Dark Prompt

Leveraging LLMs in Malware Generation and Analysis

Supervised by

Mr Laeeq uz Zaman Khan Niazi

Submitted by

Maryam Mueen	2022-CS-05
Abdullah Fayyaz	2022-CS-21
Ali Tariq	2022-CS-22
Abdullah Iftikhar	2022-CS-35

Department of Computer Science
University of Engineering and Technology, Lahore

Date

Contents

1 Introduction

1.1 Project Overview

In this research and development project we are going to develop a Large Language Model (LLM) based framework which will generate different malwares capable of bypassing anti-viruses. The project aims to create an evolutionary malware generation system that adapts to various situations based on malware detection and malware eleminating mechanisms used against it, serving as a research tool for strengthening cybersecurity defenses through proactive vulnerability assessment.

1.2 Background and Motivation

As we know, the threat of malware attacks in recent years has grown by a huge margin, especially with the advancement in artificial intelligence. Hence there is a growing need for better tools which can help in cyber security professionals understand the extent of how malware behave and develop effective and faster detection mechanisms. In such situations, traditional approaches are not feasible against rapid growth of malware threats. This project probe the potential of Generative AI and focuses in filling this gap by developing LLM-driven malware which can evolve itself to its environment and bypass the detection of anti-malware programs, thereby providing valuable insights for developing more robust attack and defence mechanism.

2 Problem Statement

The cyber security in recent times faces a huge challenges in keeping pace with increasingly evolving malware threats. Many of its limitations include reactive approach to a virus threat, rather than proactive defence, because whenever a new malware threat appears, most of the defence mechanism fail to tackle with it. And since the development of a new malware can be a random event all together, so we are in need of diverse malware samples to test the anti-malware tools. We also lack the understanding of how a malware evolves which can be a huge issue in predicting the future malware variants.

3 Project Objectives

This project is aimed at developing an LLM framework that generates evolving and dangerous malwares to stress-test the defences and lay the ground work for improving future malware eradicating mechanisms.

It includes following steps:

- We are required to develop a comprehensive database of assembly codes from executables
- Use a variety of tools like cuckoo-sandbox to extract static and dynamic features from executables

- Prepare the collected data for LLM to train on
- Fine tune the LLM to generate malwares with evolving obfuscation techniques
- Evaluate the malwares against a variety of malware detection and elemination methods
- Collect all the usefull information for future anti-malware softwares to work on

4 Methodology

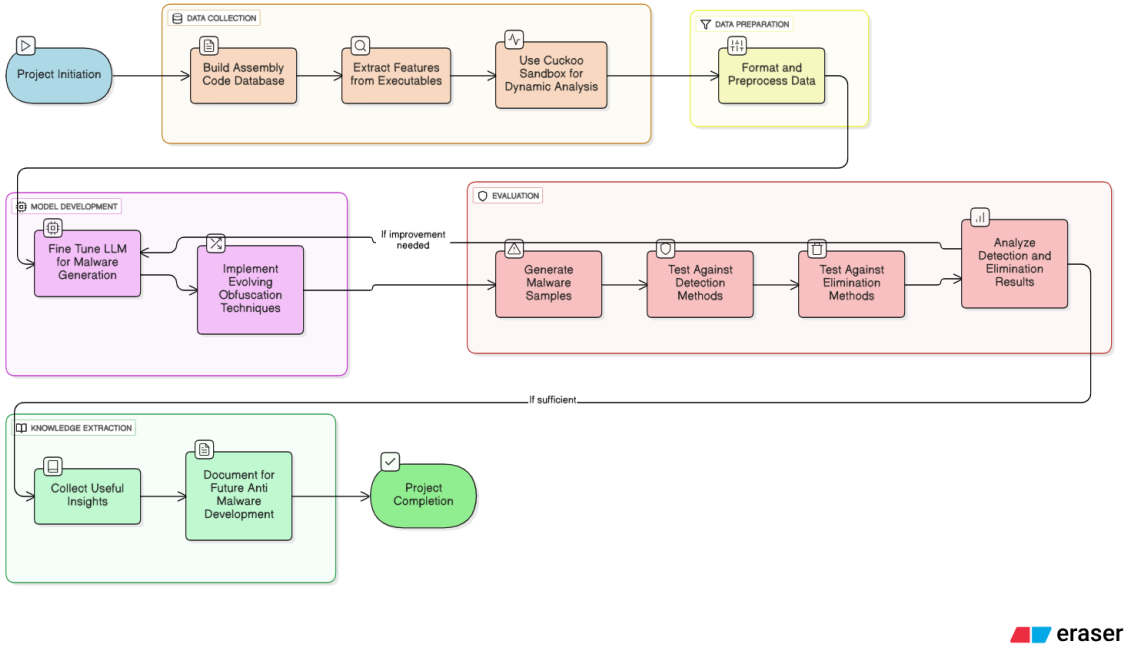

Figure 1: Methodology Diagram

5 Research Methodology

- Trained generative AI models based on different patterns of malware and their detection.
- Reinforcement learning using malware eliminating tools in a controlled environment
- Continuous feedback mechanism based on malware detection and elimination, for continuous evolution of malwares and their solutions

6 Ethical Considerations

- All of the process is being conducted in secluded and isolated environments to avoid malware threats of any kind from spreading
- All actions are strictly in accordance with academic ethics and guidelines
- Malware will not be deployed to any server or used outside of isolated environment in any way possible
- Collaboration with the best cyber security experts to avoid any mishap

7 Expected Outcomes

- A working LLM framework, for generating evolving malwares
- Finding and identifying weaknesses in current malware eliminating techniques
- Research on how to handle situations with unseen malwares and possible techniques to avert them.
- Academic publication on hostile malwares and defence stratigies
- Contribution to proactive cybersecurity defense strategies

8 Conclusion

This project is aimed at advancing cybersecurity defense systems, with the help of malwares. Main goal is to find vulnerabilities in present antiviruses, providing them with valuable information on what steps to be taken next for improvement. Malwares would be generated through LLM which will ultimately help in improving the current status of malware eliminating softwares by a huge margin