

**Máster en Ingeniería de Sistemas de Toma de Decisiones + Titulación
universitaria**



ÍNDICE

1 | Somos Educa Business School

2 | Rankings

3 | Alianzas y acreditaciones

4 | By EDUCA EDTECH Group

5 | Metodología LXP

6 | Razones por las que elegir Educa Business School

7 | Programa Formativo

8 | Temario

9 | Contacto

SOMOS EDUCA BUSINESS SCHOOL

EDUCA Business School es una institución de formación online especializada en negocios. Como miembro de la Comisión Internacional de Educación a Distancia y con el prestigioso Certificado de Calidad AENOR (normativa ISO 9001) nuestra institución se distingue por su compromiso con la excelencia educativa.

Nuestra **oferta formativa**, además de **satisfacer las demandas del mercado laboral** actual, puede bonificarse como formación continua para el personal trabajador, así como ser homologados en Oposiciones dentro de la Administración Pública. Las titulaciones de EDUCA Business School se pueden certificar con la Apostilla de La Haya dotándolos de validez internacional en más de 160 países.

Más de

18

años de
experiencia

Más de

300k

estudiantes
formados

Hasta un

98%

tasa
empleabilidad

Hasta un

100%

de financiación

Hasta un

50%

de los estudiantes
repite

Hasta un

25%

de estudiantes
internacionales

RANKINGS DE EDUCA BUSINESS SCHOOL

Educa Business School se engloba en el conjunto de EDUCA EDTECH Group, que ha sido reconocido por su trabajo en el campo de la formación online.

Todas las entidades bajo el sello EDUCA EDTECH comparten la misión de democratizar el acceso a la educación y apuestan por la transferencia de conocimiento, por el desarrollo tecnológico y por la investigación. Gracias a ello ha conseguido el reconocimiento de diferentes rankings a nivel nacional e internacional.



ALIANZAS Y ACREDITACIONES



FONDO
SOCIAL
EUROPEO



BY EDUCA EDTECH

Educa Business School es una marca avalada por **EDUCA EDTECH Group**, que está compuesto por un conjunto de experimentadas y reconocidas instituciones educativas de formación online. Todas las entidades que lo forman comparten la misión de democratizar el acceso a la educación y apuestan por la transferencia de conocimiento, por el desarrollo tecnológico y por la investigación.



ONLINE EDUCATION



METODOLOGÍA LXP

La metodología **EDUCA LXP** permite una experiencia mejorada de aprendizaje integrando la AI en los procesos de e-learning, a través de modelos predictivos altamente personalizados, derivados del estudio de necesidades detectadas en la interacción del alumnado con sus entornos virtuales.

EDUCA LXP es fruto de la **Transferencia de Resultados de Investigación** de varios proyectos multidisciplinares de I+D+i, con participación de distintas Universidades Internacionales que apuestan por la transferencia de conocimientos, desarrollo tecnológico e investigación.



1. Flexibilidad

Aprendizaje 100% online y flexible, que permite al alumnado estudiar donde, cuando y como quiera.



2. Accesibilidad

Cercanía y comprensión. Democratizando el acceso a la educación trabajando para que todas las personas tengan la oportunidad de seguir formándose.



3. Personalización

Itinerarios formativos individualizados y adaptados a las necesidades de cada estudiante.



4. Acompañamiento / Seguimiento docente

Orientación académica por parte de un equipo docente especialista en su área de conocimiento, que aboga por la calidad educativa adaptando los procesos a las necesidades del mercado laboral.



5. Innovación

Desarrollos tecnológicos en permanente evolución impulsados por la AI mediante Learning Experience Platform.



6. Excelencia educativa

Enfoque didáctico orientado al trabajo por competencias, que favorece un aprendizaje práctico y significativo, garantizando el desarrollo profesional.

RAZONES POR LAS QUE ELEGIR EDUCA BUSINESS SCHOOL

1. FORMACIÓN ONLINE ESPECIALIZADA

Nuestros alumnos acceden a un modelo pedagógico innovador **de más de 20 años de experiencia educativa con Calidad Europea.**



2. METODOLOGÍA DE EDUCACIÓN FLEXIBLE

Con nuestra metodología estudiarán **100% online** y nuestros alumnos/as tendrán acceso los 365 días del año a la plataforma educativa.



3. CAMPUS VIRTUAL DE ÚLTIMA TECNOLOGÍA



Contamos con una **plataforma avanzada** con material adaptado a la realidad empresarial, que fomenta la participación, interacción y comunicación con alumnos de distintos países.

4. DOCENTES DE PRIMER NIVEL

Nuestros docentes están acreditados y formados en **Universidades de alto prestigio en Europa**, todos en activo y con una amplia experiencia profesional.



5. TUTORÍA PERMANENTE



Contamos con un **Centro de Atención al Estudiante CAE**, que brinda atención personalizada y acompañamiento durante todo el proceso formativo.

6. DOBLE MATRICULACIÓN

Algunas de nuestras acciones formativas cuentan con la llamada **Doble matriculación**, que te permite obtener dos formaciones, ya sean de masters o curso, al precio de una.



Máster en Ingeniería de Sistemas de Toma de Decisiones + Titulación universitaria



DURACIÓN
1500 horas



**MODALIDAD
ONLINE**



**ACOMPANIAMIENTO
PERSONALIZADO**



CREDITOS
8 ECTS

Titulación

Doble Titulación: - Titulación de Máster en Ingeniería de Sistemas de Toma de Decisiones con 1500 horas expedida por EDUCA BUSINESS SCHOOL como Escuela de Negocios Acreditada para la Impartición de Formación Superior de Postgrado, con Validez Profesional a Nivel Internacional - Titulación Universitaria de Curso Universitario en Creación y Gestión de Base de Datos SQL con 200 horas y 8 ECTS expedida por UTAMED - Universidad Tecnológica Atlántico Mediterráneo.



EDUCA BUSINESS SCHOOL

como centro acreditado para la impartición de acciones formativas
expide el presente título propio

NOMBRE DEL ALUMNO/A

con número de documento XXXXXXXXX ha superado los estudios correspondientes de

Nombre del curso

con una duración de XXX horas, perteneciente al Plan de Formación de Educa Business School.

Y para que surta los efectos pertinentes queda registrado con número de expediente XXXX/XXXX/XXXX/XXXX/XXXX.

Con una calificación XXXXXXXXXXXXXXX.

Y para que conste expido la presente titulación en Granada, a (día) de (mes) del (año).

Firma del Alumno/a
NOMBRE ALUMNO/A

La Dirección Académica
NOMBRE DE AREA MANAGER



Con el aval de la Comisión, Categoría Especial del Consejo Económico y Social de la URBEOCI (Plan Propiedad 2016)

Descripción

El Máster en Ingeniería de Sistemas de Toma de Decisiones es tu pasaporte hacia el futuro de la gestión inteligente. En un mundo donde la información es poder, saber cómo analizar y utilizar datos para tomar decisiones estratégicas es fundamental. Este máster te prepara para enfrentar desafíos complejos mediante el aprendizaje de técnicas avanzadas de decisión y modelización de incertidumbre. Con un enfoque en herramientas como el análisis de datos, simulación y metaheurísticas, te convertirás en un experto en transformar datos en decisiones sólidas. La creciente demanda de profesionales capacitados en Big Data y minería de datos hace que esta formación sea especialmente valiosa. Optar por este máster significa estar a la vanguardia de la innovación, mejorando tus habilidades en áreas clave como la optimización y la gestión de bases de datos. A través de un formato online flexible, tendrás la oportunidad de aprender desde cualquier lugar, adaptando tus estudios a tus necesidades personales y profesionales.

Objetivos

- Aplicar modelos y técnicas avanzadas de decisión para optimizar procesos complejos. - Desarrollar sistemas de apoyo a la decisión utilizando algoritmos heurísticos. - Implementar métodos bayesianos en situaciones de incertidumbre para mejorar la precisión. - Utilizar simulación de Monte Carlo para evaluar riesgos en la toma de decisiones. - Emplear técnicas de minería de datos para transformar datos en decisiones estratégicas. - Diseñar soluciones de Big Data para el procesamiento eficiente de grandes volúmenes de datos. - Integrar el método Six Sigma para mejorar la calidad y eficiencia en proyectos de ingeniería.

Para qué te prepara

El Máster en Ingeniería de Sistemas de Toma de Decisiones está dirigido a profesionales y titulados en ingeniería, matemáticas o ciencias afines interesados en profundizar en modelos de decisión, análisis de datos y técnicas de optimización. Ideal para quienes buscan actualizarse en big data, minería de datos y técnicas avanzadas como la simulación de Monte Carlo o Six Sigma.

A quién va dirigido

El Máster en Ingeniería de Sistemas de Toma de Decisiones te prepara para enfrentar desafíos complejos en la toma de decisiones aplicando técnicas avanzadas como la optimización y los algoritmos heurísticos. Aprenderás a utilizar sistemas de apoyo a la decisión y modelar la incertidumbre mediante métodos bayesianos. Además, dominarás la gestión y minería de datos para extraer información valiosa, y aplicarás herramientas de Big Data y Seis Sigma para mejorar procesos y proyectos, optimizando resultados en entornos dinámicos.

Salidas laborales

'- Analista de datos en sectores avanzados - Consultor en sistemas de decisión - Especialista en Big Data para ingeniería - Experto en optimización y algoritmos - Ingeniero de sistemas de soporte a la decisión - Desarrollador de soluciones de minería de datos - Gestor de proyectos Six Sigma - Consultor en seguridad y gestión de bases de datos - Responsable de innovación tecnológica en decisiones empresariales

TEMARIO

PARTE 1. INGENIERÍA DE LA DECISIÓN

UNIDAD DIDÁCTICA 1. INTRODUCCIÓN A LA INGENIERÍA DE LA DECISIÓN

UNIDAD DIDÁCTICA 2. MODELOS Y TÉCNICAS DE DECISIÓN

UNIDAD DIDÁCTICA 3. ANÁLISIS DE DATOS Y TOMA DE DECISIONES

UNIDAD DIDÁCTICA 4. TEORÍA DE LA DECISIÓN Y RIESGO

UNIDAD DIDÁCTICA 5. SISTEMAS DE APOYO A LA DECISIÓN (DSS)

UNIDAD DIDÁCTICA 6. OPTIMIZACIÓN Y ALGORITMOS HEURÍSTICOS

UNIDAD DIDÁCTICA 7. ÉTICA Y RESPONSABILIDAD EN LA TOMA DE DECISIONES

UNIDAD DIDÁCTICA 8. DECISIONES EN ENTORNOS DINÁMICOS Y COMPLEJOS

UNIDAD DIDÁCTICA 9. COMUNICACIÓN Y GESTIÓN DE DECISIONES

UNIDAD DIDÁCTICA 10. TENDENCIAS FUTURAS Y NUEVAS TECNOLOGÍAS EN LA TOMA DE DECISIONES

PARTE 2. MODELIZACIÓN Y TRATAMIENTO DE LA INCERTIDUMBRE

UNIDAD DIDÁCTICA 1. INTRODUCCIÓN A LA INCERTIDUMBRE EN LA TOMA DE DECISIONES

UNIDAD DIDÁCTICA 2. PROBABILIDAD Y ESTADÍSTICA BÁSICA

UNIDAD DIDÁCTICA 3. MODELOS ESTOCÁSTICOS

UNIDAD DIDÁCTICA 4. MÉTODOS BAYESIANOS

UNIDAD DIDÁCTICA 5. TÉCNICAS DE OPTIMIZACIÓN EN CONDICIONES DE INCERTIDUMBRE

UNIDAD DIDÁCTICA 6. ANÁLISIS DE SENSIBILIDAD Y SIMULACIÓN

UNIDAD DIDÁCTICA 7. TEORÍA DE LA DECISIÓN

UNIDAD DIDÁCTICA 8. HERRAMIENTAS Y SOFTWARE PARA MODELAR Y TRATAR LA INCERTIDUMBRE

PARTE 3. SIMULACIÓN Y METAHEURÍSTICAS

UNIDAD DIDÁCTICA 1. INTRODUCCIÓN A LA SIMULACIÓN Y METAHEURÍSTICAS

UNIDAD DIDÁCTICA 2. CONCEPTOS BÁSICOS Y TÉCNICAS DE SIMULACIÓN

UNIDAD DIDÁCTICA 3. SIMULACIÓN DE MONTE CARLO

UNIDAD DIDÁCTICA 4. METAHEURÍSTICAS BASADAS EN BÚSQUEDA LOCAL

UNIDAD DIDÁCTICA 5. ALGORITMOS GENÉTICOS Y EVOLUTIVOS

UNIDAD DIDÁCTICA 6. OPTIMIZACIÓN BASADA EN ENJAMBRE Y COLONIAS

UNIDAD DIDÁCTICA 7. METAHEURÍSTICAS BASADAS EN NATURALEZA Y BIOLOGÍA

UNIDAD DIDÁCTICA 8. ANÁLISIS Y VISUALIZACIÓN DE RESULTADOS

PARTE 4. GESTIÓN DE BASES DE DATOS

UNIDAD DIDÁCTICA 1. INTRODUCCIÓN A LAS BASES DE DATOS

1. Ventajas e inconvenientes de las bases de datos
2. Conceptos generales
3. El modelo entidad-relación
4. El modelo entidad-relación extendido
5. Restricciones de integridad

UNIDAD DIDÁCTICA 2. EL MODELO DE BASES DE DATOS RELACIONAL

1. Estructura del modelo relacional
2. Claves en el modelo relacional
3. Restricciones de integridad
4. Teoría de la normalización
5. Diseño de una base de datos relacional
6. Tipos de lenguajes relacionales

UNIDAD DIDÁCTICA 3. LENGUAJE DE CONSULTA SQL

1. Características de SQL
2. Sistemas de Gestión de Bases de Datos con soporte SQL
3. Sintaxis en SQL
4. Especificación de restricciones de integridad

UNIDAD DIDÁCTICA 4. MYSQL COMO SISTEMA GESTOR DE BASES DE DATOS RELACIONALES

1. Características de MySQL
2. Tipos de datos
3. Sintaxis SQL para MySQL

UNIDAD DIDÁCTICA 5. SALVAGUARDA Y RECUPERACIÓN DE DATOS

1. Posibles fallos en una base de datos
2. Elementos de recuperación
3. Tipos de soporte
4. RAID

5. Servidores remotos de salvaguarda de datos
6. Diseño de un plan de salvaguarda y protocolo de recuperación de datos
7. Tipos de salvaguardas de datos
8. RTO (Recovery Time Objective) y RPO (Recovery Point Objective)
9. Mecanismos de verificación de la integridad de las copias de seguridad

UNIDAD DIDÁCTICA 6. BASES DE DATOS DISTRIBUIDAS

1. Definición de SGBD distribuido. Principales ventajas y desventajas
2. Características esperadas en un SGBD distribuido
3. Clasificación de los SGBD distribuidos
4. Enumeración y explicación de las reglas de DATE para SGBD distribuidos
5. Replicación de la información en bases de datos distribuidas
6. Procesamiento de consultas
7. Descomposición de consultas y localización de datos

UNIDAD DIDÁCTICA 7. SEGURIDAD DE LOS DATOS

1. Conceptos de seguridad de los datos: confidencialidad, integridad y disponibilidad
2. Normativa legal vigente sobre datos
3. Supuestos prácticos

UNIDAD DIDÁCTICA 8. TRANSFERENCIA DE DATOS

1. Herramientas para importar y exportar datos
2. Clasificación de las herramientas
3. Ejemplo de ejecución de una exportación e importación de datos
4. Migración de datos entre diferentes SGBD
5. Inconvenientes al traspasar datos entre distintos SGBD

PARTE 5. MINERÍA DE DATOS

UNIDAD DIDÁCTICA 1. MINERÍA DE DATOS O DATA MINING Y EL APRENDIZAJE AUTOMÁTICO

1. Introducción a la minería de datos y el aprendizaje automático
2. Proceso KDD
3. Modelos y Técnicas de Data Mining
4. Áreas de aplicación
5. Minería de textos y Web Mining
6. Data mining y marketing

UNIDAD DIDÁCTICA 2. R COMO HERRAMIENTA PARA BIG DATA

1. Introducción a R
2. ¿Qué necesitas?
3. Tipos de datos
4. Estadística Descriptiva y Predictiva con R
5. Integración de R en Hadoop

UNIDAD DIDÁCTICA 3. PRE-PROCESAMIENTO & PROCESAMIENTO DE DATOS

1. Obtención y limpieza de los datos (ETL)
2. Inferencia estadística
3. Modelos de regresión
4. Pruebas de hipótesis

UNIDAD DIDÁCTICA 4. ANÁLISIS DE LOS DATOS

1. Inteligencia Analítica de negocios
2. La teoría de grafos y el análisis de redes sociales
3. Presentación de resultados

UNIDAD DIDÁCTICA 5. PROCESAMIENTO DISTRIBUIDO DE DATOS CON HADOOP

1. ¿Qué es Hadoop?
2. El sistema de archivos HDFS
3. Algunos comandos de referencia
4. Procesamiento MapReduce con Hadoop
5. El concepto de los clusters en Hadoop

UNIDAD DIDÁCTICA 6. WEKA Y DATA MINING

1. ¿Qué es Weka?
2. Técnicas de Data Mining en Weka
3. Interfaces de Weka
4. Selección de atributos

PARTE 6. BIG DATA PARA INGENIERÍAS

UNIDAD DIDÁCTICA 1. INTRODUCCIÓN

1. ¿Qué es Big Data?
2. Paradigmas de procesamiento en Big Data
3. Las 8 V de Big Data (Volumen, Volatilidad, Variedad, Valor, Velocidad, Variabilidad, Veracidad, Validez).

UNIDAD DIDÁCTICA 2. BATCH PROCESSING

1. MapReduce
 1. - Entorno MapReduce
 2. - Función Map y función Reduce
 3. - Flujo de datos
 4. - Características de MapReduce
 5. - Uso de MapReduce
 6. - Ventajas e inconvenientes de Map Reduce
 7. - Ejercicios y ejemplos con MapReduce
2. Hadoop
 1. - Entorno Hadoop
 2. - Almacenamiento: HDFS
 3. - Características de HDFS
3. Apache Hadoop YARN

1. - Funciones de Framework computacionales
2. - YARN: El gestor de recursos del cluster
3. - Conceptos de Apache Spark
4. - Ejecución de Computational Frameworks en YARN
5. - Exploración de las aplicaciones de YARN Applications a través de la Web UIs y de Shell
4. Agregación de los logs de YARN
 1. - Configuración de Hadoop y registros de Daemon
 2. - Localizar configuraciones y aplicar cambios de configuración
 3. - Gestión de instancias de Role y añadir servicios
 4. - Configuración del servicio HDFS
 5. - Configuración de los logs de Hadoop Daemon
 6. - Configuración del servicio YARN
5. Obtención de datos en HDFS
 1. - Ingestión de datos desde fuentes de recursos externos con Flume
 2. - Ingestión de datos desde bases de datos relacionales con Sqoop
 3. - REST Interfaces
 4. - Buenas prácticas para la importación de datos
6. Planificación de un cluster Hadoop
 1. - Consideraciones generales de planificación
 2. - Elección correcta de Hardware
 3. - Opciones de Virtualización
 4. - Consideraciones de red
 5. - Configuración de nodos
7. Instalación y configuración de Hive, Pig e Impala
8. Clientes Hadoop incluidos en Hue
 1. - ¿Qué es un cliente de Hadoop?
 2. - Instalación y configuración de clientes Hadoop
 3. - Instalación y configuración de Hue
 4. - Autorizaciones y autenticación Hue
9. Configuración avanzada de un cluster
 1. - Parámetros avanzados de configuración
 2. - Configuración de puertos Hadoop
 3. - Configuración de HDFS para la organización en rack
 4. - Configuración de HDFS para obtención de alta disponibilidad
10. Seguridad Hadoop
 1. - ¿Por qué es importante la seguridad en Hadoop?
 2. - Conceptos del sistema de seguridad de Hadoop
 3. - Qué es Kerberos y cómo funciona
 4. - Securitización de un clúster Hadoop Cluster con Kerberos
 5. - Otros conceptos de seguridad
11. Gestión de recursos
 1. - Configuración de cgroups con Static Service Pools
 2. - El Fair Scheduler
 3. - Configuración de Dynamic Resource Pools
 4. - Configuraciones de CPU y memoria YARN
 5. - Impala Query Scheduling
12. Mantenimiento de un cluster
 1. - Chequeo del estado de HDFS

2. - Copia de datos entre clústers
 3. - Añadir y eliminar de nodos en el clúster
 4. - Rebalanceo del Cluster
 5. - Directorio de Snapshots
 6. - Actualización del clúster
13. Solución de problemas y monitorización de un cluster
1. - Sistema general de monitorización
 2. - Monitorización de clústers Hadoop
 3. - Solución de problemas habituales en el clúster de Hadoop
 4. - Errores habituales en la configuración

UNIDAD DIDÁCTICA 3. CIENCIA DE DATOS

1. Data Science
 1. - Que hacen los data scientists, herramientas y procesos que utilizan
 2. - Aplicación de lo aprendido en módulo 2: Uso de Hue
2. Apache Spark
 1. - Cómo trabaja Apache Spark y que capacidades nos ofrece
 2. - Que formatos de ficheros populares puede usar Spark para almacenar datos
 3. - Que lenguajes de programación puedes utilizar para trabajar con Spark
 4. - Cómo empezar a utilizar PySpark y Sparklyr
 5. - Cómo comparar PySpark y Sparklyr
3. Machine Learning
 1. - ¿Qué es machine learning?
 2. - Algunos conceptos y términos importantes
 3. - Diferentes tipos de algoritmos
 4. - Librerías que se utilizan
4. Apache Spark MLlib
 1. - Que capacidades de machine learning nos proporciona MLlib
 2. - Cómo crear, validar y utilizar modelos de machine learning con MLlib
 3. - Ejecución de trabajos Apache Spark
 4. - Cómo un trabajo de Spark se compone de una secuencia de transformaciones seguida de una acción
 5. - Cómo Spark utiliza la ejecución lenta
 6. - Cómo Spark divide los datos entre las particiones
 7. - Cómo ejecuta Spark operaciones limitadas y grandes
 8. - Cómo Spark ejecuta un trabajo en tareas y fases

UNIDAD DIDÁCTICA 4. DESARROLLO PARA SPARK Y HADOOP

1. Datasets y Dataframes
2. Operaciones en Dataframe
3. Trabajar con Dataframes y Schemas
4. Crear Dataframes a partir de Data Sources
5. Guardar DataFrames en Data Sources
6. DataFrame Schemas
7. Rapidez y lentitud de ejecución
8. Análisis de datos con consultas de DataFrame
 1. - Consultar DataFrames con el empleo de expresiones de columna

- 2. - Agrupación y agregación de consultas
- 3. - Unión de DataFrames
- 9. RDD
 - 1. - Introducción RDD
 - 2. - RDD Data Sources
 - 3. - Creando y guardando RDDs
 - 4. - Operaciones con RDDs
- 10. Transformación de datos con RDDs
 - 1. - Escritura y paso de funciones de transformación
 - 2. - Ejecuciones de transformación
 - 3. - Conversión entre RDDs y DataFrames
- 11. Agregación de datos con Pair RDDs
 - 1. - Key-Value Pair RDDs
 - 2. - Map-Reduce
 - 3. - Otras operaciones Pair RDD
- 12. Consulta y vistas de tablas con Spark SQL
 - 1. - Datasets y DataFrames
 - 2. - Creación de Datasets
 - 3. - Ejecución y guardado de Datasets
 - 4. - Operaciones de Dataset
- 13. Creación, configuración y ejecución de aplicaciones Spark
 - 1. - Creación de una aplicación Spark
 - 2. - Compilar y ejecutar la aplicación
 - 3. - Application Deployment Mode
 - 4. - La interfaz Spark Application Web UI
 - 5. - Configuración de las propiedades de la aplicación
- 14. Procesamiento distribuido
 - 1. - Apache Spark en un Clúster
 - 2. - Particiones RDD
 - 3. - Ejemplo: Particionamiento en consultas
 - 4. - Etapas y Tareas
 - 5. - Planificación de tareas de ejecución
- 15. Persistencia de datos distribuidos
 - 1. - Persistencia en Datasets y DataFrames
 - 2. - Persistencia en niveles de almacenamiento
 - 3. - Visualización de RDDs persistentes
- 16. Patrones comunes al procesar datos con Spark
 - 1. - Casos comunes de uso de Spark
 - 2. - Algoritmos de iteración en Apache Spark
 - 3. - Machine Learning
- 17. Spark Streaming: Introducción a DStreams
 - 1. - Vista general de Spark Streaming
 - 2. - DStreams
 - 3. - Desarrollo de aplicaciones en Streaming
- 18. Spark Streaming: procesamiento de múltiples lotes
 - 1. - Operaciones Multi-Batch
 - 2. - Time Slicing
 - 3. - Operaciones de estado

4. - Operaciones Sliding Window
5. - Vista previa: Streaming estructurado
19. Apache Spark Streaming: Data Sources
 1. - Vista general de Streaming Data Source
 2. - Apache Flume y Apache Kafka Data Sources
 3. - Ejemplo: uso de un Kafka Direct Data Source

UNIDAD DIDÁCTICA 5. ANÁLISIS DE DATOS

1. Introducción a Pig
 1. - ¿Qué es Pig?
 2. - Características de Pig
 3. - Casos de empleo de Pig
 4. - Interacción con Pig
2. Análisis de datos básico con Pig
 1. - Sintaxis Pig Latin
 2. - Carga de datos
 3. - Tipos simples de datos
 4. - Definición de campos
 5. - Datos de salida
 6. - Vistas y esquemas
 7. - Filtrado y ordenación de datos
 8. - Funciones habituales
3. Procesado de datos complejos con Pig
 1. - Formatos de almacenamiento
 2. - Tipos de datos complejos y anidados
 3. - Agrupaciones
 4. - Funciones predefinidas para datos complejos
 5. - Iteración de datos agrupados
4. Operaciones con multiconjuntos de datos con Pig
 1. - Técnicas para combinar conjuntos de datos
 2. - Unión de conjuntos de datos con Pig
 3. - Conjunto de operaciones
 4. - División de conjuntos de datos
5. Troubleshooting y optimización de Pig
 1. - Troubleshooting en Pig
 2. - Inicio de sesión
 3. - Empleo de UI web Hadoop
 4. - Muestreo de datos y depuración
 5. - Visión general del rendimiento
 6. - Comprensión del plan de ejecución
 7. - Consejos para mejorar el rendimiento de Jobs en Pig
6. Introducción a Hive e Impala
 1. - ¿Qué es Hive?
 2. - ¿Qué es Impala?
 3. - ¿Por qué utilizar Hive e Impala?
 4. - Schema y almacenamiento de datos
 5. - Comparación entre Hive y bases de datos tradicionales
 6. - Casos de uso

7. Consultas con Hive e Impala
 1. - Tablas y bases de datos
 2. - Sintaxis básica en consultas Hive e Impala
 3. - Tipos de datos
 4. - Empleo de Hue para ejecutar consultas
 5. - Empleo de Beeline (la Shell de Hive)
 6. - Empleo de la Shell de Impala
8. Administración de datos
 1. - Almacenamiento de datos
 2. - Creación de bases de datos y tablas
 3. - Carga de datos
 4. - Alteración de bases de datos y tablas
 5. - Simplificación de consultas con vistas
 6. - Almacenamiento de resultados de consultas
9. Almacenamiento y datos de rendimiento
 1. - Partición de tablas
 2. - Carga de datos en tablas particionadas
 3. - Cuándo utilizar el particionamiento
 4. - Elección de formato de almacenamiento
 5. - Gestión de metadatos
 6. - Control de acceso a datos
10. Análisis de datos relacional con Hive e Impala
 1. - Unión de conjuntos de datos
 2. - Funciones predefinidas habituales
 3. - Agregaciones y Windowing
11. Datos complejos con Hive e Impala
 1. - Datos complejos con Hive
 2. - Datos complejos con Impala
12. Análisis de texto con Hive e Impala
 1. - Empleo de expresiones regulares
 2. - Procesamiento de texto con SerDes en Hive
 3. - Análisis de los sentimientos y N•Grams
13. Optimización Hive
 1. - Rendimiento de las consultas
 2. - Bucketing
 3. - Indexación de datos
 4. - Hive en Spark
14. Optimización de Impala
 1. - Ejecución de consultas
 2. - Mejorar el rendimiento de Impala
15. Extendiendo Hive e Impala
 1. - Customizar SerDes y formatos de fichero en Hive
 2. - Transformación de datos con Scripts personalizados en Hive
 3. - Funciones definidas por el usuario
 4. - Consultas parametrizadas
 5. - Comparación entre MapReduce, Pig, Hive, Impala, y bases de datos relacionales. ¿Cuál elegir?

PARTE 7. ALGORITMOS DE BIGDATA PARA INGENIERÍAS

UNIDAD DIDÁCTICA 1. MINERÍA DE DATOS O DATA MINING Y EL APRENDIZAJE AUTOMÁTICO

1. Introducción a la minería de datos y el aprendizaje automático
 1. - Proceso KDD
 2. - Modelos y Técnicas de Data Mining
 3. - Áreas de aplicación
 4. - Minería de textos y Web Mining
 5. - Data mining y marketing

UNIDAD DIDÁCTICA 2. R COMO HERRAMIENTA PARA BIG DATA

1. Instalación de R y RStudio
 1. - Introducción al lenguaje
 2. - Historia e Introducción a R
 3. - Operaciones Básicas y Números
 4. - Atributos, Entrada y Coerción
 5. - Matrices
 6. - Precedencia Operaciones Vectoriales
 7. - Manejo de fechas y tiempo
 8. - Listas, Factores, Valores Faltantes y Dataframes
 9. - Subconjuntos de Datos
 10. - Leer y Escribir Datos
2. Uso del lenguaje
 1. - Estructuras de Control
 2. - Funciones
 3. - Reglas de Alcance
3. Sistema de gráficos
 1. - Funciones *apply: apply
 2. - Funciones *apply: lapply / sapply
 3. - Funciones *apply: mapply / rep
 4. - Graficación con el Sistema de Base de Gráficos
 5. - Algunas Funciones Gráficas de Alto Nivel
 6. - Parámetros en el Sistema de Gráficos
 7. - Colores en el Sistema de Gráficos
 8. - Graficación con Notación Matemática
 9. - Graficación con texto y notación matemática
 10. - Creación de Gráficas en 3D
4. Expresiones regulares. Gráficas con ggplot2 y Simulación
 1. - Expresiones Regulares
 2. - Paquete de gráficos ggplot2
 3. - Simulación
5. R en el mundo real
 1. - Estadística Descriptiva y Predictiva con R
 2. - Integración de R en Hadoop

UNIDAD DIDÁCTICA 3. PRE-PROCESAMIENTO &, PROCESAMIENTO DE DATOS

1. Obtención y limpieza de los datos (ETL)
2. Inferencia estadística
3. Pruebas de hipótesis
4. Modelos de regresión
5. Árboles de Decisión
6. Algoritmos de Clasificación / Regresión (J48/C5.0, M5P)
7. Normalización, Tipos de distancia, Correlación
8. Machine Learning
9. Comparar Artículos (k-NN)
10. Modelo predictivo de profit (k-NN, M5P...)
11. Modelo predictivo de clasificación (J48, k-NN)

UNIDAD DIDÁCTICA 4. ANÁLISIS DE LOS DATOS

1. Inteligencia Analítica de negocios
2. La teoría de grafos y el análisis de redes sociales
3. Presentación de resultados

UNIDAD DIDÁCTICA 5. PROCESAMIENTO DISTRIBUIDO DE DATOS CON HADOOP

1. ¿Qué es Hadoop?
2. El sistema de archivos HDFS
3. Algunos comandos de referencia
4. Procesamiento MapReduce con Hadoop
5. El concepto de los clusters en Hadoop

UNIDAD DIDÁCTICA 6. WEKA Y DATA MINING

1. ¿Qué es Weka?
2. Técnicas de Data Mining en Weka
3. Interfaces de Weka
4. Selección de atributos

PARTE 8. CALIDAD: SEIS SIGMA

UNIDAD DIDÁCTICA 1. INTRODUCCIÓN AL SIX SIGMA

1. ¿Qué es el Six Sigma?
2. Historia y Aplicación del Six Sigma
3. Otros Métodos de Mejora de los Procesos de Calidad
4. Conceptos de Lean
5. Conceptos Básicos de Six Sigma
6. Definición de los Problemas

UNIDAD DIDÁCTICA 2. PROYECTOS Y PROCESOS

1. ¿Qué es un proceso?
2. La Gestión de la Calidad
3. Seleccionar los Proyectos Adecuados
4. Principios de Gestión Básica del Equipo Six Sigma

5. Introducción a los métodos DMAIC y DMADV

UNIDAD DIDÁCTICA 3. EL MÉTODO DMAIC

1. Definir
2. Medir
3. Analizar
4. Mejorar
5. Controlar

UNIDAD DIDÁCTICA 4. MÉTODOS DE ESTADÍSTICA BÁSICA APLICADA AL SIX SIGMA

1. Análisis Gráfico
2. Distribución Normal de la Probabilidad
3. Correlación y Regresión

UNIDAD DIDÁCTICA 5. MÉTODOS DE ESTADÍSTICA AVANZADA APLICADOS AL SIX SIGMA

1. Distribución No-Normal de la Probabilidad
2. Evaluación de la Hipótesis
3. El Tamaño de la Muestra
4. Gráficos de Control Avanzados
5. Estadística en Aplicaciones de Negocios a Través del Six Sigma

UNIDAD DIDÁCTICA 6. CONTROL AVANZADO

1. Introducción a Minitab
2. Gráficos y Herramientas de Calidad de Minitab
3. El Menú Estadísticas en Minitab

UNIDAD DIDÁCTICA 7. EXPERIMENTOS

1. Análisis de Varianza (ANOVA)
2. Diseño de Experimentos
3. Interacciones, Factores Multinivel y Creación de Experimentos

UNIDAD DIDÁCTICA 8. MINITAB

1. Tormenta de Ideas y otras Herramientas de Mejora de Procesos
2. Mapas de Procesos
3. Monitoreo de la Cadena de Valor

