



Posudek oponenta závěrečné práce

Student: Bc. Tomáš Pajurek
Oponent práce: Ing. Daniel Vašata, Ph.D.
Název práce: Online Anomaly Detection in Time-Series
Obor: Znalostní inženýrství

Datum vytvoření: 1. 6. 2018

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 splněno bez výhrad.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	85 (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: Práce je logicky dobře strukturovaná. Použité zdroje jsou relevantní a správně citované. Po obsahové stránce na mě poněkud rušivě působilo velmi často se opakující shrnutí toho, co bude v dané sekci (podsekci) uvedeno. Po typografické stránce bych vytkl nesprávné odkazování v sekci 1.2.1 (chybí závorky u odkazů na rovnice) a nekonzistentní užívání velkého resp. malého písmene při odkazování na číslované sekce, obrázky atd.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	90 (A)
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: Nepísemnou částí práce bylo experimentální porovnání vybraných metod detekce anomálií v časových řadách. V příloze práce jsou uvedeny zdrojové kódy, které jsou relativně přehledné a umožňují případnému zájemci všechny provedené experimenty zreplicovat.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
4. Hodnocení výsledků, jejich využitelnost	85 (B)
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:

Výsledky práce jsou jak v teoretické tak v praktické rovině na velmi slušné úrovni. V teoretické rovině se jedná o rešerši významných nesupervizovaných metod pro detekci anomálií v časových řadách. Dále byly navrženy dvě vylepšení současných metod pro detekci konceptuálního driftu (tj. posunu rozdělení dat v čase) a jedno vylepšení algoritmu učení izolujícího lesa. Zde je třeba dodat, že podle informací uvedených v textu práce, spočívá první z těchto vylepšení pouze ve změně jednotek hranice pro rozhodnutí, zda nastal konceptuální drift. S původní metodou je to však přímo ekvivalentní (stačí vzít původní parametr ρ jako hodnotu, kterou dostaneme aplikací inverzní funkce přežití příslušného rozdělení na nový parametr τ). V experimentální části pak byly na šesti datasetech testovány 4 vybrané metody detekce anomálií v různých variantách, které zahrnovali i první z předchozích tří uvedených vylepšení (které je podle mého názoru pouze interpretační). V zásadě z nich vyplynulo, že toto navržené vylepšení je reálně využitelné.

Hodnotící kritérium:

Způsob hodnocení – nehodnotí se

5. Otázky k obhajobě

Popis kritéria:

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

Otázky:

Proč nebylo možné otestovat zbylá dvě navrhovaná vylepšení? Tj. vylepšení metody pro detekci konceptuálního driftu a vylepšení algoritmu učení izolujícího lesa.

Jaká data / modely by bylo třeba využít, aby toto testování mohlo být provedeno?

Hodnotící kritérium:

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

6. Celkové hodnocení

85 (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:

Práce je celkově na velmi dobré úrovni. Vzhledem k výše uvedeným drobným nedostatkům hodnotím stupněm B.

Podpis oponenta práce: