



Posudek oponenta závěrečné práce

Oponent práce:	Ing. Ivo Petr, Ph.D.
Student:	Mikuláš Kočí
Název práce:	Srovnání interpretovatelných modelů s post-hoc vysvětlitelnými black-box modely
Obor / specializace:	Umělá inteligence 2021
Vytvořeno dne:	8. června 2024

Hodnotící kritéria

1. Splnění zadání

- [1] zadání splněno
- ▶ [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Práce se zabývá aktuální problematikou interpretovatelnosti modelů využívaných ve strojovém učení a srovnáním výkonu interpretovatelných modelů vůči výkonu tzv. black-box modelů. Co se týče rešeršní části práce a praktického porovnání modelů považuji zadání za splněné. Výhrady mám k písemné části práce, kdy některé koncepty nejsou dostatečně vysvětleny a některé výsledky nejsou řádně komentovány. Celkově je text velmi stručný, rozsahem minimální přípustný pro bakalářskou práci.

2. Písemná část práce

60/100 (D)

Text práce je na velmi vysoké úrovni a dobře se čte. Obsahuje minimum překlepů. Odkazy na reference by měly být v textu odděleny mezerou od předchozího slova, jinak student pracuje s referencemi korektně. Hlavním nedostatkem práce je její přílišná stručnost. Uvážíme-li větší množství obrázků a grafů v některých částech, je text na dolní hranici přípustnosti.

Modely představené v teoretické části nejsou dostatečně komentovány. Jak například určíme význam příznaků u GAM? Jsou funkce a příznaky normalizovány? Jak interpretovat interakci v GA2M? Metoda SHAP sloužící jako hlavní nástroj pro vysvětlení CatBoost je představena velmi neurčitě takže čtenář neví jak rozumět obrázku 2.6. Další obrázky a grafy by rovněž potřebovaly rozsáhlejší komentáře.

3. Nepísemná část, přílohy

80 /100 (B)

Po praktické stránce považuji práci za zdařilou. Student použil předpřipravené implementace několika modelů na několika vybraných datasetech a porovnal jejich výkonnost. Poněkud matoucí je zařazení regresních modelů do teoretické části, jelikož tyto modely nejsou dále použity a student zkoumá výhradně úlohu binární klasifikace. Při zpracování dat mi u kategorických příznaků nepřijde vhodné doplňovat chybějící údaje mediánem. V textu jsou uvedeny "optimální" hodnoty hyperparametrů jednotlivých modelů, není ale uvedeno jaké hodnoty hyperparametrů byly zkoumány nebo jaká metoda hledání hyperparametrů byla zvolena.

4. Hodnocení výsledků, jejich využitelnost

80 /100 (B)

Práce se zabývá velmi zajímavou oblastí strojového učení a snaží se vyvrátit tvrzení, že využití interpretovatelných modelů vede ke snížení přesnosti vůči tzv. black-box modelům. Může sloužit jako základ dalšímu výzkumu v této oblasti.

Celkové hodnocení

70 /100 (C)

Téma práce je zajímavé a až na drobnosti je praktická část v pořádku. Hlavní výhrady mám k textu práce, který je příliš stručný a chybí v něm důležité informace a podrobnější představení výsledků.

Otázky k obhajobě

Můžete stručně objasnit metodu SHAP a význam hodnot zobrazených v obrázku 2.6?

Instrukce

Splnění zadání

Posudte, 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, posudte 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.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, 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. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, 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.

Nepísemná část, přílohy

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.

Hodnocení výsledků, jejich využitelnost

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.

Celkové hodnocení

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.