

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Bc. Tomáš Frýda
Oponent práce: Ing. Karel Klouda, Ph.D.
Název práce: Scalability of Predictive Modeling Algorithms
Obor: Znalostní inženýrství

Datum vytvoření: 3. 2. 2017

Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 5:
1. Náročnost a další komentář k zadání	1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání
Popis kritéria: 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.)	
Komentář: Pro splnění úkolů ze zadání bylo potřeba pochopit a zvládnout netriviální software H2O, pochopit a vstřebat (byť ne třeba úplně do hloubky) mnoho modelů a algoritmů strojového učení a v neposlední řadě bylo potřeba také docela dost programování (zejm. v Pythonu). Z toho důvodu považuji zadání za náročnější.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
2. 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 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.	
Komentář: Všechny body zadání byly splněny, a to i přes to, že vše nešlo udělat tak přímočaře, jak se (asi) při zadávání úkolů zdálo (anytime properties a H2O).	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
3. Rozsah písemné zprávy	1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.	
Komentář: Práce je rozhodně dostatečně rozsáhlá. Možná by místy zasloužila trochu pozvolnější výklad, protože se na čtenáře valí informace velkou rychlostí, ale takto pojatá by pak byla asi zase rozsáhlá příliš.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
4. Věcná a logická úroveň práce	95 (A)
Popis kritéria: 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.	
Komentář: Pořadí kapitol a dalších částí textu na mě působí promyšleně. Jak píšou výše, informace jsou hodně zahuštěné, ale to je důsledek zadání, nikoli chyba autora.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
5. Formální úroveň práce	95 (A)
Popis kritéria: Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 14/2015, článek 3.	

Komentář:

Text práce je pečlivě zpracován i když nějaké překlepy a formátovací podivnosti (např. na rovnici (1.1) jsem koukal hodně dlouho) by bylo možné najít. Rovnice by měly být uváděny tak, že jsou součástí textu (k nalezení skoro všude), to samé by mělo platit pro seznamy (např. str. 28). Angličtina je na dobré úrovni, nějaké nesrovnalosti se vyskytují výjimečně a nebrání pochopení textu. Na diplomovou práci je celkově text velice vyzrálý. Text je také přeplněný zkratkami, které nejsou vždy vysvětleny (co je např. ETL na straně 2, FAKE GAME je naopak rozepsána 3x), ale to je spíše oborem, který se v tomto vyžívá.

Hodnotící kritérium:

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

6. Práce se zdroji

100 (A)

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:

Práce obsahuje 76 relevantních referencí na zdroje všeho typu. Způsob odkazování odpovídá solidní vědecké práci.

Hodnotící kritérium:

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

7. Hodnocení výsledků, publikační výstupy a ocenění

100 (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:

Jestli jsem správně pochopil disclaimer na str. 5, část práce byla poslána jako článek do velmi dobrého časopisu (Machine learning). Obsahuje dobře podepřené porovnání hojně využívaných metod strojového učení (implementované v hojně užívaném nástroji), a proto myslím může být (nejen) pro vědeckou komunitu zajímavá.

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. Uvedte možnosti využití výsledků ZP v praxi.

Komentář:

Viz předchozí bod.

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 odřázkami).

Otázky:

1. Mohl byste vlastními slovy definovat pojem "anytime properties"? V práci jste se definici trochu vyhnul a myslím, že by tam alespoň nějaký pokus měl být.
2. Chystáte se na doktorské studium?

Hodnotící kritérium:

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

10. Celkové hodnocení

98 (A)

Popis kritéria:

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

Text hodnocení:

Práci považuji za nadprůměrně povedenou a proto ji navrhuji hodnotit známkou A.

Podpis oponenta práce: