



## Posudek oponenta závěrečné práce

**Student:** Bc. Jana Mašková  
**Oponent práce:** Ing. Simona Buchovecká  
**Název práce:** Klasifikace internetového provozu  
**Obor:** Počítačová bezpečnost

**Datum vytvoření:** 23. 1. 2020

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – následující škálou 1 až 4:</i>
<b>1. Splnění zadání</b>	<b><u>1=zadání splněno,</u> 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno</b>
<i>Popis kritéria:</i> 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.	
<i>Komentář:</i> Studentka splnila všechny pokyny ze zadání - ať již nastudování relevantních postupů strojového učení pro klasifikaci internetového provozu a relevantních atributů v rámci obsáhlého teoretického úvodu a rešerše, tak i v praktické části, kde autorka porovnávala vlastnosti jednotlivých technik. Srovnání technik je obsáhlé, autorka hodnotí vlivy normalizace a předzpracování vstupních dat dle jednotlivých algoritmů, a v práci najdeme i pokus o vylepšení známých postupů.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>2. Písemná část práce</b>	<b>90 (A)</b>
<i>Popis kritéria:</i> 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.	
<i>Komentář:</i> Práce je obsáhlá, osvětluje čtenáři teoretické pozadí strojového učení, součástí je i rešerše existujících prací pro klasifikaci provozu. Autorka využívá poměrně velké množství informačních zdrojů a ukazuje tak, že téma studovala důkladně.  Drobnou výtku mám ke kapitole 1 (Přehled teorie k počítačovým sítím a Internetu), která je poměně strohá v porovnání s následujícími kapitoly. Zvážila bych rozšíření a detailnější vysvětlení pojmů používaných v práci, tak, aby byly srozumitelné i pro čtenáře bez hlubších znalostí počítačových sítí na úkor zkrácení některých částí druhé kapitoly, která je místy až příliš detailná.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>3. Nepísemná část, přílohy</b>	<b>90 (A)</b>
<i>Popis kritéria:</i> 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	
<i>Komentář:</i> Nepísemnou část přílohy tvoří zejména zdrojové kódy testovaných a porovnávaných algoritmů. Zdrojové kódy by místy zasloužili víc/detailnější komentáře pro lepší srozumitelnost a orientaci, ale jinak je považuji za adekvátní zadání.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>

#### 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:**

Využití strojového učení v informační bezpečnosti je velmi aktuální a diskutované téma. Práce shrnuje známé algoritmy strojového učení, porovnává je a hodnotí, a může tak sloužit jako podklad pro celou škálu dalších prací, jako vodítko vhodnosti/efektivity jednotlivých algoritmů.

**Hodnotící kritérium:**

*Způsob hodnocení – nehodnotí se*

#### 5. Otázky k obhajobě

**Popis kritéria:**

Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odřádkami).

**Otázky:**

- Pro extrakci příznaků z datasetu byl zvolen nástroj Argus. Co bylo rozhodující pro jeho výběr?

- Jaký algoritmus/algoritmy byste zvolila, pokud byste navrhovala nový nástroj na klasifikaci síťového provozu a jaký algoritmus/algoritmy byste zvolila pro návrh nástroje na detekci anomálií? Jaká kritéria byste zohledňovala?

**Hodnotící kritérium:**

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

#### 6. Celkové hodnocení

90 (A)

**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:**

I přesto, že problematika strojového učení není triviální, autorka ji dobře zvládla, a splnila pokyny ze zadání. Téma je obšírně rozpracované, a to jak v teoretické, tak praktické části a může sloužit i jako podklad pro další výzkum v této oblasti. I proto práci doporučuji k obhajobě a kloním se k hodnocení stupněm A.

Podpis oponenta práce: