

# Obsah

Slovo úvodem .....	9
Předmluva .....	11
<b>1</b> Základy .....	13
1.1 Big data.....	13
1.2 Věda o datech .....	19
1.3 Umělá inteligence.....	24
1.4 Data mining.....	25
1.5 Strojové učení.....	26
1.6 Business Intelligence.....	27
1.7 Datové inženýrství.....	28
1.8 Průnik disciplín.....	28
1.9 Metodologie CRISP DM.....	30
1.10 Strategie používání dat .....	32
1.11 Výuka vědy o datech.....	34
1.12 Problematika COVID-19 .....	36
1.13 Souhrn.....	38
<b>2</b> Stručně o umělé inteligenci .....	40
2.1 Vybrané principy umělé inteligence .....	42
2.2 Expertní systémy.....	46
2.3 Pravidla v expertním systému.....	48
2.4 Souhrn.....	54
<b>3</b> Uspořádání dat .....	55
3.1 Plochý soubor (Flat file).....	55
3.2 HTML a XML soubory.....	56
3.3 JSON formát.....	57
3.4 SQL databáze.....	58
3.5 NoSQL databáze .....	60
3.6 Hadoop, MapReduce, Spark, Mahout, HBase .....	61
3.7 Datový sklad.....	65
3.8 Jezero dat.....	66

3.9	Organizace dat a datová matice .....	66
3.10	Souhrn .....	68
<b>4</b>	<b>Strojové učení – přehled .....</b>	<b>70</b>
4.1	Principy algoritmů strojového učení .....	71
4.2	Typy proměnných a příprava dat.....	74
4.2.1	Volba proměnných .....	75
4.2.2	Chybějící hodnoty .....	75
4.3	Učení s učitelem.....	76
4.3.1	Lineární regrese a její modifikace .....	77
4.3.2	Logistická regrese .....	81
4.3.3	Složitější typy regresí (GLM a GAM) .....	81
4.3.4	K-nejbližších sousedů (regrese a klasifikace) .....	82
4.3.5	Lineární diskriminační analýza .....	82
4.3.6	Klasifikace podpůrnými vektory .....	83
4.3.7	Bayesova metoda .....	84
4.3.8	Neuronové sítě .....	84
4.3.9	Rozhodovací strom .....	85
4.4	Učení bez učitele.....	87
4.4.1	Shlukování .....	87
4.4.2	Redukce dimenzionality .....	90
4.4.3	Analýza asociací .....	91
4.4.4	Detekce anomálie .....	93
4.4.5	Autokódování .....	94
4.5	Učení posilováním.....	94
4.5.1	Q-učení .....	95
4.5.2	TD-učení .....	95
4.6	Algoritmy strojového učení podle podobnosti .....	95
4.7	Evaluace algoritmů strojového učení.....	97
4.7.1	Koeficienty kvality klasifikačních modelů .....	97
4.7.2	Ukazatele kvality regresních modelů .....	100
4.7.3	Interpreovatelnost prediktivních modelů .....	100
4.8	Souhrn .....	104
<b>5</b>	<b>Metoda podpůrných vektorů .....</b>	<b>106</b>
5.1	Algoritmus metody SVM.....	106
5.2	Klasifikace pomocí SVM do více tříd.....	109
5.3	Aplikace SVM pro COVID-19 data .....	109
5.4	Souhrn.....	112
<b>6</b>	<b>Bayesova metoda .....</b>	<b>113</b>
6.1	Modelová aplikace.....	114
6.2	Bayes a ohrožení COVID-19 pandemií.....	116
6.3	Souhrn.....	120



<b>7</b>	Umělé neuronové sítě .....	121
7.1	Popis jednoduchého perceptronu .....	121
7.2	Aktivační funkce .....	124
7.3	Vytvoření architektury neuronové sítě .....	125
7.4	Problémy procesu učení a interpretace .....	130
7.5	Typy neuronových sítí .....	131
7.6	Neuronové sítě a COVID-19 .....	135
7.7	Souhrn .....	138
<b>8</b>	Rozhodovací stromy a lesy .....	139
8.1	Rozhodovací stromy – principy .....	139
8.2	Ensemble algoritmy .....	148
8.3	Stromy, lesy a COVID-19 .....	150
8.4	Souhrn .....	153
<b>9</b>	Analýza sociálních sítí .....	154
9.1	Základy teorie grafů .....	154
9.2	Sociální síť .....	158
9.3	Shlukování uzlů – Louvainská metoda .....	158
9.4	Girvan-Newman algoritmus .....	159
9.5	PageRank algoritmus .....	160
9.6	Sociální sítě v pandemii COVID-19 .....	163
9.7	Souhrn .....	164
<b>10</b>	Analýza textů .....	166
10.1	Základní definice .....	167
10.2	Analýza vztahů slov .....	168
10.3	Shlukování textů .....	169
10.4	Klasifikace textů .....	170
10.5	Shrnutí textů .....	172
10.6	Vyhledávání a analýza témat .....	173
10.7	Analýza sentimentu a názorů .....	176
10.8	Kombinovaná analýza .....	178
10.9	Zavádějící zprávy .....	180
10.10	Zavádějící texty o COVID-19 .....	185
10.11	Souhrn .....	187

<b>11</b>	<b>Programovací jazyky, systémy a nástroje</b>	
	zpracování dat .....	189
11.1	Programovací jazyky .....	189
11.1.1	Python .....	189
11.1.2	R .....	190
11.1.3	Julia .....	193
11.2	Prostředky pro scraping .....	193
11.3	Programové systémy pro vytváření modelů .....	195
11.4	Oblasti datového inženýrství .....	198
11.5	Souhrn .....	199
	Závěr .....	201
	Slovníček anglických termínů .....	203
	Literatura .....	216
	Rejstřík .....	221