

Introduzione

Gli algoritmi rappresentano la forza trainante del mondo dei computer, dalle teorie fondamentali alle applicazioni pratiche. Questa edizione aggiornata approfondisce ulteriormente il mondo dinamico degli algoritmi, ampliando la portata per affrontare questioni attuali e concrete. Partendo dai rudimenti degli algoritmi, viaggiamo attraverso una miriade di tecniche di progettazione, fino ad arrivare ad aree complesse come la programmazione lineare, il posizionamento delle pagine, i grafi e un'esplorazione più profonda del machine learning. Per garantire di essere all'avanguardia nei progressi tecnologici, abbiamo incorporato discussioni approfondite su reti sequenziali, LLM, LSTM, GRU e sulla crittografia e l'implementazione di algoritmi a larga scala negli ambienti di cloud computing.

Anche l'importanza degli algoritmi nei sistemi di raccomandazione, un elemento fondamentale nell'era digitale di oggi, è stata meticolosamente dettagliata. Per gestire in modo efficace questi algoritmi, è fondamentale comprendere le basi matematiche e la logica sottostanti. I casi di studio concreti, che vanno dalle previsioni meteorologiche all'analisi dei tweet, ai suggerimenti per i film e all'approfondimento delle sfumature dei modelli LLM, esemplificano le loro applicazioni pratiche.

Grazie agli approfondimenti di questo libro, il nostro obiettivo è quello di rafforzare la vostra fiducia nell'implementazione degli algoritmi per affrontare le moderne sfide computazionali. Entrate in questo grande viaggio nella decifrazione e nell'utilizzo degli algoritmi nel panorama digitale in evoluzione di oggi.

A chi è rivolto questo libro

Se siete programmatori o sviluppatori interessati a sfruttare gli algoritmi per risolvere problemi e creare codice efficiente, questo libro fa certamente per voi. Dagli algoritmi classici e più utilizzati alle ultime novità nel campo della scienza dei dati, del machine learning e della crittografia, questa guida copre un ventaglio completo. Sebbene la familiarità con la programmazione Python sia utile, non è obbligatoria.

Una base in qualsiasi linguaggio di programmazione vi sarà comunque utile. Inoltre, anche se non siete programmatori ma avete una passione per la tecnica, da questo libro acquisirete informazioni sul vasto mondo degli algoritmi per la risoluzione dei problemi.

Argomenti trattati

La Parte 1 introduce i fondamenti e i principali algoritmi.

Il Capitolo 1 fornisce informazioni sui fondamenti degli algoritmi. Si parte dai concetti di base degli algoritmi, come si è iniziato a utilizzare gli algoritmi per formulare problemi e i limiti dei diversi algoritmi. Poiché, per scrivere gli algoritmi, in questo libro viene utilizzato Python, spieghiamo come impostare un ambiente Python per svolgere gli esempi. Vediamo quindi come quantificare le prestazioni di un algoritmo e confrontarle con altri algoritmi.

Il Capitolo 2 tratta le strutture di dati nel contesto degli algoritmi. Poiché in questo libro utilizziamo Python, il capitolo si concentra sulle strutture di dati di Python, ma i concetti presentati possono essere generalizzati ad altri linguaggi, come Java e C++. Questo capitolo vi mostrerà come Python gestisce strutture di dati complesse e quali strutture dovrebbero essere utilizzate per i vari tipi di dati.

Il Capitolo 3 inizia presentando diversi tipi di algoritmi di ordinamento e vari approcci per progettarli. Successivamente, seguendo alcuni esempi pratici, vengono discussi anche gli algoritmi di ricerca.

Il Capitolo 4 tratta le scelte a nostra disposizione per la progettazione di algoritmi, trattando l'importanza di caratterizzare il problema che ci accingiamo a risolvere. Successivamente, utilizza il famoso problema del commesso viaggiatore come caso d'uso e applica le tecniche di progettazione che presentiamo. Inoltre, introduce la programmazione lineare e ne discute le applicazioni.

Il Capitolo 5 illustra i modi in cui possiamo impiegare i grafi per rappresentare le strutture di dati. Tratta alcune teorie, tecniche e metodi fondamentali relativi agli algoritmi per i grafi, come l'analisi della teoria delle reti e gli attraversamenti dei grafi. Esaminiamo un caso di studio utilizzando algoritmi per i grafi, così da approfondire l'analisi delle frodi. La Parte 2 si occupa degli algoritmi di machine learning.

Il Capitolo 6 spiega come l'apprendimento senza supervisione può essere applicato ai problemi concreti. Impariamo a conoscere i suoi algoritmi e le sue metodologie di base, come gli algoritmi di clustering, la riduzione della dimensionalità e l'estrazione delle regole associative.

Il Capitolo 7 approfondisce gli elementi essenziali del machine learning con supervisione, presentando i classificatori e i regressori. Esploriamo le loro capacità utilizzando problemi concreti. Presentiamo sei diversi algoritmi di classificazione, seguiti da tre tecniche di regressione. Infine, confrontiamo i loro risultati per riassumere i principali punti salienti di questa discussione.

Il Capitolo 8 introduce i concetti e i componenti principali di una tipica rete neurale. Presenta i vari tipi di reti neurali e le funzioni di attivazione in esse utilizzate. Discutiamo in dettaglio l'algoritmo di backpropagation, che è l'algoritmo più utilizzato per addestrare una rete neurale. Infine, impariamo a utilizzare il deep learning per segnalare i documenti fraudolenti tramite un'applicazione di esempio reale.

Il Capitolo 9 introduce gli algoritmi per l'elaborazione del linguaggio naturale. Introduce i fondamenti dell'elaborazione del linguaggio naturale e la preparazione dei dati. Successivamente, spiega i concetti di vettorizzazione dei dati testuali e del word embedding. Infine, presentiamo un caso d'uso dettagliato.

Il Capitolo 10 esamina l'addestramento delle reti neurali per i dati sequenziali. Tratta i principi fondamentali dei modelli sequenziali, fornendo una panoramica introduttiva delle loro tecniche e metodologie. Considera poi come il deep learning possa migliorare le tecniche di elaborazione del linguaggio naturale.

Il Capitolo 11 considera i limiti dei modelli sequenziali e il modo in cui la modellazione sequenziale si è evoluta per superare tali limiti. Approfondisce gli aspetti avanzati dei modelli sequenziali per comprendere la creazione di configurazioni complesse. Inizia suddividendo gli elementi chiave, come gli autoencoder e i modelli Seq2Seq. Successivamente, esamina i meccanismi dell'attenzione e i transformer, che sono fondamentali nello sviluppo dei modelli LLM (*Large Language Models*).

La Parte 3 è dedicata ad alcuni argomenti avanzati.

Il Capitolo 12 tratta i principali tipi di motori di raccomandazione e il loro funzionamento. Questi sistemi sono abili nel suggerire articoli o prodotti su misura, ma non sono esenti da problemi. Discutiamo sia i loro punti di forza sia i limiti che presentano. Infine, impariamo a utilizzare i motori di raccomandazione per risolvere un problema concreto. Il Capitolo 13 introduce gli algoritmi per i dati e i concetti di base della classificazione dei dati. Esaminiamo gli algoritmi di archiviazione e compressione utilizzati per gestire in modo efficiente i dati, per comprendere meglio i compromessi coinvolti nella progettazione e implementazione di algoritmi incentrati sui dati.

Il Capitolo 14 presenta gli algoritmi di crittografia. Iniziamo presentando le basi della crittografia prima di discutere gli algoritmi di crittografia simmetrica. Impariamo a conoscere gli algoritmi MD5 (*Message-Digest 5*) e SHA (*Secure Hash Algorithm*), presentando i limiti e i punti deboli di ciascuno di essi. Successivamente, discutiamo degli algoritmi di crittografia asimmetrica e di come vengono utilizzati per creare certificati digitali. Infine, presentiamo un esempio pratico che riassume tutte queste tecniche.

Il Capitolo 15 inizia introducendo gli algoritmi a larga scala e l'efficiente infrastruttura necessaria per supportarli. Esploriamo varie strategie per la gestione dell'elaborazione multi-risorsa. Esaminiamo i limiti dell'elaborazione parallela, dettati dalla legge di Amdahl, e investighiamo l'uso delle unità di elaborazione grafica, le GPU. Al termine di questo capitolo avrete acquisito solide basi sulle strategie fondamentali essenziali per la progettazione di algoritmi a larga scala.

Il Capitolo 16 presenta le questioni relative alla spiegabilità di un algoritmo, ovvero il grado in cui i meccanismi interni di un algoritmo possono essere spiegati in termini comprensibili. Successivamente, presentiamo gli aspetti etici dell'utilizzo di un algoritmo e la possibilità che essi possano sviluppare pregiudizi durante la loro implementazione. Successivamente discutiamo le tecniche per gestire problemi NP-difficili. Infine, esaminiamo i fattori che dovrebbero essere considerati prima di scegliere un algoritmo.

Codice di esempio

Il codice del libro è ospitato anche su GitHub all'indirizzo <https://github.com/cloudanum/50Algorithms>. Potete trovare lo stesso pacchetto di codici anche sul sito di Apogee all'indirizzo <https://bit.ly/apo-50al>.

Le immagini a colori

Le immagini a colori degli screenshot e dei grafici utilizzati in questo libro sono disponibili all'indirizzo: <https://bit.ly/apo-50a1>.

Convenzioni utilizzate

In questo libro vengono utilizzate numerose convenzioni testuali.

Codice nel testo: indica la presenza di parti di codice nel testo, nomi di tabelle di database, nomi di cartelle, nomi di file, estensioni di file, nomi di percorso, URL, input dell'utente. Ecco un esempio: "Proviamo a creare un semplice grafico utilizzando il pacchetto `networkx` in Python".

Corsivo: indica un nuovo termine, una parola importante o le parole visualizzate sullo schermo. Per esempio, i nuovi termini sono introdotti in questo modo: "Python è uno dei linguaggi che potete utilizzare in varie infrastrutture di cloud computing, come AWS (*Amazon Web Services*) e GCP (*Google Cloud Platform*)".

NOTA

Le note e gli avvertimenti hanno questo aspetto.

SUGGERIMENTO

I suggerimenti e i trucchi hanno questo aspetto.

L'autore

Imran Ahmad è data scientist dell'A2SC (*Advanced Analytics Solution Center*) del governo federale canadese, dove sfrutta algoritmi di machine learning per applicazioni critiche. Nella sua tesi di dottorato del 2010, ha introdotto un algoritmo basato sulla programmazione lineare su misura per l'assegnazione ottimale delle risorse nel panorama del cloud computing. Successivamente, nel 2017, ha aperto la strada allo sviluppo di un framework di analisi in tempo reale, *StreamSensing*. Questo strumento è diventato la pietra angolare di molti dei suoi articoli di ricerca ed è stato sfruttato per elaborare dati multimediali all'interno di vari paradigmi di machine learning.

Al di là del suo ruolo governativo, Ahmad è docente presso la Carleton University di Ottawa. Negli ultimi anni è diventato anche istruttore sia per Google Cloud sia per AWS.

I revisori

Aishwarya Srinivasan ha lavorato in precedenza come data scientist nel team dei servizi AI di Google Cloud, dove ha contribuito a creare soluzioni di machine learning per i casi d'uso dei clienti. Ha conseguito una laurea in Scienza dei dati presso la Columbia University e ha oltre 450.000 follower su LinkedIn. È stata segnalata come la voce mi-

gliore di LinkedIn fra gli influencer nel campo della scienza dei dati (2020) ed è stata riconosciuta *Women in AI Trailblazer*.

Tarek Ziadé è un programmatore che opera in Borgogna, Francia. Ha lavorato presso diverse importanti società di software, tra cui Mozilla ed Elastic, dove ha creato servizi web e strumenti per sviluppatori. Tarek ha fondato il gruppo francese di utenti Python, *Afpy*, e ha scritto diversi libri di successo su Python e sui servizi web.

Vorrei ringraziare la mia famiglia: Freya, Suki, Milo, Amina e Martine, che mi hanno sempre supportato.

Brian Spiering ha iniziato la sua carriera di programmatore nel laboratorio informatico della scuola elementare, hackerando il linguaggio BASIC per creare programmi che intrattenessero i suoi coetanei e facessero infuriare le figure di riferimento. Molto più tardi, Brian ha conseguito un dottorato di ricerca in psicologia cognitiva presso l'Università della California, a Santa Barbara. Brian attualmente insegna programmazione e intelligenza artificiale.

Ringraziamenti

Sono profondamente grato a mia moglie Naheed, a mio figlio Omar e a mia figlia Anum per il loro incrollabile sostegno. Un cenno speciale va ai miei genitori, in particolare a mio padre, Inayatullah, per il suo incessante incoraggiamento a continuare a imparare. Un ulteriore apprezzamento va a Karan Sonawane, Rianna Rodrigues e Denim di Packt, per il loro prezioso contributo.