

Posudek oponenta bakalářské práce

Student: Marko Július
Téma: Predikce vlivu mutace na rozpustnost proteinů (id 22864)
Oponent: Smatana Stanislav, Ing., UPSY FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**
Zadanie považujem za náročnejšie, pretože si vyžaduje naštudovanie netriviálnej problematiky z oblasti proteínového inžinierstva, štúdium metód strojového učenia, ich praktickú aplikáciu a prípravu veľmi špecifickej dátovej sady.
- 2. Splnění požadavků zadání** **zadání splněno**
Všetky body zadania boli splnené v plnom rozsahu.
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**
Rozsah technickej správy je primeraný.
- 4. Prezentací úroveň předložené práce** **80 b. (B)**
Prezentačnú úroveň práce hodnotím ako nadpriemernú, no mám k nej niekoľko výhrad. V prvom rade študent neuvádza niektoré dôležité informácie týkajúce sa vyhodnotenia klasifikátorov a návrhu riešenia. Napríklad v kapitole 6 študent uvádza presnosť klasifikátora v krížovej validácii, no nešpecifikuje či ide o priemernú presnosť naprieč iteráciami alebo je presnosť zhrnutá pomocou inej štatistiky. Rovnako nie je jasne popísaná transformácia dát pred vstupom do klasifikátora. Pri popise návrhu nástroja napríklad študent neuvádza, že sú zložky vektorov trénovacej sady normalizované. Normalizácia je spomenutá iba v kapitole o implementácii a naznačená v diagrame na strane 20. Chýbajú však akékoľvek bližšie informácie o zvolenej metóde normalizácie a pre ich zistenie je nutné nahliadnúť do zdrojového kódu riešenia. Ďalej študent vyjadruje mieru rozpustnosti proteínu a jej práh pomocou desatinného čísla, no nikde nevysvetľuje čo presne táto veličina vyjadruje.
Celkovo však kapitoly práce pokrývajú problematiku primerane a logicky na seba nadväzujú. Práca ako celok je pre čitateľa pochopiteľná.
- 5. Formální úprava technické zprávy** **90 b. (A)**
Formálna stránka práce je na vysokej úrovni. Použitý jazyk je zrozumiteľný a gramaticky správny. Po typografickej stránke ma práca len drobné nedostatky, akými sú napríklad ojedinelé osamotené citácie na začiatku riadka.
- 6. Práce s literaturou** **100 b. (A)**
Použitá literatúra pozostáva z kvalitných knižných a časopiseckých publikácií. Rozsah literatúry je primeraný a študent v texte práce dôsledne cituje použité zdroje.
- 7. Realizační výstup** **80 b. (B)**
Priložený realizačný výstup je plne funkčný. Kód implementácie je logicky členený do primerane veľkých funkcií a tried. Prehľadnosti v niektorých častiach kódu by ale prospela väčšia miera komentovania.
- 8. Využitelnost výsledků**
Práca rieši aktuálny problém z oblasti proteínového inžinierstva a prináša zaujímavé porovnanie nového prístupu k problematike s existujúcimi metódami. Preto si myslím, že ide o veľmi relevantnú prácu, ktorej výsledky sú z vedeckého pohľadu zaujímavé.
- 9. Otázky k obhajobě**
 - Vo vyhodnotení uvádzate, že dátová sada, na ktorej bol trénovaný nástroj OptSolMut, zdieľala značné množstvo príkladov s dátovou sadou na ktorej bol tento nástroj testovaný. Neskúšal ste tento nástroj znovu trénovať, tak aby k prieniku trénovacej a testovacej sady nedochádzalo? Ak nástroj nie je možné pretrénovať, akým spôsobom by bolo možné sa s týmto problémom vyrovnáť a získať porovnanie Vášho riešenia s metódou OptSolMut?
- 10. Souhrnné hodnocení** **90 b. výborně (A)**
Súhrnne hodnotím prácu študenta ako kvalitnú. Práca síce vykazuje mierne nedostatky, najmä z pohľadu prezentačnej úrovne, no prihliadnúc k obtiažnosti riešenia, ktoré hraničí s obtiažnosťou diplomovej práce, konštatujem, že študent vytvoril veľmi dobrý výstup a navrhujem komisii hodnotenie známku A.

V Brně dne: 23. června 2020

Smatana Stanislav, Ing.
oponent