

Posudek oponenta bakalářské práce

Autor práce: **Matěj Zeman**

Název práce: **Automatické rozpoznávání dialogových aktů v tištěných textech**

Bakalářská práce Matěje Zemana se zabývala klasifikací dialogových aktů v podobě rozpoznávaného textu z rastrových obrázků. Práce obsahuje tři části: vytvoření obrázkového datasetu, rozpoznání textu pomocí metod OCR a poté samotná klasifikace dialogových aktů.

Proces vytváření datasetu je automaticky prováděn vytvořením MS Word dokumentů a jejich následným převodem do obrázků jednotlivých stránek ve formátu JPG. Oceňuji, že je zde snaha o vytvoření více datových sad a obrázků z různou úrovní šumu. Pro rozpoznávání textu je použit OCR systém Tesseract a pro klasifikaci dialogových aktů jsou použity tři typy neuronových sítí. Jako původní zdroj pro texty slouží dataset Verbmobil, který obsahuje přepis dialogů v angličtině a němčině.

Kvalita řešení a dosažených výsledků

Programy a skripty jsou psány v programovacím jazyce Python a při jejich spuštění jsem nenarazil na žádné problémy. K dodanému programovému vybavení mám jednu menší výhradu. Při spuštění dojde k vytvoření všech 8 datasetů s různou úrovní šumu. Bohužel nelze využít parametry příkazové řádky a cíleně spustit skript pouze na vytvoření jednoho datasetu s požadovanou úrovní šumu.

Formální úroveň

Rozsah textu 37 stran bych považoval za adekvátní bakalářské práci. Text je psán česky, ale na str. 23 se z nějakého důvodu objeví část textu anglicky! Ačkoliv je možné použít „spellchecker“, tak se v textu vyskytují překlepy a autorovi se nepodařilo vyvarovat se i pravopisných chyb (např. str. 22 sekce 5.2.4 poslední věta).

K textu obecně **nemám větší výhrady**, ale některé sekce by si zasloužily dopracovat. Např. v sekci 4.1.2 *Binarizace* se autor omezuje pouze na globální prahování i když existují jiné sofistikované metody. Dále bych upozornil na špatnou kvalitu diagramu na obrázku 5.1 (str. 19).

Dále je v textu na str. 23 uvedeno: „... budou použity dvě konvoluční neuronové sítě a jedna síť typu LSTM...“. Chybí mi zde aspoň stručný rozbor nebo rešerše, zda se tento typ neuronových sítí v této oblasti používá a zda jsou tyto modely vhodné. Čtenář bez předchozí zkušenosti může tuto informaci postrádat.

V textu mi obecně schází lepší provázanost jednotlivých sekcí v rámci kapitol. Chybí informace o tom proč daná sekce je tam, kde je a co zrovna vysvětluje.

Práce s literaturou

Práce disponuje 20 zdroji, které obsahují online zdroje, publikované odborné články a také tištěné knihy. U některých obrázků chybí citace a celkově je prostor pro zlepšení při práci se zdroji v textu.

Splnění zadání

Zadání považuji za splněné.

Interpretace výsledků

Výsledky experimentů jsou poskytnuty formou tabulek s přesností (*Accuracy*) jednotlivých modelů. V experimentech postrádám srovnání s případem, kdy se použije přímo původní text. Nevíme tak, o kolik procent se úspěšnost zhorší použitím rozpoznávaného textu z obrázků.

Na závěr konstatuji, že autor se s **relativně těžkým a rozsáhlým zadáním vypořádal poměrně dobře**. Prokázal, že je schopen řešit problémy, úlohu dekomponovat a provést sadu experimentů. Prostor ke zlepšení vidím v prezentaci práce a text by si rozhodně zasloužil více péče (viz poznámky výše).

Dotazy k práci

1. Proč jste pro binarizaci obrázků použil globální prahování, když jsou k dispozici vhodnější metody (např. lokální Sauvolovo prahování), který právě pro tyto případy dosahuje výrazně lepších výsledků?
2. Proč jste pro vytvoření obrázkové datové sady použil nejprve konverzi do MS Word dokumentu (resp. poté ještě do PDF souboru) a až pak do jednotlivých obrázků JPG? Co bránilo přímému „vykreslení“ textu do obrázků?

Navrhuji hodnocení známkou **velmi dobře** a práci doporučuji k obhajobě.

V Plzni 18.5.2021

Ing. Jiří Martínek