

Posudek oponenta diplomové práce

Student: Dekrét Lukáš, Bc.
Téma: Techniky klasifikace proteinů (id 21685)
Oponent: Zendulka Jaroslav, doc. Ing., CSc., UIFS FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**

Zadání považuji za obtížnější a to jak v oblasti studijní, tak experimentálního vyhodnocení. V oblasti studijní to byla zejména problematika klasifikace proteinů, která vyžaduje určité znalosti související s proteiny, jejich vlastnostmi, způsoby reprezentace a dostupnými databázemi, které lze pro klasifikaci proteinů využít. Zadání rovněž vytvářelo značný prostor pro experimentální vyhodnocování při výběru vhodných metod extrakce příznaků a typu klasifikační metody.
- 2. Splnění požadavků zadání** **zadání splněno s podstatným rozšířením**

Zadání student splnil v plném rozsahu a zejména v části přípravy datových souborů pro experimentování, provedení a vyhodnocení experimentů představují podle mne jeho výsledky významná rozšíření oproti "standardnímu" splnění.
Přestože jde o téma bioinformatické a pro studenta představovalo nutnost nastudovat věci, se kterými se ve svém oboru studia přímo nesetkal, zvládl jeho zpracování výborně.
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**
- 4. Prezentací úroveň předložené práce** **98 b. (A)**

Práce má logickou strukturu, jednotlivé kapitoly na sebe dobře navazují a odpovídají posloupnosti bodů zadání. Práce je psána srozumitelně. V části popisující návrh a implementaci vytvořených aplikací pro experimentování a použití natrénovaných modelů student použil nejen slovní popis, ale i diagramy tříd popisující objektovou strukturu. Zde má některé drobné nedostatky, které uvádím podrobněji v komentáři pro studenta. Kladně hodnotím i skutečnost, že některé důležité informace týkající se vytvořených datových souborů, nastavení parametrů metod extrakce příznaků a klasifikátorů, které jsou poměrně rozsáhlé, ale ne podstatné pro čtenáře z hlediska pochopení řešení zadání, jsou umístěny ve velmi dobře zpracovaných přílohách.
- 5. Formální úprava technické zprávy** **98 b. (A)**

Po formální, jazykové i grafické stránce je zpracována technická zpráva kvalitně.
- 6. Práce s literaturou** **100 b. (A)**

Seznam použité literatury je na diplomovou práci obdivuhodně rozsáhlý. Je vidět, že student pracoval skutečně s relevantní literaturou, včetně článků z posledních let. Použití je na příslušných místech odkazováno. Vlastní výsledky a úvahy jsou jasně odlišené.
- 7. Realizační výstup** **95 b. (A)**

Realizační část práce tvoří dvě konzolové aplikace - pro experimentování při výběru metod a trénování klasifikátorů a pro hierarchickou klasifikaci. Obě jsou pak použity v rozsáhlé experimentální části zpracování zadání zaměřené na vytvoření datových sad pro experimentování, vlastní experimentování pro výběr metody extrakce příznaků a klasifikátorů pro hierarchický klasifikátor a zjištění přesnosti vytvořeného hierarchického klasifikátoru. Tuto část, včetně zpracování v technické zprávě považuji za velice zdařilou.
Mám drobnou připomínku ke komentářům v kódu aplikací v tom smyslu, že v souborech s kódem by na začátku mělo být nejen, kdo je autorem a že jde o DP a kdy, ale i stručně, co (jakou funkčnost/zodpovědnost) kód k souboru zajišťuje.
- 8. Využitelnost výsledků**

Práce je inspirována existujícími přístupy ke klasifikaci proteinů a existujícími metodami extrakce příznaků. Metody extrakce příznaků student sám implementoval. Existující práce a jejich výsledky rozšiřuje jednak o nově vytvořené datové soubory, které lze pro porovnávání metod klasifikace proteinů využít, jednak o aplikaci, která umožňuje experimentovat s metodami extrakce příznaků a klasifikačními metodami a o nový hierarchický klasifikátor. V neposlední řadě pak jsou přínosné i poznatky z rozsáhlých experimentů.
- 9. Otázky k obhajobě**
 1. V závěru uvádíte jako jedno z možných pokračování práce kombinaci několika metod extrakce příznaků. Jak by taková kombinace mohla vypadat?
- 10. Souhrnné hodnocení** **97 b. výborně (A)**

Pan Lukáš Dekrét vypracoval pěknou diplomovou práci, jejíž výsledky považuji za nadstandardní. Podle mne při zpracování plně prokázal inženýrské schopnosti.

Doporučuji zvážit návrh na nominaci do soutěže IT SPY 2020.

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

V Brně dne: 17. června 2020

Zendulka Jaroslav, doc. Ing., CSc.
oponent