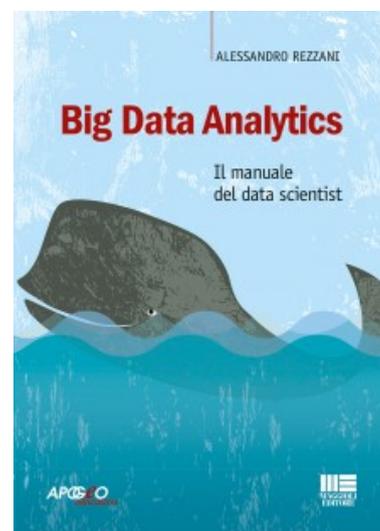


BIG DATA ANALYTICS Il manuale del data scientist

Prezzo: **Prezzo di listino**
34,00 € **Prezzo a te riservato**
32,30 €



Codice	9788891621856
Tipologia	Libri
Data pubblicazione	9 giu 2017
Reparto	LIBRI
Autore	Rezzani Alessandro
Editore	Maggioli

Descrizione

Nell'era dei big data e agli albori della data driven economy, emerge una figura professionale in grado di analizzare, gestire, elaborare e comunicare i dati. È il data scientist, lo "scienziato dei dati", che ha solide competenze in informatica, statistica, economia ed è in grado di far fronte alla sempre crescente complessità dei dati. Questo volume si propone come una guida completa sia per chi intenda intraprendere questa professione emergente, sia per chi, già esperto, desidera approfondire alcune tematiche. L'autore illustra le principali conoscenze relative alla gestione e all'analisi avanzata dei dati; descrive i big data e gli strumenti e le architetture che permettono di gestirli (Hadoop in particolare) e presenta i temi della data ingestion e dell'elaborazione con alcuni tool di analisi (Hive, Pig, Spark e R) le cui funzionalità sono illustrate anche tramite esempi commentati. Una parte è dedicata alla predictive analytics e mostra le tecniche per la creazione di modelli predittivi: dalla preparazione dei dati, alla scelta dell'algoritmo più adatto, alla valutazione delle performance. Il testo è un valido supporto per la comprensione dei concetti relativi all'analisi dei dati (big data o dati tradizionali), anche da parte del management aziendale che, dall'analisi avanzata, può trarre le informazioni utili ad assumere decisioni, a valutare rischi e disegnare strategie.

Parte I – Creare valore dai dati tradizionali e dai big data

Capitolo 1 Dati: un patrimonio importante

Introduzione

1.2 I dati aziendali

1.2.1 Le fonti

1.2.2 Tipi di supporto

1.2.3 I tipi di struttura

1.2.4 La provenienza

1.3 Attori aziendali e dati

1.3.1 I manager

1.3.2 Il personale esecutivo

1.3.3 I data scientist

Capitolo 2 Big data

2.1 Introduzione

2.2 Definizione di big data

2.3 Tipologie di big data

2.4 Tecnologie big data in breve

2.4.1 Acquisizione

2.4.2 Immagazzinamento e organizzazione

2.4.3 Trasformazione e analisi

2.4.4 Non solo Hadoop!

2.4.5 Servizi Cloud

2.5 Casi interessanti

2.5.1 Il settore banking

2.5.2 Industry 4.0

2.5.3 IOT - Internet of Things

2.5.4 Le smart city

2.6 L'architettura "data lake"

Capitolo 3 Le tecniche di analisi: come estrarre valore dai dati

3.1 L'analisi dei dati e la data monetization

3.2 Analisi descrittiva (o descriptive analytics)

3.3 Analisi predittiva (predictive analytics)

3.4 Analisi prescrittiva (o prescriptive analytics)

3.5 Applicazioni della predictive (e prescriptive) analytics

3.5.1 CRM

3.5.2 Ricerca di anomalie

3.5.3 Marketing

3.5.4 Altri utilizzi

Parte II – Hadoop

Capitolo 4 Hadoop: piattaforma e motori di calcolo

4.1 Introduzione

4.2 Hadoop

4.3 L'ecosistema Hadoop

4.4 HDFS

4.5 MapReduce

4.6 YARN

4.7 Altre componenti

4.8 Hadoop 3.x

Capitolo 5 Hadoop: basi di dati

5.1 Introduzione

5.2 Hive

5.3 HCatalog

5.4 Apache Accumulo

5.5 HBase

5.6 Apache Phoenix

5.7 Apache Kylin: OLAP su hadoop

Capitolo 6 Gli strumenti per la “data ingestion”

6.1 La data ingestion

6.2 Strumenti ETL tradizionali e funzionalità per i big data

6.3 Sqoop

6.4 Flume

6.5 Storm

6.6 Kafka

6.7 Oozie

6.8 Nifi

Parte III – Tecnologie per l’analisi dei dati

Capitolo 7 Elaborazione e analisi dei dati con Pig

7.1 Introduzione

7.2 Le caratteristiche principali di Pig

7.3 Introduzione al linguaggio Pig Latin

7.4 Funzioni di caricamento e salvataggio

7.5 Comandi e funzioni per l’elaborazione dei dati

7.6 User Defined Functions

7.7 Ottimizzazioni

7.8 Conclusioni

Capitolo 8 SQL per l’analisi dei dati strutturati

8.1 Introduzione

8.2 SQL per la preparazione e l’analisi dei dataset

8.3 Hive

8.4 IMPALA

8.5 U-SQL

8.6 Apache Drill

Capitolo 9 Spark

- 9.1 Introduzione a Spark
- 9.2 Spark Unified Stack: le componenti integrate di Spark
- 9.3 Architettura
- 9.4 Scrivere ed eseguire codice Spark
- 9.5 Breve introduzione al linguaggio Scala
- 9.6 RDD: Resilient Distributed Dataset
- 9.7 Operazioni comuni sugli RDD
- 9.8 Spark SQL, Dataframe e Dataset
- 9.9 Operazioni comuni su Dataset e DataFrame
- 9.10 Monitoring
- 9.11 Spark machine learning (MLib e ML)
- 9.12 Streaming
- 9.13 GraphX

Capitolo 10 R e l'analisi dei dati

- 10.1 Introduzione a R
- 10.2 Strumenti di sviluppo e help on-line
- 10.3 I package
- 10.4 Strutture dati di base
- 10.5 I dataframe
- 10.6 Funzioni built-in
- 10.7 Funzioni custom
- 10.8 Espressioni condizionali e cicli
- 10.9 Le funzioni apply
- 10.10 Grafici
- 10.11 Creazione di script
- 10.12 R e i grandi dataset
- 10.13 Conclusioni

Parte IV – La Predictive Analytics

Capitolo 11 Metodologia per la predictive e prescriptive analytics

- 11.1 Introduzione
- 11.2 CRISP-DM e una proposta metodologica
- 11.3 Le fasi della metodologia
- 11.4 Conclusioni

Capitolo 12 Preparazione dei dati

- 12.1 Introduzione
- 12.2 Le variabili
- 12.3 Esplorazione dei dati
- 12.4 Operazioni sulle variabili
- 12.5 Creazione di variabili
- 12.6 Trasformazioni dei valori
- 12.7 Codifica di variabili categoriche
- 12.8 Selezione delle variabili
- 12.9 Trattamento dei valori mancanti
- 12.10 Gli outliers
- 12.11 Classi sbilanciate
- 12.12 Campionamento
- 12.13 Errori comuni nella preparazione dei dati

Capitolo 13 Gli algoritmi

- 13.1 Introduzione
- 13.2 Problematiche comuni
- 13.3 Algoritmi di classificazione
- 13.4 Algoritmi di regressione
- 13.5 Algoritmi semi supervisionati
- 13.6 Algoritmi di associazione (cenni)
- 13.7 Algoritmi di clustering
- 13.8 Reti neurali
- 13.9 Deep learning
- 13.10 Model ensembles

Capitolo 14 La valutazione dei modelli

- 14.1 Il test e la valutazione dei modelli
- 14.2 Valutazione dei modelli di classificazione
- 14.3 Valutazione dei modelli di regressione
- 14.4 Valutazione dei modelli di clustering
- 14.5 Conclusioni

Bibliografia

Rimaniamo a disposizione per qualsiasi ulteriore chiarimento allo 0461.232337 o 0461.980546

oppure via mail a : servizioclienti@libriprofessionali.it

www.LibriProfessionali.it è un sito di Scala snc Via Solteri, 74 38121 Trento (Tn) P.Iva 01534230220

