

Obsah

1.	Co je Data Mining	5
2.	Metodika CRISP-DM	7
2.1.	Hierarchický rozklad metodologie CRISP-DM	7
2.2.	Fáze CRISP-DM	10
2.3.	Etapy a úkoly metodologie CRISP-DM	12
2.3.1.	Porozumění problému (Business Understanding)	12
2.3.2.	Porozumění datům (Data Understanding)	13
2.3.3.	Příprava dat (Data Preparation)	14
2.3.4.	Modelování (Modeling)	15
2.3.5.	Vyhodnocení výsledků (Evaluation)	16
2.3.6.	Využití výsledků (Deployment)	17
2.4.	Shrnutí	18
3.	Porozumění problému	20
3.1.	Analýza profilů	22
3.2.	Segmentace	23
3.3.	Odezva	23
3.4.	Rizika	24
3.5.	Aktivace	25
3.6.	Křížový a následný prodej	26
3.7.	Odliv zákazníků	26
3.8.	Čistá současná hodnota	27
3.9.	Hodnota po dobu existence	27
3.10.	Shrnutí	27
4.	Porozumění datům	29
4.1.	Typy dat	32
4.2.	Zdroje dat	33
4.2.1.	Interní zdroje	34
4.2.2.	Externí zdroje	38
4.3.	Výběr dat pro modelování	39

4.3.1.	Data pro získávání zákazníků.....	39
4.3.2.	Data pro zákaznické modely	43
4.3.3.	Data pro modely rizika.....	45
4.4.	Vytváření datových sad pro modelování.....	47
4.4.1.	Jak velký má být můj vzorek?.....	47
4.4.2.	Vzorkovací metody	49
4.4.3.	Vytváření modelů z modelovaných dat.....	50
4.4.4.	Kombinace dat z mnoha nabídek	50
4.5.	Shrnutí	51
5.	Příprava dat pro modelování	52
5.1.	Přístup k datům.....	55
5.1.1.	Klasifikace dat.....	58
5.1.2.	Typy dat v systému Clementine	59
5.1.3.	Čtení zdrojových dat	60
5.2.	Vytvoření sady dat pro modelování	65
5.2.1.	Spojování souborů.....	66
5.2.2.	Výběr záznamů.....	81
5.2.3.	Agregace.....	86
5.2.4.	Přidávání případů a vážení	89
5.2.5.	Změna pořadí a časový posuv	95
5.2.6.	Generování odvozených proměnných.....	98
5.2.7.	Výběr proměnných a změna typu proměnné.....	104
5.2.8.	Ošetření vynechaných hodnot a chyb.....	111
5.2.9.	Reorganizace datové matice.....	129
6.	Závěr.....	130
7.	Použité programové vybavení	131
7.1.	SPSS	131
7.2.	Clementine	131
8.	Seznam použité literatury	133
9.	Seznam obrázků	134
10.	Seznam tabulek	139