

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Active Learning for NLP
Jméno autora:	Anton Kretoř
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Vedoucí práce:	Ing. Jan Drchal, PhD.
Pracoviště vedoucího práce:	AIC

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Student musel detailně nastudovat problematiku dvou oblastí: moderních jazykových modelů a aktivního učení. Experimenty byly časově náročné, proto bylo třeba dobře plánovat.	

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	B - velmi dobře
<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 mimořádně samostatný: velkou část doby, kdy na tématu pracoval, trávil v zahraničí. Alespoň ve finální fázi bych ocenil více schůzek (telefonních hovorů) – mohli bychom asi lépe doladit strukturu textu. Každopádně s výsledkem jsem spokojen.	

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>	
Práce staví na detailní přípravě: rešerši provedené na zadané téma kombinace aktivního učení a modelů strojového učení pro NLP. Výstupem je pak 1) kvalitní knihovna implementující state-of-the-art metody aktivního učení, 2) rozsáhlá experimentální část, kde jsou jednotlivé metody a jejich meta-parametry porovnány z mnoha hledisek na vhodně vybrané množině datových sad.	

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 kvalitní angličtinou s výjimečnými překlepy. Text je rozsáhlý, ale v některých pasážích je možná až příliš rozvlákný. Struktura textu je v pořádku. Líbí se mi detailní popis metodologie, ale například sekci 4.1.6 bych připojil přímo k popisu experimentů – oddělení těchto textů mě nutilo k častému listování zpět. Metoda CoreSet a zejména Kmeans++, použitý v rámci studentem testované metody BADGE, by si zasloužily o něco více detailů, nepovažuji to ale za významný problém. K formálnímu zápisu nemám výhrady. Typografie textu lehce trpí přímým kopírováním grafů z Weights and Biases – jsou sice vhodné pro interaktivní prohlížení výsledků, ale jejich popisky jsou hůře čitelné v dokumentech určených k tisku.	

Výběr zdrojů, korektnost citací	A - výborně
--	--------------------

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.

Reference jsou citovány správně. Rozsah odkazované literatury je nadstandardní.

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.

Výsledky práce jsou velmi důležité pro výzkum metod NLP a zejména sběru datových sad, které provádíme v rámci NLP skupiny AIC. Práce jako bonus obsahuje velmi pěkný přehled nástroje Weights and Biases.

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

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

S prací studenta jsem byl velmi spokojen. I přes drobné výhrady k textu předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 3.2.2023

Podpis: