

# Indice generale

<b>Autori e collaboratori .....</b>	<b>xv</b>
Gli autori.....	xv
I revisori.....	xv
<b>Ringraziamenti .....</b>	<b>xvii</b>
<b>Introduzione .....</b>	<b>xix</b>
Primi passi con il machine learning.....	xix
Teoria e pratica .....	xix
Perché Python?.....	xx
Esplorare il campo del machine learning .....	xx
A chi è rivolto questo libro .....	xx
Argomenti trattati .....	xxi
Dotazione software necessaria .....	xxiii
Download dei file di codice degli esempi .....	xxiii
Convenzioni impiegate .....	xxiv
<b>Capitolo 1 Dare ai computer la capacità di apprendere dai dati .....</b>	<b>1</b>
Costruire macchine intelligenti per trasformare i dati in conoscenza.....	2
I tre diversi tipi di machine learning.....	2
Fare predizioni sul futuro tramite l'apprendimento con supervisione.....	2
Risolvere problemi interattivi con il reinforcement learning.....	6
Individuare strutture nascoste con l'apprendimento senza supervisione.....	7
Introduzione alla terminologia e alle notazioni di base.....	8
Notazione e convenzioni impiegate in questo libro .....	9
Terminologia per il machine learning.....	10

Una roadmap per la costruzione di sistemi di machine learning .....	11
Pre-elaborazione: uniformazione dei dati.....	12
Addestramento e scelta di un modello predittivo .....	12
Valutazione dei modelli e predizioni su istanze dei dati mai viste.....	13
Usare Python per il machine learning .....	13
Installazione di Python e dei pacchetti dal Python Package Index .....	14
Uso della distribuzione di Python e del gestore di pacchetti Anaconda.....	14
Pacchetti per il calcolo scientifico, la data science e il machine learning .....	15
Riepilogo .....	15

## **Capitolo 2 Addestrare gli algoritmi a compiti di classificazione ..... 17**

Neuroni artificiali: breve introduzione agli alberi del machine learning.....	17
La definizione formale di neurone artificiale .....	18
La regola di apprendimento del perceptron .....	20
Implementazione di un algoritmo di apprendimento perceptron in Python.....	23
Un'API object-oriented per un perceptron .....	23
Addestramento di un modello di perceptron sul dataset Iris.....	26
Neuroni adattativi lineari e convergenza dell'apprendimento .....	32
Minimizzare le funzioni di costo con la discesa del gradiente.....	33
Implementare Adaline in Python.....	35
Migliorare la discesa del gradiente tramite la riduzione in scala delle caratteristiche .....	39
Machine learning su larga scala e discesa stocastica del gradiente.....	41
Riepilogo .....	46

## **Capitolo 3 Tour dei classificatori di machine learning con scikit-learn ..... 47**

Scelta di un algoritmo di classificazione .....	47
Primi passi con scikit-learn: addestramento di un perceptron .....	48
Modellazione delle probabilità delle classi tramite la regressione logistica .....	54
Regressione logistica e probabilità condizionali .....	54
Apprendimento dei pesi della funzione di costo logistico.....	58
Conversione di un'implementazione Adaline in un algoritmo di regressione logistica .....	60

Addestramento di un modello a regressione logistica con scikit-learn.....	64
Risoluzione dell'overfitting tramite la regolarizzazione .....	66
Classificazione a massimo margine con le macchine a vettori di supporto .....	69
L'idea della massimizzazione del margine .....	70
Soluzione di un caso non separabile linearmente impiegando variabili slack .....	71
Implementazioni alternative in scikit-learn.....	74
Risoluzione di problemi non lineari impiegando una SVM kernel .....	74
Metodi kernel per dati che non sono separabili linearmente .....	74
Uso del trucco kernel per trovare iperpiani di separazione in uno spazio a elevata dimensionalità.....	76
Apprendimento ad alberi decisionali .....	80
Massimizzazione del guadagno informativo: minima spesa, massima resa .....	81
Costruzione di un albero decisionale.....	84
Uso combinato di più alberi decisionali a formare foreste casuali.....	88
K-Nearest Neighbor: un algoritmo di apprendimento "pigro".....	91
Riepilogo .....	94

<b>Capitolo 4</b>	<b>Costruire buoni dataset di addestramento: la pre-elaborazione .....</b>	<b>97</b>
	Il problema dei dati mancanti.....	97
	Identificazione dei valori mancanti nei dati in formato tabulare .....	98
	Eliminazione degli esempi di addestramento o delle caratteristiche con valori mancanti.....	99
	Attribuzione dei valori mancanti.....	100
	L'API degli estimator di scikit-learn .....	101
	Gestione di dati categorici .....	102
	Codifica di dati categorici con pandas.....	103
	Mappatura delle caratteristiche ordinali .....	103
	Codifica delle etichette delle classi .....	104
	Applicazione della codifica one-hot alle caratteristiche nominali.....	105
	Partizionamento di un dataset nei set di addestramento e di test...108	
	Adattamento di scala delle caratteristiche.....	111
	Selezione delle caratteristiche significative .....	113
	Regolarizzazione L1 e L2 per introdurre penalità contro la complessità del modello .....	114
	Un'interpretazione geometrica della regolarizzazione L2.....	114
	Soluzioni sparse con la regolarizzazione L1 .....	116

Algoritmi sequenziali per la selezione delle caratteristiche .....	120
Valutazione dell'importanza delle caratteristiche con le foreste casuali .....	125
Riepilogo .....	128

## **Capitolo 5 Comprimere i dati tramite riduzione della dimensionalità.....129**

Riduzione della dimensionalità senza supervisione tramite l'analisi dei componenti principali .....	130
I passi generali dell'analisi dei componenti principali .....	130
Estrazione dei componenti principali passo dopo passo.....	132
Varianza totale e spiegata.....	134
Trasformazione delle caratteristiche.....	135
Analisi dei componenti principali con scikit-learn.....	138
Compressione dei dati con supervisione tramite l'analisi dei discriminanti lineari .....	141
Analisi dei componenti principali vs. analisi dei discriminanti lineari .....	142
Il funzionamento interno dell'analisi dei discriminanti lineari .....	143
Calcolo delle matrici di dispersione.....	143
Selezione dei discriminanti lineari per il nuovo sottospazio delle caratteristiche .....	146
Proiezione degli esempi sul nuovo spazio delle caratteristiche .....	148
Analisi dei discriminanti lineari con scikit-learn .....	149
Uso della kernel PCA per mappature non lineari .....	150
Le funzioni kernel e il trucco kernel .....	152
Implementazione in Python di un'analisi a kernel dei componenti principali .....	156
Proiezione di nuovi punti dei dati.....	162
Analisi a kernel dei componenti principali in scikit-learn .....	166
Riepilogo .....	167

## **Capitolo 6 Valutare i modelli e ottimizzare gli iperparametri.....169**

Semplificazione dei flussi di lavoro tramite le pipeline.....	169
Caricamento del dataset Breast Cancer Wisconsin .....	170
Combinare i transformer e gli estimator in una pipeline .....	171
Uso della convalida incrociata k-fold per valutare le prestazioni di un modello .....	173
Il metodo holdout .....	173
Convalida incrociata k-fold.....	175
Debugging degli algoritmi con le curve di apprendimento e convalida.....	178

Diagnosi dei problemi di bias e varianza con le curve di apprendimento .....	179
Affrontare l'overfitting e l'underfitting con le curve di convalida .....	182
Ottimizzazione dei modelli di machine learning con la ricerca a griglia .....	184
Ottimizzazione degli iperparametri tramite la ricerca a griglia.....	184
Selezione dell'algoritmo con la convalida incrociata annidata..	186
Le diverse metriche di valutazione delle prestazioni .....	187
Lettura di una matrice di confusione.....	187
Ottimizzazione di precisione e recall di un modello di classificazione.....	189
Tracciamento di un grafico ROC (Receiver Operating Characteristic) .....	192
Metriche di valutazione per i classificatori multi-classe .....	194
Il problema dello squilibrio fra le classi.....	195
Riepilogo .....	197

## **Capitolo 7    Combinare fra loro più modelli: ensemble learning .....199**

Machine learning con gli ensemble .....	199
Combinazione di più classificatori tramite il voto a maggioranza.....	203
Implementazione di un semplice classificatore con voto a maggioranza.....	204
Uso del principio di votazione a maggioranza per eseguire predizioni.....	209
Valutazione e ottimizzazione del classificatore ensemble .....	212
Bagging: costruire un assieme di classificatori da campioni di bootstrap.....	217
Il bagging in poche parole .....	218
Applicazione del bagging per classificare gli esempi del dataset Wine.....	219
Impiego di modelli di apprendimento deboli tramite l'adaptive boosting .....	222
Come funziona il boosting .....	223
Applicazione di AdaBoost con scikit-learn .....	227
Riepilogo .....	230

## **Capitolo 8    Applicare il machine learning all'analisi del sentiment .....231**

Preparazione dei dati delle recensioni di film in IMDb per scopi di elaborazione del testo .....	232
Download del dataset di recensioni di film .....	232

Pre-elaborazione del dataset dei film in un formato più comodo.....	232
Introduzione al modello bag-of-words .....	234
Trasformazione di parole in vettori delle caratteristiche .....	235
Valutazione della rilevanza delle parole tramite la tecnica Term Frequency – Inverse Document Frequency.....	236
Pulizia dei dati testuali .....	238
Elaborazione dei documenti per estrarne i token.....	240
Addestramento di un modello a regressione logistica per la classificazione dei documenti .....	242
Lavorare sui big data: algoritmi online e out-of-core learning .....	245
Modellazione degli argomenti con l’allocazione latente di Dirichlet.....	248
Decomposizione di documenti testuali con l’allocazione latente di Dirichlet .....	248
Allocazione latente di Dirichlet con scikit-learn.....	249
Riepilogo .....	252

## **Capitolo 9 Embedding di un modello in un’applicazione web.....253**

Serializzazione di estimator di scikit-learn già configurati .....	254
Configurazione di un database SQLite per l’archiviazione dei dati .....	257
Sviluppo di un’applicazione web con Flask.....	259
La nostra prima applicazione web in Flask.....	259
Convalida e rendering di form.....	261
Trasformazione del classificatore di recensioni in un’applicazione web .....	266
File e cartelle: l’aspetto dell’albero delle directory .....	268
Implementazione dell’applicazione principale come app.py .....	269
Configurazione del form per la recensione .....	271
Creazione di un template per la pagina dei risultati .....	272
Pubblicazione (deploy) dell’applicazione web su un server pubblico .....	274
Creazione di un account PythonAnywhere.....	274
Caricamento sul server dell’applicazione del classificatore di film .....	275
Aggiornamento del classificatore di film.....	276
Riepilogo .....	279

## **Capitolo 10 Predire variabili target continue con l’analisi a regressione .....281**

Introduzione alla regressione lineare .....	282
Regressione lineare semplice.....	282

Regressione lineare multipla .....	283
Esplorazione del dataset Housing .....	284
Caricamento del dataset Housing in un data frame .....	284
Visualizzazione delle caratteristiche importanti di un dataset...286	
Osservazione delle relazioni impiegando una matrice di correlazione.....	287
Implementazione di un modello a minimi quadrati ordinari .....	289
Risoluzione dei parametri della regressione con la discesa del gradiente .....	290
Stima del coefficiente di un modello a regressione tramite scikit-learn.....	293
Configurazione di un solido modello a regressione con RANSAC .....	295
Valutazione delle prestazioni dei modelli a regressione lineare .....	297
Utilizzo di metodi regolarizzati per la regressione .....	300
Trasformazione di un modello a regressione lineare in una curva: la regressione polinomiale.....	302
Aggiunta di termini polinomiali con scikit-learn .....	302
Modellazione delle relazioni non lineari contenute nel dataset Housing.....	304
Cattura di relazioni non lineari tramite foreste casuali.....	307
Regressione ad alberi decisionali.....	307
Regressione a foresta casuale .....	309
Riepilogo .....	312

<b>Capitolo 11 Lavorare con dati senza etichette: l'analisi dei cluster .....</b>	<b>313</b>
Raggruppamento di oggetti per similarità con k-means.....	313
Clustering k-means con scikit-learn.....	314
Un modo più intelligente per collocare i centroidi iniziali dei cluster impiegando k-means++ .....	318
Clustering hard e soft.....	319
Impiego del metodo elbow per trovare il numero ottimale di cluster .....	321
Quantificazione della qualità del clustering tramite grafici a silhouette.....	322
Organizzazione dei cluster come un albero gerarchico.....	326
Raggruppamento di cluster in stile bottom-up.....	327
Esecuzione del clustering gerarchico su una matrice delle distanze .....	328
Collegamento dei dendrogrammi a una mappa termica.....	332
Applicazione del clustering agglomerativo tramite scikit-learn .....	333
Individuazione delle regioni di alta densità tramite DBSCAN .....	334
Riepilogo .....	339

<b>Capitolo 12 Implementare una rete neurale artificiale a layer multipli .....</b>	<b>341</b>
Modellazione di funzioni complesse con reti neurali artificiali .....	341
Ripasso sulle reti neurali a un solo layer .....	343
Introduzione all'architettura a layer multipli .....	345
Attivazione di una rete neurale tramite propagazione in avanti .....	348
Classificazione di cifre scritte a mano .....	350
Download e preparazione del dataset MNIST .....	351
Implementazione di un perceptron a layer multipli .....	356
Addestramento di una rete neurale artificiale .....	367
Calcolo della funzione di costo logistico .....	367
Approfondimento sulla retropropagazione .....	369
Addestramento di reti neurali tramite retropropagazione .....	371
La questione della convergenza nelle reti neurali .....	374
Conclusioni sull'implementazione della rete neurale .....	375
Riepilogo .....	376

<b>Capitolo 13 Parallelizzare l'addestramento di reti neurali con TensorFlow .....</b>	<b>377</b>
TensorFlow e le prestazioni dell'addestramento .....	378
Sfide prestazionali .....	378
Che cos'è TensorFlow? .....	379
Come studieremo TensorFlow .....	380
Primi passi con TensorFlow .....	380
Installazione di TensorFlow .....	381
Creazione di tensori in TensorFlow .....	382
Manipolazione del tipo di dati e della forma di un tensore .....	382
Applicazione di operazioni matematiche ai tensori .....	383
Split, stack e concatenamento di tensori .....	385
Costruzione di pipeline di input con tf.data:	
l'API per Dataset di TensorFlow .....	386
Creazione di un dataset di TensorFlow partendo dai tensori esistenti .....	387
Combinazione di due tensori a formare un unico dataset .....	388
Shuffle, batch e repeat .....	389
Creazione di un dataset a partire dai file contenuti sui dischi locali .....	392
Lettura dei dataset disponibili dalla libreria tensorflow_datasets .....	394
Costruzione di un modello a rete neurale in TensorFlow .....	400
L'API Keras per TensorFlow (tf.keras) .....	400
Costruzione di un modello a regressione lineare .....	401
Addestramento del modello tramite i metodi .compile() e .fit() .....	405



Costruzione di un perceptron a layer multipli per classificare i fiori del dataset Iris.....	406
Valutazione del modello addestrato sul dataset di test.....	410
Salvataggio e ricarica del modello addestrato.....	410
Scelta delle funzioni di attivazione per le reti neurali a layer multipli.....	411
Ripasso sulla funzione logistica.....	411
Stima della probabilità delle classi nella classificazione multi-classe tramite la funzione softmax.....	413
Ampliamento dello spettro di output con una tangente iperbolica.....	414
Attivazione a unità lineare rettificata (ReLU).....	416
Riepilogo.....	417

## **Capitolo 14 Approfondimenti: come funziona TensorFlow.....419**

Le funzionalità chiave di TensorFlow.....	420
I grafi di calcolo di TensorFlow: migrazione a TensorFlow v2.....	421
Che cosa sono i grafi di calcolo.....	421
Creazione di un grafo in TensorFlow v1.x.....	422
Migrazione di un grafo a TensorFlow v2.....	423
Caricamento dei dati di input in un modello: lo stile TensorFlow v1.x.....	423
Caricamento dei dati di input in un modello: lo stile TensorFlow v2.....	424
Miglioramento delle prestazioni computazionali con i decorator di funzioni.....	424
Gli oggetti Variable TensorFlow per la memorizzazione e l'aggiornamento dei parametri del modello.....	426
Calcolo dei gradienti tramite differenziazione automatica e GradientTape.....	429
Calcolo dei gradienti della funzione loss rispetto a delle variabili addestrabili.....	430
Calcolo dei gradienti rispetto a tensori non addestrabili.....	431
Risparmio di risorse per il calcolo di più gradienti.....	432
Semplificazione delle implementazioni delle architetture più comuni tramite l'API Keras.....	433
Risoluzione di un problema di classificazione XOR.....	435
Come rendere più flessibile la costruzione del modello con l'API funzionale Keras.....	440
Implementazione dei modelli basati sulla classe Model di Keras.....	441
Scrittura di layer Keras personalizzati.....	442
Gli estimator di TensorFlow.....	446
Lavorare sulle colonne delle caratteristiche.....	446
Machine learning con estimator pronti all'uso.....	450

Usò degli estimator per la classificazione delle cifre scritte a mano del dataset MNIST .....	454
Creazione di un estimator personalizzato a partire da un modello Keras .....	456
Riepilogo .....	458

## **Capitolo 15 Classificare immagini con le reti neurali convoluzionali profonde .....459**

Gli elementi costitutivi delle reti neurali convoluzionali.....	459
Le reti neurali convoluzionali e le gerarchie di caratteristiche .....	460
Esecuzione di convoluzioni discrete .....	462
I layer di subcampionamento .....	471
Ricapitolando: implementazione di una rete neurale convoluzionale.....	472
Utilizzo di più input o canali per i colori.....	473
Regolarizzazione di una rete neurale con il dropout.....	476
Funzioni loss per la classificazione .....	479
Implementazione di una rete neurale convoluzionale profonda con TensorFlow.....	481
L'architettura a layer multipli di una rete neurale convoluzionale.....	481
Caricamento e pre-elaborazione dei dati.....	482
Implementazione di una rete neurale convoluzionale impiegando l'API Keras per TensorFlow.....	483
Classificazione del genere dalle immagini di volti con una rete neurale convoluzionale.....	489
Caricamento del dataset CelebA .....	489
Trasformazione delle immagini e data augmentation .....	490
Addestramento di una rete neurale convoluzionale per la classificazione del genere .....	496
Riepilogo .....	501

## **Capitolo 16 Modellare dati sequenziali con le reti neurali ricorrenti .....503**

Introduzione ai dati sequenziali .....	504
Modellazione di dati sequenziali: l'ordine è importante .....	504
Rappresentazione delle sequenze .....	505
Le varie categorie di modellazione delle sequenze.....	505
Reti neurali ricorrenti per la modellazione delle sequenze.....	506
Il meccanismo a ciclo della rete neurale ricorrente .....	507
Calcolo delle attivazioni in una rete neurale ricorrente.....	509
Ricorrenza nascosta vs. ricorrenza nell'output .....	511

Le sfide legate all'apprendimento di interazioni a distanza.....	514
Celle di memoria Long short-term memory .....	515
Implementazione in TensorFlow di reti neurali ricorrenti per la modellazione di sequenze .....	517
Progetto 1: predizione del sentiment delle recensioni di film su IMDb .....	518
Progetto 2: modellazione del linguaggio naturale a livello dei caratteri in TensorFlow .....	531
Comprensione del linguaggio naturale con il modello Transformer .....	542
Funzionamento del meccanismo di self-attention .....	543
Multi-Head Attention e il blocco Transformer .....	545
Riepilogo .....	547

## **Capitolo 17 Reti generative avversarie per la sintesi di nuovi dati .....549**

Introduzione alle reti generative avversarie .....	549
Cominciamo dagli autoencoder .....	550
Modelli generativi per la sintesi di nuovi dati .....	552
Generazione di nuovi campioni tramite reti generative avversarie.....	554
Le funzioni loss delle reti generatrice e discriminatrice in un modello a reti generative avversarie.....	555
Implementazione di una rete generativa avversaria partendo da zero .....	556
Addestramento di modelli a reti generative avversarie su Google Colab.....	557
Implementazione delle reti generatrice e discriminatrice .....	560
Definizione del dataset di addestramento.....	563
Addestramento del modello a reti generative avversarie .....	565
Miglioramento della qualità delle immagini sintetizzate grazie a una GAN profonda convoluzionale o di Wasserstein .....	572
Convoluzione trasposta.....	573
Normalizzazione batch .....	574
Implementazione del generatore e del discriminatore .....	576
Misurazione della dissimilarità fra due distribuzioni .....	582
Uso pratico della distanza EM per le reti generative avversarie.....	586
Penalizzazione dei gradienti .....	586
Implementazione della WGAN-GP per addestrare il modello DCGAN.....	587
Mode collapse .....	591
Altre applicazioni delle reti generative avversarie .....	592
Riepilogo .....	593

<b>Capitolo 18</b>	<b>Reinforcement learning per decisioni in ambienti complessi.....</b>	<b>595</b>
	Introduzione: imparare dall'esperienza.....	596
	Che cos'è il reinforcement learning .....	596
	Definizione dell'interfaccia agente-ambiente in un sistema di reinforcement learning.....	598
	Le basi teoriche del reinforcement learning .....	599
	I processi decisionali di Markov .....	599
	La formulazione matematica dei processi decisionali di Markov.....	600
	Terminologia del reinforcement learning: ritorno, policy e funzione valore.....	603
	Programmazione dinamica con l'equazione di Bellman .....	606
	Algoritmi di reinforcement learning.....	607
	Programmazione dinamica.....	608
	Reinforcement learning con Monte Carlo .....	610
	Temporal Difference .....	612
	Implementazione di un primo algoritmo di reinforcement learning.....	615
	Introduzione al toolkit OpenAI Gym .....	615
	Soluzione del problema grid world con Q-learning .....	624
	Panoramica sul deep Q-learning .....	627
	Riepilogo del capitolo e del libro .....	635
	<b>Indice analitico.....</b>	<b>639</b>