

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Zkvalitňování fakticity generovaných sumarizací
Jméno autora:	Halama Václav
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	KP
Oponent práce:	Gustav Šír
Pracoviště oponenta práce:	KP

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zamerení na fakticnost sumarizaci je docela nova, a tedy narocna, vec.	

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>	
Bez vyhrad splneno.	

Zvolený postup řešení	vynikající
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Postup je vyborny, i kdyz evaluace fakticnosti neni asi uplne samostatnym modulem, jak bylo mozna puvodne zamysleno.	

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>	
Student pracuje korektnim zpusobem se state-of-the-art metodami.	

Formální a jazyková úroveň, rozsah práce	A - výborně
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Textova cast prace je opravdu povedena, profesionalni, a dobre se cte.	

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>	
Bez problemu.	

Další komentáře a hodnocení
<i>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.</i>
Viz celkove hodnoceni.

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.

Jedná se o velmi peknou, silně implementačně zaměřenou, práci testující nejrůznější přístupy k sumarizaci textu. Student vhodně staví nad existujícími předtřetovanými modely pro extraktivní a abstraktivní sumarizaci a detailně zkouma jejich použití se zaměřením na faktičnost sumarizace, k čemuž navrhuje i vlastní metriku BERTSource.

Otázka:

Faktičnost sumarizaci je zjevně relativně specifická a obtížně hodnotitelná věc, kdy ultimátním ohodnocením je stejně nakonec člověk – nemelo by smysl ho nějak „aktivněji“ zapojit rovnou do procesu trénování, třeba stylem RLHF, a ne jen pasivně na konci, či na začátku ve fázi filtrování?

Zaroveň by mi tak přišlo vhodně nehodnotit 1/0 ale na nějaké relativní skále (user ranking), a detailněji na úrovni jednotlivých problematických částí/slov shrnutí pro detailnější feedback do modelu.

Take by bylo velmi zajímavé zjistit zero-shot performance chatGPT pro těch 100 článků s použitím vhodného promptingu s ukázkami těch faktografických chyb jako priming.

Fig. 2.2

- A_B_E -> A_C_E ?

Fig. 3.2

- Je to beam 2 nebo 3?

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

Datum: 5.6.2023

Podpis: