

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Automatic Intron Detection in Metagenomes Using Neural Networks
Jméno autora:	Bc. Martin Indra
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Oponent práce:	Doc. RNDr. Petr Baldrian, Ph.D.
Pracoviště oponenta práce:	Mikrobiologický ústav AVČR, v. v. i.

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	mimořádně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Řešený úkol představuje velkou výzvu v oblasti genomiky mikroorganismů, o čemž svědčí zejména skutečnost, že přes potenciální význam pro environmentální mikrobiologii nebyl problém dosud úspěšně vyřešen.	

Splnění zadání	splněno
<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 splnila zadání ve všech jeho bodech. Základem práce je důkladná rešerše zvolených přístupů, dostatečné seznámení s biologickým podkladem problému, správná volba přístupů k řešení a jejich implementace. Nejsem odborníkem na algoritmické a matematické aspekty práce, ale pohledem biologa, tedy adresáta výsledků práce, je překvapivé, že došlo k vypracování robustního řešení problému, které ve srovnání s jinými přístupy (SVM) je jednak přesnější, jednak méně náročné na počítačový čas.	

Zvolený postup řešení	vynikající
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
V práci byly vhodně zvoleny jednotlivé komponenty přístupu a jejich postupným splněním a kombinací pak došlo k řešení problému, tedy kompletaci pipeline, kde je implementován celý optimalizovaný postup. Strategie testování výkonnosti algoritmu je z biologického pohledu vhodně volena. Za cenné považuji části, vyhodnocující vliv chyb v okolí vybraných elementů intronu, na kvalitu predikce. Vlastní algoritizaci nedovedu posoudit, nicméně prezentované výsledky predikce považuji za vysoce kvalitní co se týče odhalení potenciálních intronů.	

Odborná úroveň	A - výborně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Diplomová práce je psána na dostatečné odborné úrovni. Co se týká analýzy dostupných pramenů, řešících analogické problémy, zdá se, že jsou pokryty všechny zásadní zdroje, což svědčí o velmi podrobném seznámení s problémem. Velmi pozitivně hodnotím logickou výstavbu struktury diplomové práce i fakt, že ji lze snadno sledovat i pro neoborníka. Biologické fenomény jsou v práci popsány správně, ale poměrně stručně, což nečiní problémy biologovi, ale mohlo by komplikovat pochopení textu čtenáři bez základních znalostí molekulární biologie.	

Formální a jazyková úroveň, rozsah práce	B - velmi dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je psána přehledně, anglickým jazykem, který lze snadno sledovat. Jazykové chyby se nevyskytují často, ale přesto jsou v textu patrné, časté je nadužívání počátečních velkých písmen v pojmech. Typografická úprava a zápis vzorců se zdá být v pořádku.	

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.

Zdroje pro práci byly vybrány vhodně a jsou vždy korektně citovány. Student podle mého názoru použil relevantních zdrojů. Převzaté prvky a citace jsou v práci snadno identifikovatelné a oddělitelné od vlastních výsledků.

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.

Z pohledu potenciálního uživatele algoritmu konstatuji, že došlo k vytvoření a komplexní validaci algoritmického přístupu k řešení komplexního problému, který má potenciál k bezprostřední aplikaci v oblastech analýzy metagenomů a jednoznačně zvyšuje potenciál správné identifikace biologických funkcí v genomickém zdrojovém materiálu.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Předložená diplomová práce podle mého názoru splnila všechny body zadání díky vhodné volbě metodického přístupu a jeho aplikací (nakolik jsem toto schopen jako biolog posoudit) se podařilo dosáhnout výsledku, který má potenciál usnadnit popis genetické informace hub. Cenné je srovnání s předchozími přístupy i vyhodnocení vlivu sekvenačních chyb na kvalitu výsledků. Práce je psána přehledně slušnou angličtinou.

Otázka: Práce porovnává přístup k problému aplikací neural networks a support vector machines. Bylo by možné uvažovat o kombinaci těchto přístupů tak, aby byla ještě zvýšena přesnost predikce?

Datum: 20.1.2021

Podpis:

