



# Posudek oponenta závěrečné práce

**Oponent práce:** Ing. Karel Klouda, Ph.D.  
**Student:** Herman Tiumentsev  
**Název práce:** Rozpoznávání textu v historických archiváliích ze 17.-19. století  
**Obor / specializace:** Znalostní inženýrství  
**Vytvořeno dne:** 27. srpna 2021

## Hodnotící kritéria

### 1. Splnění zadání

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

Zadání práce bylo na bakalářskou práci poměrně náročné a bylo vesměs splněno. Jedna z největších výtek, které k práci mám, je ta, že z ní není vždy úplně jasné, co student udělal a proč. Jednou z věcí, která není moc vysvětlená, je, jakými typy textů se měla práce zabývat. Jestli pouze těmi tištěnými, tedy podobnými textům dnešním, nebo i ručně psanými, navíc fonty ze 17. století. Tak jako tak mi přijde zadání pro bakalářku dostatečné a z pohledu jednotlivých bodů i splněné.

### 2. Písemná část práce

65/100 (D)

Práce je na bakalářskou práci poměrně obsáhlá. Velmi mi v ní ale chybí úvod a představení toho, čeho má práce dosáhnout. Ukázka textu, který se má zpracovávat není nikde v úvodu k dispozici, přitom by velmi pomohla pochopení a představě toho, co stojí před autorem za úkol. Popisy jednotlivých kroků image processingu a letmý popis neuronových sítí mi naopak přijde často nadbytečný a pro čtenáře neužitečný. Vše je povrchní a skutečné porozumění nelze na základě takových encyklopedických hesel získat. Jako názornou ukázkou zmíním část 1.5.2 na str. 14 o gradientu, kde se bez komentáře objeví různé (a navíc neobvyklé značky) a pak se prostě řekne, že se bere gradient obrázku, který je ale diskrétní a vesměs nespojitý.

Po formální stránce má text také mnohé vady. Zápis rovnic je většinou chybný v tom smyslu, že rovnice jsou odtrženy od textu a nejsou jeho součástí. Navíc velmi často obsahují symboly, které nejsou vysvětleny. V některých částech je text členěn tak, že každá věta má svůj odstavec. Angličtina je většinou srozumitelná, ale místy by potřeboval text ještě zkontrolovat. V některých částech, např. při popisu experimentů, se

text podobá spíše vyprávění a neobsahuje důležité detaily a nějaká přehledná shrnutí. Někde mi chybí citace, není třeba vždy jasné, odkud pochází některé obrázky.

### **3. Nepísemná část, přílohy**

70/100 (C)

Přiložené soubory, stejně jako text práce, nejsou moc dobře a přehledně strukturovány. Obsahují jak datasety, tak i fragmenty ze zpracování (segmentace textu). Stěžejní kód je nejspíše v Jupyter noteboocích, které obsahují hodně kódu, ale opět téměř bez komentářů a nějakého srozumitelného členění.

### **4. Hodnocení výsledků, jejich využitelnost**

70/100 (C)

Výsledky ukazují, s jakou asi úspěšností lze použít současné nástroje pro strojové čtení starých textů. To je samo osobě zajímavé. V práci je také implementován celý proces zpracování nasčenovaného obrázku starého tištěného textu.

## **Celkové hodnocení**

70/100 (C)

K práci má mnoho výhrad a nepovažuji ji za příliš podařenou. Hodnocení ale navrhuji na stupeň C, neboť naopak pozitivně hodnotím to, že se jedná na bakalářskou práci o poměrně rozsáhlé dílo, se kterým měl student asi dost práce.

## **Otázky k obhajobě**

1. Jakým způsobem jste nastavoval váhy ve vzorci (4.6) na str. 40? Jaké bylo kritérium pro jeho volbu?
2. Mohl byste popsat postup, jak přidat do Vašeho testovacího frameworku další metodu (např. novou neuronovou síť)?

## **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.