

Introduzione

Questo libro vi insegna due cose: i modelli di machine learning e come usarli. Esistono diversi tipi di modelli di machine learning. Alcuni restituiscono una risposta deterministica, sì o no; altri restituiscono la risposta come una probabilità. Alcuni usano equazioni; altri usano istruzioni `if`. Una cosa che hanno in comune è che restituiscono tutti una risposta, o una previsione. Il ramo del machine learning che comprende i modelli che restituiscono una previsione è giustamente chiamato *machine learning predittivo*. Questo è il tipo di machine learning su cui ci concentreremo in questo libro.

Come è organizzato il libro: una roadmap

Tipi di capitoli

Questo libro contiene due tipi di capitoli. La maggior parte di essi (i Capitoli 3, 5, 6, 8, 9, 10, 11 e 12) tratta uno specifico modello di machine learning. Il modello corrispondente in ogni capitolo viene studiato in dettaglio, inclusi esempi, formule, codice ed esercizi da risolvere. Altri capitoli (i Capitoli 4, 7 e 13) contengono tecniche da utilizzare per addestrare, valutare e migliorare i modelli di machine learning. In particolare, il Capitolo 13 contiene un esempio completamente svolto su un dataset reale, in cui potrete applicare tutte le conoscenze acquisite nei capitoli precedenti.

Percorsi consigliati

Potete utilizzare questo libro in due modi. Quello che vi consiglio è di affrontarlo in modo lineare, capitolo per capitolo, perché scoprirete che l'alternanza tra modelli di apprendimento e tecniche di apprendimento per addestrarli è gratificante. Tuttavia, un altro percorso di apprendimento consiste nell'imparare prima tutti i modelli (Capitoli 3, 5, 6, 8, 9, 10, 11 e 12) e poi le tecniche per addestrarli (Capitoli 4, 7 e 13). E, ovviamente, poiché tutti impariamo in modi differenti, siete liberi di creare il vostro personale percorso di apprendimento.

Appendici

Questo libro ha tre appendici. L'Appendice A contiene le soluzioni degli esercizi di ciascun capitolo. L'Appendice B contiene alcuni calcoli matematici formali, utili ma più tecnici rispetto al resto del libro. L'Appendice C contiene un elenco di riferimenti e risorse che vi consiglio, se desiderate approfondire la vostra comprensione.

Requisiti e obiettivi di apprendimento

Questo libro fornisce una solida conoscenza del machine learning predittivo. Per ottenere il massimo da questo libro, dovrete adottare una mentalità visiva e avere una certa conoscenza della matematica elementare: grafici, equazioni e calcolo delle probabilità. È utile (ma non obbligatorio) saper programmare, soprattutto in Python, perché in tal modo avrete l'opportunità di implementare e applicare diversi modelli su dataset reali. Dopo aver letto questo libro, sarete in grado di fare quanto segue.

- Descrivere i modelli più importanti del machine learning predittivo e il loro funzionamento, tra cui quello di regressione lineare e logistica, naive Bayes, alberi decisionali, reti neurali, macchine a vettori di supporto e metodi ensemble.
- Identificare i loro punti di forza e di debolezza, e i parametri che essi utilizzano.
- Scoprire come questi modelli vengono utilizzati nel mondo reale e formulare potenziali modi per applicare il machine learning a qualsiasi problema desiderate risolvere.
- Imparare a ottimizzare questi modelli, a confrontarli e a migliorarli per creare i migliori modelli di machine learning possibili.
- Programmare i modelli, manualmente o utilizzando un pacchetto, e utilizzarli per fare previsioni su dataset reali.

Se avete in mente un determinato dataset o problema, vi invito a riflettere su come applicargli ciò che imparerete in questo libro e a utilizzarlo come punto di partenza per implementare e sperimentare i vostri modelli.

Sono davvero entusiasta di iniziare questo viaggio con voi e spero di rendervi altrettanto entusiasti.

Altre risorse

Questo libro è autonomo. Ciò significa che, a parte i requisiti descritti in precedenza, nel libro viene descritto ogni concetto di cui abbiamo bisogno. Tuttavia, includo molti riferimenti, che vi consiglio di consultare se desiderate comprendere i concetti a un livello più profondo o se desiderate esplorare ulteriori argomenti. I riferimenti sono tutti nell'Appendice C e anche a questo indirizzo: <http://serrano.academy/grokking-machine-learning>. In particolare, molte delle risorse che suggerisco accompagnano il materiale di questo libro. Nella mia pagina su <http://serrano.academy> potete trovare molti materiali sotto forma

di video, post e codice. I video sono anche nel mio canale YouTube www.youtube.com/c/LuisSerrano, che vi consiglio di visitare. La maggior parte dei capitoli di questo libro ha un video corrispondente, che vi consiglio di guardare mentre leggete il capitolo.

Scriveremo del codice

In questo libro scriveremo codice Python. Tuttavia, se pensate di apprendere i concetti senza usare il codice, potete comunque seguire il libro ignorando le implementazioni. In ogni caso, vi consiglio di dare almeno un'occhiata al codice, per familiarizzare col suo funzionamento.

Questo libro è dotato di un repository di codice e la maggior parte dei capitoli vi darà l'opportunità di programmare gli algoritmi da zero o di utilizzare alcuni pacchetti Python molto noti per creare modelli adatti a specifici dataset. Il repository GitHub è www.github.com/luisguiserrano/manning e in tutto il libro presento i riferimenti ai notebook corrispondenti. Nel file `README` del repository troverete le istruzioni per i pacchetti da installare per eseguire correttamente il codice.

I principali pacchetti Python che utilizziamo in questo libro sono i seguenti.

- NumPy: per memorizzare array ed eseguire calcoli matematici complessi.
- Pandas: per archiviare, manipolare e analizzare dataset di grandi dimensioni.
- Matplotlib: per tracciare i dati.
- Turi Create: per archiviare e manipolare i dati e addestrare i modelli di machine learning.
- Scikit-Learn: per addestrare i modelli di machine learning.
- Keras (TensorFlow): per addestrare le reti neurali.

Il codice nel testo

Questo libro contiene molti esempi di codice sorgente formattato con un carattere `mono-spaziato`, per distinguerlo rispetto al testo normale. A volte il codice è anche in grassetto, per evidenziare le parti modificate rispetto ai passaggi precedenti del capitolo, per esempio quando a una riga di codice esistente aggiungo una nuova funzionalità.

In molti casi, il codice sorgente originale è stato riformattato; abbiamo dovuto aggiungere interruzioni di riga e rielaborare le indentazioni a causa dei vincoli della pagina del libro. Inoltre, i commenti nel codice sorgente sono stati spesso rimossi dai listati, in quanto il codice è già descritto nel testo. Vi sono solo alcune annotazioni che accompagnano alcuni listati, per evidenziare i concetti più importanti.

Il codice degli esempi contenuti in questo libro è disponibile per il download sul sito web di Apogeo all'indirizzo <https://bit.ly/apo-mlmf> e su GitHub all'indirizzo www.github.com/luisguiserrano/manning.

L'autore



Luis G. Serrano è uno scienziato ricercatore nel campo dell'intelligenza artificiale quantitativa presso Zapata Computing. Ha lavorato in precedenza come Machine Learning Engineer presso Google, come Lead Artificial Intelligence Educator presso Apple e come Head of Content in Artificial Intelligence and Data Science presso Udacity. Luis ha conseguito un dottorato di ricerca in Matematica presso l'Università del Michigan, ha una laurea e un master in Matematica presso l'Università di Waterloo e ha lavorato come ricercatore post-dottorato presso il Laboratoire de Combinatoire et d'Informatique Mathématique presso l'Università del Quebec a Montreal. Luis gestisce un noto canale YouTube sul machine learning con oltre 85.000 iscritti e oltre 4 milioni di visualizzazioni e tiene spesso conferenze sull'intelligenza artificiale e sulla scienza dei dati.

Ringraziamenti

Innanzitutto, vorrei ringraziare la mia editor, Marina Michaels, senza la quale questo libro non esisterebbe. La sua organizzazione, il suo editing accurato e i suoi preziosi input hanno contribuito a dare forma a questo libro. Ringrazio Marjan Bace, Bert Bates e tutti quelli del team Manning per il loro supporto, la loro professionalità, le grandi idee che mi hanno dato e la pazienza che hanno manifestato. Ringrazio chi si è occupato della revisione tecnica, Shirley Yap e Karsten Stræk, e della revisione del codice, Kris Athi e tutti i revisori per avermi dato ottimi feedback e per aver corretto molti dei miei errori. Ringrazio la redattrice di produzione, Keri Hales, la redattrice, Pamela Hunt, la redattrice grafica, Jennifer Houle, il correttore di bozze, Jason Everett, e l'intero team di produzione per il loro meraviglioso lavoro nel rendere questo libro una realtà. Ringrazio Laura Montoya per il suo aiuto nell'uso di un linguaggio inclusivo e nel trattare l'etica dell'IA, Diego Hernandez per le preziose aggiunte al codice e Christian Pic per il suo immenso aiuto con gli aspetti tecnici del repository e dei pacchetti.

Sono grato a Sebastian Thrun per il suo eccellente lavoro di democratizzazione dell'istruzione. Udacity è stata la piattaforma che per prima mi ha dato voce, per permettermi di insegnare al mondo, e vorrei ringraziare i meravigliosi colleghi e studenti che vi ho incontrato. Alejandro Perdomo e il team di Zapata Computing meritano un ringraziamento per avermi fatto conoscere il mondo del machine learning quantistico. Grazie anche ai tanti meravigliosi dirigenti e colleghi che ho incontrato in Google e Apple e che sono stati determinanti per la mia carriera. Un ringraziamento speciale a Roberto Cipriani e al team di Paper Inc. per avermi permesso di far parte della famiglia e per lo splendido lavoro che svolgono nella comunità educativa.

Vorrei ringraziare i miei numerosi mentori accademici, che hanno plasmato la mia carriera e il mio modo di pensare: Mary Falk de Losada e il suo team alle Olimpiadi della matematica colombiana, dove ho iniziato ad amare la matematica e ho avuto la possibilità di incontrare grandi personaggi e creare amicizie che durano da una vita; il mio relatore per il dottorato, Sergey Fomin, è stato determinante nella mia formazione matematica e nel mio stile di insegnamento; l'assistente del mio professore, Ian Goulden; Nantel e François Bergeron, Bruce Sagan e Federico Ardila, e i numerosi professori e colleghi con cui ho avuto l'opportunità di lavorare, in particolare quelli delle Università di Waterloo, Michigan, Quebec a Montreal e York; e infine Richard Hoshino, il team e gli studenti della Quest University, che mi hanno aiutato a sottoporre a test e migliorare il materiale di questo libro.

A tutti i revisori: Al Pezewski, Albert Nogués Sabater, Amit Lamba, Bill Mitchell, Borko Djurkovic, Daniele Andreis, Erik Sapper, Hao Liu, Jeremy R. Loscheider, Juan Gabriel Bono, Kay Engelhardt, Krzysztof Kamyczek, Matthew Margolis, Matthias Busch, Michael Bright, Millad Dagdoni, Polina Keselman, Tony Holdroyd e Valerie Parham-Thompson, i vostri suggerimenti hanno contribuito a rendere migliore questo libro.

Vorrei ringraziare mia moglie, Carolina Lasso, che mi ha supportato in ogni fase di questo processo con amore e gentilezza; mia mamma, Cecilia Herrera, che mi ha cresciuto con amore e mi ha sempre incoraggiato a seguire le mie passioni; mia nonna Maruja, per essere l'angelo che mi guarda dal cielo; il mio migliore amico, Alejandro Morales, per essere sempre stato disponibile; e i miei amici, che hanno illuminato il mio cammino e la mia vita, vi ringrazio e vi amo con tutto il cuore.

YouTube, blog, podcast e social media mi hanno dato la possibilità di entrare in contatto con migliaia di menti brillanti in tutto il mondo. Menti curiose con una passione infinita per l'apprendimento, colleghi educatori che condividono generosamente le loro conoscenze e intuizioni, formano una tribù elettronica che mi ispira ogni giorno e mi dà l'energia per continuare a imparare e insegnare. Ringrazio chiunque condivida le proprie conoscenze con il mondo o si sforzi di imparare ogni giorno.

Ringrazio chiunque, là fuori si stia impegnando per rendere questo mondo un luogo più giusto e pacifico. A chiunque combatta per la giustizia, per la pace, per l'ambiente e per le pari opportunità per ogni essere umano sulla Terra, indipendentemente dalla razza, dal sesso, dal luogo di nascita, dalle condizioni e dalle scelte. Li ringrazio dal profondo del cuore.

E infine, questo libro è dedicato a voi, lettori. Avete scelto la via dell'apprendimento, la via del miglioramento, la via del sentirvi a vostro agio anche nelle situazioni scomode, e questo è ammirevole. Spero che questo libro sia un passo positivo nel vostro percorso, per seguire le vostre passioni e creare un mondo migliore.