

# Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

**Student:** Bc. Filip Mudruněk  
**Oponent práce:** Ing. Jiří Smítka  
**Název práce:** Detekce phishingu  
**Obor:** Webové a softwarové inženýrství

**Datum vytvoření:** 8. 6. 2016

<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - následující škálou 1 až 5:</b>
<b>1. Náročnost a další komentář k zadání</b>	<b>1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání</b>
<b>Popis kritéria:</b> Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.)	
<b>Komentář:</b> Náročnější zadání.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - následující škálou 1 až 4:</b>
<b>2. Splnění zadání</b>	<b>1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno</b>
<b>Popis kritéria:</b> Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.	
<b>Komentář:</b> Zadání splněno.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - následující škálou 1 až 4:</b>
<b>3. Rozsah písemné zprávy</b>	<b>1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky</b>
<b>Popis kritéria:</b> Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.	
<b>Komentář:</b> Zpráva splňuje požadavky.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>4. Věcná a logická úroveň práce</b>	<b>85 (B)</b>
<b>Popis kritéria:</b> 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.	
<b>Komentář:</b> Práce je po věcné stránce v pořádku. Kapitola 1.2.3 by měla obsahovat zmínku o zneužití slabín DNS komunikace a podstrčení závadné odpovědi na uživatelův dotaz. Autor bez zdůvodnění radikálně omezil množství zkoumaných klasifikátorů na 3 druhy - Naive Bayes, rozhodovací strom (resp. Random Forest) a SVM. Chybí diskuse o klasifikátorech založených např. na HMM nebo neuronových sítích. Implementace je postavená nad aplikací RapidMiner, též bez nějaké širší diskuse o případných alternativách. Práce má jinak dobrou logickou strukturu, kapitoly na sebe navazují.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>5. Formální úroveň práce</b>	<b>95 (A)</b>
<b>Popis kritéria:</b> Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 12/2014, článek 3.	
<b>Komentář:</b> Formální úroveň je dobrá, seznam literatury by měl obsahovat u online zdrojů datum. Text je jasný a dobře čitelný a téměř neobsahuje typografické a jazykové prohřešky (proxy se obvykle píše tvrdě y - viz str.11).	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>6. Práce se zdroji</b>	<b>100 (A)</b>

**Popis kritéria:**

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

**Komentář:**

Autor vybral relevantní zdroje, na které se v textu odkazuje. Oblast tématu je vybranými prameny pokryta.

*Hodnotící kritérium:*

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

**7. Hodnocení výsledků, publikační výstupy a ocenění**

90 (A)

**Popis kritéria:**

Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

**Komentář:**

Autor provedl měření prototypové implementace. Toto měření slibuje relativně dobré výsledky při případném nasazení algoritmu v praxi.

*Hodnotící kritérium:*

*Způsob hodnocení - nehodnotí se*

**8. Komentář o využitelnosti výsledků**

**Popis kritéria:**

Uvedte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uveďte možnosti využití výsledků ZP v praxi.

**Komentář:**

Autor neměl k dispozici větší trénovací i testovací data a tak jsou výsledky zkrácené, nicméně i tak poměrně nadějně.

*Hodnotící kritérium:*

*Způsob hodnocení - nehodnotí se*

**9. Otázky k obhajobě**

**Popis kritéria:**

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

**Otázky:**

1) Měřil jste časovou náročnost jednotlivých klasifikátorů ?

2) V práci není příliš diskutováno nasazení v praxi. Jak byste zapojil Vaši implementaci phishingového detektoru na nějakém standardním poštovním serveru ? Je možné (a nutné) využít RapidMiner nebo je třeba vytvořit klasifikátory jiným způsobem ?

*Hodnotící kritérium:*

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

**10. Celkové hodnocení**

85 (B)

**Popis kritéria:**

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nesmí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

**Text hodnocení:**

Autor splnil zadání práce. Zpracované výsledky jsou přehledné a mají dobrou vypovídací hodnotu. Výběr klasifikátorů, což je stěžejní část, však nijak nekomentoval, jiné druhy klasifikátorů nejsou ani zmíněny. Nevíme tedy, jaké by byly výsledky např. při použití HMM. tato skutečnost bohužel práci mírně degraduje.

Podpis oponenta práce: