



POSUDEK VEDOUCÍHO DIPLOMOVÉ PRÁCE

Jméno studenta: Bc. Radek Salay

Název práce: Zpracování nestrukturovaných dat pomocí IBM Watson

Autor posudku: Ing. Barbora Tesařová, Ph.D.

Cíl práce: Diplomová práce seznámí čtenáře s problematikou zpracování nestrukturovaného textu. Dále práce představí produkt Watson pro zpracovávání anglického jazyka. V závěru práce bude ukázané praktické využití těchto poznatků v knihovní aplikaci.

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

Vyjádření k výsledku anti-plagiátorské kontroly

Anti-plagiátorská kontrola zaznamenala 1% shody. Jde o řádně označené citace.

Díličí připomínky a náměty:

V prvních kapitolách teoretické části práce nás autor seznamuje s pojmy týkající se velkých dat a historií zpracování textu.

Pátá kapitola se věnuje díličím postupům pro zpracování textu. Celá tato kapitola by měla být hlubší, obsáhlejší, podrobnější.

Uvítala bych mimo popisu jednotlivých postupů při zpracování textu nějaký rozbor celé problematiky. Byl by zde vhodný přehledný popis problémů a souvislostí, které se musí řešit např. v různých jazycích.

Kapitola šestá a sedmá nás seznamuje s produkty IBM pro zpracování textu.

Praktická část práce popisuje návrh a implementaci aplikace Knihovna 2.0, ve které autor využil postupů NLP.

Při psaní diplomové práce autor narážel na celou řadu úskalí, se kterými se musel potýkat. Původním záměrem bylo také vyzkoušet si vytvořit vlastní model, což se ukázalo vzhledem k potřebě velmi

robustního modelu jako nereálné. Autor tedy převzal model, který je vyvíjen společností IBM a je v jeho produktech k dispozici. Tento model je ovšem velmi obecný a bez přidání dalších vazeb se pro některé analýzy textu ukázal jako ne příliš vhodný. Pro seznámení s danou problematikou byl však dostačující.

Text občas ruší neobratné věty. Autor se nevyhnul řadě pravopisných a gramatických chyb, které ale nepřekračují únosnou míru.

Celkové posouzení práce a zdůvodnění výsledné známky:

Téma práce je aktuální. Strojové zpracování přirozeného jazyka se stále více uplatňuje v různých oblastech.

Celkově lze konstatovat, že autor prokázal znalost dané problematiky. Cíle diplomové práce byly splněny. Po formální stránce splňuje posuzovaná práce požadavky pro diplomovou práci.

Otázky k obhajobě:

V práci jste pracoval s anglickými texty. Jaké další rozdíly a problémy by přinesl rozbor textu psaného v češtině?

Práci doporučuji k obhajobě.

Navržená výsledná známka: C

V Hradci Králové, dne 29. prosince 2020



podpis