

Tartalomjegyzék

Bevezetés	1
A könyvről	1
Feltételezéseink rólad	3
A könyvben használt ikonok.....	3
A könyvön túl.....	4
Hol is kezdjük?	5
1. rész: HOGYAN TANULNAK A GÉPEK?	7
1. fejezet: Az AI igaz története	9
Lépjünk túl a felhajtáson	10
Álmodjunk elektronikus báránnyal	11
Az AI és a gépi tanulás története.....	12
Mit képes nyújtani a gépi tanulás az AI számára?	13
A gépi tanulás céljai	14
A gépi tanulás határai a hardver alapján.....	14
Emelkedjünk felül az AI-vágyálmokon	15
Az AI és a gépi tanulás felhasználásainak divathóbortjai	16
Az AI és a gépi tanulás valódi felhasználási módjai	17
Legyünk hasznosak; legyünk hétköznapiak.....	19
Milyen kapcsolat van az AI és a gépi tanulás között?	19
Az AI és a gépi tanulás specifikációi	20
Mi választja el a művészetet a mérnöki munkától?.....	21
Eljön-e a következő AI-tél?	22
2. fejezet: Tanulás a big data korában	25
A gépi tanulás alapjai	26
Definiáljuk a big datát	27
A big data adatok forrásai	28
Készítsünk új adatforrást	29
Használjunk meglévő adatforrásokat	31
Keressünk tesztadatforrásokat.....	32
A statisztika szerepe a gépi tanulásban.....	32
Ismerjük meg az algoritmusok szerepét.....	34
Mit csinálnak az algoritmusok?	34
Tekintsük át az öt fő technikát.....	34
Mit jelent a tanítás?	36
3. fejezet: Pillantsunk a jövőbe	39
Hozzunk létre hasznos technológiákat a jövő számára.....	40
A gépi tanulás szerepe a robotoknál.....	40
Gépi tanulás az egészségügyben	41
Intelligens rendszerek különböző igényekhez.....	41
Gépi tanulás ipari környezetben	42
A processzorok és más hardverek frissítésének jelentősége	43

A gépi tanulással megjelenő új munkalehetőségekről.....	43
Hogyan dolgozzunk egy gépnek?	44
Hogyan dolgozzunk a gépekkel?.....	45
Hogyan javítsuk a gépeket?	45
Hogyan hozzuk létre új gépi tanulási feladatokat?	46
Hogyan hozzuk létre új gépi tanulási környezeteket?	47
Kerüljük el a jövő technológiáinak lehetséges buktatóit	47

2. rész: KÉSZÍTSÜK ELŐ AZ ESZKÖZÖKET A TANULÁSHOZ..... 49

4. fejezet: Telepítsünk egy Python-disztribúciót..... 51

Használjuk az Anacondát a gépi tanuláshoz.....	52
Szerezzük be az Anacondát.....	52
Miért az Anacondát használjuk a könyvben?	53
Telepítsük az Anacondát Linuxon	54
Telepítsük az Anacondát Mac OS X-en	55
Telepítsük az Anacondát Windowson.....	56
Töltsük le az adathalmazokat és a példakódokat.....	59
Használjuk a Jupyter Notebookot	59
Határozzuk meg a kódtárat	61
Ismerjük meg a könyvben használt adathalmazokat.....	66

5. fejezet: Lépünk túl az alapszintű Python-programozáson 69

Határozzuk meg az alapvető tudnivalókat	70
Tekintsük át a Python alapjait	70
Használjunk függvényeket.....	75
Tároljunk adatokat halmazok, listák és értéksorok használatával.....	81
Hozzuk létre halmazokat	81
Végezzünk műveleteket halmazokon	81
Használjunk listákat	82
Hozzuk létre és használjunk értéksorokat	85
Definiáljunk hasznos iterátorokat	87
Használjunk tartományokat	87
Járjunk be több listát a zip használatával.....	87
Használjunk generátorokat a yield paranccsal.....	88
Indexeljünk adatokat szótárakkal.....	89
Készítsünk szótárakat.....	89
Tároljunk adatokat szótárakban és nyerjük ki adatokat szótárakból	89

6. fejezet: Használjuk a Google Colabot..... 91

Mi a Google Colab?.....	92
Mire képes a Google Colab?	92
Miben tér el az online kódolás?	94
Használjuk a helyi futtatókörnyezetes támogatást	95
Használjuk a Google Colab funkcióit	96
Regisztráljunk egy Google-fiókot	99
Hozzuk létre a fiókot	99
Jelentkezzünk be	99
Kezeljük jegyzetfüzeteket.....	101
Hozzuk létre egy új jegyzetfüzetet	101
Nyissunk meg meglévő jegyzetfüzeteket.....	102

Töltsünk fel jegyzetfüzeteket	104
Mentsünk el jegyzetfüzeteket	105
Töltsünk le jegyzetfüzeteket	107
Végezzünk el gyakori feladatokat	107
Hozzunk létre kódcellákat	108
Hozzunk létre szöveges cellákat.....	111
Hozzunk létre speciális cellákat	111
Szerkesszünk cellákat	112
Mozgassunk cellákat	112
Használjunk hardveres gyorsítást	113
Tekintsük meg a kódot	114
Jelenítsük meg a tartalomjegyzéket.....	114
Kérjünk le információkat a jegyzetfüzetről	114
Ellenőrizzük a kód végrehajtását.....	115
Futtassuk a kódot	115
Osszuk meg a jegyzetfüzetet.....	116
Kérjünk segítséget.....	117

3. rész: ISMERKEDJÜNK MEG A MATEMATIKAI ALAPOKKAL..... 119

7. fejezet: Tisztázzuk a gépi tanulás matematikai alapjait 121

Kezeljünk adatokat.....	122
Sajátítsuk el a terminológiát	123
Ismerjük meg a skalár- és vektorműveleteket	124
Szorozzunk össze vektorokat.....	127
Hozzunk létre mátrixokat	127
Ismerkedjünk meg az alpműveletekkel.....	129
Szorozzunk össze mátrixokat	130
Vessünk egy pillantást a haladó szintű mátrixműveletekre.....	132
Használjuk a vektorizálást hatékonyan	133
Fedezzük fel a valószínűségek világát.....	134
Mi az a valószínűség?.....	135
Végezzünk műveleteket valószínűségekkel.....	136
Vizsgáljuk a valószínűségek feltételeit a Bayes-tétellel	137
Hogyan használjuk a statisztikát?	140

8. fejezet: Ereszkedjünk le a gradiensen..... 145

Ismerjük meg a tanulás különböző fajtáit	146
Felügyelt tanulás	146
Felügyelet nélküli tanulás.....	147
Megerősítéses tanulás	147
A tanulási folyamat.....	148
Képezzünk le egy ismeretlen függvényt	149
Ismerkedjünk meg a költségfüggvényekkel.....	151
Ereszkedjünk le az optimalizálási görbén	154
Végezzünk optimalizálást big data adatokkal	155
Alkalmazzunk mintavételezést	156
Használjunk párhuzamosságot.....	157
Végezzünk tanulást a központi memórián kívül	158

9. fejezet: Ellenőrizzük a gépi tanulást	161
Hogyan használjuk a példaadatokat?	162
Ellenőrizzük a mintán kívüli hibákat	163
Ismerjük meg a minták fogalmát	163
Keressük meg az általánosítás Szent Grálját	164
Kísérletezzünk a torzítás és a variancia működésével	166
Vegyük figyelembe a modell összetettségét	169
Tartsuk egyensúlyban a megoldásokat	170
Rajzoljunk tanulási görbéket	172
Végezzünk betanítást, ellenőrzést és tesztelést	174
Tekintsük át az adatok felosztását	174
Alkalmazzunk keresztvalidációt	175
Keressünk alternatív módszereket az ellenőrzésre	177
Optimalizáljunk keresztvalidációval	178
Az előrejelzési teljesítmény forrásai	178
Fedezzük fel a hiperparaméterteret	179
Válasszuk ki a lényeges jellemzőket	181
Kerüljük el a minta torzítását és a szivárgással kapcsolatos csapdákat	182
10. fejezet: Egyszerű tanuló algoritmusok	185
Fedezzük fel a hihetetlen perceptront	186
A csoda elmarad	186
Érjük el a szétválaszthatatlanság határát	189
Neveljünk mohó osztályozó fákat	191
Jelezzük előre az eredményeket az adatok felosztásával	191
Metsszük meg a túlbujánzott fákat	196
Forduljunk a valószínűségi számítás felé	200
Ismerjük meg a naiv Bayes-algoritmust	201
Becsüljük meg a választ a naiv Bayes-algoritmussal	204
4. rész: TANULJUNK INTELLIGENS ÉS BIG DATA ADATOKBÓL	209
11. fejezet: Az adatok előfeldolgozása	211
Gyűjtsünk be adatokat, és tisztítsuk meg őket	212
Javítsuk a hiányzó adatokat	213
Azonosítsuk a hiányzó adatokat	214
Válasszuk ki a megfelelő helyettesítési stratégiát	215
Alakítsunk át eloszlásokat	218
Hozzunk létre saját jellemzőket	220
Derítsük ki, hogy miért szükséges jellemzőket létrehozni	220
Hozzunk létre jellemzőket automatikusan	221
Tekintsük át az SVD alapjait	223
Szervezzük át az adatokat	225
Határoljuk körül a rendellenes adatokat	227
Alkalmazzunk egyváltozós stratégiát	228
Alkalmazzunk többváltozós modelleket	230
12. fejezet: Használjuk ki a hasonlóságokat	233
Mérjük meg a vektorok közötti hasonlóságot	234
Értelmezzük a hasonlóságot	234
Számítsuk ki a távolságokat a tanuláshoz	235

Használjunk távolságokat a klaszterek megkereséséhez	237
Tekintsük át a feltételezéseket és az elvárásokat.....	238
Nézzünk be a K-közép algoritmus gépházába	240
Finomhangoljuk a K-közép algoritmust	241
Kísérletezzünk a K-közép megbízhatóságával	243
Kísérletezzünk a középpontok konvergálásával.....	246
Keressünk hasonlóságokat a k-legközelebbi szomszéddal	251
Ismerjük meg a k paramétert	251
Kísérletezzünk egy rugalmas algoritmussal.....	253
13. fejezet: A könnyebb út: használjunk lineáris modelleket.....	257
Kezdjük el jellemzőket egyesíteni	258
Tekintsük át, hogy mi az a regresszió	258
Problémamegoldás gépi tanulási megközelítéssel.....	261
Értelmezzük az R-négyzetet	264
Keverjük össze különböző típusú jellemzőket	266
Váltsunk át valószínűségekre	270
Adjunk meg egy bináris választ	270
Kezeljük több osztályt	274
Találjuk ki a megfelelő jellemzőket.....	275
Határozzuk meg az eredményt együtt nem használható jellemzők esetén.....	275
Oldjuk meg a túlillesztés problémáját mohó kiválasztással	276
Kezeljük a túlillesztést regularizálással	278
Tanuljunk meg egyszerre egy példát.....	280
Használjuk a gradiens csökkentést	280
Tisztázzuk, hogy miben különbözik az SGD.....	281
14. fejezet: Neurális hálózatokkal kezelünk bonyolult problémákat.....	287
Térjünk vissza a perceptronhoz	288
Haladjunk előre előreccsatolással.....	290
Menjünk még mélyebbre a rókalyukban	293
Forduljunk vissza a visszaterjesztéssel.....	297
Szemléltessük, hogy egy hálózat hogyan tanul	300
Értelmezzük a túlillesztés problémáját	300
Válasszunk keretrendszert	302
Szerezzünk be egy TensorFlow- és Keras-példányt.....	303
Nyissuk ki a fekete dobozt	307
Ismerkedjünk meg a mély tanulással.....	312
Tisztázzunk néhány alapvető dolgot a mély tanulással kapcsolatban	313
Ismerjük meg a konvolúciók varázslatát	314
Ismerkedjünk meg az ismétlődő neurális hálózatokkal.....	319
15. fejezet: Tegyük egy lépést tartóvektor-gépekkel	325
Térjünk vissza a szétválasztási problémához.....	326
Ismerjük meg az algoritmust.....	328
Kerüljünk el a szétválaszthatatlanság buktatóit.....	329
Alkalmazzunk nemlinearitást.....	331
Értelmezzük a kerneltrükköt egy példa alapján	332
Végezzünk osztályozást és becslést SVM-mel.....	335

16. fejezet: Alkalmazzunk együttes tanuló algoritmusokat	339
Használjunk döntési fákat	340
Növesszünk erdőt a fákból.....	342
Ismerkedjünk meg a fontossági mértékekkel.....	348
Használjunk majdnem véletlenszerű becsléseket.....	350
Zsákoljuk be a prediktorokat Adaboosttal.....	351
Gyorsítsunk intelligens előrejelzőket.....	354
Idézzük fel újra a gradiens csökkentést.....	355
Nézzük meg, melyek a legmodernebb technológiák táblázatos adatokhoz.....	357
Átlagoljunk különböző előrejelzőket	357
Keverjük össze megoldásokat	358
Halmozzunk különböző megoldásokat	359

5. rész: ALKALMAZZUK A TANULÁST VALÓS PROBLÉMÁKRA..... 361

17. fejezet: Osztályozunk képeket.....	363
Használjuk a képek egy készletét	364
Tekintsük át a legmodernebb számítógépes látási technológiákat	369
Nyerjük ki vizuális jellemzőket	372
Ismerjük fel arcokat sajátarcok használatával	374
Osztályozunk képeket	378
18. fejezet: Pontozunk véleményeket és érzelmeket.....	383
Ismerkedjünk meg a természetes nyelvi feldolgozással.....	384
Tekintsük át a legmodernebb NLP-technológiákat	385
Ismerjük meg, hogyan olvasnak a gépek	386
Határozzuk meg a bemeneti adatokat	387
Dolgozzunk fel és javítsunk szöveget	388
Nyerjük ki szöveges adathalmazokat az internetről.....	394
Kezeljük a nyers szöveggel kapcsolatos problémákat	398
Használjunk pontozást és osztályozást.....	398
Végezzünk osztályozási feladatokat.....	399
Elemizzünk e-kereskedelmi véleményeket	402
19. fejezet: Ajánljunk termékeket és filmeket.....	405
Hogyan valósul meg az e-kereskedelem forradalma?	406
Töltsünk le értékelési adatokat	408
Verekedjük át magunkat a MovieLens adathalmazon.....	408
Navigáljunk az anonim webes adatok között.....	412
Ismerjük meg az értékelési adatok korlátait	414
Ismerkedjünk meg a kollaboratív szűréssel.....	415
Ismerjük fel a viselkedési adatok korlátait.....	420
Integráljuk a szöveget és a viselkedést.....	422
Nézzük meg az attribútumokat.....	422
Szerezzünk be statisztikai adatokat	423
Alkalmazzunk SVD-t	423
Vegyük át, hogy kerül a képbe az SVD	424
Nézzük meg az SVD-t működés közben	425

6. rész: TOP 10.....	429
20. fejezet: Tíz módszer, amellyel javíthatsz a gépi tanulási modelljeiden	431
Tanulmányozzuk a tanulási görbéket	432
Használjuk megfelelően a keresztvalidációt	433
Válasszuk ki a megfelelő hiba- vagy pontszámmetrikát	434
Keressük meg a legjobb hiperparamétereket	434
Teszteljünk több modellt.....	435
Átlagoljunk modelleket	436
Halmozzunk modelleket	436
Alkalmazzunk jellemzőtervezést	437
Válasszuk ki jellemzőket és példákat	438
Kutassunk további adatok után.....	438
21. fejezet: Tíz irányelv az adatok etikus felhasználásához	441
Szerezzünk engedélyt.....	442
Használjunk technikákat az adatok megtisztítására	442
Kerüljük el az adatok kikövetkeztetését	443
Használjuk helyesen az általánosítást	444
Kerüljük el a diszkriminatív gyakorlatokat	445
Ismerjük fel a fekete hattyúkat a kódban	445
Legyünk tisztában a folyamattal.....	446
Vegyük figyelembe a tettek következményeit	446
Tegyük kiegyensúlyozottá a döntéshozatalt.....	447
Ellenőrizzük az adatforrásokat	448
22. fejezet: Tíz gépi tanulási csomag	449
Gensim	449
imbalanced-learn.....	450
OpenCV	450
SciPy.....	451
SHAP.....	452
Statsmodels.....	452
Modin	453
PyTorch.....	453
Poetry	454
Snorkel.....	454
Tárgymutató	455
Végjegyzetek	467