



Posudek oponenta závěrečné práce

Oponent práce:	Mgr. Vojtěch Rybář
Student:	Tomáš Machorek
Název práce:	Dotazování nad relačními databázemi s pomocí velkých jazykových modelů (LLM)
Obor / specializace:	Umělá inteligence 2021
Vytvořeno dne:	11. 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

Zadání bylo formálně splněno.

2. Písemná část práce

65 /100 (D)

Práce je dobře strukturovaná a spíše kratšího rozsahu. Obsahuje samonosný popis problémů a nástin řešení, které nastávají při dotrénování jazykového modelu pro specifickou aplikaci s omezeným výpočetním výkonem.

Z textu není jasná motivace některých kroků, např. proč v třetí kapitole jsou diskutovány pouze dva modely.

Čtvrtá kapitola obsahující popis zlepšování výstupu modelu pro specifickou aplikaci, by mohla být podrobnější a konkrétnější.

3. Nepísemná část, přílohy

60 /100 (D)

Příložený kód obsahuje Jupyter notebooky s kódem demonstrujícím jak lze dotrénovat jazykové modely v prostředí Google Colab.

Webová aplikace velmi dobře slouží zamýšlenému účelu, tedy evaluaci upraveného modelu na vlastních databázích.

Vlastní testování dotrénovaných modelů v dodané aplikaci s jednoduchou databází ukazuje prostor pro možná zlepšení.

4. Hodnocení výsledků, jejich využitelnost

75 /100 (C)

Získané poznatky jak zvyšovat přesnost jazykových modelů při omezené výpočetní kapacitě a vzniklá aplikace by mohly tvořit dobrý základ pro další úsilí v této oblasti, která prochází aktuálně bouřlivým vývojem.

Celkové hodnocení

69 /100 (D)

Práce obsahuje určité nedostatky jak v písemné, tak experimentální části. Ty jsou ale vzhledem ke komplexitě řešeného problému očekávatelné a nesnižují kvalitu práce.

Otázky k obhajobě

1. Popište prosím podrobněji postup, kterým jste zlepšoval kvalitu modelu pro uvažovanou aplikaci.
2. Bylo by podobných výsledků možné dosáhnout metodou Few-Shot Learning?
3. Bylo by srovnatelných výsledků možné dosáhnout zapojením zpětné vazby, tedy např. zahrnutím neúspěšného SQL dotazu s chybovým hlášením do následného promptu pro model?

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.