POSUDEK VEDOUCÍHO ZÁVĚREČNÉ PRÁCE

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Detekce malwaru ze slabě označených URL pomocí metod hlubokého učení
Jméno autora:	Vít Zlámal
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Vedoucí práce:	Jan Brabec
Pracoviště vedoucího práce:	Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání **mimořádně náročné**

Hodnocení náročnosti zadání závěrečné práce.

Zadání považuji za mimořádně náročné, protože kromě samotné práce s daty vyžaduje i velmi silné softwarově inženýrské schopnosti. Datové sady se kterými se v rámci práce pracuje jsou ohromné. Navíc, samotná doména je z hlediska datové analýzy složitá a správná interpretace výsledků vyžaduje odborné znalosti z oblasti kybernetické bezpečnosti. Nakonec, zadání práce je výzkumného charakteru a vyžaduje kreativní přínos, protože neexistují účinné state-of-the-art přístupy, které by se daly na daný problém přímo aplikovat.

Splnění zadání **splněno s menšími výhradami**

Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.

Součástí zadání byl i návrh metody pro kombinaci několika labelů do jednoho slabého labelu. Tato část zadání nebyla splněna. Z časových důvodů jsme se v průběhu zpracovávání rozhodli tento problém neřešit v rámci této práce a místo toho řešit problém učení neuronové sítě se slabými labely, které jsme však vyrobili synteticky. Tuto odchylku nepovažuji za překážku k obhájení práce, protože všechny ostatní body zadání jsou splněny a pouze tento bod byl v průběhu upraven do méně ambiciozní verze, která však stále obtížností odpovídá požadavkům na diplomovou práci.

Aktivita a samostatnost při zpracování práce **C - dobře**

Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posuďte schopnost studenta samostatné tvůrčí práce.

Student na práci pracoval pravidelně po dobu celého roku a výsledky pravidelně konzultoval. V některých fázích byl však postup pomalejší a méně samostatný než bych očekával. Nebyl jsem příliš spokojený s organizací výsledků experimentů a myslím že pořádek v kódu mohl být lepší.

Odborná úroveň **B - velmi dobře**

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

Student pronikl hluboko na mezioborovém rozhraní strojového učení, kybernetické bezpečnosti a big-data inženýrství. Během zpracovávání kladl správné otázky a modely vyhodnocuje dle kritérií, která jsou důležitá pro jejich praktickou užitečnost.

Formální a jazyková úroveň, rozsah práce **B - velmi dobře**

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Práce je psaná anglickým jazykem. Některé formulace jsou o něco méně formální než je u tohoto typu textu zvykem, avšak je to stále s jistotou v mezích daného stylu. Navíc, dle mého pozorování se tímto, méně formálním avšak snadněji čitelným, směrem v posledních letech ubírají i publikace na nejlepších konferencích v oboru. Práce splňuje všechny formální požadavky, obsahuje několik originálních vizualizací vysoké kvality a esteticky působí velmi dobře. Rozsah je standardní.

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

POSUDEK VEDOUCÍHO ZÁVĚREČNÉ PRÁCE

Výběr zdrojů, korektnost citací	**A - výborně**
Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.	
Bez výhrad.	

Další komentáře a hodnocení
Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.
Navržené klasifikátory fungují dle běžně užívaných metrik jako je klasifikační chyba nebo plocha pod ROC křivkou výborně. Snahou však bylo vyvinout klasifikátor, který bude skutečně prakticky použitelný pro detekci závažného malwaru v síťové komunikaci, což klade extrémně vysoké nároky na precision a recall. Výsledkem práce je skutečně prakticky použitelný klasifikátor a infrastruktura která ho dokáže velmi snadno přetrénovat a vyhodnotit na nových datech. Toto je samo o sobě skvělý přínos.

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.

Vzhledem k vysoké obtížnosti zadání a dosaženým výsledkům, které jsou prakticky užitečné, hodnotím předloženou závěrečnou práci klasifikačním stupněm B - velmi dobře.

Datum: 10.6.2020 Podpis: