records from the dataset.
- Build a decision tree based on these N records.
- Choose the number of trees you want in your algorithm and repeat steps 1 and 2.
- In case of a regression problem, for a new record, each tree in the forest predicts a value for Y (output). The final value can be calculated by taking the average of all the values predicted by all the trees in forest. Or, in case of a classification problem, each tree in the forest predicts the category to which the new record belongs. Finally, the new record is assigned to the category that wins the majority vote.

Support Vector Machines:

A classifier that categorizes the data set by setting an optimal hyper plane between data. I chose this classifier as it is incredibly versatile in the number of different kernelling functions that can be applied and this model can yield a high predictability rate. Support Vector Machines are perhaps one of the most popular and talked about machine learning algorithms. They were extremely popular around the time they were developed in the 1990s and continue to be the go-to method for a high-performing algorithm with little tuning.

- How to disentangle the many names used to refer to support vector machines.
- The representation used by SVM when the model is actually stored on disk.
- How a learned SVM model representation can be used to make predictions for new data.
- How to learn an SVM model from training data.
- How to best prepare your data for the SVM algorithm.
- Where you might look to get more information on SVM.

VII. RESULT

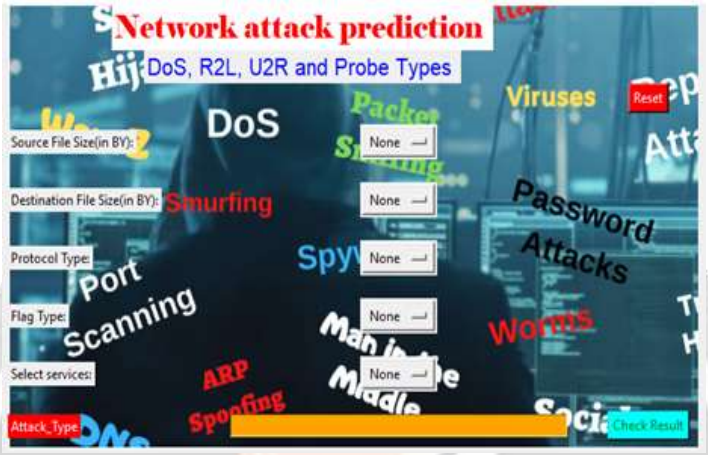


Fig.4.Display screen

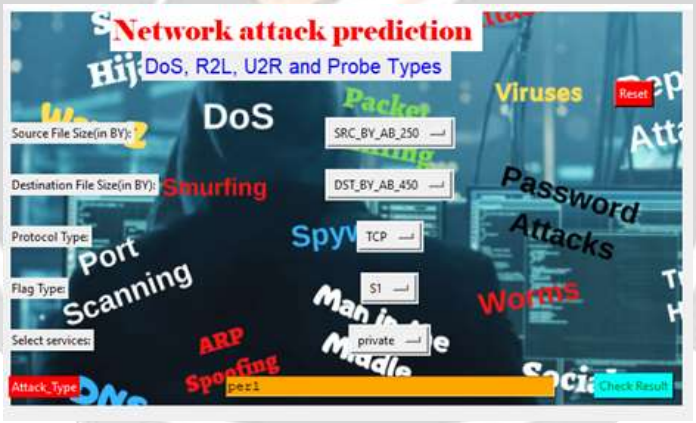


Fig.5. Display screen with predicted attack

VIII. CONCLUSION

The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation. The best accuracy on public test set is higher accuracy score will be find out by comparing each algorithm with type of all network attacks for future prediction results by finding best connections. This brings some of the following insights about diagnose the network attack of each new connection. To presented a prediction model with the aid of artificial intelligence to improve over human accuracy and provide with the scope of early detection. It can be inferred from

this model that; area analysis and use of machine learning technique is useful in developing prediction models that can helps to network sectors reduce the long process of diagnosis and eradicate any human error.

IX . REFERENCES

[1] Preetish Ranjan and Abhishek Vaish "Apriori viterbri model for prior detection of socio-Technical Attacks in social Network" In 2014 IEEE international conference.

[2] Mohamad Syahir Abdullah;Anazida Zainal;Mohd Aizaini Maarof;Mohamad Nizam Kassim "Cyber-Attack Features for Detecting Cyber Threat Incidents from Online News" In 2018 IEEE international conference

[3] Xiaoyong Yuan, Pan He, Qile Zhu and Xiaolin Li"Adversarial Examples: Attacks and Defenses for Deep Learning" In 2019 IEEE international conference

[4] Wenying Xu and Guoqiang Hu "Distributed Secure Cooperative Control Under Denial-of-Service Attacks From Multiple Adversaries" In 2019 IEEE international conference

[5] Ziwen Sun , Shuguo Zhang" Modeling of Security Risk for Industrial Cyber-Physics System under Cyber-attacks" In 2021 IEEE international conference

[6] Z. Liu, C. Chen, X. Yang, J. Zhou, X. Li, and L. Song, "Heterogeneous graph neural networks for malicious account detection," in Proc. 27th ACM Int. Conf. Inf. Knowl. Manage., Oct. 2018, pp. 2077–2085.

[7] L. Tang and H. Liu, "Relational learning via latent social dimensions," in Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD), 2009, pp. 817–826

.

[8] Y. Zhang et al., "Your style your identity: Leveraging writing and photography styles for drug trafficker identification in darknet markets over attributed heterogeneous information network," in Proc. World Wide Web Conf. (WWW), 2019, pp. 3448–3454.

[9] R . Ying, R. He, K. Chen, P. Eksombatchai, W. L. Hamilton, and J. Leskovec, "Graph convolutional neural networks for Web-scale recommender systems," in Proc. 24th ACM SIGKDD Int. Conf. Knowl.Discovery Data Mining, Jul. 2018, pp. 974–983.

[10] W. Fan et al., "Graph neural networks for social recommendation," in Proc. World Wide Web Conf. (WWW), 2019, pp. 417–426.

[11] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in Proc. Int. Conf. Learn. Represent., 2017,pp. 1–14.

[12] P. Veli˘ckovi´c, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," in Proc. Int. Conf. Learn.Represent., 2018, pp. 1–12.