

Review of Bachelor's Thesis

Student: Lacko Dávid
Title: Platform for Biological Sequence Analysis Using Machine Learning (id 25037)
Reviewer: Burgetová Ivana, Ing., Ph.D., DIFS FIT BUT

- 1. Assignment complexity** **more demanding assignment**
Zadání bakalářské práce považuji za obtížnější, protože jeho splnění vyžaduje velmi dobrou orientaci v oblasti proteinového inženýrství a v oblasti hodnocení prediktivních modelů.
- 2. Completeness of assignment requirements** **assignment fulfilled with enhancements**
Zadání bakalářské práce bylo splněno v plném rozsahu. Oproti zadání je práce rozšířena o testování statistické významnosti výsledků vytvořeného nástroje.
- 3. Length of technical report** **in usual extent**
- 4. Presentation level of technical report** **91 p. (A)**
Logická struktura předložené technické zprávy je velmi dobrá, jednotlivé kapitoly mají přiměřený rozsah a dobře na sebe navazují.
- 5. Formal aspects of technical report** **90 p. (A)**
Z typografického i jazykového pohledu se jedná o kvalitní práci, která obsahuje pouze minimální množství překlepů a gramatických chyb.
- 6. Literature usage** **90 p. (A)**
Výběr studijních pramenů je dobrý, použité zdroje dobře pokrývají řešenou problematiku a převzaté prvky jsou řádně odlišeny od vlastních výsledků a úvah.
- 7. Implementation results** **94 p. (A)**
V rámci bakalářské práce vznikl nástroj, který umožňuje trénování a testování nástroje využívajícího PLS metodu pro predikci vlivu mutací na zvolenou cílovou vlastnost proteinu. Nástroj také umožňuje predikci změny vybrané vlastnosti proteinu způsobenou mutacemi v proteinu. Vytvořený nástroj je plně funkční. Na implementaci oceňuji kromě funkčnosti také to, že se student pokusil vytvořit efektivní nástroj s využitím mezipaměti a paralelizace.
- 8. Utilizability of results**
V rámci bakalářské práce byla implementována metoda innov'SAR. Principy této metody byly publikovány, ale zdrojové kódy nejsou veřejně dostupné, proto student vytvořil vlastní implementaci této metody. Vytvořený nástroj byl dále rozšířen o druhý způsob hodnocení vytvořeného modelu, paralelizaci a mezipaměť pro testované kombinace vlastností. Vytvořený nástroj bude pravděpodobně v budoucnu využíván pro potřeby proteinových inženýrů v Loschmidtových laboratořích.
- 9. Questions for defence**
 - Jsou v uloženém modelu uchovány i vybrané vlastnosti z databáze AAindex, které byly při trénování modelu vyhodnoceny jako nejvhodnější?
- 10. Total assessment** **92 p. excellent (A)**
Předloženou bakalářskou práci považuji za nadstandardní. V rámci práce byl vytvořen funkční a kvalitní nástroj pro predikci vlivu mutací na zkoumanou vlastnost proteinu, který pravděpodobně bude použit v praxi. Tento nástroj je doplněn kvalitní technickou zprávou. Proto navrhuji hodnocení stupněm A.

In Brno 1 June 2022

Burgetová Ivana, Ing., Ph.D.
reviewer