

# Tartalomjegyzék

---

<b>Bevezetés</b> .....	<b>1</b>	
A könyvről .....	2	
A könyvben használt ikonok .....	3	
Hol is kezdjük? .....	3	
<b>1. fejezet: Mit jelent az adatelemzés?</b> .....	<b>5</b>	
Ki tudja hasznosítani az adatelemzést? .....	6	
Vizsgáljuk meg az adatelemzési kirakós darabjait .....	9	
Adatok gyűjtése, lekérdezése és értelmezése .....	11	
Matematikai modellezés adatelemzési feladatokra .....	12	
Statisztikai módszerekből levezethető információk .....	13	
Kódolás, kódolás, kódolás – ez is a játék része .....	14	
Adatelemzés alkalmazása adott témakörben .....	14	
Az adatokból nyert mögöttes információk kommunikálása	16	
<b>2. fejezet: Az adattechnika</b> .....	<b>17</b>	
Mi az a három V? .....	17	
Megküzdés az adatmennyiséggel .....	18	
Az adatsebesség kezelése .....	18	
Az adatok sokféleségének kezelése .....	19	
A fontos adatforrások azonosítása .....	21	
Az adatmegközelítések közötti különbségek .....	21	
Az adatelemzés meghatározása .....	22	
A gépi tanulási technika .....	23	
Az adattechnika .....	24	
Gépi tanulási mérnökök; adatelemzők; adatelemzők ..	25	
Adatok tárolása és feldolgozása .....	26	
Adatok tárolása és elemzése közvetlenül a felhőben .....	26	
Adatfeldolgozás valós időben .....	32	
A generatív MI-hatás felismerése .....	33	
Az adattechnika átalakulása .....	33	
Az MI-alapú számítási feladatokat támogató eszközök és	keretrendszerek .....	33

**3. fejezet: Adatfeldolgozás gépi tanulással .....35**

A gépi tanulás és folyamatai .....	35
A gépi tanulás lépései .....	36
A gépi tanulást leíró fogalmak .....	36
Tanulási stílusok .....	38
Tanulás felügyelt algoritmusokkal .....	38
Tanulás felügyelet nélküli algoritmusokkal.....	38
Megerősítéssel tanulás .....	39
A lehetőségek.....	39
Algoritmusok választása funkció alapján.....	39
Valós idejű elemzések generálása Sparkkal .....	44

**4. fejezet: Matematika, valószínűség és statisztikai modellezés.....47**

A valószínűség és a következtető statisztika .....	48
Valószínűségi eloszlások .....	51
Feltételes valószínűség a naiv Bayes-módszerrel .....	53
A korrelációk számszerűsítése .....	54
Korreláció kiszámítása a Pearson-féle $r$ -értékkel .....	55
Változópaárok rangsorolása a Spearman-féle rangkorrelációval .....	57
Az adatok dimenziószámának csökkentése lineáris algebrával.....	58
Adatok felbontása a dimenziószám csökkentése érdekében.....	59
Dimenziószám-csökkentés faktoranalízissel.....	63
A dimenziószám csökkentése és a kiugró értékek eltávolítása PCA-val .....	64
Döntések modellezése több feltételen alapuló döntéshozatallal .....	66
A hagyományos MCDM .....	67
A puha MCDM.....	69
A regressziós módszerek .....	69
Lineáris regresszió .....	70
Logisztikus regresszió.....	72
Legkisebb négyzetes regressziós módszerek .....	73
Kiugró értékek észlelése.....	74
Szélsőséges értékek elemzése .....	74
Kiugró értékek észlelése egyváltozós elemzéssel.....	75