# Vasu Bansal

# 20BCE1072

# CSE4003

# Digital Assignment-1

# Research Papers related to our Topic:

1. https://www.semanticscholar.org/paper/7abdea0006ef4751117581807456634f2ce4e513
2. <https://www.semanticscholar.org/paper/Analysis-of-Machine-learning-Techniques-Used-in-Firdausi-Lim/4e45a3f474bdb4b8263e3027bd0bf3197d920a9d>
3. <https://www.semanticscholar.org/paper/9b106cacc5f245d3a61dd20668a674e314a56c0c>
4. https://ieeexplore.ieee.org/document/9673465/
5. https://ieeexplore.ieee.org/document/9117547/
6. https://ieeexplore.ieee.org/document/5352759
7. [https://typeset.io/papers/malware-detection-using-machine-learning-pihyogavbu](https://typeset.io/papers/malware-detection-using-machine-learning-pihyogavbu%20) .
8. <https://www.semanticscholar.org/paper/11462de04b00e89b0a2db2abc58e14734ecedd3f>
9. <https://www.semanticscholar.org/paper/92cfb00cc74818cb5c35a533acc778db8a95a0da>
10. https://www.semanticscholar.org/paper/39174e16ea4295200dcb83985c4396f1e33c0667

# Literature Survey:

1. “Dynamic Analysis of Executables to Detect and Characterize Malware” - This paper discusses the use of neural algorithms to detect malware using system calls generated by executables. Several deep learning techniques and liquid state machines are examined and compared against a random forest. The results suggest that each of the examined machine learning algorithms is a viable solution to detect malware, achieving between 90% and 95% class-averaged accuracy (CAA). The induced models can also be used as a forensics tool to provide directions for investigation and remediation.
2. “Analysis of Machine learning Techniques Used in Behaviour-Based Malware Detection” - This paper discusses the use of automated behavior-based malware detection using machine learning techniques as a solution to the increase of malware exploiting the Internet. The behavior of each malware on an emulated environment is automatically analyzed and generates behavior reports which are preprocessed into sparse vector models for further machine learning classification. The classifiers used in this research are k-Nearest Neighbors (kNN), Naïve Bayes, J48 Decision Tree, Support Vector Machine (SVM), and Multilayer Perceptron Neural Network (MlP). The overall best performance was achieved by J48 decision tree with a recall of 95.9%, a false positive rate of 2.4%, a precision of 97.3%, and an accuracy of 96.8%.
3. “Accurate Malware Detection by Extreme Abstraction” - This paper presents a novel approach to malware analysis that uses an extreme abstraction of the operating system that intentionally strays from real behavior. The key insight is that the presence of malicious behavior is sufficient evidence of malicious intent, even if the path taken is not one that could occur during a real run of the sample. The system, TAMALES (The Abstract Malware Analysis LEarning System), aggregates features from multiple paths and uses a funnel-like configuration of machine learning classifiers to achieve high accuracy without incurring too much of a performance penalty. The results show an FPR (False Positive Rate) of 0.10% with a TPR (True Positive Rate) of 99.11%, demonstrating that extreme abstraction can be extraordinarily effective in providing data that allows a classifier to accurately detect malware.
4. "Malware Detection Using Machine Learning": This paper discusses the exponential growth of malware over the last decade and its significant financial impact on organizations. It highlights the importance of detecting malware in files to protect data and information. The proposed method for malware detection uses different machine learning algorithms such as decision tree, random forest etc. The algorithm with the maximum accuracy is selected to provide a great detection ratio for the system. The performance of the system is detected by calculating the false positive and false negative rates using the confusion matrix.
5. "Malware Detection & Classification using Machine Learning": This paper discusses how malware is one of the major digital dangers nowadays due to fast development of the web. It explains how attackers design polymeric malware that continuously changes its recognizable feature to fool detection techniques that use typical signature-based methods. This is why there is a need for Machine Learning based detection. Behavioral-patterns are obtained through static or dynamic analysis and different ML techniques are applied to identify whether it's malware or not. Behavioral based Detection methods will be discussed to take advantage from ML algorithms so as to frame social-based malware recognition and classification model.
6. "Malware detection using machine learning" : This paper proposes a versatile framework in which one can employ different machine learning algorithms to successfully distinguish between malware files and clean files while aiming to minimize the number of false positives.
7. "Malware detection using machine learning ": This paper provides a method for delaying malicious attacks on machine learning models that are trained using input captured from a plurality of users. The method includes deploying a model designed to be used with an application for responding to requests received from users, receiving input from one or more users, determining if the received input comprises malicious input using a malicious input detection technique, removing the malicious input from the input to be used to retrain the model if it is determined to be malicious, retraining the model using received input that is determined not to be malicious input, and providing a response to a received user query using the retrained model which delays the effect of malicious input on provided responses by removing malicious input from retraining input.
8. “Windows Malware Detector Using Convolutional Neural Network Based on Visualization Images” - This paper discusses the evolution of malware and the need for malware analysis to defend against its sophisticated behavior. The text proposes a Convolutional Neural Network (CNN) based Windows malware detector that uses the execution time behavioral features of Portable Executable (PE) files to detect and classify obscure malware. The proposed approach was effective in uncovering malware PE files and attained a detection accuracy of 97.968 percent.
9. “Dynamic malware analysis using machine learning algorithm” - This paper discusses the importance of malware detection for the security of personal computer systems. It mentions that signature-based strategies are not effective in detecting zero-day attacks and polymorphic viruses, leading to the need for machine learning-based detection. The text presents suggested methods for machine learning-based malware classification and detection and provides guidelines for its implementation. The study can serve as a base for further research in the field of malware analysis using machine learning methods.
10. “Malware detection in mobile environments based on Autoencoders and API-images” - This text discusses the need for effective tools to detect malware on Android devices due to their popularity and vulnerability to attacks. The authors propose a method that represents the sequences of API calls invoked by apps as sparse matrices (API-images) and uses autoencoders to extract the most representative features from these matrices. These features are then provided to an artificial neural network-based classifier to detect malware. Experimental results show that this framework outperforms more complex machine learning approaches in malware classification.