

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	NLP Methods for Word Problems
Jméno autora:	Ronald Krist
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Vedoucí práce:	Ing. Jan Drchal, PhD.
Pracoviště vedoucího práce:	Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Práce vyžaduje dobrou znalost moderních metod strojového učení pro zpracování přirozeného jazyka, které nejsou systematicky pokryty výukou na FEL. Experimenty jsou výpočetně náročné, proto je nutno je dobře naplánovat.	

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 bylo splněno bez výhrad.	

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>	
Student byl po celou mimořádně aktivní a na všechny konzultace byl vždy připraven. Chci jej pochválit za samostatnost – automatizované řešení slovních úloh se odchyluje od problémů, které běžně řešíme v rámci výzkumné skupiny, proto si musel student velkou část řešené domény nastudovat zcela sám. Tohoto úkolu se zhostil výborně.	

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 nemám námitek. Práce obnášela zejména předzpracování (hlavně překlad) řady cizojazyčných datových sad a dále pak experimenty s vybranými neuronovými modely. Oceňuji systematickosti, s níž si student stanovil hypotézy a volbu experimentální metodologie. Rozsah experimentální části je značný a diskuze experimentů je detailní. Za velmi zajímavé považuji, mimo jiné, pozorování, které naznačuje, že dříve publikovaný model „Gri“ byl testován na nereprezentativní datové sadě, což vedlo k příliš optimistickým závěrům. Důležitým výsledkem je potvrzení hypotézy, že je strojový překlad datových sad dobře použitelným přístupem pro řešení slovních úloh zadaných v českém jazyce.	

Formální a jazyková úroveň, rozsah práce	B - velmi dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je psána výbornou angličtinou s minimem překlepů. Text je dobře strukturovaný a čtivý. Kromě již zmíněných částí týkajících se metodologie a experimentů chci pochválit i úvodní sekce s velmi pěkným popisem problematiky řešení slovních úloh: student pokrývá historii domény, přehled datových sad i moderních solverů. Menší výhradu mám k odkazům na tabulky a obrázky, kdy není vždy použito příslušné označení – v některých případech to komplikuje čtení. Rovněž bych doporučil v tabulkách zvýrazňovat důležité výsledky, a to například tučným řezem písma.	

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</i>	

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.

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.

Zadání práce vychází z problémů, které jsme zaregistrovali při vývoji automatizovaného systému pro ověřování faktů.

V současné době nejrozšířenější NLP modely typu Transformer nasazované na úlohy typu Natural Language Understanding/Inference mají často problémy se zpracováním číselných údajů a obecně numerických vztahů mezi entitami. Předpokládáme, že nám datové sady a zkušenosti získané trénováním modelů v rámci této práce, pomohou stávající stav zlepšit.

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

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

S přístupem studenta i výstupy jeho práce jsem celkově velmi spokojen. Drobnou výhradu jsem měl jen k textu. Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 8.6.2022

Podpis: