

**I. IDENTIFIKAČNÍ ÚDAJE**

<b>Název práce:</b>	<b>Zkvalitňování modelů Natural Language Inference pro numerické dotazy</b>
<b>Jméno autora:</b>	<b>Korladinov Viktor Emil</b>
<b>Typ práce:</b>	bakalářská
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra kybernetiky
<b>Oponent práce:</b>	Ing. Jiří Kubalík, Ph.D.
<b>Pracoviště oponenta práce:</b>	CIIRC, ČVUT

**II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ**

<b>Zadání</b>	<b>náročnější</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Jedná se o náročnější zadání bakalářské práce. Aby student mohl splnit požadavky zadání, musel nejprve nastudovat principy a fungování současných velkých jazykových modelů (LLM - Large Language Model), specializovaných modelů pro inferenci v přirozeném jazyce (Natural Language Inference) a modelů pro řešení matematických slovních úloh (Math Word Problem solvers). Všechno to jsou velice pokročilé metody, pro jejichž pochopení je třeba obsáhnout množství netriviálních znalostí.	

<b>Splnění zadání</b>	<b>splněno</b>
<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 splněno bez výhrad.	

<b>Zvolený postup řešení</b>	<b>vyňikající</b>
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Ke zvolenému postupu nemám výhrad. Student nejprve nastudoval hlavní zástupce ze všech sledovaných tříd modelů. Oceňuji, jak se mu podařilo stručně a věcně vystihnout hlavní rysy popisovaných modelů a metod. Dále zpracoval přehled několika nejvýznamnějších datasetů, které se používají pro učení a testování studovaných modelů. Nakonec navrhl tři experimenty pro otestování LLM a pro zkvalitnění datasetu Ape210K. Všechna svá rozhodnutí při návrhu experimentů správně motivoval a výsledky experimentů podrobně analyzoval.	

<b>Odborná úroveň</b>	<b>A - výborně</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Po odborné stránce je práce na velice dobré úrovni. Zvláště oceňuji, jak se student vypořádal s četnými problémy, které se během řešení vyskytly.	

<b>Formální a jazyková úroveň, rozsah práce</b>	<b>A - výborně</b>
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Formální i jazyková úroveň je opět nadprůměrná. Nemám žádných výhrad. Práce je dobře strukturovaná, myšlenky a postřehy jsou jasně formulovány.	

<b>Výběr zdrojů, korektnost citací</b>	<b>A - výborně</b>
<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>	
Výběr i počet citovaných zdrojů je nadprůměrný. V naprosté většině případů se jedná o články z posledních několika málo let. To opět mluví ve prospěch aktuálnosti předložené práce.	

**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.*

**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.*

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

Do diskuze mám následující dotaz:

- Pročištění datové množiny Ape210K je jistě krokem, který umožní ladit přesnější MWP modely. Mohl byste uvést i jiný příklad toho, jak by šlo rozšířit stávající data s cílem zlepšit MWP modely?

Datum: 12.6.2023

Podpis: