

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Kompenzace vlivu matoucích faktorů při klasifikaci RNASeq dat.
Jméno autora:	Karina Balagazova.
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky.
Vedoucí práce:	Doc. Ing. Jiří Kléma, PhD.
Pracoviště vedoucího práce:	Katedra počítačů FEL ČVUT

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Úlohu v širším kontextu pokládám za velmi náročnou. Vyžaduje základní znalost molekulární biologie a velmi dobrou znalost statistiky a strojového učení. Vzhledem k tomu, že jde o bakalářskou práci, jsme od počátku po dohodě pracovali s celou řadou silných zjednodušení, pro která práci nakonec hodnotím za průměrně náročnou.	

Splnění zadání	splněno s menšími výhradami
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Práce splňuje všechny psané požadavky zadání. Na počátku jsem si myslel, že se nám podaří testovat širší škálu vstupních parametrů (např. měnit počty vzorků ve třídách OT a CR) a více pracovat s parametry generátoru vstupních dat (změny distribucí a jejich parametrů) či klasifikačními algoritmy (širší využití našla jen logistická regrese). To bylo objektivně komplikováno poměrně velkou časovou náročností experimentů, některých výše uvedených rozšíření ale v obvyklé časové dotaci pro bakalářskou práci pravděpodobně dosáhnout šlo.	

Aktivita a samostatnost při zpracování práce	D - uspokojivě
<i>Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posuďte schopnost studenta samostatně tvůrčí práce.</i>	
Karina Balagazova odevzdává svoji bakalářskou práci po standardním termínu. V průběhu naší spolupráce byla delší období, kdy se bakalářské práci nevěnovala, či pokroky byly velmi malé. Postup prací tak byl spíše nárazový, včetně významného pokroku před závěrečným odevzdáním. Pokud jde o vlastní návrhy řešení, držela se studentka převážně mých pokynů. Na druhou stranu, studentka byla nepochybně schopna samostatného studia literatury, v problematice se globálně orientovala, provedla poměrně časově rozsáhlé experimenty a uměla je odpovídajícím způsobem jednoduše shrnout.	

Odborná úroveň	C - dobře
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Práce není dokonalá, některé odhady by musely jít daleko za úroveň bakalářské práce, aby byly odborně široce přijatelné. Jako příklad lze uvést použití F-rozdělení pro modelování středních hodnot exprese transkriptů v kapitole 4.1. Použitý model nemá oporu v teorii, v literatuře ale ustálený model neexistuje a typ modelu velmi pravděpodobně nemá hlubší vliv na konečný výsledek studie. Líbilo se mi zavedení heatmap pro souhrnnou interpretaci experimentů i názorný výklad související s obrázkem 4.8. Jinde je ale úroveň kolísavá (nepřesnosti v teoretickém úvodu, jak byla skutečně regularizovaná logistická regrese v provedených experimentech apod.).	

Formální a jazyková úroveň, rozsah práce

D - uspokojivě

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Jazyková úroveň práce je ovlivněna tím, že studentka není rodilou mluvčí. Optimální by bylo, kdyby rodilý mluvčí celou práci korigoval, což jsem také studentce několikrát doporučoval, jazyk ale vzhledem k dříve uvedenému nepokládám za klíčovou věc pro mé hodnocení. Typograficky má práce rezervy, není ale pod našimi standardy.

Výběr zdrojů, korektnost citací

A - výborně

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

K rozsahu literatury a jejímu výběru nemám výhrady. Studentka odpovídajícím způsobem doporučenou literaturu zapracovala do své práce, byla schopna sama dohledat další zdroje. Práce s literaturou odpovídá tomu, že jde o první větší práci autorky.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Vložte komentář (nepovinné hodnocení).

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.

Studentka splnila všechny hlavní požadavky zadání. Její práce na základní úrovni zodpovídá klíčovou otázku, zda je možné u RNASeq dat s podobnými parametry, jaké pozorujeme v praxi dosáhnout kompenzace vlivu matoucích faktorů, zde konkrétně imunosupresiv. Současně zhruba naznačuje, jak by se RNASeq data musela změnit, aby kompenzace mohla být provedena. Současně má práce řadu výše popsaných nedostatků, pro které ji nemohu hodnotit lépe než jako uspokojivou.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm D - uspokojivě.

Datum: 12.1.2022

Podpis: