*Machine Learning Classifier for Mobile Malware Detection*

Dissertation Chapter # 1: Introduction

Student Name: Harjinder

# Chapter # 1 Introduction

Machine learning model used to predict detect any vulnerability attack on mobile device, (Narudin, F.A., Feizollah, 2016) android phone various malware has been detected due to running of unsafe application. Using the deep learning mode application in the android phone which essentially required to safe inside and outside mobile application. Custom neural network deep learning algorithm classify vulnerability and malware attack and it produces the high accuracy, deep learning model scans 1000 of running applications in second and if any malware detected, it fix it automatically. (Mercaldo, F. Santone, A., 2020) deep learning model learns itself the subtle cues that malware stand out from safe applications, this removes human bias and preconceptions from the decision process, the networks can find pattern that are not obvious to human analysts, which makes harder for malicious apps to evade detection, for each app, the evidence from the machine code can be combined with additional information such as the permissions required by the app, its API calls, and its manifest to reach and even more accurate to reach resilient decision. Machine learning neural network can stand alone on the end user’s mobile device so it does not need an internet network connection to work, deep learning model approach refine and classify the applications more efficiently to identify and predict the malware from the applications, malware could be able to harm the user applications as well hardware storage system, and its robust against the methods used to obscure and hide malicious code in the form of encryption. Machine learning Technology is platform in which malware virus can be detected and fixed on scanning the android phone network. Deep learning model has been working on prototype methods because various industrial IOT applications essentially improved the security and data protection. It enables enterprise user secure IOT devices across the network and its able to scan the individual mobile application as well, finally cloud application also scan during the uploading and downloading from the cloud computing. Android phone users which has total of 72.2% of market share in world mobile survey in recent research 2020. Hackers wants to attack android phone to steal and capture important data & information from the cell phone, with various method such as credential theft, surveillance, malicious advertising, so machine learning algorithm effectively works well to detect and identify the threats and malware attack in the mobile phone. This research going to provides the brief & systematic approach to detect and defined the android mobile user malware attack system. (Liu, Y., 2020) experiences landing of machine learning technique onto the market scale using the mobile malware detection & prevention. The better understanding of mobile android phone implementing by the various machine learning techniques & advanced machine learning algorithm. Failure of android devices due to instability of android phone, the mobile user increasing day by day due to emerging technology of software application. In previous research the research gap exist due to the insecure mobile applications version and insecure third party applications and non-reliable applications which does not provides the security features. The implementation of T-market for over one year using single commodity server to vet of 10K apps every day. The evaluation result presented in this research archive the overall precision call score of 98% accuracy by implementing advance machine learning methods such as genetic algorithm, support vector machine, naïve bays & decision tree classifier technique. By selecting the feature selection of machine learning of training and testing the dataset, the feature selection is main component of removing malware from the android phone and use of machine learning techniques to fix them by scanning the mobile applications in real time, Google provides the best state of the art security feature to the android phone because of advance machine learning methods which was essentially used by the operating to system to scan the android applications. The scanning applications has been detected & predicted which essentially helpful for the user to secure the android applications. Android user used various third party application which was not secured due to less security features and the security feature does not support third party applications so the previous research address them to fix the android malware by implementing the machine learning algorithm to predict and prevent the machine learning algorithms. The machine learning prediction technique briefly describe in literature review section and methods chapter discussed the research methods in which the problem has been solved.

## Problem Statement

The research problem is highlight the mobile malware, which needs to remove and fixed from the android phone. Various machine learning methods and advance programming techniques has been used in previous research, this research going to present and merge them the existing and new machine learning technique, literature review section solved the android malware problem by addressing the research question. In chapter 4 the result & analysis part combine the previous outcomes and new advance improved machine learning technique has been addressed briefly. The common android phone problem exist in previous mobile android version which was known as malware as follows:

1. Adware
2. Backdoor
3. Ransomware
4. Riskware
5. Scareware
6. Spyware
7. Trojan
8. Trojan-sms
9. Trojan spy
10. PUA potentially unwanted application
11. File infector

Most common type of malware is known as adware which disturbs the user by sending unwanted spam advertisements, which might be attracted the user and redirect them to capture sensitive information from mobile phone. The malware detection &prevention issues has been resolved in chapter 2 and chapter 4 by using the advance machine learning methods. Previous releases of android version contains security issues & malware issues, but the current research presented the state of the art security problem which ask security permission from the user to access any mobile app from other user, if user ask for any app to access from other user app android grant the permission if user grant them to executes. Android phone is commonly used by user with with market shares of 86%, because android phone choice of every community including student and business men.

## Research Gap

Negativity & issues has been found in previous android operating system releases, various security issues has been exist in android version 4, 5 & 6. Google updated the security features in android 11 and 12 version. It is noted that in old android version there is lack of screen touch issues found and Google map navigation issue exist. Google search engine not displaying on android screen in previous android releases. Research gap exist from 2011-2017 due to instability of android operating system, but after release of updated android version 10 and 11 the security features and malware scanning feature has been updated by Google. This research address the research gap from 2011 to 2017 in literature review section, negativity and common problems & security issues has been address in literature review chapter and chapter 3 address the research methods by using the advance machine learning technique.

## Aims & Objective of Thesis

Aim of this research to produce machine learning algorithm for android mobile operating system to detect malware in the operating system. By implementing bays network and random forest classifier algorithm which produced the highest accuracy rates of malware detection.

Objectives of this research is to plan latest machine learning algorithm which is based on malware attack detection method. This research objectives is going to address the previous android phone negativity and latest android phone positive points. The negativity issues has been addressed in literature review section, and solution has been addressed from the Google android latest version with enhance security feature.

## Challenges & Contribution

(Sahib, N.M., 2021) mobile phone users increasing on daily basis due to increase demands of online communication and online business, various educational institutes planned online lecture session for distance learning and training. Challenge remain exist due to various vulnerable attack happened on mobile phone devices, android phone security feature updated but the challenge and issue remain questionable to get the secret info from the mobile device, mobile apps provides the great convenient feature to user for instant access of weather update, sports update and ongoing communication over the internet by sending instant messaging.

## Machine learning challenges

Machine learning programming facing the following challenges as follows

1. Deprived superiority of data
2. Underfitting of training data
3. Overfitting of training data
4. Absence of training data by using the differ sample size
5. Immaterial feature selection from dataset
6. Trash in and trash out info

Non representative of training dataset, the common issues mainly deal by the machine learning contains to select & predict the irrelevant machine learning dataset. Select the more advanced model one and more parameter selection. Reducing the number of training dataset attributes from the given dataset. The major issue in machine by selecting the dataset and cleaning the dataset, because various data scientist, the solution has been provided by the Kaggle and machine learning UCI repository.

# References

Gong, L., Li, Z., Qian, F., Zhang, Z., Chen, Q.A., Qian, Z., Lin, H. and Liu, Y., 2020, April. Experiences of landing machine learning onto market-scale mobile malware detection. In *Proceedings of the Fifteenth European Conference on Computer Systems* (pp. 1-14).

Wei, S., Zhang, Z., Li, S. and Jiang, P., 2021. Calibrating Network Traffic with One-Dimensional Convolutional Neural Network with Autoencoder and Independent Recurrent Neural Network for Mobile Malware Detection. Security and Communication Networks, 2021.

Omer, M.A., Zeebaree, S.R., Sadeeq, M.A., Salim, B.W., x Mohsin, S., Rashid, Z.N. and Haji, L.M., 2021. Efficiency of malware detection in android system: A survey. Asian Journal of Research in Computer Science, pp.59-69.

Casolare, R., De Dominicis, C., Iadarola, G., Martinelli, F., Mercaldo, F. and Santone, A., 2021. Dynamic Mobile Malware Detection through System Call-based Image representation. J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl., 12(1), pp.44-63.

Kouliaridis, V., Barmpatsalou, K., Kambourakis, G. and Chen, S., 2020.A survey on mobile malware detection techniques. IEICE Transactions on Information and Systems, 103(2), pp.204-211.