Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Student: **BC. JAN ZDRAŽIL** **Oponent:** **Ing. Ladislav Vyskočil**

Studijní program: **Informační technologie**
Studijní obor/Specializace: **Kybernetická bezpečnost**
Akademický rok: **2022/2023**

Téma diplomové práce: **Detekce malwaru bežícího pod operačním systémem Android s využitím metod strojového učení**

Hodnocení práce:

Cílem diplomové práce bylo popsat problematiku detekce malwaru běžícího pod operačním systémem Android s využitím metod strojového učení, k jehož dosažení bylo třeba naplnit několik bodů, jejichž přesná specifikace byla součástí zásad uvedených v zadání práce. Diplomová práce je napsána v anglickém jazyce, je přehledně strukturována a jednotlivé části na sebe logicky navazují. Text práce je zpracován srozumitelně. Po jazykové stránce nebyly nalezeny žádné pravopisné, nebo stylistické chyby. Po formální stránce je práce vhodným způsobem řazena do logických celků a doplněna komentáři i odkazy na odpovídající literární či elektronické zdroje. Diplomová práce obsahuje přiměřené množství obrázků, tabulek a příloh.

V teoretické části byl nejdříve popsán výzkum zabývající se současným stavem detekce malwaru pro Android, jehož cílem bylo poskytnout základní znalosti o dané problematice a vybrat vhodné metody extrakce příznaků z aplikací systému Android a techniky strojového učení. Další část se zabývá celkovým popisem architektury operačního systému Android ve vztahu k možnému zneužití malwarem. Následně byly popsány základy umělých neuronových sítí (ANN) od nejjednodušší verze perceptron, až po pokročilejší konvoluční neuronové sítě (CNN) a jejich možnosti a použití při detekci malware. Také bylo přestaveno použití transformátorů se zaměřením na verzi Vision Transformer (ViT), určenou pro analýzu obrazu, která zde bude využita i pro detekci malwaru. Závěr teoretické části práce je věnován hyperoptimalizaci parametrů k testování modelů strojového učení a dosažení vysoké přesnosti.

Úvod praktické části práce je zaměřen na sběr dat a jejich následné rozdělení na trénovací, testovací a validační množinu. Následně jsou popsány čtyři přístupy extrakcí charakteristik z androidových aplikací a jejich vizualizace, kdy jsou získané charakteristiky testovány pomocí metod strojového učení a hyperoptimalizace, k vyhodnocení jejich detekčního potenciálu. Ze zjištěných výsledků byla vybrána nejlepší metoda založena na přístupu "DEX to RGB image" spolu s pomocnou metodou "Extraction of AndroidManifest.xml properties". Dále je popsán na míru vytvořený model neuronové sítě MD-NNM, založený na CNN a také další experimentální model MD-ViTNNM, který využívá transformátor typu ViT. Následně bylo provedeno jejich porovnání na testovací množině pomocí hyperoptimalizace parametrů, při kterém byly použity i další renomované modely DenseNet a ResNet. Dále byly prezentovány výsledky testů se zjištěním, že modely vytvořené na zakázku (MD-NNM, MD-ViTNNM) dosahovaly výrazně lepších výsledků (až 98%) než modely importované. Závěr praktické části se zabývá implementací zvolených modelů MD-NNM a MD-ViTNNM do aplikace s grafickým uživatelským rozhraním, kterou lze používat k testování aplikací OS Android.

Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

Diplomová práce je na vysoké odborné úrovni a její zpracování obsahuje řadu unikátních řešení. Všechny body zadání diplomové práce byly splněny v plném rozsahu. Diplomant popisované problematice velmi dobře rozumí.

Přínosem práce je přehledný a ucelený popis rozsáhlé problematiky detekce malwaru OS Android s využitím metod strojového učení.

Diplomová práce se jeví jako velmi zdařilá a splňující svůj cíl, a proto ji doporučuji předložit k obhajobě.

Celkové hodnocení práce:
Známku uvede oponent dle svého uvážení dle klasifikační stupnice ECTS:
A – výborně, B – velmi dobře, C – dobře, D – uspokojivě, E – dostatečně, F – nedostatečně.
Stupeň F znamená též „nedoporučuji práci k obhajobě“.

Předloženou diplomovou práci doporučuji k obhajobě a navrhuji hodnocení

A - výborně.

V případě hodnocení stupněm „F – nedostatečně“ uveďte do připomínek a slovního vyjádření hlavní nedostatky práce a důvody tohoto hodnocení.

Datum 1. 6. 2023 Podpis oponenta diplomové práce