



Posudek oponenta závěrečné práce

Oponent práce: Ing. Karel Hynek
Student: Lukáš Jančíčka
Název práce: Klasifikace komunikace uvnitř Tor spojení
Obor / specializace: Teoretická informatika
Vytvořeno dne: 6. června 2021

Hodnotící kritéria

1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Zadání považuji za složitější, jelikož je spíše výzkumného charakteru. Student navíc nad rámec zadání provedl vyhodnocování kvality dostupné datové sady.

2. Písemná část práce

80 /100 (B)

Písemná část práce je psaná v angličtině, je rozumně členěná a dobře se čte. Nicméně jsem zaznamenal překlepy, typografické chyby a jiné drobnosti. Text cituje celkem 49 zdrojů a nemalá část z nich jsou odborné publikace či články. Bohužel, používání citací na konci odstavce je velice nešťastné, protože není zřejmé k jaké informaci se citovaná publikace váže. Naopak, na některých místech mi chyběly citace nebo alespoň reference umístěné pod čarou. Místy také chybí informace, které jsou podle mého názoru důležité. Například metodika výběru hyperparametrů nebo jejich výsledné hodnoty nejsou vůbec popsány.

3. Nepísemná část, přílohy

95 /100 (A)

Nepísemná část práce je opravdu kvalitní. Výstupem jsou jupyter notebooky obsahující experimenty a prototyp napsaný v jazyce python, který je schopen pracovat s reálnými síťovými daty. Zdrojové kódy jsou čitelné a okomentované. Jupyter notebooky obsahují poměrně hodně vysvětlujících markdown polí, nicméně bych ocenil jejich přiložení i ve formátu pdf nebo html tak, aby je bylo možné zobrazit i bez specializovaného softwaru.

4. Hodnocení výsledků, jejich využitelnost

100 /100 (A)

Výsledky považuji za velice pěkné a využitelné v rámci infrastruktury sdružení CESNET.

Celkové hodnocení

90 /100 (A)

Celkově na mě práce působí pěkným dojmem. Student si nastudoval velké množství materiálů ohledně monitorování, technologie strojového učení a technologie Tor. Tyto znalosti dokázal následně aplikovat na reálný problém síťové klasifikace. Navíc nad rámec zadání student dokázal provést důkladnou analýzu existující datové sady a zhodnotit jejich použitelnost v rámci bakalářské práce, což nebývá obvyklé ani u odborných publikací. Z výše popsaného důvodu doporučuji práci k obhajobě a hodnotím ji stupněm A.

Otázky k obhajobě

Jaké jsou největší rozdíly vašeho přístupu oproti původní metodě Kanadského Institutu pro Kyberbezpečnost (CIC), ve které použili dataset ISCXTor2016?

Jak jste určoval hyperparametry pro jednotlivé testované klasifikátory?

Která charakteristika síťového provozu byla nejdůležitější pro rozeznání Tor provozu a která pro rozeznání typů provozu?

Instrukce

Splnění zadání

Posudte, 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, posudte 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 posudte, 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. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3.

Posudte, 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.