

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Improving Performance of Natural Language Inference Models for Numerical Queries
Jméno autora:	Viktor Emil Korladinov
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Vedoucí práce:	Ing. Jan Drchal, PhD.
Pracoviště vedoucího práce:	Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
NLP není v rámci bakalářského studia vyučováno. Strojové překlady dat, které práce vyžaduje jsou časově náročné a je tudíž třeba dobře plánovat rezervaci výpočetního času.	

Splnění zadání	splněno
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Zadání je splněno bez výhrad. Původně jsme plánovali v rámci práce naučit (případně ladit) specifické existující modely typu NLI. Ukázalo se, že nejrozsáhlejší datová sada Ape210k je zatížena problémem tzv. <i>external knowledge</i> , kdy zadání slovních úloh vyžaduje informace neobsažené přímo v textu. Rozhodli jsme se tedy naučit jiný typ modelu, který tuto zhruba třetinu dat odfiltruje. Vzhledem ke značnému pokroku v oblasti Large Language Models (LLMs) jsme se navíc rozhodli otestovat schopnosti modelu GPT3.5 v režimu zero-shot learning.	

Aktivita a samostatnost při zpracování práce	A - výborně
<i>Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posuďte schopnost studenta samostatně tvůrčí práce.</i>	
Se studentem jsem byl velmi spokojen. Na pravidelné konzultace byl vždy výborně připraven	

Odborná úroveň	A - výborně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
K odborné úrovni práce nemám výhrady. Všechny kroky jsou dobře motivovány. Vzhledem k neexistenci vhodně anotovaných dat, se student rozhodl pro ruční anotaci a evaluaci nezanedbatelného rozsahu. K tomu účelu si implementoval specializovaný nástroj.	

Formální a jazyková úroveň, rozsah práce	A - výborně
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je psána velmi kvalitní angličtinou. Text je čtivý a dobře strukturovaný. Úvodní kapitoly velmi pěkně shrnují existující datové sady i metody automatizovaného řešení slovních (numerických) úloh. Formální zápis není příliš zastoupen, ale to je dáno tím, že těžištěm práce je hlavně analýza datových sad. Z typografického hlediska je práce na vysoké úrovni.	

Výběr zdrojů, korektnost citací	A - výborně
<i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.</i>	

Zdroje jsou citovány správně.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Hlavními výstupy práce jsou přeložené datové sady a jejich analýza. Velkým, a přitom neplánovaným, přínosem je studentem vytvořená datová sada a na ní naučený model pro filtraci úloh, které pro své vyřešení potřebují dodatečné informace (*external knowledge*). Se studentem budeme dále spolupracovat – výstupy jeho práce plánujeme využít pro vylepšení posledních verzí Natural Language Inference modelů, které vyvíjíme v rámci automatizovaného systému pro fact-checking.

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.

K práci nemám výhrady, proto ji hodnotím klasifikačním stupněm **A - výborně**.

Datum: 10.6.2023

Podpis: