



## POSUDEK VEDOUCÍHO DIPLOMOVÉ PRÁCE

**Jméno studenta:** Tomáš Kratochvíl

**Název práce:** Shluková analýza datových toků

**Autor posudku:** Antonín Slabý

**Cíl práce:** Tato diplomová práce se zabývá problematikou shlukové analýzy datových toků. V práci jsou popsány vybrané algoritmy pro shlukovou analýzu datových toků se zaměřením na kategorická data. V práci je navrženo řešení pro shlukovou analýzu e-mailů, využitelné antisпамovým systémem, které vychází z vybraných algoritmů. Navržené řešení je testováno na datovém souboru vytvořeném z e-mailů, které byly vyhodnoceny antisпамem za jeden den.

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 type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	<input type="checkbox"/>	<input checked="" 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

Dle antiplagiátorské kontroly nevykazuje práce shodu s jinými texty (podobnost je 0 procent).

### Dílejší připomínky a náměty:

### Celkové posouzení práce a zdůvodnění výsledné známky:

Tématika práce je obtížná a intenzivně zkoumaná. Práce obsahuje mj.: Popis vybraných algoritmů pro shlukovou analýzu datových toků a představení podstatných algoritmů se zaměřením na kategorická data (představeno 5 algoritmů aproximace četností),

Dokumentaci vytvořeného vlastního autorova řešení pro shlukovou analýzu e mailů využitelnou antisпамovým systémem. Vybrány k řešení, využity a upraveny jsou k tomu 2 algoritmy pro shlukovou analýzu datových toků s kategorickými atributy ( Algoritmus StreamCluCD, Algoritmus CSketch) , je vytvořen algoritmus Sliding Lossy Counting.. Otestování aplikace je na datovém souboru vytvořeném z e mailů které byly antisпамem vyhodnoceny antisпамem za jeden den.

Cíle práce byly splněny. Práce obsahuje pěkný příspěvek k teorii, metodologii (algoritmus) i praxi (autorovo řešení a dokonalé otestování). Formálně a jazykově má práce dobrou úroveň. Čtivý, přesný a zajímavý text práce je vhodně doplněn o ukázky zdrojových kódů a obrázky, je vhodně strukturován.

### **Otázky k obhajobě:**

Diskuse se může věnovat řadě oblastí. Vyberme např.:

Představení podstaty autorova řešení.

Zkušenosti získané testováním a možnosti zlepšení.

**Práci doporučuji k obhajobě.**

**Navržená výsledná známka: A**

**V Hradci Králové, dne 28. dubna 2023**

---

**podpis**