



## POSUDEK VEDOUCÍHO BAKALÁŘSKÉ PRÁCE

**Jméno studenta:** Andrea Kubcová

**Název práce:** Data mining a možnosti nekomerčního softwaru

**Autor posudku:** Hana Skalská

**Cíl práce:** Vyhledat, popsat a případně porovnat nekomerční typy softwarových nástrojů, které lze využít pro data mining. Pro zvolený nekomerční nástroj provést podrobnější zhodnocení jeho možností a řešit jeho pomocí ukázkovou úlohu.

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 checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování teoretických aspektů tématu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Využitelnost výsledků práce v teorii (v praxi)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### Vyjádření k výsledku anti-plagiátorské kontroly

Práce byla vyhodnocena v Odevzdej.cz (celková podobnost 0%).

### Dílčí připomínky a náměty:

V kapitole 3 (Cíle práce a metodika) by neměly být popisované konkrétní provedené postupy, ale podrobněji by měl být vysvětlený cíl a očekávané kroky, které k cíli povedou. V aplikaci nejsou pro modely v tabulce 3 (Porovnání metrik klasifikátorů) jednoznačně popsány kategorie, mezi kterými se klasifikace provádí.

Formální a pouze výjimečné jsou drobné gramatické prohřešky.

### Celkové posouzení práce a zdůvodnění výsledné známky:

Bakalářská práce prokazuje mimořádný talent a schopnost studentky orientovat se v literatuře k danému tématu, vytvořit si vlastní náhled na řešení problematiky a aplikovat nové typy metod pomocí softwarových nástrojů, se kterými se seznámila.

Studentka zmapovala relevantní množinu nekomerčních softwarových produktů a pro vybranou množinu porovnávala funkcionality, implementované metody a algoritmy jednotlivých produktů.

Jednotlivé kroky jsou věcně a srozumitelně popsány.

V závěrečné části jsou pro vybrané reprezentanty porovnána zvolená kvalitativní hlediska, relevantní ve vztahu k cíli práce.

Aplikační část řeší úlohu klasifikace textu pomocí softwaru Orange. Řeší se úlohy klasifikace pro data českých internetových zpravodajských médií (shluková analýza a kategorie sentimentu).

Je popsán postup volby, získání a přípravy vlastních trénovacích a testovacích dat, jsou porovnány dvě úlohy klasifikace a pomocí tří různých metrik jsou porovnány dva algoritmy klasifikace s učícím souborem.

### **Otázky k obhajobě:**

Jakou konkrétní metodu křížové validace jste použila ve své aplikaci?

### **Práci doporučuji k obhajobě.**

Navrhuji zařadit práci do soutěže o nejlepší Bc. práci.

**Navržená výsledná známka: A**

**V Hradci Králové, dne 30. prosince 2020**

---

**podpis**