INTERNATIONAL JOURNAL OF TECHO-ENGINEERING

ISSN : 2057-5688

CYBER ATTACKS IN VIRTUAL CONNECTIVITY AS AN EMERGING THREATS IN CYBER SECURITY

[1]K Manoj Kumar Reddy, [2]Ks Nikil, [3]T Surya, [4]Dr .Vijay Kumar

[123]Under Graduate Student, Department of Computer Science, Jain Global Campus, Kanakapura Taluk Ramanagara District, Karnataka, India -562112

[4]Assistant Professor, Department of Computer Science, Jain Global Campus, Kanakapura Taluk Ramanagara District, Karnataka, India -562112

Abstract:

Network intrusion detection systems (NIDS) is a crucial technology to ensure cyber security. Recently, machine-learning-based NIDSs are currently being researched in the context of various machine learning methods that are being proposed. However, the existing NIDSs aren't as robust in generality due to the fact that they are using specific characteristics derived by analyzing a few partial datasets. In addition the NIDS datasets exhibit an unbalanced ratio of the normal data and those with abnormal characteristics. This causes the minority class issue which must be solved in order to create solid and reliable NIDSs operating in different contexts. This paper proposes a new technique that utilizes service-aware partitioning of datasets that allows for the ability to handle large and growing networks with ease and allows the classifier improve its classification efficiency in terms of speed and accuracy. Our approach was tested using the Kyoto2016 dataset known as a dataset that has a high degree of imbalance and various classification algorithms and parameters to achieve the best performance . We then evaluated it against the most recent techniques. Our experiments revealed that our method can categorize network traffic quickly and accurately when dealing with massive data sets that are imbalanced. We conclude that this approach can solve the major issues associated with insufficient data sets that are used in modern machine-learning-based NID solutions.

Keywords: Cybersecurity, artificial intelligence, big data, switching virtualization, data security.

1. INTRODUCTION:

In the past two decades, the virtual realm known as Cyberspace has exploded in both reach and size. The world is experiencing an increase in Internet activities, bringing residents of remote areas closer via the numerous platforms that are connected to the Internet. The rapid growth of commercial Internet has led to the world becoming highly interconnected, facilitating the expansion of services such as social media, e-mail such as telemedicine, online retailing, and banking. These services aren't just accessible, but they're also quite accessible. The world's modern nations have moved the control of vital processes like manufacturing utilities transportation, banking, transportation, and communications to computers that are networked without weighing the potential risks that come with transition to a digital. The dramatic growth in access to information and increased connectivity has empowered people and organisations on the one hand and presented new challenges for government officials and the citizens to the contrary.

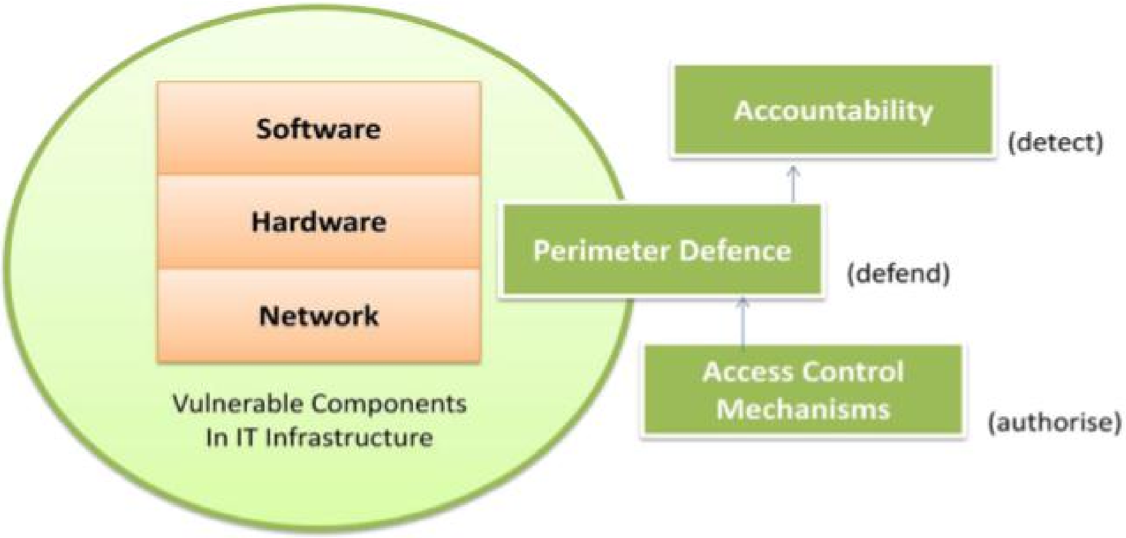


Fig 1 : Vulnerabilities and defense strategies in existing systems.

Cyberspace is a virtual area that is not bound by any borders and has evolved into a completely new realm, completely distinct from physical space that has borders. The cyberspace of today is now an invisible fabric that binds society as well as the entire around the world operate in the new cyberspace. Cyberspace technology is being considered to be of great importance for advancement by nations because of the benefits accrued. The rapid growth of networks has significantly contributed to the development of society. Open networks have helped facilitate the seamless circulation of information and sharing and created an environment for the development of startups and innovation with lower costs, and has significantly improved the health of people around the world, their prosperity and wealth. Thanks to the growing number of networks around the world, community and individuals are able to interact, socialize, and coordinate their lives via cyberspace. Cyberspace is rapidly becoming an integral part of the daily lives of the people living all over the world.

The cyber security issue is a major concern for every country and business across the globe and is not tied to any specific nation or company. Thus, cybersecurity issue is an issue that is global and calls for risk-based strategies as well as best practices, and international collaboration between both industries and government to tackle the issue. As with any other country it is the Indian Critical Infrastructure which includes banking and finance, energy as well as transportation, communications and the defense industrial base, is also dependent on cyberspace industrial control systems, as well as information technology that could be vulnerable to disruptions or misuse.

II RELATED WORK:

In the early days malware was created as experiments frequently to highlight security flaws or in some instances to showcase technical skills. Nowadays, malware is used mostly to steal sensitive financial, personal or business data for the benefit of other. For instance malware is commonly targeted at websites of corporations or government agencies to steal confidential information, or to interfere with their operations. In some cases malware can also be used against individuals in order to obtain personal information , such as Social Security numbers as well as number of credit cards. Due to the widespread availability of broadband Internet access, which is more affordable and more efficient malware has been created more often than just for protection of information, but for profit. As an example, the majority of the malware that is widely used have been created to control of computers used by users to exploit black markets, like sending spam emails or monitoring web surfing habits and then displaying unwelcome ads.

A large portion of the spiritual and material capital of nations is invested in this area, and a large portion of the economic and achievements of the citizens come from or have significant impact on the place (Amir and Givargis 2020). Also, various aspects of the lives of citizens are inextricably linked to this space. Any issues, instability, or challenges that occur in the space directly impact the various aspects of citizens' lives (Li and Li. 2020).

Network infrastructure and protocol vulnerabilities

The first network protocol was designed to work in a completely different environments on a smaller size and is not a good fit in the vast majority of scenarios the way it is utilized today. The weaknesses in network protocols can be difficult when administrators and users are not equipped with understanding of the network infrastructure. For example, administrators of the system don't use an the most efficient encryption method or apply the recommended patches in time, or do not remember the application of security policies or filters.

One of the most well-known network attacks is by exploiting the weaknesses of the widely used protocols for network communication, such as Internet Protocol (IP), Transmission Control Protocol (TCP) or Domain Name System (DNS). IP IP is the principal protocol for the layer of network. It is the primary protocol used to distribute packets between computers and routers in the network.

The initial IP protocol did not include any method to ensure the authenticity or privacy of the data being transmitted. This meant that data could be changed or intercepted as they were being transmitted through an unidentified networks among two different devices. To stop this from happening, IPSec was developed to ensure the security of IP traffic. For a long time, IPSec has been used as a key techniques for the development of the virtual private network (VPN) which provides an encrypted connection over the Internet between computers located in remote locations and an established network (i.e. an intranet for a company).

III DATASET AND METHODOLOGY:

The research study was carried out in both public as well as private companies responsible for the maintaining and protecting their information systems and infrastructure. It was conducted with the CIOs Security Officers (CISOs) as well as senior staff of the organizations were contacted to collect primary information. Survey instrument through questionnaires was employed to gather primary data, in addition to interviews. The Delphi method was employed to collect data that was not possible to obtain through the questionnaire method since the data sought was highly sensitive and, when leaks occur, they could cause damage to the image of the organization.

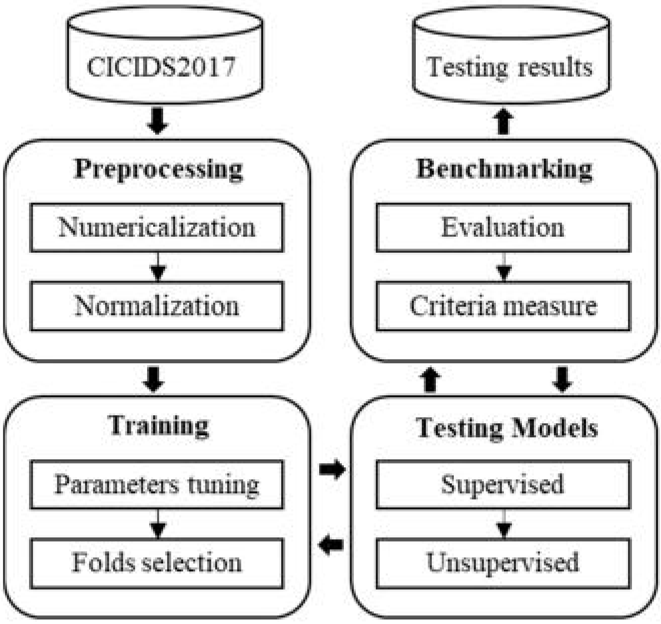
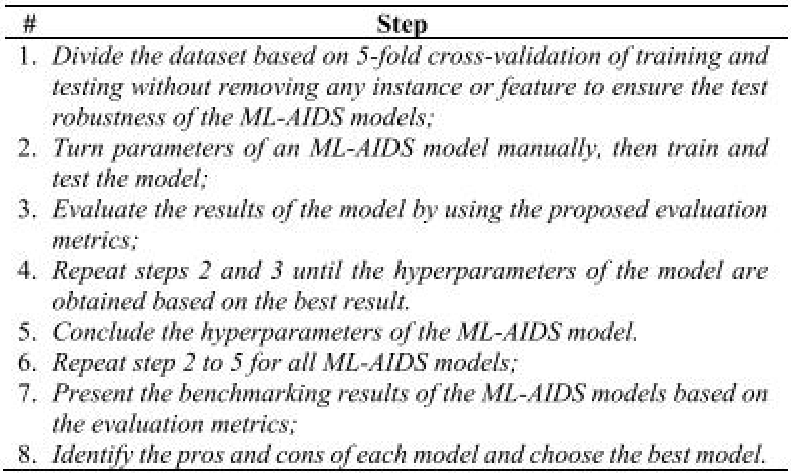


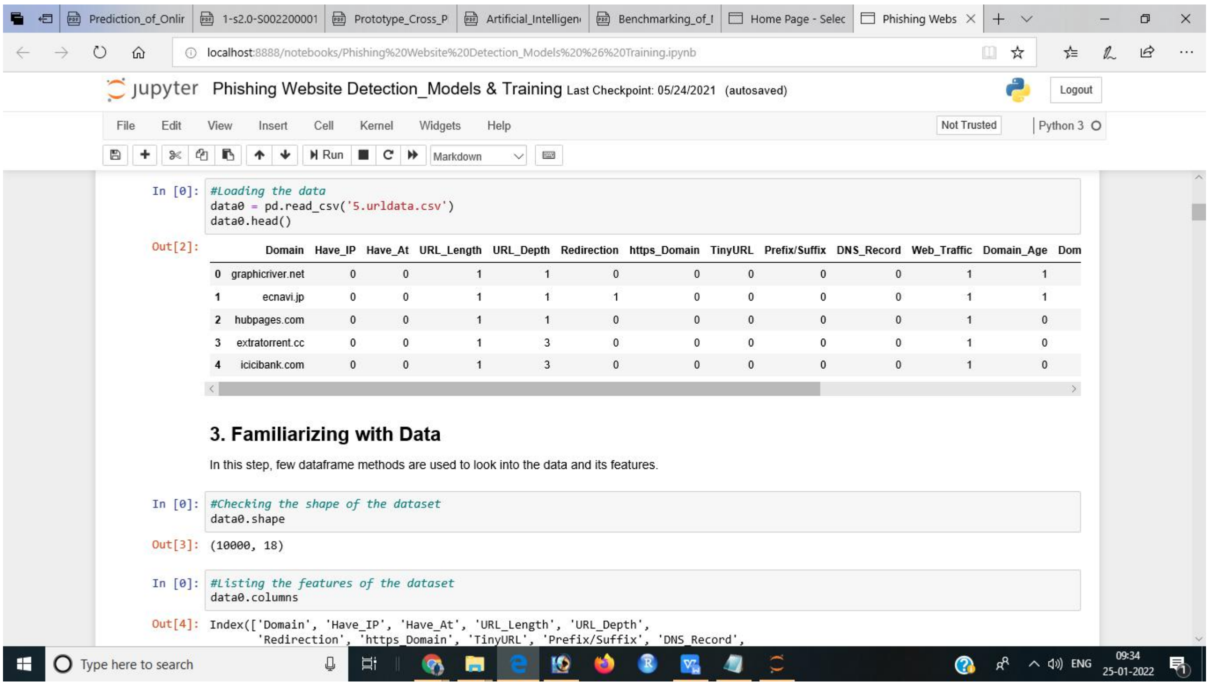
Fig : The benchmarking methodology



#	Step
1.	*Divide the dataset based on 5-fold cross-validation of training and testing without removing any instance or feature to ensure the test robustness of the ML-AIDS models;*
2.	*Turn parameters of an ML-AIDS model manually, then train and test the model;*
3.	*Evaluate the results of the model by using the proposed evaluation metrics;*
4.	*Repeat steps 2 and 3 until the hyperparameters of the model are obtained based on the best result.*
5.	*Conclude the hyperparameters of the ML-AIDS model.*
6.	*Repeat step 2 to 5 for all ML-AIDS models;*
7.	*Present the benchmarking results of the ML-AIDS models based on the evaluation metrics;*
8.	*Identify the pros and cons of each model and choose the best model.*

Fig : Benchmarking the ML-AIDS algorithm.

IV RESULTS AND DISCUSSION:



Phishing Website Detection_Models & Training Last Checkpoint: 05/24/2021 (autosaved)

```
In [0]: #Loading the data
        data0 = pd.read_csv('5.urldata.csv')
        data0.head()
```

Out[2]:

	Domain	Have_IP	Have_At	URL_Length	URL_Depth	Redirection	https_Domain	TinyURL	Prefix/Suffix	DNS_Record	Web_Traffic	Domain_Age	Dom
0	graphicriver.net	0	0	1	1	0	0	0	0	0	1	1	
1	ecnavi.jp	0	0	1	1	1	0	0	0	0	1	1	
2	hubpages.com	0	0	1	1	0	0	0	0	0	1	0	
3	extratorrent.cc	0	0	1	3	0	0	0	0	0	1	0	
4	icicibank.com	0	0	1	3	0	0	0	0	0	1	0	

3. Familiarizing with Data

In this step, few dataframe methods are used to look into the data and its features.

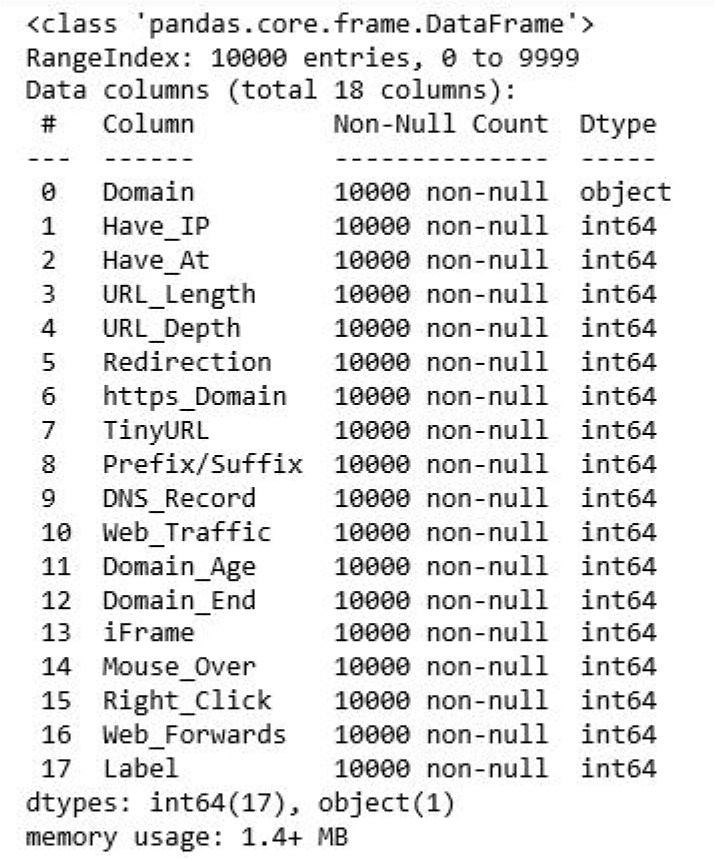
```
In [0]: #Checking the shape of the dataset
        data0.shape

Out[3]: (10000, 18)

In [0]: #Listing the features of the dataset
        data0.columns

Out[4]: Index(['Domain', 'Have_IP', 'Have_At', 'URL_Length', 'URL_Depth',
               'Redirection', 'https_Domain', 'TinyURL', 'Prefix/Suffix', 'DNS_Record',
```

Fig : The dataset loaded to the python platform



```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 18 columns):
 #   Column         Non-Null Count  Dtype 
---  ------         --------------  ----- 
 0   Domain         10000 non-null  object
 1   Have_IP        10000 non-null  int64 
 2   Have_At        10000 non-null  int64 
 3   URL_Length     10000 non-null  int64 
 4   URL_Depth      10000 non-null  int64 
 5   Redirection    10000 non-null  int64 
 6   https_Domain   10000 non-null  int64 
 7   TinyURL        10000 non-null  int64 
 8   Prefix/Suffix  10000 non-null  int64 
 9   DNS_Record     10000 non-null  int64 
 10  Web_Traffic    10000 non-null  int64 
 11  Domain_Age     10000 non-null  int64 
 12  Domain_End     10000 non-null  int64 
 13  iFrame         10000 non-null  int64 
 14  Mouse_Over     10000 non-null  int64 
 15  Right_Click    10000 non-null  int64 
 16  Web_Forwards   10000 non-null  int64 
 17  Label          10000 non-null  int64 
dtypes: int64(17), object(1)
memory usage: 1.4+ MB
```

Fig : Null value in the data

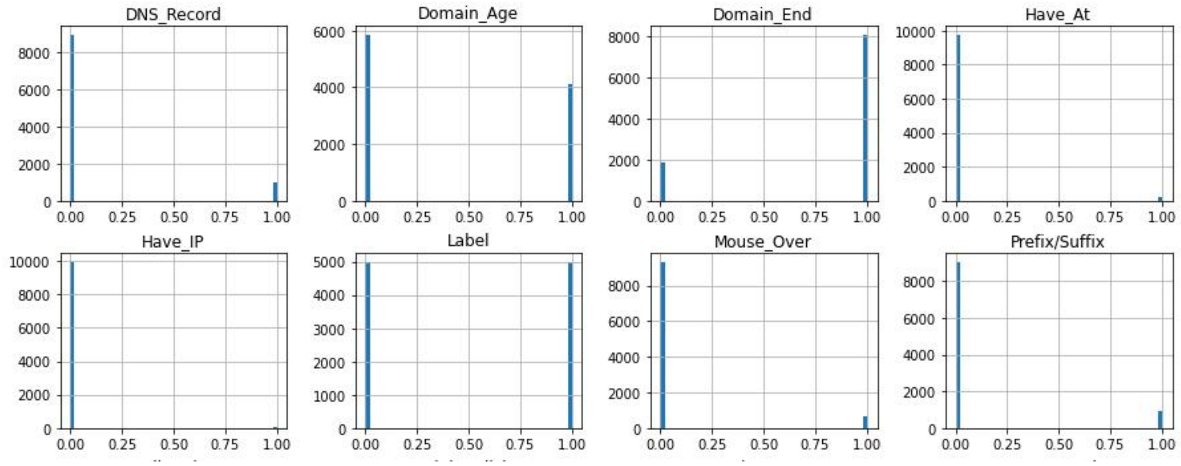


Fig : Features from the data

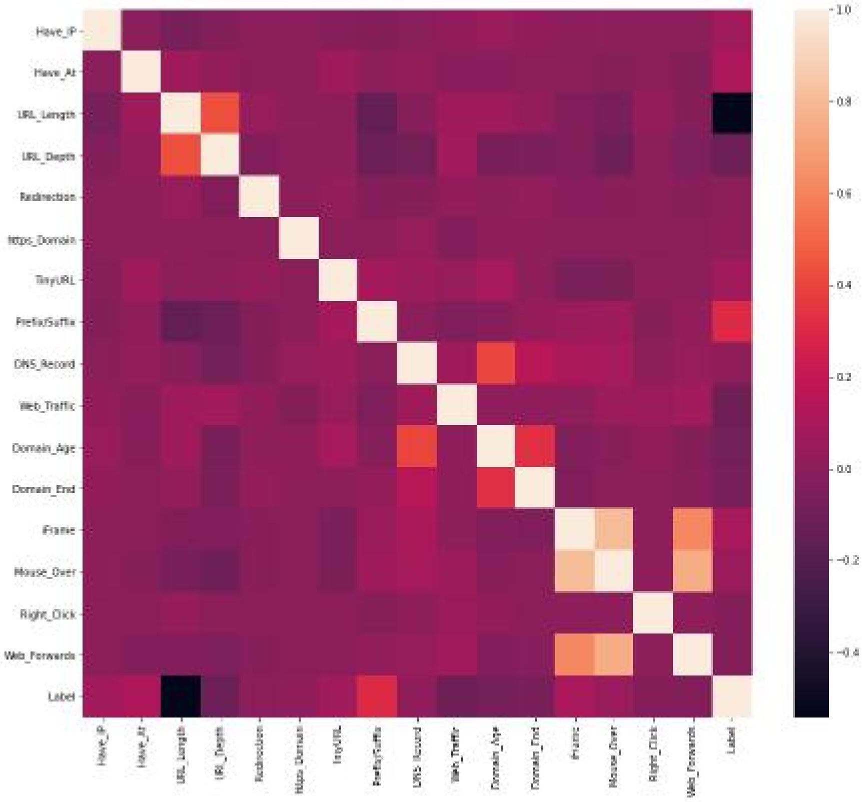


Fig : Correlation between the columns

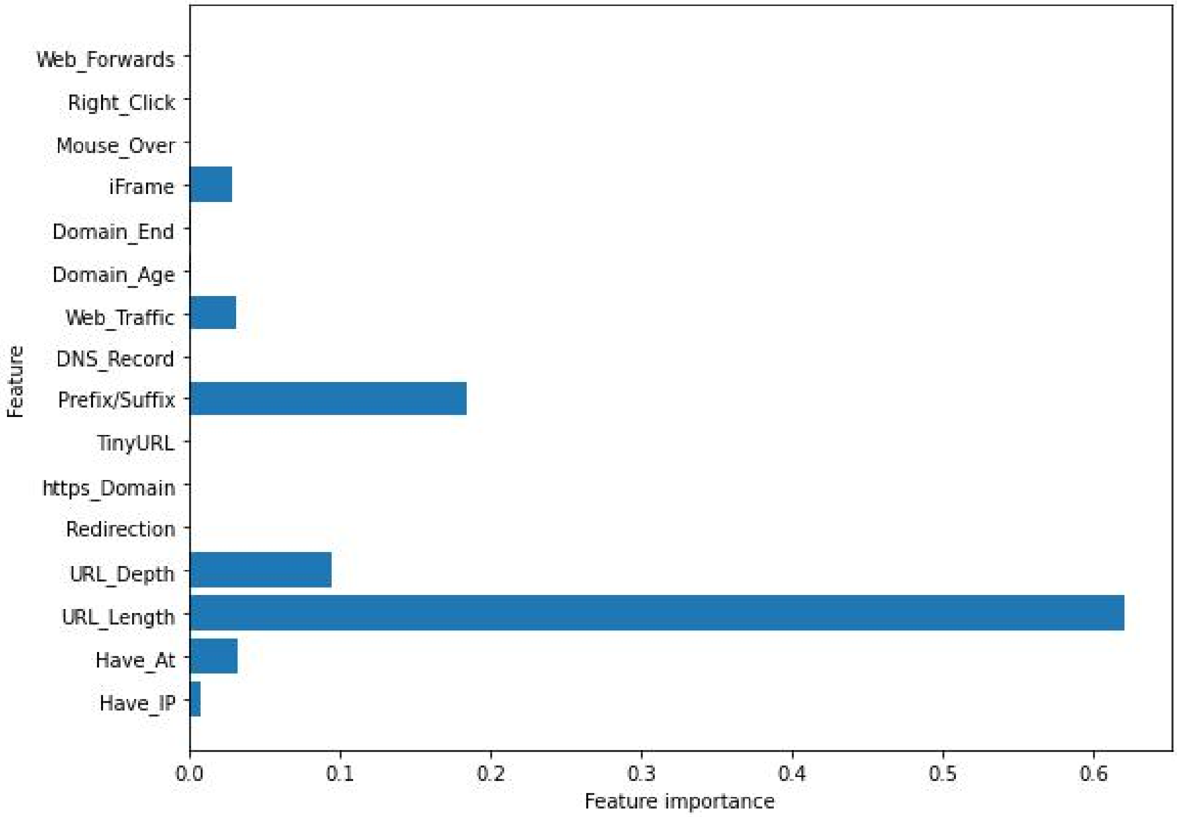
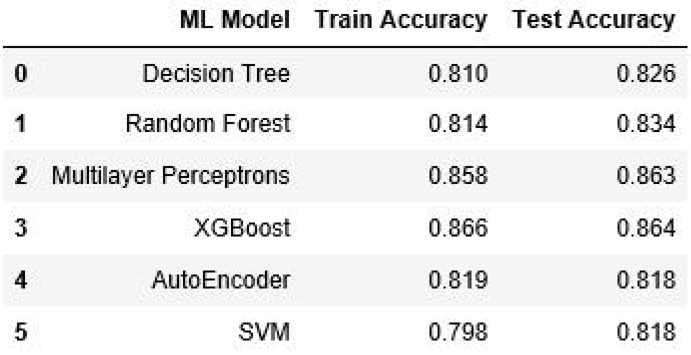
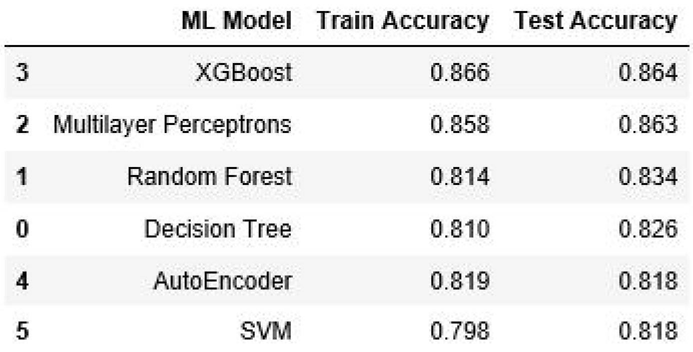


Fig : Features which are been used for the intrusion functionality



	ML Model	Train Accuracy	Test Accuracy
0	Decision Tree	0.810	0.826
1	Random Forest	0.814	0.834
2	Multilayer Perceptrons	0.858	0.863
3	XGBoost	0.866	0.864
4	AutoEncoder	0.819	0.818
5	SVM	0.798	0.818

Fig : Comparison of the model on detecting the final accuracy



	ML Model	Train Accuracy	Test Accuracy
3	XGBoost	0.866	0.864
2	Multilayer Perceptrons	0.858	0.863
1	Random Forest	0.814	0.834
0	Decision Tree	0.810	0.826
4	AutoEncoder	0.819	0.818
5	SVM	0.798	0.818

Fig : Result after the performance tuning factors

Everywhere in the world, there are attacks on critical infrastructure by non-state or state actors, as if these infrastructures aren't in use the essential services cannot be provided to citizens. Cyberattacks on critical infrastructure can threaten national security of the country and aid in helping attackers achieve their goal of causing instability in the nation. With a plethora of threats, ranging from nation-states and hackers to cybercriminals and terrorists Our infrastructure that is globally connected is in need of protection. The security measures implemented by responsible authorities are either insufficient or incorrect, which allows attackers to exploit weaknesses that exist. Also, it seems that the management of the organization isn't carrying out risk assessments, leaving gaps in addressing the vulnerability in the process or system. Risk assessment within an organisation can provide transparency to management on the current danger that must be reduced by distributing the necessary resources, including human resources and sufficient funding towards an IT department. These requirements must be updated regularly taking into consideration the constantly growing cyber-security threats as a result of the rapid rise in sophisticated malware, as well as the emergence of rogue characters.

VI CONCLUSION

It's a major problem for security experts in the field of information to follow the steps necessary to establish a security framework , and then implement policies to help organizations become more secure from ever-changing cyber-attacks. Every organization must have a strategy in place to regularly assess its security position and to continue to invest in and upgrade its equipment and software to mitigate the risk of cyber-attacks. They should also provide instruction in the area of cyber security to information security professionals , and run regular cyber awareness programs for their employees. Cyber security is going to pose a threat, and top management should be aware, implement strategies to mitigate the risk and be vigilant. This should be an integral part of the business's plan of action and management must be able to play a crucial role in establishing a cyber-security policy. Every organization should utilize its IT resources under the assumption that they are in a vulnerable state. This is why it should focus on response and detection controls as well as security controls for prevention. The framework was based is based on three pillars: People, Process & Technology which can be further developed with specific implementation and guidelines to each control that are listed. The framework could be applied across all kinds of basic infrastructures for information and can be tailored to critical infrastructures for information. Critical infrastructures for information should be shielded from the Internet to the extent that is possible . Very rigorous security measures must be in place to prevent any cyberattacks because if the services are inaccessible, it could compromise the security of the nation.

VII REFERENCES

[1] The Measurement of Scientific, Technological and Innovation ActivitiesGuidelines for Collecting and Reporting Data on Research and Experimental Development, Frascati, Italy.

[2] M. Rossberg and G. Schaefer, ''A survey on automatic configuration of virtual private networks,'' Comput. Netw., vol. 55, no. 8, pp. 1684–1699, Jun. 2011, doi: 10.1016/j.comnet.2011.01.003.

[3] A. Galán-Jiménez and J. Gazo-Cervero, ''Overlay networks: Overview, applications and challenges,'' Int. J. Comput. Sci. Netw. Secur., vol. 10, no. 12, pp. 40–49, 2010.

[4] K. Han, J. Liu, D. Yang, and Q. Yuan, ''The design of secure embedded VPN gateway,'' in Proc. IEEE Workshop Adv. Res. Technol. Ind. Appl. (WARTIA), Sep. 2014, pp. 350–353, doi: 10.1109/WARTIA.2014. 6976267.

[5] Y. Q. Fan, L. Lv, M. L. Liu, and F. Xie, ''Improvements based on the IPSec VPN secuirity,'' Adv. Mater. Res., vols. 756–759, pp. 2693–2697, Sep. 2013, doi: 10.4028/www.scientific.net/AMR.756-759.2693.

[6] T. Yu and S. Jajodia, Eds., Secure Data Management in Decentralized Systems, vol. 33. Boston, MA, USA: Springer, 2007.

[7] A. Ali and M. M. Afzal, ‘‘Database Security: Threats and Solutions,’’ Int. J. Eng. Invent., vol. 6, no. 2, pp. 25–27, 2017. [Online]. Available: http://www.ijeijournal.com/papers/Vol.6-Iss.2/D06022527.pdf

[8] J. O. Chan, ‘‘An architecture for big data analytics,’’ Commun. IIMA, vol. 13, no. 2, p. 1, 2013.

[9] M. Malik and T. Patel, ‘‘Database security–attacks and control methods,’’ Int. J. Inf. Sci. Techn., vol. 6, nos. 1–2, pp. 175–183, Mar. 2016, doi: 10.5121/ijist.2016.6218.

[10] H. Wirkuttis and N. Klein, ‘‘Artificial intelligence in cybersecurity,’’ Cyber., Intell. Secur., vol. 1, no. 1, pp. 103–119, 2017.

[11] H. M. Rajan, S. Dharani, and V. Sagar, ‘‘Artificial intelligence in cyber security—An investigation,’’ Int. Res. J. Comput. Sci., vol. 4, no. 9, pp. 28–30, 2017.

[12] K. Demertzis and L. Iliadis, ‘‘A bio-inspired hybrid artificial intelligence framework for cyber security,’’ in Computation, Cryptography, and Network Security. Cham, Switzerland: Springer, 2015, pp. 161–193.

[13] G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, and M. Marchetti, ‘‘On the effectiveness of machine and deep learning for cyber security,’’ in Proc. 10th Int. Conf. Cyber Conflict (CyCon), May 2018, pp. 371–390, doi: 10.23919/CYCON.2018.8405026.

[14] A. L. Buczak and E. Guven, ‘‘A survey of data mining and machine learning methods for cyber security intrusion detection,’’ IEEE Commun. Surveys Tuts., vol. 18, no. 2, pp. 1153–1176, 2nd Quart., 2016, doi: 10.1109/COMST.2015.2494502.

[15] J. Kim, J. Kim, H. L. Thi Thu, and H. Kim, ‘‘Long short term memory recurrent neural network classifier for intrusion detection,’’ in Proc. Int. Conf. Platform Technol. Service (PlatCon), Feb. 2016, pp. 1–5, doi: 10.1109/PlatCon.2016.7456805.

[16] P. Torres, C. Catania, S. Garcia, and C. G. Garino, ‘‘An analysis of recurrent neural networks for botnet detection behavior,’’ in Proc. IEEE Biennial Congr. Argentina (ARGENCON), Jun. 2016, pp. 1–6, doi: 10.1109/ARGENCON.2016.7585247.

[17] G. E. Dahl, J. W. Stokes, L. Deng, and D. Yu, ‘‘Large-scale malware classification using random projections and neural networks,’’ in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., May 2013, pp. 3422–3426, doi: 10.1109/ICASSP.2013.6638293.

[18] M. Kalash, M. Rochan, N. Mohammed, N. D. B. Bruce, Y. Wang, and F. Iqbal, ‘‘Malware classification with deep convolutional neural networks,’’inProc.9thIFIPInt.Conf.NewTechnol.,MobilitySecur.(NTMS), Feb. 2018, pp. 1–5, doi: 10.1109/NTMS.2018.8328749.

[19] G. D. Hill and X. J. A. Bellekens, ‘‘Deep learning based cryptographic primitive classification,’’ 2017, arXiv:1709.08385. [Online]. Available: http://arxiv.org/abs/1709.08385

[20] R. Pascanu, J. W. Stokes, H. Sanossian, M. Marinescu, and A. Thomas, ‘‘Malwareclassificationwithrecurrentnetworks,’’inProc.IEEEInt.Conf. Acoust., Speech Signal Process. (ICASSP), Apr. 2015, pp. 1916–1920, doi: 10.1109/ICASSP.2015.7178304.