

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Hodnocení obsahu mediálních zpráv
Jméno autora:	Ondřej Mězl
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačové grafiky a interakce
Oponent práce:	Ing. Jan Drchal, PhD.
Pracoviště oponenta práce:	Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Řešení NLP části zadání vyžaduje zvládnutí poměrně běžných typů klasifikačních modelů. Vizualizační část může být náročnější v případě implementace dynamických vizualizací.	

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>	
Všechny body zadání byly splněny.	

Zvolený postup řešení	částečně vhodný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Zvolené řešení je v zásadě správné, přesto musím vytknout použití starších, méně robustních metod. Současné state-of-the-art klasifikátory textu jsou typicky postaveny na vekých předučených jazykových modelech, které jsou doladěny na cílové doméně, typicky s použitím již relativně malého datasetu. Student sice v textu krátce zmiňuje právě modely typu BERT, ale bohužel je nijak netestuje. Předpokládám, že s těmito novějšími modely postavenými na kontextových reprezentacích by mohlo být dosaženo lepších výsledků. U vizualizační části práce postrádám nějaké dynamické (interaktivní) vizualizace, myslím, že by z uživatelského hlediska byly přínosné. Prezentované zobrazení dat je většinou postaveno na základních grafech vykreslených knihovnou <i>matplotlib</i> . Nástroj Doccano je spíše zamýšlen pro anotaci textu, než pro jeho vizualizaci, proto mě jeho využití překvapilo.	

Odborná úroveň	D - uspokojivě
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Práce má nižší odbornou úroveň. Práce takřka nepopisuje použité datové sady, stejně tak nejsou jasné detaily experimentů, chybí např. zmínka o učícím algoritmu, jeho parametrech, není uvedena ztrátová funkce, nejsou definovány hodnotící metriky, není jasné, co je přesně vstupem jednotlivých architektur. Reprodukovatelnost výsledků je tudíž nízká. Některé použité datové sady jsou veřejné a známé, postrádám srovnání vlastních výsledků s publikovanými daty.	

Formální a jazyková úroveň, rozsah práce	D - uspokojivě
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je psána češtinou s občasnými překlepy. Stylistika je na dosti nízké úrovni (viz např. úvod páté kapitoly). Formální zápis není využit. Struktura práce je často chaotická: např. sekce 2.1 „Rešerše“ řeší hned ve druhém a třetím odstavci přeučení a ve čtvrtém přeskakuje na informaci o tom, že nejběžnějším jazykem používaným pro strojové učení je Python. Podobně v rešerši vizualizační části (sekce 3.2) student přeskakuje mezi metodami vizualizace a vizualizačními nástroji. Text rešeršních částí nejde nikdy dostatečně do hloubky. V experimentální části postrádám nějaké ukázky zpracovávaných textů – bylo by například vhodné ukázat témata, která vedla k výraznějšímu kolísání sentimentu.	

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.

Zdroje jsou citovány správně. Počet citací odpovídá diplomové práci.

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.

Největší výtky mám k textu práce a její nízké reprodukovatelnosti. Přesto musím konstatovat, že objem odvedené práce není zanedbatelný. Práci hodnotím klasifikačním stupněm **D - uspokojivě**.

K práci mám následující otázku:

Jak jste postupoval při volbě architektury modelů – v textu jsem nenašel žádnou zmínku?

Datum: 19.6.2023

Podpis: