

Posudek oponenta bakalářské práce

Student: Kačur Ján
Téma: Generování kódu z textového popisu funkcionality (id 24065)
Oponent: Ondřej Karel, Ing., UPGM FIT VUT

- 1. Náročnost zadání** průměrně obtížné zadání
Práce vyžaduje nastudovat principy strojového učení z oblasti generování textu, implementaci jednoduchého demonstračního systému a jeho vyhodnocení.
- 2. Splnění požadavků zadání** zadání splněno s vážnými výhradami
Výhradu mám k vyhodnocení systému, které by mělo být podrobnější a jasnější. Chybí diskuze o kvalitě navržených funkcí systémem na základě podrobnosti, resp. složitosti, jejího popisu.
- 3. Rozsah technické zprávy** je v obvyklém rozmezí
Práce obsahuje 47 vysázených stran, nicméně velké množství obrázků nepřináší žádnou přidanou hodnotu.
- 4. Prezentací úroveň předložené práce** 65 b. (D)
Jednotlivé části na sebe většinou logicky navazují, nicméně některé části by zasloužily rozšířit a jiné obsahují nezáživné implementační detaily.
Kapitola o generování zdrojových kódů, resp. obecně generování textu, včetně aktuálního stavu řešení, by měla být výrazně rozsáhlejší a autor by měl srovnat jednotlivé řešení a vyhradit se proti nim.
Kapitola popisující implementaci připomíná spíše podrobnou dokumentaci jednotlivých funkcí a datových struktur, včetně častých ukázek zdrojových souborů. Naopak samotný model je popsán velmi nedostatečně a bylo by velmi obtížné reprodukovat výsledky pouze na základě textu práce.
- 5. Formální úprava technické zprávy** 75 b. (C)
Mluvnickou stránku nejsem schopen posoudit, jelikož je práce psána ve slovenském jazyce, nicméně obsahuje větší množství zřejmých překlepů a několik typografických chyb:
 - Malé závorky v rovnici (2.1), funkce kurzívou, např. logaritmus v rovnici (2.6), znak * místo násobení v rovnici (2.6).
 - Chybějící interpunkční znaménka v seznamech nebo popiscích obrázků.
 - Obrázky (strana 25, obrázek 5.9) a útržky kódu (strana 28, obrázek 6.1) v malém rozlišení, často jako print screen.
- 6. Práce s literaturou** 75 b. (C)
Použitá literatura obsahuje 27 referencí, včetně několika online blogů. Samotný formát referencí není jednotný a často jsou uvedeny pouze nejzákladnější informace.
- 7. Realizační výstup** 60 b. (D)
Student implementoval systém pro generování zdrojových kódů na základě dokumentačního řetězce a deklarace funkce v jazyce Python, včetně jednoduché demo aplikace a modulu do textového editoru Sublime.
Při vyhodnocení systému chybí srovnání s jinými modely, takže je obtížné interpretovat co znamená "přesnost 85%". Samotná metrika "přesnost" není formálně definována, pouze slovně a pomocí ukázky části kódu, a její vhodnost v dané oblasti je též sporná. Zdrojové soubory nejsou v podstatě zdokumentované a není jasné, které části jsou převzaty a které jsou vlastním dílem autora.
- 8. Využitelnost výsledků**
Student sice položil v práci hypotézu ohledně přínosu přidání informace o syntaktické analýze dokumentačního řetězce, nicméně ji při vyhodnocení vhodně neověřuje. Odezva systému při generování relativně krátkých funkcí je na průměrném laptopu v řádech několika minut, což ho činí v praxi nepoužitelný.
- 9. Otázky k obhajobě**
 1. Diskutujte kvalitu navržených funkcí na základě jejich popisu.
 2. Výsledky v kapitole 7.1 jsou na trénovací nebo testovací datové sadě?
 3. Srovnajte architekturu GPT-3 s Vámi implementovaným modelem.
- 10. Souhrnné hodnocení** 60 b. uspokojivě (D)
Největší nedostatky práce vidím v nedostatečném vyhodnocení systému, včetně analýzy výstupů systému. Zároveň není jasné, zda jsou výsledky reportované na trénovací nebo testovací datové sadě. Na základě odpovědí na otázky zmíněných výše navrhuji komisi hodnotit práci známkou D-E.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 12. srpna 2021

Ondřej Karel, Ing.
oponent