

## Hodnocení vedoucího diplomové práce

Bc. Jindřich Pouba

## Klasifikace uživatelů přistupujících k webovým službám

Diplomová práce Bc. Jindřicha Pouby vychází z oblasti hlubokého odborného zájmu diplomanta a zabývá se klasifikací uživatelů, kteří přistupují k webovým službám s cílem odlišit lidské uživatele od robotů, a to pomocí technik z oblasti strojového učení, což je cesta, která zatím rozhodně nepředstavuje v této oblasti hlavní proud. To také dodává práci její atraktivitu a slibuje možnost dosažení zajímavých odborně-vědeckých závěrů.

Diplomant pracoval svědomitě, samostatně a prokazatelně velmi aktivně. Evidentně prostudoval značné množství relevantní literatury a online dostupných zdrojů a navrhl a vyzkoušel řadu technik klasifikace uživatelů. Problém konzultoval také s odborníky v oblasti počítačových sítí a zejm. síťové bezpečnosti. Rozsah vykonané práce je značný a její kvalita svědčí mj. i o disciplíně autora.

Nicméně vedoucí může aktivitu diplomanta posuzovat pouze nepřímo, z viditelných výsledků jeho práce, protože osobních kontaktů mezi vedoucím a diplomantem bylo velmi málo.

Na konzultace s vedoucím docházel diplomant jen zřídka a zejména v posledních dvou měsících před odevzdáním práce. Nelze však usoudit, nakolik to bylo ke škodě práce, protože ta je velmi dobrá. Když už diplomant na konzultaci přišel, byl bezchybně připraven a měl jen několik málo velmi konkrétních, relevantních dotazů. Jelikož oblast počítačových sítí není předmětem odborného zájmu vedoucího, nemohl mu tento ani poskytnout výraznější pomoc. Naopak v oblasti strojového učení bych asi ocenil intenzivnější kontakt.

V zásadě mám jen jedinou podstatnější výtku ke spolupráci s vedoucím: Konečná podoba průvodního textu práce, zejm. teoretická část, byla dle mého subjektivního názoru konzultována nedostatečně a vysloveně na poslední chvíli.

Práce je zcela původní, při návrhu řešení autor vycházel z řady knih, časopiseckých a konferenčních článků a dalších odborných publikací v dané oblasti z poslední doby. Bibliografie je přiměřeně rozsáhlá (27 titulů) a relevantní. Bohužel ale zcela chybí literatura na téma rozpoznávání, klasifikace a strojového učení – domnívám se, že celkem ke škodě práce; zejména v teoretické části zabývající se klasifikací je znát, že autorovi poněkud chybí formální školení v předmětné problematice.

Citace v textu i bibliografie na konci práce jsou provedené v souladu s požadavky.

Programové řešení, přestože charakter práce je spíše výzkumný a software má tedy charakter prototypu, je na velmi vysoké úrovni. Řešení je plně funkční a stabilní. Při návrhu i implementaci se autor držel všech zásad a doporučení, následkem čehož je např. dekompozice úlohy přímo vzorná. Přestože se při testování ukázalo, že některé z implementovaných metod klasifikace uživatelů webových služeb nedosahují spolehlivosti potřebné pro praktické nasazení, hodnotu práce to nijak nesnižuje. Autor jednak samostatně navrhl a implementoval inovativní metodu („Vylepšený Markovův model“), která funguje velmi dobře a potřebné spolehlivosti dosahuje, jednak provedl velice cenné vyhodnocení vhodnosti řady metod řešení daného problému a jejich vzájemné porovnání.

K vývoji byl použit jazyk Java. Objem zdrojového kódu je úměrný řešené problematice a kód je zapsán pečlivě, ve shodě se zvyklostmi, je celkem dobře čitelný a přehledný. Je ovšem poměrně málo komentovaný, navíc některé komentáře pochopení komentovaných konstrukcí spíše znesnadňují.

Textová část díla je poměrně rozsáhlá – má 82 stran bez příloh. Autorovo odborné vyjadřování je velice dobré, text je srozumitelný, dobře čitelný, autorovy úvahy je snadné sledovat. Dokument napsán vynikající technickou češtinou a vysázen v L<sup>A</sup>T<sub>E</sub>Xu. Grafická úroveň práce je zcela profesionální a celkově působí velmi harmonickým dojmem. Gramatické chyby se v práci prakticky nevyskytují.

Práce je dobře logicky strukturovaná a poměr jednotlivých částí je celkem vyvážený, snad až na (již zmíněný) menší rozsah obecné teoretické části popisující techniky klasifikace a klasifikátory. Text velice dobře popisuje jednotlivé implementované metody klasifikace, uvedené úvahy a argumentace jsou konzistentní a validní.

Text je doplněn množstvím kvalitně vysázených vzorců, tabulek a diagramů, které text žádoucím způsobem doplňují.

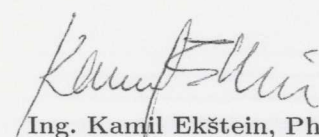
Jelikož se vedoucí práce necítí být odborníkem v oblasti bezpečnosti počítačových sítí, není schopen objektivně posoudit praktickou využitelnost předloženého díla. Subjektivně (a z hlediska strojového učení) se mi jeví autorem vyvinutý klasifikátor jako velice významný příspěvek k praktickému zvládnutí úlohy odlišení lidského uživatele webu od robotů.

Spolehlivost klasifikace metodou vylepšeného Markovova modelu je navíc tak vysoká, že i při testování odhalené nevýhody (citlivost na změnu struktury webu) nebrání nasazení této techniky do reálného provozu. Nakolik je to ale možné (např. z důvodů možného omezení propustnosti webového serveru) musí zřejmě posoudit odborníci lépe orientovaní v dané problematice.


Všechny body zadání byly splněny. Autor při řešení dané problematiky jednoznačně prokázal schopnosti řešit inženýrským způsobem náročné problémy a nasazovat do praxe a i vhodným způsobem modifikovat metody a techniky, s nimiž se seznámil studiem odborné literatury.

Práci **doporučuji k obhajobě** a hodnotím klasifikačním stupněm

„výborně“.

  
Ing. Kamil Ekštejn, Ph.D.  
KIV FAV ZČU

V Plzni dne 30. května 2017

  
**SOUHLASÍ  
S ORIGINÁLEM**

Západočeská univerzita v Plzni  
Fakulta aplikovaných věd  
katedra informatiky a výpočetní techniky  
①