



Hodnocení vedoucího závěrečné práce

Student: Bc. Marek Mařík
Vedoucí práce: Ing. Karel Klouda, Ph.D.
Název práce: Identifikace webového obsahu v šifrovaném provozu
Obor: Znalostní inženýrství

Datum vytvoření: 25. 1. 2021

Hodnotící kritérium:	Způsob hodnocení – následující škálou 1 až 4:
1. Splnění zadání	1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: 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.	
Komentář: Zadání bylo bezezbytku splněno. Splnění všech úkolů navíc bylo dost pracné. Bylo třeba vymyslet nástroj pro vytvoření datasetu a implementovat jej. Nad tímto datasetem bylo třeba udělat již standardní proces vytvoření modelu, i to ale díky specifické doméně a poměrně neprobádané problematice nebylo úplně triviální a přímočaré. Velmi ale pomohly konzultace s kolegy z výzkumné skupiny Tomáše Čejky, kteří poskytly svou doménovou znalost a další dobré rady.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	80 (B)
Popis kritéria: 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. 26/2017, 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.	
Komentář: Text je psán poměrně srozumitelně a bylo poznat, že se v průběhu psaní autor zlepšuje a novější texty jsou srozumitelnější. Je ale škoda, že kvůli časové tísni nebylo možné všechny části pečlivěji zkontrolovat a hlavně se zamyslet nad jejich uspořádáním. V textu podle mého nic významného nechybí ani nepřebývá, ale šlo by výsledky prezentovat přehledněji.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	85 (B)
Popis kritéria: 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	
Komentář: Autor odvedl opravdu hodně práce. Výstupy jsou dva. První je generátor datasetu, který simuluje šifrovaný provoz na virtuálním stroji (různé kombinace operačních systémů a prohlížečů). Ten je napsán převážně v jazyce Python, ale vyžadoval i mnoho konfigurace nástroje Joy a virtualizačního nástroje. Druhým výstupem je pak zpracování dat, feature engineering a experimenty s různými modely strojového učení. Tyto experimenty mají formu Jupyter notebooků a několika skriptů, vše též v jazyce Python. Jak u generátoru datasetu tak experimentů platí, že jsou výsledkem hodným velmi dobré diplomové práce, ale trochu jim chybí dokumentace či přehlednost.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

4. Hodnocení výsledků, jejich využitelnost

90 (A)

Popis kritéria:

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.

Komentář:

Provedené experimenty jasně prokazují, že zkoumané problémy jsou uspokojivě řešitelné nástroji strojového učení. Věřím, že při troše další práce a ladění modelů (možná i využití některých neprozkoumaných), je možné přetvořit získané výsledky v solidní publikaci.

Hodnotící kritérium:

Způsob hodnocení – následující škálou 1 až 5:

5. Aktivita a samostatnost studenta

5a:

1=výborná aktivita,
2=velmi dobrá aktivita,
3=průměrná aktivita,
4=slabší, ale ještě dostatečná aktivita,
5=nedostatečná aktivita

5b:

1=výborná samostatnost,
2=velmi dobrá samostatnost,
3=průměrná samostatnost,
4=slabší, ale ještě dostatečná samostatnost,
5=nedostatečná samostatnost

Popis kritéria:

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 (5a). Posuďte schopnost studenta samostatně tvůrčí práce (5b).

Komentář:

Pan Mařík byl velice samostatný a se vším si poradil v podstatě sám. Někdy byla samostatnost až přílišná a bylo tak těžší mít o postupu prací přehled. Celkově se mi ale s panem Maříkem spolupracovalo dobře.

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

6. Celkové hodnocení

87 (B)

Popis kritéria:

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.

Text hodnocení:

Autor odvedl mnoho dobré práce na poměrně složitém tématu. Výsledky trochu sráží jistá nedotaženost a to zejména v jejich prezentaci. Proto navrhuji hodnotit předkládanou práci známkou B jako velmi dobrou.

Podpis vedoucího práce: