



# Hodnocení vedoucího závěrečné práce

**Vedoucí práce:** Mgr. Martin Mareš  
**Student:** Bc. Justína Kušpálová  
**Název práce:** Odezírání  
**Obor / specializace:** Znalostní inženýrství  
**Vytvořeno dne:** 5. června 2024

## Hodnotící kritéria

### 1. Splnění zadání

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

Práce popsala současné přístupy a modely pro strojové odezírání. Architektura modelu s nejvyšší úspěšností byla pak využita pro několik experimentů, které vedly na vytvoření prvního českého modelu pro strojové odezírání. Při vytváření modelu nebyly příliš využity existující jazykové nástroje pro češtinu a experimenty s architekturou byly pouze základní. V práci se podařilo vytvořit automatický proces pro vytváření anotovaných dat pro potřeby strojového učení, kvalita procesu byla ověřena pomocí porovnávání s existujícím anglickým datasetem. Tento automatický proces byl využit pro vytvoření prvního veřejného českého testovacího datasetu.

### 2. Písemná část práce

72 / 100 (C)

Práce je psaná v angličtině. Text je čitelný, práce dodržuje dobré stylistické zásady, tvrzení jsou podložena zdroji. Text poutavou formou popisuje odezírání z pohledu obecné lingvistiky, a široce popisuje teorii zvolených metod strojového učení a odezírání. Množství teorie v textu v některých případech upozaduje experimenty, které by v některých případech zasloužily více popsat (např. výsledná architektura modelů). Některá rozhodnutí by pak mohla být více zdůvodněna (např. odstranění čísel z datasetu).

### 3. Nepísemná část, přílohy

85 / 100 (B)

Přílohy práce obsahují český testovací dataset, natrénované modely, výsledky experimentů a zdrojové kódy použité pro vytváření testovacího datasetu, a zdrojové kódy pro učení a vyhodnocování modelů. Zdrojový kód je čitelný, dobře členěný a jednotlivé

komponenty jsou dostatečně popsány. Výsledný kód by měl být více konzistentní v používání komentářů.

#### 4. Hodnocení výsledků, jejich využitelnost

75 /100 (C)

Práce vytvořila dostupný testovací dataset pro potřeby odezírání, zároveň vytvořila i nástroj, který do budoucna umožní tento dataset automaticky rozšiřovat. Došlo k vytvoření českého modelu, který byl založený na současných poznacích pro odezírání z jiných jazyků a tento model byl vyhodnocen. Výsledky českého modelu sice nedosahují anglických konkurentů, ale za dobrých podmínek umožňují chápat kontext odezírání vět.

#### 5. Aktivita studenta

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

Studentka pravidelně chodila na schůzky a na schůzkách prezentovala svůj postup. Studentka samostatně komunikovala s Českou televizí pro potřeby vytváření testovacího datasetu.

#### 6. Samostatnost studenta

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

Studentka samostatně vyhledávala související práce a teoretické podklady, v průběhu dosahovala menší samostatnosti při určování jednotlivých dílčích kroků pro řešení práce.

### Celkové hodnocení

75 /100 (C)

Odezírání je složitější téma, které bylo navíc komplikované nedostupností českého datasetu. V práci se podařilo vytvořit nástroj pro automatické anotování a tím vznikl první český veřejný testovací dataset. S pomocí získaných dat vznikl český model, který vyhodnotil úspěšnost několika základních přístupů k odezírání z jiných jazyků pro odezírání v českém jazyce.

## 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.