



Hodnocení vedoucího závěrečné práce

Vedoucí práce: Ing. Jakub Žitný
Student: Bc. Tomáš Detko
Název práce: Clustering a datová analýza Jupyter notebooků na GitHubu
Obor / specializace: Znalostní inženýrství
Vytvořeno dne: 3. června 2022

Hodnotící kritéria

1. Splnění zadání

- [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- ▶ [3] **zadání splněno s většími výhradami**
- [4] zadání nesplněno

The work covers the direction proposed in the assignment properly, but the goals of clustering and analysis are not finished - preparations have been made, insights are clear, but there is a lot of work that can be done to consider it successful. This could be a problem of incorrectly formulating the assignment - some parts are vague and a large chunk of work was needed to prepare data. Details are listed below.

2. Písemná část práce

69 /100 (D)

The thesis is written in solid English, theoretical part covers important topics in detail and it is clear the student understands them properly. There are, however, missing pieces that would help better understand some decisions from the implementation phase - such as choosing LSTM model, or loss function. Furthermore, some of the in-depth descriptions of time series analysis seem to be redundant. It is clear the implementation section was written in a big rush and more clarity would be needed to better understand some choices or the overall achievement of goals. All in all, the structure is chosen correctly and if one had the missing pieces of the story, the clarity would be much better.

3. Nepísemná část, přílohy

67 /100 (D)

The experiments from student, preprocessed data, trained models and attached code is somewhat reusable for continuing this work. The code is of a good quality, uses common idioms from industry and is readable. There are, however, missing experiments and some choices seem random.

4. Hodnocení výsledků, jejich využitelnost

60 /100 (D)

The direction of the work is good, several solid concepts and discoveries have been made, but much more experimentation and data wrangling would be needed to achieve the ultimate goal of having a good set of clusters.

5. Aktivita studenta

- [1] výborná aktivita
- [2] velmi dobrá aktivita
- ▶ [3] **průměrná aktivita**
- [4] slabší, ale ještě dostatečná aktivita
- [5] nedostatečná aktivita

In the beginning, student was very active and eager to work on data collection, preprocessing and training time series, brought a lot of ideas on what can we try next, etc. But after a while, after trying several directions, the activity diminished and none of the original goals have been met 100%.

6. Samostatnost studenta

- [1] výborná samostatnost
- ▶ [2] **velmi dobrá samostatnost**
- [3] průměrná samostatnost
- [4] slabší, ale ještě dostatečná samostatnost
- [5] nedostatečná samostatnost

The student is very independent in general, understands the advanced topics and came up with solutions on his own.

Celkové hodnocení

63 /100 (D)

Considering all points above, and the student's activity at the beginning, I am proposing grade D.

Instrukce

Splnění zadání

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Aktivita studenta

V souvislosti s průběhem a výsledkem práce posudte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven.

Samostatnost studenta

V souvislosti s průběhem a výsledkem práce posudte schopnost studenta samostatně tvůrčí práce.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.