

**SOUHLASÍ  
S ORIGINÁLEM**

## HODNOCENÍ DIPLOMOVÉ PRÁCE

Západočeská univerzita v Plzni  
Fakulta aplikovaných věd  
Katedra kódermetiky

Vedoucí DP

Jméno diplomanta: Bc. Jaromír NOVOTNÝ

Garantující katedra: KKY

Název diplomové práce: Multi-dokumentová sumarizace novinových článků

	Předmět hodnocení	Nadprůměrné	Průměrné	Podprůměrné
1	Jazyková a grafická úprava	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
2	Samostatnost zpracování tématu DP	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Vhodnost použitých metod	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	Způsob zpracování a vyhodnocení	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Správnost získaných výsledků	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	Vlastní přínos	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### Doplnění hodnocení, připomínky, dotazy:

Cílem této práce bylo otestovat metody pro automatickou sumarizaci textu z více dokumentů, zejména novinových článků. Práce obsahuje detailní popis teoretických základů sumarizace, dělení sumarizačních metod a možností jejich využití. Práce se také zabývá metodami pro automatické vyhodnocení výsledků sumarizace. Hlavní část práce se zabývá multi-dokumentovou sumarizací, zde student detailně popisuje tři vybrané metody a jejich implementaci. Všechny metody byly otestovány na českých datech z ACL workshopu MultiLing 2013 a výsledné souhrny jsou následně vyhodnoceny pomocí metod ROUGE.

Po obsahové stránce je práce velice kvalitní, důkladně podložená citovanou literaturou. Student zcela samostatně nastudoval velké množství literatury k tématu a vybral tři v současné době nejpoužívanější metody pro otestování v této konkrétní úloze. Všechny metody student implementoval a otestoval na ACL korpusu článků. Výsledky všech metod jsou velmi detailně a důkladně vyhodnoceny pomocí různých metod ROUGE. Výsledky jsou pak popsány a rozebrány, nejlepší metodu pak student otestoval na reálných datech a různých úlohách (souhrn z určitého tématu, souhrn z určitého časového období).

Po formální stránce je kvalita textu práce spíše průměrná až podprůměrná, práce obsahuje překlepy, nesprávně formulované věty. Formátování textu není jednotné v celé práci, například na stranách 39 a 40 jsou odstavce formátované jinak než zbytek práce. Nedostatky ve formálním zpracování práce jsou převáženy jejím obsahem, jako výsledné hodnocení navrhuji výborně.

### Doplňující dotazy:

Jaký je rozdíl mezi LSA a NMF rozkladem výchozí matice? Myslíte, že jsou lepší výsledky NMF+K-means metody způsobeny následným zpracováním (K-means), nebo už samotným rozkladem vstupní matice?

Splnění bodů zadání	<input checked="" type="checkbox"/> úplně	<input type="checkbox"/> částečně	<input type="checkbox"/> nesplněno	
Doporučení práce k obhajobě	<input checked="" type="checkbox"/> ano		<input type="checkbox"/> ne	
Celkové hodnocení práce	<input checked="" type="checkbox"/> výborně	<input type="checkbox"/> velmi dobře	<input type="checkbox"/> dobře	<input type="checkbox"/> nevyhověl
Jméno, příjmení, titul vedoucího DP: Lucie Skorkovská, Ing.				
Pracoviště vedoucího DP: KKY				

11.6.2014

Datum



Podpis