



## POSUDEK VEDOUCÍHO DIPLOMOVÉ PRÁCE

**Jméno studenta:** Norbert Snášel

**Název práce:** Data mining: Analýza klientů retailového bankovníctví

**Autor posudku:** Jan Draessler

**Cíl práce:** Cílem práce je s použitím vybraných nástrojů analyzovat datový soubor týkající se využívání služeb na běžných účtech bank v ČR.

Povinná kritéria hodnocení práce	Stupeň hodnocení (známka)			
	1	2	3	4
Práce svým zaměřením odpovídá studovanému oboru	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vymezení cíle a jeho naplnění	<input type="checkbox"/>	<input checked="" 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"/>
Zpracování praktických aspektů tématu	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input type="checkbox"/>	<input checked="" 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"/>
Formální úprava a náležitosti práce	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	<input type="checkbox"/>	<input checked="" 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"/>

### Dílejší připomínky a náměty:

Diplomová práce sleduje analýzu datového souboru pořízeného ze serveru bankovnípoplatky.com, který obsahuje počty transakcí a využívání služeb na bankovních účtech klientů českých bank.

Struktura celé práce je členěna do pěti kapitol (bez úvodu) kde první tři kapitoly lze považovat za převážně teoretické a poslední dvě za praktické. Autor v teoretické části seznamuje s historií data miningu (DM), jeho využíváním, s metodikou CRISP-DM a vybranými metodami DM. V praktické části potom provádí analýzy, které autor s ohledem na stanovené cíle a data sám zvolil.

Praktickou část celé práce lze považovat za poměrně náročnou. Velký časový prostor musel autor věnovat zejména přípravě dat, která byla dodána v surové podobě, čítající cca 90 tis. případů a více než 40 proměnných. Celý dodaný soubor byl nekonzistentní, obsahoval velké množství chybějících hodnot a chyb, které bylo třeba opravit či odstranit. Velké úsilí bylo věnováno i vlastní analýze, hledání vhodných modelů, studiu vybraných metod, a studium práce s produktem SPSS Modeler. Autor navíc testoval i některé další modely, které později nebyly využity, a je škoda, že v práci je zobrazen už jen výsledný stream a tyto modely nejsou dále zmíněny.

Po jazykové stránce je text celkem čistý, konzistentní, bez pravopisných chyb. Jako nedostatek lze vnímat časté používání holých vět a opakování podobných frází.

Po stránce stavby celé práce považuji podíl teoretické a praktické části za celkově vyvážený, i když v některých partiích se teoretický popis prolíná s praktickou prací s daty, což nepovažuji za nejšťastnější.

Za nedostatek považuji absenci některých výstupů z vytvořeného streamu např. z explorační dat, jejíž výstup bych očekával v příloze práce. Také zvolená formulace hypotéz vzhledem k cílům není úplně ideální. Závažnější problém shledávám v absenci interpretace konkrétních výsledků vzhledem ke stanovené hypotéze. Interpretace jsou umístěny až do závěru práce a to spíše obecně, celkově tak výsledky práce vyznívají poněkud hluše, což celkový dojem z práce poněkud kazí.

Za největší přínos studenta lze považovat především vytvořený stream, který je dále, buď přímo, a nebo po drobných úpravách, využitelný pro další analýzy. Přínosem jsou bezesporu i doporučení pro kalkulátory bankovních poplatků, byť to primárně nebylo cílem práce.

### **Celkové posouzení práce a zdůvodnění výsledné známky:**

Autor se ve své závěrečné práci pustil do problematiky, která vyžaduje nejen dobré znalosti, ale také určitou zkušenost. Ne vše se podařilo tak, jak si předsevzal a mnohé nedostatky jsou dílem nezkušenosti a dílem opomenutím. Nicméně cíl práce lze považovat za splněný a závěrečnou práci schopnou obhajoby. Po formální stránce text splňuje požadavky na závěrečnou práci.

**Práci doporučuji k obhajobě.**

### **Otázky pro diskusi:**

V práci postrádám diskusi o využitelnosti výsledků analýz. Mohl by měl vysvětlit, jak lze využít výsledků v praxi?

**Navržená výsledná známka: C**

V Hradci Králové, dne 12.5.2016

---

**podpis**