



# Posudek oponenta závěrečné práce

Oponent práce: Ing. Karel Klouda, Ph.D.  
Student: Jonáš Sirko  
Název práce: Vizualizace extrakce témat Korpusu českého verše  
Obor / specializace: Znalostní inženýrství  
Vytvořeno dne: 5. února 2024

## Hodnotící kritéria

### 1. Splnění zadání

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

Všechny body zadání jsou splněny v tom smyslu, že ke každému z nich lze najít příslušnou část práce. Celkově mi ale v práci chybí nějaká diskuze nad výsledky a v tomto smyslu ji považuji za neúplnou.

### 2. Písemná část práce

72 / 100 (C)

Text je poměrně dobře čitelný a strukturovaný. Přijde mi ale, že jak se blíží konec práce, kvalita a informační bohatost klesá. Práce má sotva 30 stránek a to je jich ještě mnoho pokryto obrázky. To lze vzhledem k povaze práce pochopit, ale jak píše výše, nějaký kontext a diskuze by si i tak obrázky a grafy zasloužily.

V práci lze najít ne úplně zanedbatelné množství formálních chyb, od překlepů po nejasnosti (např. nedefinované proměnné) v matematických formulích a definicích. Nelíbí se mi např. text u metody K-monoids (co je RSS?, citlivost záleží na volbě  $d$ , nebo ne?) příp. důležitá formule jako poznámka pod čarou na téže stránce.

V textu se také vyskytují termíny, které by zasloužily vysvětlení: např. Iris dataset, lemma a hlavně koherence.

### 3. Nepísemná část, přílohy

65 / 100 (D)

Přílohy mají formu pěti Jupyter notebooků. Notebooky obsahují veškerý kód, ale jsou velmi nepřehledné a minimálně okomentované.

#### 4. Hodnocení výsledků, jejich využitelnost

70 /100 (C)

Výsledky lze využít jako přehled vizualizačních metod pro shlukování v oblasti NLP. Bohužel vzhledem k nepřehlednému a nepřívětivému stavu Jupyter notebooků je to složitější, stejně tak nepomáhá ani neuspořádanost výstupů v textové části.

#### Celkové hodnocení

68 /100 (D)

Vzhledem k výše napsanému navrhuji hodnotit práci známkou D. Z práce mám pocit, že byla dodělávána ve spěchu a že autor tak nedotáhl práci do stavu, do jakého by byl schopen.

#### Otázky k obhajobě

- 1) Co je to koherence zmíněná v sekci 2.3.0.1?
- 2) Jde nějak např. Top2Vec donutit, aby počty dokumentů ve shlucích byly rovnoměrnější? (viz obr. 2.8)

## **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.

### **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.