

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Aktivní učení pro metody zpracování přirozeného jazyka
Jméno autora:	Anton Kretoř (474672)
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Oponent práce:	Mgr. Pavel Schlesinger (externista)
Pracoviště oponenta práce:	TRASK SOLUTIONS a.s.

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	mimořádně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Práce vyžaduje detailní nastudování tématu aktivního učení a spjatých moderních technik strojového učení, experimentální část vyžaduje orientaci v realizaci MLOps.	

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>	
Student splnil zadání ve všech bodech bez výhrad, s nadprůměrným rozsahem práce.	

Zvolený postup řešení	vynikající
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Zvolený postup řešení považuji za správný. Oceňuji systematický postup i popis v práci. Student shrnul state-of-the-art stav problematiky, provedl časově náročnou experimentální část, která obsahuje i shrnutí jeho vlastního přínosu.	

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>	
Odbornou část hodnotím velmi pozitivně. Čtenáři se dostává do ruky ucelená práce, která přehledně zachycuje nejmodernější stav a trendy. Za nadměru náročnou považuji experimentální část v těchto bodech: 1. rozsah experimentů (šíře), 2. parametrizace (hloubka), 3. textové a grafické shrnutí. Student prokázal nadhled a neztratil koncentraci v detailech, což svědčí o erudici i schopnosti samostatně pracovat.	

Formální a jazyková úroveň, rozsah práce	A - výborně
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Student práci sepsal výbornou angličtinou s minimem překlepů. Text je dobře strukturovaný a čtivý.	

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>	
Reference práce a zpracování zdrojů považuji za příkladné: 9 stránek citací, odkaz na významné state-of-the art články, zdroji jsou významné časopisy a konference, nechybí ani citace prací z posledního období včetně prací z roku 2022.	

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.

Student dosáhl state-of-the-art výsledku při evaluaci Czech NLI, tj. partikulární části experimentu. Navíc je patrná velká pracnost experimentální části jako celku, student provedl více než 161 konfigurací pro následně provedené experimenty, z kterých nakonec vybral ty nejzajímavější výsledky.

Za zásadnější považuji přínos jakým způsobem toho student dosáhl, totiž vytvořením vlastní knihovny sdílené na GitHubu. Toto workflow je v práci popsáno. Zapojuje integraci RCI clusteru, využití MLOps (Weights & Biases), dovoluje metodické vedení návrhu a realizaci experimentu pro úlohy aktivního učení.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Diplomová práce je dobře strukturovaná a mimořádně čtivá, např. v experimentální části je studentův popis transparentní, detailní a čtivý zároveň. Student dokáže vhodně zvolit formu. (Např. srovnajme Fig. 2.7. na str. 18, formálně precizní algoritmus versus ich-forma v postupech a nastaveních na str. 31 či st. 37.)

Student svým popisem a vloženými návody a zdrojovým kódem dovoluje **reproducibilitu své praktické části**. Práce zároveň dovoluje na ni navázat a přepoužít výsledky (formou kódu, workflow, metodiky nastavení).

Bude-li možno, poprosil bych studenta se vyjádřit k následujícímu:

1. V experimentální části a kap 4.2. se student věnuje menším jazykovým modelům. Jako oponent postrádám hlavní závěry a shrnutí, resp. doporučení, které z této části lze vyvodit.
2. Kap. 6 pro shrnutí a částečně kap. 5. pro budoucí využití práce částečně popisuje přínosy a možnosti, nicméně jako oponent bych ocenil shrnutí, hlavní pozorování z proběhlých experimentů a doporučení jak postupovat na začátku využití Active Learning v praxi. Existuje takové zjednodušené doporučení, může student doložit?

Mé výtky jsou ovšem zanedbatelné v porovnání s kvalitou a rozsahem, které práce ve svém celku nabízí.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 3.2.2023

Podpis: