



Posudek vedoucího bakalářské práce

Jméno studenta:

Jan Šlahora

Téma práce:

Statistika v lexikální analýze

Cíl práce:

Cílem práce bylo vytvořit aplikaci pro výpočet vybraných lingvistických jevů a zahrnující vybrané statistické metody z oblasti testování statistických hypotéz a oblasti redukce dimenze dat. Získané výsledky pak slouží pro porovnávání autorství textů.

Náročnost zadání bakalářské práce na:

teoretické znalosti

vyšší

praktické zkušenosti

vyšší

podkladové materiály (vstupní data) a jejich zpracování

vyšší

A: Slovní hodnocení:

Naplnění cíle práce:

Nastudování lingvistických pojmů a vybraných statistických metod bylo velmi náročné. Práce svým obsahem značně překračuje požadavky na bakalářskou práci.

Aplikace umožňuje v textu sledovat četnost hlásek, délku slova, délku věty, agregaci, asonanci a aliteraci.

Pro porovnávání středních hodnot těchto veličin jsou naprogramovány: Bartlettův test, ANOVA, Scheffého metoda a Kruskalův-Wallisův test.

Dále je implementována metoda hlavních komponent, která umožňuje znázornit redukovaná měření v rovině.

Výpočty autor provádí u vybraných autorů na více básních a může pro odhadnuté střední hodnoty prvních dvou hlavních komponent vykreslit z určené kovarianční matice konfidenční elipsu. Tato elipsa vyjadřuje nejistotu odhadnutých hlavních komponent zkoumaných autorů.

V průběhu zpracování tématu se autor věnoval zejména fónickým jevům a poezii. Tato skutečnost vede k tomu, že není provedeno porovnání textů Václava Hanky navržené v zadání.

Získané výsledky prokázali, že lze textovou analýzu pro klasifikaci autorů užívat. Mohu konstatovat, že se autor s obrovským vyvinutým úsilím dostal se svou aplikací na práh možností deskriptivních charakteristik textu a statistických metod.

Logická stavba a stylistická úroveň práce:

plně vyhovující

Využití záměrů, námětů a návrhů v praxi:

vytvořená aplikace umožňuje lingvistům okamžité použití

Případné další hodnocení (připomínky k práci):

v teoretické části práce se neobjevuje popis testů dobré shody, přestože je aplikace užívá při ověřování Poissonova rozdělení

B: Kriteriaální hodnocení:

Nápovědu k vyplnění vybraného pole je možné zobrazit klávesou F1, stručně je uvedena i ve stavovém řádku.

Kriteria hodnocení práce:	Úroveň	Připomínky
Úroveň dokumentu		
logická stavba práce	nadprůměrné	
stylistická úroveň	nadprůměrné	
práce s literaturou včetně citací	průměrné	
formální úprava práce (text, grafy, tabulky)	nadprůměrné	
Teoretická část		
rozsah a úroveň zpracování rešerše	průměrné	
formulace teoretických východisek pro praktickou část	nadprůměrné	
odborné zvládnutí problematiky	nadprůměrné	
Praktická část – produkt (řešení)		
adekvátnost použitých metod, SW, postupů	nadprůměrné	
kvalita návrhu řešení	nadprůměrné	
komplexnost řešení	komplexní	
návrh datových struktur	nadprůměrné	
uživatelské rozhraní	nadprůměrné	
odborné zvládnutí problematiky	nadprůměrné	
rozpracovanost	dokončeno, otestováno	
využitelnost praktické části v praxi	ve větší míře	
Praktická část - popis		
popis řešení v bakalářské práci	nadprůměrné	
ostatní přílohy (tabulky, grafy, výpočty, ...)	průměrné	
uživatelská příručka	průměrné	
Uložení dokumentu/ů bakalářské práce na CD	ano	
Uložení výsledku praktické části na CD	ano	
Stupeň splnění cíle práce	splněn	

C: Otázky k obhajobě (max 2):

1. Prosim uvést porovnání jevů u českých překladů básně E. A. Poe: The Raven.
2. Jaké veličiny nejvíce ovlivňovali u vybraných básní první a druhou komponentu?
Které jevy by nemusely být měřeny?

Doporučení práce k obhajobě: ano

Navržený klasifikační stupeň: výborně

Posudek vypracoval:

Jméno, tituly: Mgr. Jaroslav Marek, Ph.D.
Zaměstnavatel: FEI, Univerzita Pardubice

V Pardubicích dne: 31. 5. 2012

Podpis: