

Posudek oponenta bakalářské práce

Student: Brezánský Tomáš
Téma: Přehled současných přístupů ke klasifikacím (id 22419)
Oponent: Šústek Martin, Ing., UITS FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**
Dle mého názoru se jedná o mírně obtížnější zadání, protože se předpokládá pochopení fungování různých klasifikátorů, což považuji za obtížnější látku pro bakalářskou práci.
- 2. Splnění požadavků zadání** **zadání splněno**
Formálně bylo zadání splněno.
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**
- 4. Prezentační úroveň předložené práce** **10 b. (F)**
Vhodnější uspořádání kapitoly 5 by bylo podle experimentů s jednotlivými datovými sadami. Neuvedení použitých algoritmů v kapitole 2 (např. RCE) a popis až v následující kapitole 3 je pro čtenáře matoucí, jinak je logická struktura práce dobrá.
Prezentační úroveň práce je však nedostatečná. Přestože student uvádí všechny použité algoritmy, zřejmě nerozumí ani základním pojmům a použité notaci. Ku příkladu vztahy 2.1 a 2.4 reprezentují totožnou věc s použitím jiné notace, přičemž student chybně uvádí symboly ve vztahu 2.1, které ani nepopíše. Vztah 2.2 využívá podmíněnou entropii bez přiblížení významu či vysvětlení symbolu T . Vztah 2.3 opět používá špatný symbol a chybí zde jedna suma. Student několikrát v textu zaměňuje význam uzlu a listu (teorie grafů). Dle popisku Obrázku 2.7 nerozumí rozdíl mezi dimenzionalitou a počtem dat. Vztahy 2.7 a 2.8 si protiřečí. V práci je použit přístup "naive bayes", který má být popsán vztahem 2.10, ovšem uvedený vztah je obecný a nejedná se o "naive bayes". Jak vztah 3.1, tak vztah 3.3 (jedná se o tentýž), jsou nepřesné (chybí symbol u sumy), student tak ani nedokáže správně přepsat citované vztahy. Obr. 2.1 nepůsobí profesionálně a způsob zobrazení výsledků Obr. 5.1, 5.2, 5.3 je nevhodný. Použitý parametr R_{dc} u RCE sítě musí být menší než 1, aby algoritmus dával smysl, student v experimentech (Tabulka 5.8) používá hodnoty větší než 1. Přestože tento výčet je dlouhý, jedná se pouze o část všech nedostatků práce.
- 5. Formální úprava technické zprávy** **35 b. (F)**
Gramatiku nedokážu hodnotit, jelikož je práce psaná ve slovenštině. V práci se vyskytuje opravdu velké množství překlepů (decision, decission, ...), chybějící, přebývající nebo špatně umístěné interpunkce a jiné, příklad: Mezera před čárkou (např. u klíčových slov), chybějící mezera za ukončením uvozovek (str. 4), mezera před otazníkem (str.9). Student zavádí nové pojmy vymyšleným/nepřesným překladem některých pojmů z angličtiny - ku příkladu trénování (trénung, vzdělávací přístup, výškolený) nebo excitační (recitačné) - používá anglický pojem, když existuje zažitý český ekvivalent - např. podmnožina (subset) a zaměňuje pojmy při překladu - noise (zhluk), continuous (nepretržité). U odkazu na obrázek se vyskytuje malé o.
- 6. Práce s literaturou** **30 b. (F)**
Mám problém s odlišením vlastních výsledků a úvah v kapitolách 2 a 3, jelikož se student odkazuje na vhodnou literaturu, ale uvádí nepřesná či dokonce nepravdivá tvrzení. Práce s literaturou je tak špatná, přestože výběr literatury je dobrý.
- 7. Realizační výstup** **85 b. (B)**
Algoritmus RCE má špatné výsledky, jednou z příčin může být, že student v implementaci RCE místo provedení bodu 8 (str. 25) zmenšuje poloměr vynásobením konstantou. Nejsem si jistý správnosti jeho provedení evaluace na testovacích datech u RCE. Kromě tohoto nedostatku působí výsledná aplikace funkčně a byla pracná.
- 8. Využitelnost výsledků**
Jelikož se student rozhodl využít starší metody, které nedosahují state-of-the-art výsledků, a program nemá vizualizační charakter (není využitelný pro výukové účely), práce je dle mého názoru v praxi nevyužitelná.
- 9. Otázky k obhajobě**
 - Vysvětlíte vztahy 2.1 a 2.2 včetně všech použitých symbolů
 - Vysvětlíte bayesův vzorec uvedený ve vztahu 2.10, kdy se bude jednat o **naivní** bayesovský klasifikátor?
- 10. Souhrnné hodnocení** **40 b. nevyhovující (F)**
Dle mého názoru je realizační výstup práce dobrý, ovšem kvalita technické zprávy (kapitoly 2 a 3) je nepřipustná, navrhuji hodnotit 40 body (F).

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 18. srpna 2020

Šustek Martin, Ing.
oponent