

Big Data

Volume
Data Size

Velocity
Speed of
Change

Variety
different
Forms of
Data Sources

Veracity
Uncertainty
of Data

Value
business
value

BIG DATA Practical Analytics



مقدمه

۹ بایت) داده نگه داری نماید، بنابراین برای ذخیره سازی همه داده های دنیا به حدود ۲۰ میلیارد PC نیاز است. مبتنی بر کلیه اطلاعات فوق ضرورت توجه به مفهوم حجیم داده در روزگار اخیر بیش از هر زمان دیگری دارای اهمیت می باشد.

اهداف دوره

دوره ی آموزشی کاربردی BIG DATA یک راهنمای کامل است برای کسانی که می خواهند درک کنند که BIG DATA چیست و و کاربرد عینی آن در صنعت IT و دیگر صنایع چیست ؟

مخاطبان دوره

این دوره برای تمامی افرادی که نقشی در ارائه، کاربری و مدیریت محصولات و خدمات فناوری اطلاعات دارند، جذاب است. این افراد شامل موارد ذیل هستند:

- ارائه دهندگان خدمات داخلی و خارجی
- مشتریان
- کاربران
- مدیران

مدت زمان دوره

۴ روز (۳۲ ساعت)

اصطلاح Big Data برای مجموعه داده های حجیم که بزرگ ، متنوع ، با ساختار پیچیده و با دشواریهایی برای ذخیره سازی ، تحلیل و تصویرسازی (نمایش) ، مواجه می باشند مورد استفاده قرار می گیرد...
پروژه تحقیق بر روی داده های حجیم جهت آشکار سازی الگوهای مخفی و راز همبستگی ها و تجزیه و تحلیل آن ها Big Data نامیده میشود. این داده ها از تراکنشهای online, email, ها، ویدئوها، صوتها، کلیک کردن ها، log ها و ارسالها، درخواستهای جستجو، یادداشتها، تعاملات شبکه های اجتماعی، داده های علمی، سنسورها و تلفنهای همراه و برنامه های کاربردی آنها تولید میشوند. این اطلاعات بر روی پایگاه داده ها که به شکل حجیم رشد میکنند، ذخیره میشوند و ضبط، شکل دهی، ذخیره سازی، مدیریت و به اشتراک گذاری، تحلیل و نمایش آنها از طریق ابزارهای نرم افزاری پایگاه داده، دشوار می گردند. پنج اگزا بایت (۱۰ به توان ۱۸ بایت) دیتا تا سال ۲۰۰۳ به وسیله انسان به وجود آمده است. امروزه این مقدار اطلاعات در دو روز تولید میشود. در سال ۲۰۱۲ داده های دنیای دیجیتال به ۲۰۷۲ زتا بایت (۱۰ به توان ۲۱ بایت) توسعه پیدا کرد. پیش بینی میشود که این مقدار هر دو سال ، دو برابر شود و به حدود ۸ زتا بایت داده در سال ۲۰۱۵ میرسیم. مطابق تحقیقاتی که توسط IBM مورد انجام قرار گرفته است روزانه ۲.۵ اگزا بایت داده تولید میشود و همچنین ۹۰% از داده ها در ۲ سال اخیر تولید شده اند. اگر فرض نماییم یک کامپیوتر شخصی حدود ۵۰۰ گیگا بایت (۱۰ به توان

Map Reduce Programming – Java Programming

- Hands on “Word Count” in Map Reduce in standalone and Pseudo Distribution Mode
- Write some Map Reduce programs to solve some real world problems

Spark Basics

- What is Apache Spark?
- Spark Installation
- Spark Configuration
- Spark Context
- Using Spark Shell
- Resilient Distributed Datasets (RDDs) – Features, Partitions, Tuning Parallelism
- Functional Programming with Spark

Working with RDDs

- RDD Operations - Transformations and Actions
- Types of RDDs
- Key-Value Pair RDDs – Transformations and Actions
- MapReduce and Pair RDD Operations
- Serialization

Spark on a cluster

- Overview
- A Spark Standalone Cluster
- The Spark Standalone Web UI
- Executors & Cluster Manager
- Spark on YARN Framework

Writing Spark Applications

- Spark Applications vs. Spark Shell
- Creating the Spark Context
- Configuring Spark Properties
- Building and Running a Spark Application
- Logging
- Spark Job Anatomy

پیش نیازهای دوره

آشنایی مقدماتی با سیستم عامل لینوکس و یکی از زبان های برنامه نویسی الزامی میباشد

سرفصل های دوره

Introduction to Hadoop

- Hadoop Distributed File System
- Comparing Hadoop & SQL
- Industries using Hadoop
- Data Locality
- Hadoop Architecture
- Map Reduce & HDFS

Hadoop Distributed File System (HDFS)

- HDFS Design & Concepts
- Blocks, Name nodes and Data nodes
- HDFS High-Availability and HDFS Federation
- Hadoop DFS The Command-Line Interface
- Basic File System Operations
- Anatomy of File Read, File Write
- Block Placement Policy and Modes
- More detailed explanation about Configuration files
- Metadata, FS image, Edit log, Secondary Name Node and Safe Mode
- How to add New Data Node dynamically, decommission a Data Node dynamically (Without stopping cluster)
- FSCK Utility. (Block report)
- How to override default configuration at system level and Programming level

Map Reduce

- Map Reduce Functional Programming Basics
- Map and Reduce Basics
- How Map Reduce Works
- Anatomy of a Map Reduce Job Run
- Shuffling and Sorting
- Splits, Record reader, Partition, Types of partitions & Combiner
- Distributed Cache and Hadoop Streaming (Python, Ruby and R)
- YARN
- Sequential Files and Map Files
- Map side Join with distributed Cache

درباره ی NISICT-IT HOUSE

مؤسسه فناوری اطلاعات راهکار نوآوران فرتاک پس از ۷ سال همکاری موفق و مستمر شرکت NIS CERT کانادا با سازمان های ایرانی، در سال ۲۰۱۱ میلادی به عنوان یکی از واحدهای کسب و کار استراتژیک (SBU) این شرکت، با نام تجاری NISICT پا به عرصه ی ظهور گذاشت. این مجموعه در طی ۸ سال پس از همکاری موفق با بیش از ۲۰۰ سازمان برتر ایرانی، از ابتدای سال 2019 میلادی با توجه به راه اندازی مراکز تخصصی پیشرفته با همکاری شبکه ی گسترده ای از شرکای معتبر داخلی و بین المللی در طیف وسیعی از بهروش ها و استانداردهای بین المللی با هدف ارائه راهکار های جامع مدیریتی در حوزه ی فناوری اطلاعات، ماهیت تجاری خود (برند) را تغییر داده و از نشان تجاری IT HOUSE در حوزه ی محصولات و خدمات خود استفاده خواهد نمود. IT HOUSE با همکاری انحصاری مؤسساتی از جمله Quint (هلند)، NTT (ایرلند)، POWER ACT (مراکش) تحت اعتبار مراجع جهانی همچون انجمن جهانی tmforum آمریکا، PCI Council آمریکا، ISACA آمریکا، AXELOS انگلستان قادر به ارائه خدمات در تزار استاندارد های بین المللی به سازمان های ایرانی می باشد

تعداد شرکت کنندگان

۶ الی ۲۰ نفر

درباره ی مدرک

به شرکت کنندگان در این دوره ی آموزشی گواهی حضور از سوی شرکت NISICT-IT HOUSE و در صورت شرکت در آزمون بین المللی این دوره ی آموزشی، گواهی موفقیت از سوی مرجع اعتبار دهی EXIN هلند اعطا خواهد شد.

درباره ی آزمون

آزمون ندارد

زبان

در صورت استفاده از اساتید بین المللی، دوره به زبان انگلیسی ارائه می گردد که امکان استفاده از سیستم ترجمه همزمان در صورت درخواست متقاضیان مقدور می باشد. در دوره های آموزشی که توسط اساتید ایرانی این موسسه ارائه می گردد زبان مبنا پارسی می باشد.

محتوای آموزشی

محتوای مورد استفاده در این دوره ی آموزشی آخرین ویرایش ارائه شده توسط مرجع اعتبار دهی بین المللی EXIN هلند می باشد که دارای درجه کیفی Golden از این انجمن جهانی است.

تماس با NISICT-IT HOUSE

آدرس:

تهران، سهروردی شمالی، کوچه تهمتن، پلاک ۶ واحد ۱

تلفن: + ۹۸ (۰) ۲۱ ۸۸۷۳۱۴۶۶

فکس: + ۹۸ (۰) ۲۱ ۸۹۷۷۹۸۵۵

وبسایت: www.it-house.me

ایمیل: info@it-house.me

