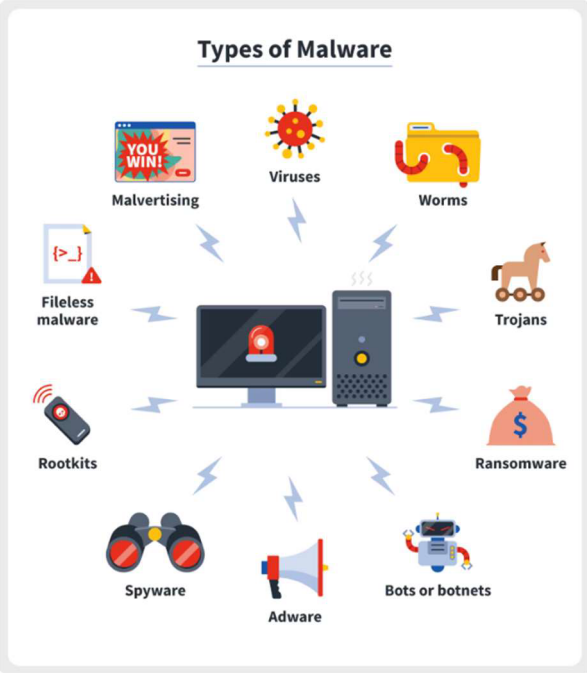
Malware Detection Using Machine Learning Techniques

Introduction

Malware attacks have become a forever growing threat to large organisations all across the globe.

Aim is to create a tool that uses machine learning techniques to determine whether a file is malicious or benign.

Tool will allow single individuals to identify the state of a small number of files and also allow large organisations to identify the state of a large data set.



66 percent of organizations worldwide were victims of a ransomware attack between March 2022 and March 2023

Process

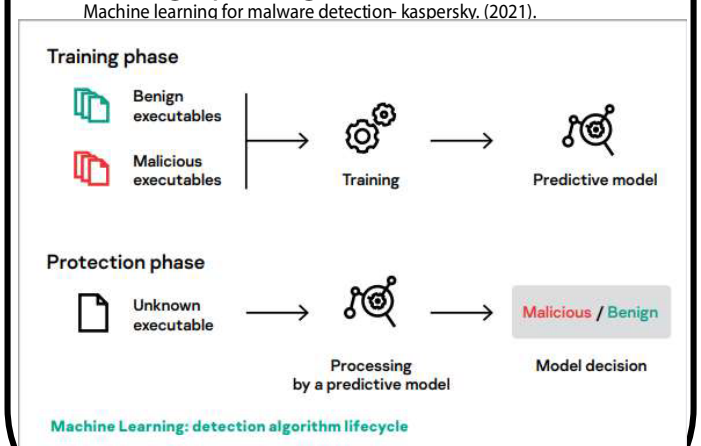
Data set with a variety of file structures that will be analysed through the machine learning.

Two possible end determinants – Malicious or Benign

Forms of Static Analysis will be used to identify malicious files.

Examples include File Format Inspection, String Extraction and Fingerprinting

Machine learning for malware detection- kaspersky. (2021).



Research Topics

Polymorphic Malware

Static vs Dynamic Analysis

'Never-seen-before malware' detection techniques.

Why a non-dynamic detection models are become more outdated as time progresses.

Classification techniques

Future Steps

Identifying a thorough and varied dataset

Pinpoint a machine learning model capable of this project

Conduct thorough training of this model to ensure it can accurately identify the state of a file.

Deliberate 'set-in stone' classification models to provide useful outputs.

References

Machine learning: detection algorithm lifecycle - Machine learning for malware detection – Kaspersky (2021)

Ransomware attack statistics - Ani Petrosyan 30, N. (2023) Ransomware attacks worldwide by country 2022, Statista.

Types of Malware image - 10 types of malware + how to prevent malware from the start (n.d) United States.