

PODĚKOVÁNÍ	xv
PŘEDMLUVA	xvii
ÚVOD	xix
Co je data mining?	xix
Jaké je zaměření této knihy?	xix
Kdo by měl tuto knihu číst?	xx
Jak je kniha uspořádána	xx
Nástroje, které budete potřebovat	xxi
Doprovodné CD	xxi
Čím se kniha nezabývá?	xxi
Shrnutí	xxii
O AUTORCE	xxiii
O SPOLUAUTORECH	xxv

ČÁST I Plánujeme menu I

KAPITOLA I Stanovení cíle	3
Definování cíle	4
Analýza profilů	6
Segmentace	6
Odezva	7
Rizika	7
Aktivace	8
Křížový a následný prodej	8
Odliv zákazníků	8
Čistá současná hodnota	9
Hodnota po dobu existence	9
Výběr techniky modelování	9
Lineární regrese	10
Logistická regrese	12
Neuronové sítě	13

	Genetické algoritmy	14
	Klasifikační stromy	14
	Adaptivní firma	16
	Najímání odborníků a týmová práce	17
	Orientace na produkt versus orientace na zákazníka	18
	Shrnutí	18
KAPITOLA 2	Výběr zdrojů dat	19
	Typy dat	19
	Zdroje dat	20
	Interní zdroje	21
	Zákaznické databáze	21
	Transakční databáze	22
	Databáze historie nabídek	22
	Oslovování poštou a telefonicky	23
	Datový sklad	23
	Externí zdroje	27
	Výběr dat pro modelování	27
	Data pro získávání zákazníků	28
	<i>Modely pro získávání nových zákazníků</i>	28
	Data pro zákaznické modely	30
	<i>Modelování křížového prodeje, následného prodeje, retence a obnovy</i>	31
	Data pro modely rizika	32
	<i>Modelování rizika</i>	32
	Vytváření datových sad pro modelování	33
	Jak velký má být můj vzorek?	33
	Vzorkovací metody	34
	Vytváření modelů z modelovaných dat	36
	Kombinace dat z mnoha nabídek	36
	Shrnutí	36
ČÁST 2	Přehledka vaření	37
KAPITOLA 3	Příprava dat pro modelování	39
	Přístup k datům	39
	Klasifikace dat	41
	Čtení zdrojových dat	42

Vytvoření sady dat pro modelování	44
Vzorkování	45
Vyčištění dat	46
Spojité proměnné	47
Odhlehlé hodnoty a chyby dat	48
Chybějící hodnoty	50
Kategoriální proměnné	54
Chyby v datech	55
Chybějící hodnoty	55
Shrnutí	56
KAPITOLA 4	
Výběr a transformace proměnných	57
Definování cílové funkce	58
Pravděpodobnost aktivace	58
Index rizika	59
Ziskovost produktu	59
Marketingové výdaje	59
Odvozené proměnné	59
Sumarizace	60
Podíly	60
Kalendářní data	61
Redukce počtu proměnných	61
Spojité proměnné	61
Metoda 1: Jeden model	61
Metoda 2: Dva modely	63
Kategoriální proměnné	63
Metoda 1: Jeden model	63
Metoda 2: Dva modely	67
Vytvoření lineárních prediktorů	68
Spojité proměnné	68
Segmentace	72
Transformace	73
Kategoriální proměnné	78
Nalezení interakcí	81
Shrnutí	82

KAPITOLA 5	Zpracování a vyhodnocení modelu	83
	Zpracování modelu	83
	Rozdělení dat	85
	Metoda 1: Jeden model	88
	<i>Předběžné vyhodnocení</i>	96
	Metoda 2: Dva modely – Odezva	98
	Metoda 2: Dva modely – Aktivace	99
	Porovnání metody 1 a metody 2	100
	Shrnutí	103
KAPITOLA 6	Ověření modelu	105
	Tabulky navýšení a grafy navýšení	105
	Metoda 1: Jeden model	105
	Metoda 2: Dva modely	107
	Použití modelů na alternativních data setech	107
	Převzorkování	113
	Jackknifing	114
	Bootstrapping	118
	Analýza klíčových proměnných podle decilů	125
	Shrnutí	129
KAPITOLA 7	Implementace a údržba modelu	131
	Hodnocení nového souboru	131
	Hodnocení ve firmě	131
	<i>Ověření dat</i>	132
	Hodnocení a audit u externí firmy	134
	Implementace modelu	140
	Výpočty finančního efektu	140
	Určení zákazníků, kterým má být zaslána nabídka	144
	Champion versus Challenger	145
	Matice dvou modelů	146
	Sledování modelu	148
	Závěrečné ověření	153
	Údržba modelu	154
	Životnost modelu	154

Benchmarking	154
Přestavění nebo obnovení?	154
Protokol modelu	155
Shrnutí	156

ČÁST 3 Recepty pro každou příležitost 157

KAPITOLA 8 Porozumění zákazníkovi: profilace a segmentace 159

Je důležité porozumět zákazníkovi?	159
Typy profilace a segmentace	160
<i>RFM: doba, četnost, hodnota</i>	160
<i>Demografická analýza</i>	161
<i>Analýza životních etap</i>	161
Profilace a analýza penetrace u zásilkového obchodu	164
Analýza RFM	164
Analýza penetrace	167
Výpočet matice hodnoty zákazníků pro společnost nabízející kreditní karty	170
Analýza hodnoty zákazníků	172
Určení segmentů zákazníků pomocí shlukové analýzy	176
Shrnutí	177

KAPITOLA 9 Zaměření na nové zákazníky: modelování odezvy 181

Definice cíle	181
Všichni stávající zákazníci nejsou stejní	182
Příprava proměnných	184
Spojité proměnné	184
Kategoriální proměnné	191
Zpracování modelu	194
Ověření dat metodou bootstrapping	198
Implementace modelu	203
Shrnutí	204

KAPITOLA 10	Vyhýbání se rizikovým zákazníkům: modelování rizika	205
	Hodnocení úvěrů a modelování rizika	205
	Definice cíle	207
	Příprava proměnných	207
	Zpracování modelu	217
	Ověření modelu	220
	Bootstrapping	222
	Implementace modelu	224
	Změna rozsahu hodnot vypočteného rizika	225
	Podvody jako jiný typ rizika	226
	Shrnutí	228
KAPITOLA 11	Udržení ziskových zákazníků: modelování pravděpodobnosti odchodu zákazníků	229
	Věrnost zákazníků	229
	Definice cíle	230
	Příprava proměnných	233
	Spojité proměnné	233
	Kategoriální proměnné	236
	Zpracování modelu	240
	Ověření modelu	241
	Bootstrapping	242
	Implementace modelu	245
	Vytvoření profilů	245
	Optimalizace ziskovosti zákazníků	247
	Aktivní metody udržování zákazníků	249
	Shrnutí	250
KAPITOLA 12	Směrování na ziskové zákazníky: modelování hodnoty produktu po dobu jeho existence	251
	Co je hodnota po dobu existence?	252

Využití hodnoty po dobu existence	252
Složky hodnoty po dobu existence	253
Využití hodnoty po dobu existence	255
Případové studie pro výpočet hodnoty po dobu existence	255
<i>Marketing v prostředí business-to-business</i>	255
<i>Je-li doba mezi opakovanými nákupy delší než rok</i>	257
<i>Analýza pravidelných návštěvníků restaurace</i>	258
Výpočet hodnoty po dobu existence pro obnovovaný produkt či službu	258
Případová studie: výpočet hodnoty po dobu existence	259
Případová studie: čisté příjmy v roce I	260
Výpočet hodnoty po dobu existence	267
Shrnutí	272

KAPITOLA 13 Rychlé občerstvení: modelování na webu 275

Získávání dat z webu a modelování	275
Definice cíle	276
<i>Případová studie pro prodej úvěrových karet</i>	276
Zdroje dat na webu	277
Příprava dat získaných z webu	278
Výběr metodiky	279
Značky na webu	284
Získávání přehledu o zákaznících v reálném čase	285
Získávání dat o využití webu – případová studie	286
Shrnutí	290

DODATEK A Analýza spojitých proměnných 291

Čistá hodnota domu	291
Odhadnutý věk	293
Celkový počet účtů	295
Počet účtů otevřených za posledních 6 měsíců	297
Vyhodnocení proměnné actopl6	299
Celková výše zůstatků	299
Počet žádostí o úvěr za posledních 6 měsíců	301
Celkový počet otevřených účtů	303

Celkový počet záznamů úvěrové historie	305
Stáří souboru úvěrové historie	307
Počet případů, ve kterých byla splátka zaslána 30 dnů po splatnosti	309
Náhrada chybějících hodnot pro příjem	311

DODATEK B

Analýza kategoriálních proměnných	315
DOPORUČENÁ LITERATURA	323

Rejstřík

325