WWW.IJRPR.COM

**International Journal of Research Publication and Reviews**

Journal homepage: www.ijrpr.com ISSN 2582-7421

# Evaluation of Malware Using Machine Learning

***Prof. Rupali Panpaliya[1], Dheeraj Paropate[2], Devendra Sonawane[3], Pooja Domate[4], Vaishnavi Budruk[5]***

*Profrssor, Department Of Computer, Siddhant College Of Engineering, Maharashtra, India[1]*
*Student, Department Of Computer, Siddhant College Of Engineering, Maharashtra, India[2345]*

A B S T R A C T

*When users enter their credentials in a public place, they risk adversaries gaining access to their credentials. 2 An attacker can get a password by observing of recording a person's authentication session. Shoulder-surfing is a well-known danger, and it's especially risky when in public venues, you must authenticate yourself. The user's main line of defense against shoulder-surfing was until recently was his own alertness. The user is protected from shoulder-surfing by 2a password authentication system that is resistant to intrusion detection. Because the user is never obliged to click directly on password symbols, it allows users to authenticate in insecure locations by typing their password graphically. The usability testing of this technique indicated that novice users were able to type their graphical password accurately and remember it 3over a period of time 1However, the protection against shoulder-surfing comes at the cost of a longer identification process.*

**Keywords**:malware prediction, machine learning algorithm, neural network, k-fold, decision tree, Microsoft malware dataset.

## 1. Introduction

A malware is a malicious code. Malware can be considered as an entity in that easily added new feature and that enhance its side effects in the form of various attacks. These malwares can be dangerous with all side effects like break the system, corrupt data, etc. IOT Applications have led to the development of modern concept of the information society. However, security concerns pose a major challenge in realizing the benefits of industrial revolution as cyber-criminals attack individual PC's and net-works for stealing confidential data for financial gains and causing DOS attacks to systems. Such attackers make use of malicious software or malware to cause serious threats and vulnerabilities of system. Malware is a computer software that is designed to harm the operating system. Deep Learning is an artificial intelligence function that mimics the human brain's functions in data processing and pattern creation for decision-making. Deep Learning) is a form of Machine Learning in Ai Technology (AI) that uses neural networks to learn unsupervised from unorganized and unlabeled data. Deep Neural Learning or Convolutional Neural Network are other terms for the same thing.

### *1.1. Construction of references*

*"Risk Prediction of Malware Victimization Based on User Behavior-2015"*

Understanding what types of users and usage are more prone to malware infestations is critical if we are to develop sufficient strategies for dealing with and minimizing the effects of computer crime in all of its manifestations. Real-time usage data is thus critical for making better evidence-based judgments that will increase user security. To that purpose, we performed a 4-month field study with 50 individuals, collecting real-time data by detecting potential illnesses and obtaining information on user behavior. We present a first attempt at forecasting the likelihood of malware victimization based on user behavior in this research. Using neural networks, we created a predictive model with an accuracy of up to 80 percent.
"Multilevel Permission Extraction in Android Applications for Malware Detection-2019"

With the widespread use of Android applications in security-critical circumstances, an increasing amount of Android malware has been detected. Existing malware detection research fails to automatically learn effective feature interactions, which are important to the operation of many prediction models. In this work, we offer Multilevel Permission Extraction, an approach to automatically identifying permission interactions that are effective in discriminating between dangerous and benign programs, in able to locate malware rapidly and reliably. The gathered information is can then use by machine learning-based classification algorithms to classify dangerous and benign programs. We test our method on a huge data set that includes 4,868 benign and 4,868 malicious applications. The experimental study suggests shown our malware detection approach may obtain a detection rate of much more than 95.8 achieve a better malware detection rate of 97.88.

## 2. System Design

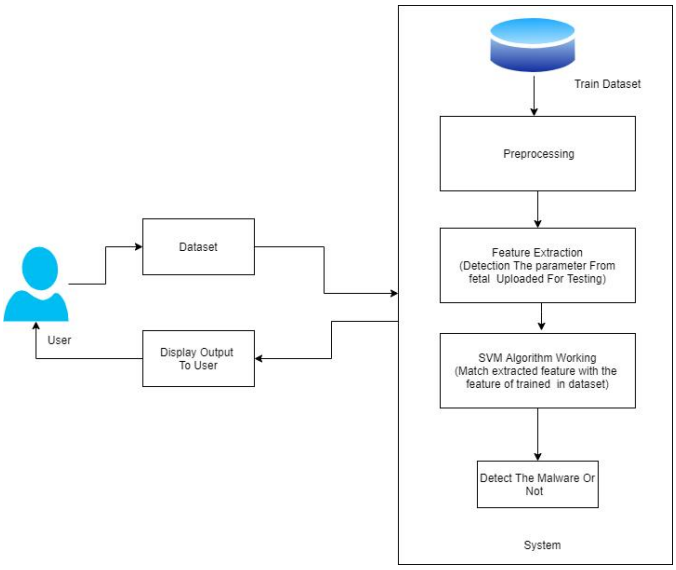


**Fig. 1–System Architecture.**

A system architecture is the conceptual model that defines the structure, behaviour, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviours of the system

## 3. Proposed System

### Preprocessing and Enhancement Images

Image sharpening consists of adding to the original dataset that is proportional to a high- pass filtered version of the original image. In contrast, enhancing the high frequency components of a data leads to an improvement in the visual quality. Data sharpening refers to any enhancement technique that highlights edges and fine details in an image. Image sharpening is widely used in printing and photographic industries for increasing the local contrast and sharpening the data.

### Segmentation Methods

Thresholding is the method which is not including the Spatial characteristics which are important for the malignant tumour detection. Also bit map images from 0 to 255 Gray scale values were not involved and the tumour cells might be wrongly interpreted.

### Feature Extraction

Feature extraction methods encompass, besides the traditional transformed and non- transformed signal characteristics and texture, structural and graph descriptors. The feature selection methods described in this chapter are the exhaustive search, branch and bound algorithm, max–min feature selection, sequential forward and backward selection, and also Fisher's linear discriminant. Advanced feature representation methods are becoming necessary when it comes to dealing with the local image content or with spatio- temporal characteristics or with the statistical image content. A review of the most important feature selection and extraction techniques for biomedical image processing is given.

### Analysis and Results

In this step, the final result is displayed that whether a malware detection data has detected or not.

## 4. Algorithm

### Support Vector Machine

SVM Classifier, or SVM, is a prominent Supervised Learning technique that is used for both classification and regression issues. However, it is mostly utilized in Machine Learning for Classification difficulties. The SVM algorithm's purpose is to find the optimum line or decision boundary for categorizing n-dimensional space so that we may simply place fresh data points in the correct category as in later. A hyperplane is the optimal choice boundary. SVM selects the extreme points/vectors that aid in the creation of the hyperplane. These extreme examples are referred to as support vectors, and the technique is known as the SVM Classifier.

### Conclusion

SVM is better classification technique which can be used for detection of malware. Needs attention to construct better feature representation for better generalization.

REFERENCES

[1] Gavrilut D., Cimpoesu M., Anton D., Ciortuz L., "Malware Prediction Using Machine Learning", International Multiconference on Computer Science and Information Technology, 2009.

[2] Rhode, M., Burnap, P., Jones, K., "Early-stage malware prediction using re- current neural networks", computers security, 2018.

[3] Baset, M, "Machine Learning For Malware Detection", 2016.

[4] Yeo, M., Koo, Y., Yoon, Y., Hwang, T., Ryu, J., Song, J., Park, C.,"Flow- based malware detection using convolutional neural network",2018 International Conference on Information Networking, 2018.

[5] "Features", Light GBM Documentation. [online] Available.https://lightgbm.readthedocs.io/en/latest/Features.html.