FAKULTA INFORMAČNÍCH TECHNOLOGIÍ ČVUT V PRAZE

Hodnocení vedoucího závěrečné práce

Vedoucí práce:	Mgr. Martin Jureček, Ph.D.
Student:	Bc. Tomáš Zvara
Název práce:	Generování signatur malwarových rodin z behaviorálních grafů pomocí nesupervizovaného učení
Obor / specializace:	Počítačová bezpečnost
Vytvořeno dne:	9. 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

Všetky body zo zadania práce považujem za splnené.

2. Písemná část práce 93 /100 (A)

Teoretická časť práce je pekne spracovaná na takmer 30 stranách. Práca je dobre členená a má odpovedajúci rozsah. Posledná časť práce týkajúca sa generovania signatúr by mohla byť rozsiahlejšia. Študent využil len jednoduchú metódu, na základe ktorej vybral ťažisko klastru - reprezentatívny graf - z ktorého sa odvodila signatúra pre celý klaster. Taktiež kvalita týchto signatúr bola vyhodnotená len manuálne.

3. Nepísemná část, přílohy 98 /100 (A)

Klastrovanie a vyhodnotenie úspešnosti klastrov bolo vykonané v jazyku Python prostredníctvom IPython notebookov za využitia známych knižníc zo strojového učenia. Takéto nástroje sú štandardné pre dátovú analýzu.

4. Hodnocení výsledků, jejich využitelnost 85 /100 (B)

Experimentálne výsledky naznačujú, že grafové embedingy, ktoré sa využívajú v práci, momentálne nedosahujú natoľko dobré výsledky, aby boli nasadené do produkcie. Táto reprezentácia malwarových rodín avšak nie je dielom študenta, študent ju len využíval vo svojej práci. Ak by mal študent k dispozícii vhodnejšiu reprezentáciu, tak by

pravdepodobne dosiahol významnejšie výsledky. Postupy z oblasti strojového učenia, ktoré študent v práci použil, sú správne.

5. Aktivita studenta

- [1] **výborná aktivita**
  [2] velmi dobrá aktivita
  [3] průměrná aktivita
  [4] slabší, ale ještě dostatečná aktivita
  [5] nedostatečná aktivita

Študent pravidelne konzultoval s vedúcim práce najnovšie výsledky a smery výskumu počas celého obdobia práce bez väčších časových okien.

6. Samostatnost studenta

- [1] **výborná samostatnost**
  [2] velmi dobrá samostatnost
  [3] průměrná samostatnost
  [4] slabší, ale ještě dostatečná samostatnost
  [5] nedostatečná samostatnost

Študent si samostatne naštudoval potrebnú teóriu zo strojového učenia a využil svoje znalosti z analýzy škodlivého kódu.

Celkové hodnocení 93 /100 (A)

Po teoretickej stránke je práca pekne spracovaná. Praktická časť uvedená v práci by mohla byť širšia, avšak priložené IPython notebooky obsahujú dostatok experimentov, pričom niektoré z nich mohli byť uvedené aj v práci. Dosiahnuté výsledky nie sú zatiaľ natoľko dobré, aby boli použité v praxi, ale práca na tomto projekte stále pokračuje. Okrem samotných embedingov behaviorálnych grafov doporučujem vylepšiť generovanie signatúr, napr. pomocou klasifikátorov založených na pravidlách. Celkovo vzhľadom k vyššie uvedeným bodom hodnotím študentovu prácu známkou A.

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.

Aktivita studenta

V souvislosti s průběhem a výsledkem práce posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven.

Samostatnost studenta

V souvislosti s průběhem a výsledkem práce posuďte schopnost studenta samostatné tvůrčí práce.

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.