FAKULTA INFORMAČNÍCH TECHNOLOGIÍ ČVUT V PRAZE

Posudek oponenta závěrečné práce

Oponent práce:	prof. Ing. Róbert Lórencz, CSc.
Student:	Bc. Tomáš Zvara
Název práce:	Generování signatur malwarových rodin z behaviorálních grafů pomocí nesupervizovaného učení
Obor / specializace:	Počítačová bezpečnost
Vytvořeno dne:	30. května 2022

Hodnotící kritéria

1. Splnění zadání

▸ [1] **zadání splněno**
[2] zadání splněno s menšími výhradami
[3] zadání splněno s většími výhradami
[4] zadání nesplněno

Vyjmenované body v zadání jsou splněné.

2. Písemná část práce 92 /100 (A)

Část práce popisující teoretický základ je rozsáhla a dobře strukturována. Její obsah odpovídá řešenému problému. Malé výtky by mohli být adresovány popisu generování signatur.

3. Nepísemná část, přílohy 95 /100 (A)

Jazyk Python byl implementační bázi pro provedení experimentu s klastrováním.

4. Hodnocení výsledků, jejich využitelnost 82 /100 (B)

Výsledky práce v současnosti nemohou být uvedený do praxe. Pravděpodobně při výhodnější reprezentaci malwarových rodin, by byly výsledky významnější.

Celkové hodnocení 91 /100 (A)

Práce má výraznější teoretickou část. Praktická část vykazuje některé nedostatky. Přesto všechno považují práci a za zdařilou a hodnotím ji stupněm A.

Otázky k obhajobě

1. Neuronové sítě použité k vytvoření 32bitových embeddingů nejsou v článku specifikovány. Je možné je blíže specifikovat, nebo je to interní tajemství? Proč jsou vložené grafy právě 32bitové?
2. Grafové embeddingy byly vytvořeny trénováním neuronových sítí s použitím pouze dvou tříd (malware vs. čistý) a poté použity k reprezentaci rodin malwaru (tj. bylo uvažováno více tříd). Nebylo by lepší, kdyby se embeddingy vytvářely trénováním pomocí více tříd pro každou rodinu malwaru?
3. Jak se definuje kmen malwaru?

Instrukce

Splnění zadání

Posuďte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.