

BIG DATA ANALYTICS

CON SPARK

DESCRIZIONE

Apache Spark sta assumendo una grande rilevanza nel mondo dei Big Data. Spark è un motore di calcolo distribuito, fault tolerant, altamente scalabile che ha la capacità di elaborare grandi moli di dati in modo efficiente e in tempi molto contenuti. Spark presenta alcune componenti fortemente integrate tra di loro e che aggiungono alla piattaforma alcune interessanti funzionalità: SparkSQL, Spark Streaming, Spark ML e MLib e GraphX.

Il corso fornisce una buona conoscenza della parte core di Spark e delle sue componenti e affronta i vari argomenti sia dal punto di vista teorico, sia attraverso esercizi pratici. Gli esempi sono realizzati in linguaggio Scala, Python o R (scelta da effettuare prima dell'erogazione del corso).

CONTENUTI: sezioni principali

- Introduzione a Spark
- L'architettura
- Lavorare con gli RDD (NB: solo Python o Scala)
- Spark SQL e i DataFrame
- Machine Learning con Spark ML e Mlib
- Spark Streaming (NB: solo Python o Scala)

PREREQUISITI

È consigliabile frequentare previamente il corso introduttivo: **“Big Data Analytics: tutto quello che c'è da sapere sui Big Data”**.

TAKEAWAYS

I partecipanti del corso apprenderanno le ultime novità per quanto riguarda le tecnologie Big Data nelle applicazioni di Business Intelligence. Il corso offre l'opportunità ai partecipanti di affrontare in prima persona dei casi reali, simili a quelli che potranno poi verificarsi in azienda.

BIG DATA ANALYTICS

CON SPARK



PREZZO

1100 euro a corso a persona*



GIORNI

2 giorni

*Sconti applicabili a gruppi di partecipanti
*I prezzi espressi non comprendono le spese di trasferta