

---

Posudek oponenta bakalářské práce

---

Jakub Fafek  
Klasifikace skenovaných dokumentů

---

Cílem bakalářské práce pana Jakuba Fafka bylo navrhnout a implementovat klasifikátor skenovaných dokumentů. Natrénovaný model má být schopen rozlišit dokumenty do celkem deseti tříd. Výsledný klasifikátor vznikl spojením dvou modelů, jejichž vstupem je textová resp. vizuální vrstva skenovaného dokumentu. Textová vrstva byla extrahována pomocí knihovny Tesseract OCR. Trénink i testování implementovaného modelu bylo provedeno nad standardním volně dostupným datasetem Tobacco-3482. Na úvod je vhodné říci, že se jedná již o druhou iteraci řešení.

Při hodnocení výsledného softwarového produktu jsem postupoval podle návodu v příloze A.1 *Příprava*. Zde autor uvádí, že ke spuštění příložených skriptů je nutné mít nainstalovaný Python verze 3.7.9, což je v současné době obrazně řečeno prehistorická verze, která je dle oficiální dokumentace téměř rok *po smrti* a pro nové projekty by neměla být použita. Po extrakci adresáře `Aplikace_a_knihovny` jsem dle textu práce spustil instalaci závislostí, která ovšem skončila chybovým hlášením o tom, že některé ze závislostí vyžadují Python alespoň verze 3.8. Nakonec se mi skripty podařilo spustit po úpravě souboru `requirements.txt` a přechodu na novější verzi jazyka Python. Vzhledem k tomu, že je použitý dataset součástí zadání práce, považuji za nešťastné po uživateli vyžadovat jeho manuální stažení a postupné předzpracování pro potřeby trénování – část uživatelské příručky by se implementací jednoduchého skriptu tak mohla značně zjednodušit. Zdrojový kód je tvořen cca 2 032 řádky rozdělenými do 32 modulů. V některých případech autor vcelku nelogicky *vynalézá kolo*, např. implementace vlastního indikátoru průběhu namísto použití standardního modulu `tqdm` apod.

Vytvořený průvodní dokument obsahuje 37 normostran (74 146 znaků) a 5 stran tištěných příloh. Práce je včetně úvodu a závěru členěna do 6 kapitol. Dokument je vysázen v typografickém systému  $\LaTeX$  a je na dobré vizuální úrovni. V práci se občas vyskytují překlepy, gramatické chyby či odchylky od typografických pravidel. Například v sekci 5.2.4 *Filnální verze* se autor kromě ponechání jednopísmenné neslabičné předložky na konci řádku pokouší o „převádění veklých písmen“. Z dalších příkladů můžeme uvést klasické nesprávné použití spojovníku a pomlčky, občasné opakování slov či neadekvátně velké obrázky vedoucí k téměř prázdným stránkám. Stejně tak diagramy 2.12 nebo 2.13 působí mírně ledabylým dojmem, graf 2.14 nemá popsání osy a obrázku 2.15 chybí adekvátní popis. Nutno říci, že počet chyb se ke konci dokumentu zvyšuje, což by mohlo napovídat pozdějšímu vypracování. Subjektivní vyjádření se v textu téměř nevyskytují. Citace v textu jsou dostatečné a provedené v souladu s požadavky. Některé pasáže by bylo vhodné podrobněji rozpracovat, konkrétně pak sekce 2.2.3 *Transformer*, 2.7.3 *BERT* nebo 2.8.1 *EfficientNet*, které popisují techniky později použité v praktické části práce.

Předložená bakalářská práce splňuje zadání ve všech bodech. Výsledný softwarový produkt i použité techniky byly dostatečně popsány. Původně jsem na svůj posudek práce z loňského roku nechtěl vůbec odkazovat, ale závěrem je na místě uznat, že došlo k výraznému zlepšení a většina připomínek byla zapracována.

I přes značné množství drobných výtek práci považuji za obstojnou, **doporučuji ji k obhajobě** a navrhuji hodnocení kvalifikačním stupněm

„dobře“.

V Plzni dne 23. května 2024

**Ing. František Pártl**  
NTIS FAV ZČU

**Doplňující otázky:**

1. Proč jste si pro praktickou část práce vybral Python verze 3.7, který je podle oficiální dokumentace již téměř rok "po smrti"?
2. Jaký má význam výraz  $[H]$  v rovnici 2.1 na straně 14?
3. Obrázek 2.15 na straně 26 jste nedostatečně popsal slovy „Model BERT“. Jaký význam mají vyobrazené popisky „Pre-training“, „Fine-Tuning“, „MNLI“, „NER“, „SQuAD“ aj.?
4. V teoretické části práce hovoříte o rozdělení datové sady na trénovací, testovací a validační část. Proč v tabulkách 5.1 až 5.6 porovnáváte výsledky dosažené na testovací sadě a ne na validační?