Univerzita Tomáše Bati ve Zlíně
Fakulta aplikované informatiky

# POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

**Student:** **BC. MARTIN KUBÍČEK** **Oponent:** **Ing. Ladislav Vyskočil**

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

Téma diplomové práce: **Analýza phishingových útoků a návrh proaktivního řešení s využitím metod umělé inteligence**

**Hodnocení práce:**

Cílem diplomové práce bylo popsat problematiku analýzy phishingových útoků a navrhnout proaktivní řešení s využitím metod umělé inteligence, k jehož dosažení bylo třeba splnit několik bodů, jejichž přesná specifikace byla součástí zásad uvedených v zadání práce. Diplomová práce je napsána ve slovenském jazyce. Po jazykové stránce nebyly nalezeny žádné pravopisné, nebo stylistické chyby. Text práce je zpracován srozumitelně. Po formální stránce je práce vhodným způsobem řazena do logických celků, které na sebe přirozeně navazují. Náročnost a rozsah diplomové práce hodnotím nadstandardně. V diplomové práci autor uvádí přiměřené množství obrázků, tabulek a příloh. Práce je doplněna komentáři i odkazy na literární či elektronické zdroje.

Teoretická část nejprve podrobně popisuje rozdělení phishingu na jeho typy a uvádí příklady takových útoků s názornou ukázkou jejich průběhů. Jsou popsány rysy a charakteristiky, díky kterým je možné rozpoznat phishing od legitimního obsahu. Toto téma je doplněno i o přehledné statistiky týkající se odvětví a domén, ve kterých se phishing nejčastěji vyskytuje. Také je prezentována rešerše aktuálního stavu řešení. Dalším tématem je dokumentace vývoje phishingu od jeho počátku až k současnosti, včetně statistických údajů, s cílem zkoumat vývoj a cílení phishingu i zdokonalování technik útočníků. Posledním tématem této části je analýza současných existujících řešení, která zahrnuje metody přímé detekce phishingu i metody které umožňují phishingu předcházet. Tyto metody jsou poté shrnuty v teoretickém postupu implementace, kde je popsané předzpracování dat, postup klasifikace i výběr features. Teoretická část popisuje celou problematiku velice rozsáhle.

Praktická část nejprve představuje phishing kit pro vytváření phishing útoků. Následuje jeho zkoušení a analýzy ke zvolení detekčních metod, návrhů features a potřeby klasifikace. Bylo také demonstrováno několik reálných útoku na známé webové servery, kdy proběhla podrobná analýza implementace těchto útoků. Dalším tématem byla implementace proaktivního řešení detekce phishingu s využitím metod umělé inteligence. Zde jsou detailně popsané metody detekce na základě URL adresy a detekce na základě obsahu e-mailu. Následuje návrh řešení pro detekci phishingu, založený právě na těchto dvou typech detekcí. U detekce na základě URL bylo popsáno odstranění šumu a duplikátů i výběr testovací a tréninkové množiny. Dále výběr 10 nových features, na základě kterých je prováděna klasifikace. Takto navržené features jsou následně extrahované z URL adres a slouží jako vstupy pro metody umělé inteligence. Pro klasifikaci bylo použito pět nejvhodnějších AI modelů, po jejímž dokončení byla popsána i přesnost jednotlivých klasifikátorů. Ke klasifikaci na základě obsahu e-mailu bylo použito také pět AI modelů, spolu s pěti natrénovanými transformátory. Natrénované modely a transforátory byly použity pro predikci.

Po další analýze bylo navrženo řešení, schopné klasifikovat phishingové URL adresy a e-mailové zprávy. Zajímavostí řešení je kombinace klasifikace více AI modely s jedním finálním výstupem. Poté bylo provedeno zhodnocení návrhu a porovnání s jinými řešeními. Posledním tématem praktické části je experiment porovnání přesnosti klasifikace, kdy bylo zjištěno, že navržené features zvyšují přesnost klasifikace o cca 2%, podle použitého modelu, avšak překonání hranice 90% již naráží na problémy false positives. V tomto ohledu je zvýšení přesnosti o 2% možné považovat za dobrý výsledek.

Diplomový práce je velmi rozsáhlá a velmi podrobně popisuje celou problematiku phishingu i navržená řešení. Všechny body zadání diplomové práce byly splněny v plném rozsahu, 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