

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Data Analysis using Knowledge Networks
Jméno autora:	Arina Lebedeva
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra telekomunikační techniky
Oponent práce:	Ing. Jan Drchal, PhD.
Pracoviště oponenta práce:	Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Témata znalostních sítí a NLP nejsou pokryta výukou. Pro úspěšné řešení tématu práce je nutno věnovat hodně času samostudiu.	

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>	
Splněno bez výhrad.	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Studentka se po dobře motivovaném vysvětlení rozhodne pro návrh algoritmu postaveném na sadě pravidel. Vzhledem k neexistujícím trénovacím datům, považuji toto řešení za správné. V opačném případě by bylo robustnější nasadit metody strojového učení. Líbí se mi detail, se kterým ke tvorbě pravidel přistupuje i množství názorných ukázek, kterými je text doplněn. Jedinou výtku mám k tomu, že výstupy navržené metody jsou testovány příliš subjektivně. Myslím, že by bylo vhodné dopředu vyrobit alespoň malou testovací množinu a spočítat přesněji, kolika chyb se implementovaná metoda dopouští.	

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>	
K odborné úrovni nemám výhrady. Je zřejmé, že se studentka v problematice dobře orientuje.	

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 obstojnou češtinou s občasnými překlady – působí na mě jako text nepsaný rodilým mluvčím, který ale věnuje zvýšenou pozornost odstraňování chyb. Celkově se práce čte velmi dobře. Rešeršní část práce je pokryta kapitolou 2. Rešerše je dobře napsaná, ale v některých pasážích příliš spoléhá na omezený počet zdrojů (zejména zdroj [7]). Text pokračuje precizní analýzou datového korpusu. Drobnou výhradu mám k tomu, že se z textu nedozvíme více o původu a zdrojích dat. Jak již zmiňuji výše, návrh a implementace metody pro tvorbu znalostních grafů jsou popsány detailně a srozumitelně. Přesto bych text ještě doplnil o pseudokód případně o vývojový diagram, které by pomohly navržený algoritmus formálněji popsat. Z typografického hlediska nemám co vytknout.	

Výběr zdrojů, korektnost citací	B - velmi dobře
<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</i>	

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, přesto bych čekal větší objem citací než 11.

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.

Celkově na mě výsledek této práce působí velmi kladně. Myslím, že se jedná o dobrý základ pro další výzkum. I současné výstupy nad cílovým korpusem novinových textů mohou být uživatelsky velmi zajímavé – jak pro novináře, tak pro různé směry výzkumu v humanitních vědách.

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.

Vzhledem k menším výhradám (viz výše) předloženou závěrečnou práci hodnotím klasifikačním stupněm
B - velmi dobře.

K práci mám následující otázku:

Máte nějaké detailnější informace o tom, jakými metodami byla předzpracována surová textová data do formátu, který uvádíte v kapitole 3?

Datum: 1.6.2023

Podpis: