



# Posudek oponenta závěrečné práce

**Oponent práce:** doc. Ing. Štěpán Starosta, Ph.D.  
**Student:** Michal Lebeda  
**Název práce:** Zpracování výpisu z účtu pomocí metod strojového učení  
**Obor / specializace:** Znalostní inženýrství  
**Vytvořeno dne:** 12. června 2023

## Hodnotící kritéria

### 1. Splnění zadání

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

Zadání považuji za splněné.

### 2. Písemná část práce

90/100 (A)

Dle mého soudu je největším místem pro zlepšení psané části kapitola 6, která se zabývá popisem implementace. Obsahem je nyní spíše vyprávění o tom, jak probíhala implementace, doplněná některými detaily. Chybí tedy přesnější struktura a jasnější popis vlastního návrhu. V kapitole 2 jsou navíc detaily o formátu PDF, protože nakonec správně byla použita knihovna, která toto řeší. Naopak bych čekal více informací o této knihovně. V kapitole 3 jsou informace k neuronovým sítím, není ale jasné, kde a jak jsou všechny dále v práci použity. Práce obsahuje menší množství překlepů a jazykových nedostatků. Práce s referencemi je v pořádku. Některé odkazy z poznámek pod čarou by měly být v seznamu literatury s datem citace. Lze nalézt i nepřesnosti, např. definice "bounding box" 4.1 na str. 19 vyžaduje velké domýšlení, ačkoliv v uvedeném zdroji [27] je mnohem přesnější. Mimo zmíněné neduhy text se vším všudy hodnotím velmi kladně.

### 3. Nepísemná část, přílohy

89/100 (B)

Výsledkem práce je funkční balík skriptů v jazyce Python. Byly použity standární balíčky a jejich volba je většinou odůvodněna a popsána v textu. Pro lepší využití chybí dokonalejší pokrytí kódu komentáři - ne vždy je rovnou jasné, co který kus dělá.

#### 4. Hodnocení výsledků, jejich využitelnost

85 /100 (B)

Další použití vzniklého díla by bylo usnadněno detailnějším popisem návrhu a více komentáři v kódu, pravděpodobně i lepší strukturou kódu. I přes to mi výsledek přijde jako hodnotný: jednak vzhledem k popsáním existujícím řešením a i k tomu, že některé banky tento úkol zdá se stále řeší zcela manuálně (část 4.4.). Jako hodnotný mi přijde i vzniklý dataset.

#### Celkové hodnocení

90 /100 (A)

Ačkoliv získané výsledky nejsou vůbec přesvědčivé, dokonce jsou horší než navržený heurtický algoritmus, práce odvedená na jejich získání mi přijde jako velmi dobrá a objemná, a to včetně většiny písemné části, která vytýkám především nedobrý popis vlastního návrhu řešení.

#### Otázky k obhajobě

Jak bylo časové náročné vytvoření vlastního datasetu? Plánujete tento dataset zveřejnit? Pokud ne, proč?

## **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.