

## Obsah

1. Co je Data Mining.....	5
2. Metodika CRISP-DM.....	8
2.1. Hierarchický rozklad metodologie CRISP-DM .....	8
2.2. Fáze CRISP-DM .....	10
2.3. Etapy a úkoly metodologie CRISP-DM.....	13
2.3.1. Porozumění problému (Business Understanding).....	13
2.3.2. Porozumění datům (Data Understanding).....	14
2.3.3. Příprava dat (Data Preparation).....	15
2.3.4. Modelování (Modeling) .....	16
2.3.5. Vyhodnocení výsledků (Evaluation).....	17
2.3.6. Využití výsledků (Deployment) .....	18
2.4. Shrnutí .....	19
3. Porozumění problému .....	22
3.1. Analýza profilů.....	24
3.2. Segmentace.....	24
3.3. Odezva.....	25
3.4. Rizika .....	26
3.5. Aktivace .....	27
3.6. Křížový a následný prodej.....	27
3.7. Odliv zákazníků.....	27
3.8. Čistá současná hodnota .....	28
3.9. Hodnota po dobu existence .....	28
3.10. Shrnutí .....	29
4. Porozumění datům.....	31
4.1. Typy dat.....	34
4.2. Zdroje dat .....	35
4.2.1. Interní zdroje .....	36
4.2.2. Externí zdroje .....	40
4.3. Výběr dat pro modelování.....	41

4.3.1.	Data pro získávání zákazníků.....	41
4.3.2.	Data pro zákaznické modely .....	45
4.3.3.	Data pro modely rizika.....	47
4.4.	Vytváření datových sad pro modelování.....	49
4.4.1.	Jak velký má být můj vzorek?.....	49
4.4.2.	Vzorkovací metody .....	51
4.4.3.	Vytváření modelů z modelovaných dat.....	52
4.4.4.	Kombinace dat z mnoha nabídek .....	52
4.5.	Shrnutí .....	53
5.	Příprava dat pro modelování .....	55
5.1.	Přístup k datům.....	58
5.1.1.	Klasifikace dat.....	60
5.1.2.	Typy dat v systému Clementine .....	61
5.1.3.	Čtení zdrojových dat .....	63
5.2.	Vytvoření sady dat pro modelování .....	68
5.2.1.	Spojování souborů.....	68
5.2.2.	Výběr záznamů.....	84
5.2.3.	Agregace.....	89
5.2.4.	Přidávání případů a vážení .....	92
5.2.5.	Změna pořadí a časový posuv .....	98
5.2.6.	Generování odvozených proměnných.....	101
5.2.7.	Výběr proměnných a změna typu proměnné.....	106
5.2.8.	Ošetření vynechaných hodnot a chyb.....	114
5.2.9.	Reorganizace datové matice.....	132
6.	Závěr.....	134
7.	Použité programové vybavení.....	135
7.1.	SPSS .....	135
7.2.	Clementine .....	135
8.	Seznam použité literatury .....	137
9.	Seznam obrázků .....	138
10.	Seznam tabulek .....	144