



POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

Jméno studenta: David Illner

Název práce: Data mining a možnosti nekomerčního softwaru

Autor posudku: Antonín Slabý

Cíl práce: Bakalářská práce řeší dolování dat pomocí úloh data mining v nekomerčním softwaru Rattle. Práce vychází z metodiky Knowledge Data Discovery. Zabývá se selekcí, archivací, předzpracováním a transformací dat a klasifikačními úlohami pro vydolování znalostí. Praktická část navazuje na poznatky z teoretické části a na vytvoření vlastního modelu logistické regrese v softwaru Rattle, který je postavený na jazyku R od developera Grahama Williamse. V tomto programu byla předzpracována a transformována k vytvoření efektivního modelu. V programu byl vytvořen model logistické regrese pro předpověď počasí na daném datasetu weather.csv z depositáře programu. Pomocí tohoto modelu byly nalezeny důležité proměnné pro předpověď, které pozitivně nebo negativně ovlivňují klasifikaci předpovědí pro každý den. Pro zhodnocení efektivnosti a kontrole správnosti předpovědí modelu byly použity metody ROC křivka a klasifikační matice.

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 type="checkbox"/>	<input checked="" 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 type="checkbox"/>	<input checked="" 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 type="checkbox"/>	<input checked="" 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

Podobnost je 0 %.

Dílčí připomínky a náměty:

Vložte připomínky k práci

Celkové posouzení práce a zdůvodnění výsledné známky:

Text práce se týká postupně následujících problémů: Pojmový aparát KDD a data miningu, typy úloh data miningu, instalace software Rattle (vytvořeného s použitím jazyka R) základy práce se softwarem a jeho prostředím, příprava dat a metody předzpracování dat použití software k řešení vybraných typických úloh data miningu (diskriminace, klasifikace). Demonstrace na praktickém existujícím modelu předpovědi počasí (využití datasetu weather.csv) ke kontrole použita klasifikační matice a ROC analýza.

Jasný dostatečně přesný a čtivý text demonstruje velmi dobře zkoumanou problematiku, využitý software, jeho rysy, funkčnosti a jeho použití v různých (i náročnějších/pokročilých statistických úlohách) a jeho uživatelskou přátelskost a efektivitu. Formálně, jazykově a stylisticky je práce na dobré úrovni. Cíle práce byly splněny. Práci doporučuji k obhajobě a vzhledem k rozsahu a přesnosti odvedené práce hodnotím stupněm A.

Otázky k obhajobě:

Které rysy software Rattle autora nejvíce příjemně a které naopak nepříjemně překvapily ?

Práci doporučuji k obhajobě.

Navržená výsledná známka: A

V Hradci Králové, dne 30. dubna 2020

podpis