



Uniempresarial

FUNDACIÓN UNIVERSITARIA EMPRESARIAL
DE LA CÁMARA DE COMERCIO DE BOGOTÁ
VIGILADA MINEDUCACIÓN / SNIES 2738

Una filial de



Cámara
de Comercio
de Bogotá



Diplomado en Big Data y Analítica

OBJETIVOS

- Reconocer los conceptos generales de los ecosistemas de datos distribuidos aplicados a la analítica de datos.
- Implementar técnicas de Map Reduce en el procesamiento de volúmenes de datos distribuidos en clusters mediante el uso de algoritmos paralelos.
- Identificar el framework con sistemas de archivos HDFS como entorno de programación para almacenar y procesar grandes cantidades de datos
- Reconocer y aplicar la arquitectura HDFS en el tratamiento de los datos, mediante el uso de conectores, procesos de seguridad y data analytics.
- Ejecutar análisis de datos con modelos básicos de estadísticas y minería de datos básica

METODOLOGIA

- Esta información se impartirá de manera presencial-virtual.
- Habrá componente teórico y práctico
- Se utilizarán ejemplos prácticos mediante Máquina Virtual instalada con todos los componentes del Eco-sistema Hadoop
- Debido al procesamiento de las herramientas tecnológicas se sugiere contar con computador personal con procesador Core i3 o superior y 8GB de ram como mínimo.

DURACIÓN

Un total de 80 horas

MODALIDAD

Mediado por la Plataforma Microsoft Teams

CONTENIDO

MÓDULO 1

Lenguajes Interpretados

Introducción a consultas SQL

- Consultas en 2 o más tablas
- Funciones básicas SQL
- Agrupamiento y having
- Consultas y subconsultas

Introducción a Python

- Variables y expresiones
- Condicionales y funciones
- Ciclos e iteración

Introducción a Python

- Strings
- Manejo de archivos

- Listas, diccionarios y tuplas

MÓDULO 2

Estadística Básica

Contexto General

- Media, promedio, mediana, moda
- Desviación estándar
- Líneas de regresión
- K-means

MÓDULO 3

Introducción a Big Data

Contexto General

- Introducción y conceptos de Big Data
- Las 7 V del Big data
- Diferencias entre Big data y datascience Fases de un proyecto de Big Data
- Eco sistema Hadoop
- Métodos de ciencias de datos
- Arquitecturas de referencia Cassandra, MongoDB, Business Intelligence Componentes de eco-sistema hadoop
- Arquitecturas de referencia Cassandra, MongoDB, Business Intelligence Componentes de eco-sistema hadoop
- Laboratorio No. 1

MÓDULO 4

Herramientas Para Big Data I – Adquisición y Almacenamiento de Datos

Eco sistema Hadoop

- Hadoop
- Map reduce
- HDFS
- YARN
- Tecnologías BBDD
- Flume
- Sqoop
- Kafka
- BBDD -Analítica, Business Intelligence, Hive, Spark
- Grand processing I y II

MÓDULO 5

Herramientas Para Big Data II – Procesamiento y Análisis

Ciclos Jupyter

- Introducción al Jupyter con Python
- Exploración estructura de datos
- Introducción a la Modelización (Arboles de regresión, arboles de clasificación, Bosques aleatorios)
- Modelización
- Calibración del modelo
- Resultados de Modelización
- Ciclos Jupyter
- Redes neuronales
- Clasificación automática
- Revisión de la ciencia de datos I y II
- introducción a patrones

MÓDULO 6

Herramientas Para Big Data II – Visualización de los Datos

Introducción a la visualización de datos:

- Ciclo de vida predictivo o Ciclo de vida iterativo o Ciclo de Vida incremental
- Ciclo de Vida ágil
- Implementación Ágil – Creación de un entorno ágil (2 horas)
- Mentalidad Ágil
- Liderazgo de un Líder de Servicio

[Más información](#)

educacioncontinua@uniempresarial.edu.co

Teléfono: 3808000 Ext.6605