

Cyber Crime and their Detection with Machine Learning: Comprehensive Study of Phishing and Cyberbullying

Atrakesh Pandey^{1*}, Neeraj Kumar Tiwari²

Abstract

In Today's world scenario cyber-crimes increased in large numbers, the criminals are now technologically advanced and very smart, and catching and preventing them is a tedious task, so this paper, analyzing of cyber-attacks cases of cyber bullying and dos attacks increased very much, this is used for stealing user data with the help of machine learning algorithm it can be countered. machines are more efficient, and accurate than humans, so we must take the help of machine learning techniques to tackle this problem, fraudster make target the innocent people so that they can harm them, but we need some methods so that we can counter them if we are unable to counter them, they keep on harming on the innocent peoples, people suffer from depression and many of them attempt suicide due to cyber-crimes like phishing and cyber bullying, many people are innocent and unaware about these frauds, these criminals needs to be stopped as soon as possible, if we don't have enough techniques to tackle them, then there willpower to harm innocent people will keep on increasing and they will continue to harm and fraud innocent people life and money, the most important work is to stop them so that they cannot continue to harm people in terms of monetary or mentally, so we must make use of modern techniques to fight with them, machine learning techniques are one of those techniques which help us to fight with the fraudster so that people can surf internet freely, many antivirus companies are taking care of that.

Keywords: Cyber crime, Cyber-attacks, dos attack, machine learning, techniques, counter, fraud, cyber bullying, phishing, antivirus

INTRODUCTION

In Today's world, criminals are Very Technological smart, it is very hard for law enforcement bodies

*Author for Correspondence

Atrakesh Pandey

E-mail: atrakesh@poornima.org

¹Assistant Professor, Department of Computer Science and Engineering, Poornima Institute of Engineering and Technology, ISI-2, RIICO Institutional Area, Goner Road, Sitapura, Jaipur, India

²Undergraduate Student, Department of Computer Science and Engineering, Poornima Institute of Engineering and Technology, ISI-2, RIICO Institutional Area, Goner Road, Sitapura, Jaipur, India

Received Date: September 02, 2022

Accepted Date: September 08, 2022

Published Date: September 25, 2022

Citation: Atrakesh Pandey, Neeraj Kumar Tiwari. Cyber Crime and their Detection with Machine Learning: Comprehensive Study of Phishing and Cyberbullying. National Journal of Cyber Security Law. 2022; 5(1): 1–6p.

to track them, in the present scenario cybercrime is increasing rapidly. The criminals are very advanced in wrongdoings. They target innocent people and trap them, normal human being who is not aware of fraud become victims of it. people don't know where to approach for help. They become the victim of online fraud, cybercrime, online cyberbullying.

Their sensitive details like credit card details and login credentials are stolen by hackers with help of machine learning tools we can analyze the data, and in the future we can prevent it, the intruders are much advanced than way before, so to tackle them we must take use of the modern techniques so that tackling them will become much easier, with modern and scientific techniques like machine learning, we can tackle those intruders.

ROLE OF MACHINE LEARNING IN CYBER CRIME DETECTION

Machine learning is a branch of artificial intelligence which helps in finding and predicting output based on previous data, it uses previous data for finding the output, it is very useful in cyber security as it finds pattern of crime, based on previous data it can detect wrong happening, it tracks similar attacks and responds in similar behavior, it detects the anomalies and tell us any wrong happening, without use of machine learning it is not possible for human (Figure 1) [1].

With help of machine learning, cybersecurity becomes affordable, more effective, we need lot of data to make our machine learning more effective, not only we need large quality data, we need quality data too.



Figure 1. Cyber crime detection [2].

Source: google images

TYPES OF CYBER ATTACKS AND THEIR PREVENTION THROUGH MACHINE LEARNING

There are various types of cyber-attacks performed by hacker. They are very much technological smart.

Phishing Attack

The hacker or criminal make fake website or webpage which looks exactly like original website, it is generally used for stealing precious data [3]. Hacker steals data, hackers uses different social engineering techniques to trap the victims, victims falls in trap of the hacker and give important details to them (Figure 2).

From machine learning we can able to make project which detect spam links, it uses the previous data of spam links and identifies them, for that we use logistic regression, where we fit logistic regression , creating confusion matrix of predicted values and real value, with support vector machine svm with a rbf kernel and using grid searchcv to predict best parameters for svm was a really good choice, and fitting the model with predicted best parameters, we can also get good results using random forest classification, it gives good accuracy around 92%.

Cyber Bullying

Nowadays many people use social media, with large number of growth in users [4], The case of cyber bullying increased exponentially, bullying is the most difficult and painful experience of people, cyber bullying is becoming more significant to children, teenagers and women, children are innocent, and females are more sociable, they easily get noticed by hackers who bullies them, Facebook and twitter have some highest rate of cyber bullying, they get hatred from internet, even you tuber gone through cyber bullying in their comments section (Figures 3 and 4).

They go through mental harassment , many people committed suicide after getting affected from becoming victim from cyber fraud, many people gets threat from cyber bullies' through personal messages, tweets and youtube comments, we have seen in recent times that after social media getting

famous nowadays , many people are now suffering from the Cyber Bullying, most of the cases are unreported to the authorities, and many people are unaware about conditions and they have zero idea how to approach this situation.



Figure 2. Phishing attack [2].

Source: google images

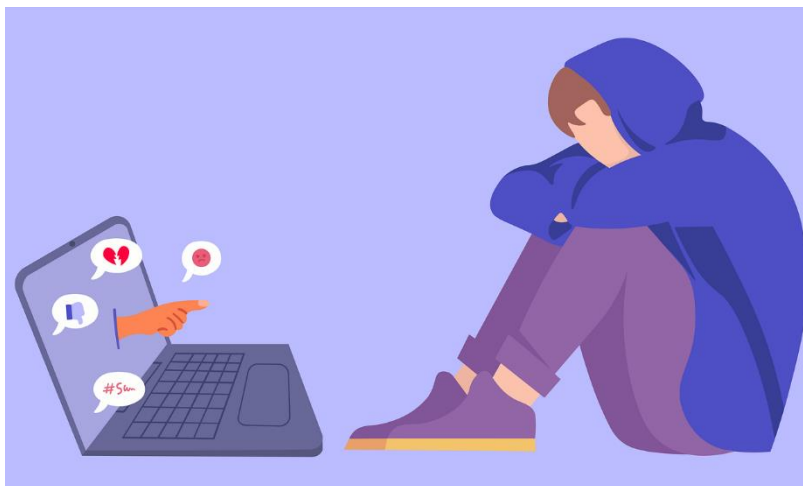


Figure 3. Cyber bullying [2].

Source: google images



Figure 4. Threat regarding cyber bullying [2].

Source: google images

This results in depression, self-harm thoughts and lack of socialness in real life, they feel lonely and self-harm thoughts are prevalent to them, we must detect these types of harm done by those bullies' and take necessary action.

With machine learning algorithm we can prevent cyber bullying, based on previous data youtube, facebook and other social media uses spam filter, which censors bad word.

The dataset of bad tweets/comments are capable of handling, The good news is that with help of modern advancements in big data have helped create sophisticated AI speech recognition systems that can identify nuances in written speech [5].

These machine learning tools detect abusive and damaging online behaviour and alert authorities. This not only helps ensure the safety of likely victims but also helps rehabilitate potential cyber bullies. Even natural language processing is helpful in finding bad words.

ADVANTAGES OF MACHINE LEARNING IN CYBER SECURITY

Cyber Threat Identification

Machine learning can identifies the threat and takes action automatically, it is helpful in large organization where it is not possible for human to track down those datas, human accuracy is not efficient in that of large dataset and human cannot able to look on large data, so we require computation power of machine to tackle that [6].

AI based Antivirus

It is recommended that we should always use antivirus and update it time to time, Nowadays antivirus is integrated with machine learning and they catches and detect any abnormal behaviour, it also identifies threat based on previous findings and datas.

User Behaviour Modelling

Hackers may steal user information and login or make transaction through credit card, even antivirus cannot detect such pattern because credentials are right [7], here comes role of machine learning, machine learning identifies user behaviour and identifies his login and logout pattern, if it detect any abnormal pattern it notifies the authorities.

Fighting Artificial Intelligence Threat

with technological advancement hackers have created artificial intelligence threats such as botnet, who attacks on user system, it is very important to fight against them, so machine learning models help us on identifying those threat and fighting against them.

Email and Monitoring

We should always monitors are email as they are high chance of phishing is there, hackers can steal the data from it, So it is very important for people for monitoring their email, here machine learning comes to help, it figures out spam email and fake links and notifies user [8].

Bad Word Filtering

With machine learning, we can filter out bad words in comments and tweets, which help us in detecting cyber attack.

CHALLENGES OF MACHINE LEARNING IN CYBER SECURITY

Hackers are Advanced and Tech Savvy

Hacker can now easily bypass the security and they can find loopholes, even machine learning models can made fools by this, they are becoming advance [9].

Hacker can Hide Themselves

Hacker use VIRTUAL private network and proxy to hide themselves at the time attacks, with these practices hacker remain anonymous and non trackable.

False Detection

Many times machine Learning gives wrong results, which is not desirable, we must make a system which accurate results.

Expensive

Artificial intelligence is very expensive technology, it is not economical that normal person can use it, we must make it economical so that even poor and middle class people can afford it [10].

Cyber Threats Evolving

Threats are evolving day by day, it is difficult for old models to detecting and preventing attack, new types of cyber threats are now getting prevalent.

Still Needs Lots of Human Effort

We still require lots of manpower, we cannot solely rely on machine learning techniques to safe us from cyber threats [7].

CONCLUSION

With many machine learning techniques we can prevent cyber attacks efficiently and effectively, still we need more improvement in the algorithms to make it more effective. This will help common human to be safe from cyber threats.

It will prevent hacker and intruder to make loss of common people, people can be safe from cheating and cyber abuse with help of machine learning techniques. This gives suggestions to the advancement of cybercrime identification model in which it is skilled to distinguish cybercrime in contrast with the current methods successfully.

This will help people from getting safe from monetary and emotionally loss, and people will be safe in our society.

FUTURE SCOPE

Machine learning is very new technique in cyber security, it will be very beneficial in coming future to detecting cyber crime, it will help us in stopping the phishing attack and cyber bullying, it will ease cyber crime detection, as it is very good in identifying crime, cases of cyber bullying „phishing can be minimized, the antivirus of mobile and computers should incorporate new machine learning techniques in their new updates so that the security can be strengthened and frauds can be minimized, the cloud should also incorporate these techniques so that they can tackle them , there should be system to rollback the transaction if the system detects the cyber crime with machine learning techniques. The social media sites must will make some system with machine learning techniques which catches and detect those harmful words which can cause mental disturbance, depression and harm to the individuals, the military system will also adopt these technologies to prevent leakage of sensitive informations of the country.

REFERENCES

1. Top 5 Applications of Machine Learning in Cyber Security - GeeksforGeeks [Internet]. GeeksforGeeks. 2020 [cited 2022 Sep 20]. Available from: <https://www.geeksforgeeks.org/top-5-applications-of-machine-learning-in-cyber-security/>
2. Peterson R. The Pros and Cons of Using AI in Cybersecurity [Internet]. Anonymania. Anonymania; 2019 [cited 2022 Sep 20]. Available from: <https://anonymania.com/artificial-intelligence-cybersecurity-pros-cons/>.

3. Dobhal-abhishek. GitHub - dobhal-abhishek/detecting-phishing-websites: detecting phishing websites using machine learning [Internet]. GitHub. 2022 [cited 2022 Sep 20]. Available from: <https://github.com/dobhal-abhishek/detecting-phishing-websites>
4. K. Reynolds, A. Kontostathis, and L. Edwards, "Using machine learning to detect cyberbullying," in 2011 10th International Conference on Machine learning and applications and workshops, 2011, vol. 2, pp. 241-244: IEEE
5. The 8 Most Common Cybersecurity Weaknesses to Watch for in Small Businesses [Internet]. ISACA. 2020 [cited 2022 Sep 20]. Available from: <https://www.isaca.org/resources/news-and-trends/isaca-now-blog/2020/the-8-most-common-cybersecurity-weaknesses-to-watch-for-in-small-businesses>
6. M. A. Al-garadi, K. D. Varathan, and S. D. J. C. i. H. B. Ravana, "Cybercrime detection in online communications: The experimental case of cyberbullying detection in the Twitter network," vol. 63, pp. 433-443, 2016.
7. Pandey, A., & Jain, R. (2023). 1–4D Protein Structures Prediction Using Machine Learning and Deep Learning from Amino Acid Sequences. In Proceedings of the Third International Conference on Information Management and Machine Intelligence (pp. 615-621). Springer, Singapore.
8. Haidar, M. Chamoun, and A. Serhrouchni, "Multilingual cyberbullying detection system: Detecting cyberbullying in Arabic content," in 2017 1st Cyber Security in Networking Conference (CSNet), 2017, pp. 1-8: IEEE.
9. Can data analytics and machine learning prevent cyberbullying? [Internet]. Selerity. 2021 [cited 2022 Sep 20]. Available from: <https://seleritysas.com/blog/2021/06/11/can-data-analytics-and-machine-learning-prevent-cyberbullying/>
10. Pandey, A., Pant, D., & Gupta, K. K. (2013). A Novel Approach on Color Image Refocusing and Defocusing. International Journal of Computer Applications, 73(3).