



POSUDEK VEDOUCÍHO DIPLOMOVÉ PRÁCE

Jméno studenta: Bc. David Illner
Název práce: Extrakce témat z nestrukturovaného textu
Autor posudku: Ing. Martina Husáková, Ph.D.
Cíl práce: Vytvoření modelu pro extrakci témat z nestrukturovaného textu s využitím analytické platformy KNIME a vyhodnocení výsledků modelu.

Povinná kritéria hodnocení práce	Stupeň hodnocení (známka)					
	A	B	C	D	E	F
Práce svým zaměřením odpovídá studovanému oboru	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vymezení cíle a jeho naplnění	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování teoretických aspektů tématu	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Využitelnost výsledků práce v teorii (v praxi)	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vyjádření k výsledku anti-plagiátorské kontroly

Antiplagiátorská kontrola vykazuje 0% podobnosti s jinými pracemi v systému Odevzdej.cz.

Díličí připomínky a náměty:

- Str. 3: úprava čísla stránky u poslední položky *Přílohy*
- Str. 6: Zdroj [2] umístit ještě blíže k článku, resp. do věty, která se o zdroji poprvé zmiňuje.
- Str. 12: „...výběr vzorků z vrstev rozdělených, tak aby se žádná...“
- Místy zvážit umístění zdrojů, viz např. na str. 24: „...předmět diskurzu nebo jeho části. [24] Ovšem v detekci témat je ...“, „...nad pevně daným slovníkem. [25] Každé téma je...“.
- Str. 35: „Platforma KNIME byla navržena se třemi hlavními objektivy...“

Celkové posouzení práce a zdůvodnění výsledné známky:

Diplomant vytvořil kvalitní práci, ve které se mu podařilo vytvořit model k detekci hlavních témat pro paletu odborných článků (v počtu 21 tis.) v platformě KNIME. Články byly získány z volně dostupného datasetu z webu Kaggle.com. Pro modelování témat využil jeden model a to model LDA. Z tohoto úhlu pohledu by bylo zajímavé použít více různých modelů a podrobit analýze jejich výstupy. Diplomant do teoretické části práce vložil klíčovou terminologii, kterou srozumitelně vysvětluje, místy vkládá doplňující příklady. Platforma KNIME, ve které diplomant vytvořil model pro

detekci témat, je názorně představena, viz str. 35 – 39. Teoreticky detailně vysvětluje také hlavní problematiku tj. modelování témat. Teoretickou část práce by bylo vhodné rozšířit o další modely, např. o model LSA. Z pohledu praktické části práce pozitivně hodnotím řádně a detailně okomentovaný každý krok tvorby modelu. Diplomová práce splňuje metodické pokyny kladené na zpracování tohoto typu práce.

Otázky k obhajobě:

-

Práci doporučuji k obhajobě.

Navržená výsledná známka: A

V Hradci Králové, dne 14. května 2024

podpis