



EUROINNOVA
INTERNATIONAL ONLINE EDUCATION



FORMACIÓN ONLINE

Titulación certificada por EUROINNOVA INTERNATIONAL ONLINE EDUCATION

Curso de Big Data para Ingenierías

www.euroinnovaformazione.it



Euroinnova International Online Education

Especialistas en **Formación Online**

SOMOS
**EUROINNOVA
INTERNATIONAL
ONLINE
EDUCATION**



Euroinnova International Online Education inicia su actividad hace más de 20 años. Con la premisa de revolucionar el sector de la educación online, esta escuela de formación crece con el objetivo de dar la oportunidad a sus estudiantes de experimentar un crecimiento personal y profesional con formación eminentemente práctica.

Nuestra visión es ser una escuela de **formación online reconocida en territorio nacional e internacional** por ofrecer una educación competente y acorde con la realidad profesional en busca del reciclaje profesional. Abogamos por el aprendizaje significativo para la vida real como pilar de nuestra metodología, estrategia que pretende que los nuevos conocimientos se incorporen de forma sustantiva en la estructura cognitiva de los estudiantes.



**CERTIFICACIÓN
EN CALIDAD**

Euroinnova International Online Education es miembro de pleno derecho en la **Comisión Internacional de Educación a Distancia**, (con estatuto consultivo de categoría especial del Consejo Económico y Social de NACIONES UNIDAS), y cuenta con el **Certificado de Calidad de la Asociación Española de Normalización y Certificación (AENOR)** de acuerdo a la normativa ISO 9001, mediante la cual se Certifican en Calidad todas las acciones formativas impartidas desde el centro.

Descubre Euroinnova International Online Education

Nuestros **Valores****ACCESIBILIDAD**

Somos **cercanos y comprensivos**, trabajamos para que todas las personas tengan oportunidad de seguir formándose.

**HONESTIDAD**

Somos **claros y transparentes**, nuestras acciones tienen como último objetivo que el alumnado consiga sus objetivos, sin sorpresas.

**PRACTICIDAD**

Formación práctica que suponga un **aprendizaje significativo**. Nos esforzamos en ofrecer una metodología práctica.

**EMPATÍA**

Somos **inspiracionales** y trabajamos para **entender al alumno** y brindarle así un servicio pensado por y para él

A día de hoy, han pasado por nuestras aulas **más de 300.000 alumnos** provenientes de los 5 continentes. Euroinnova es actualmente una de las empresas con mayor índice de crecimiento y proyección en el panorama internacional.

Nuestro portfolio se compone de **cursos online, cursos homologados, baremables en oposiciones y formación superior de postgrado y máster.**

DESCRIPCIÓN

Este Curso de Big Data para Ingenierías le ofrece una formación especializada en la materia. Con la realización de este curso online de la Familia Profesional de Sistemas, Redes y Comunicaciones y más concretamente del área de conocimiento Sistemas y Telemática el alumno estudiara los conceptos básicos y específicos de dicha materia. Realiza esta formación y especialícese en Big Data para Ingenierías.

OBJETIVOS

Los objetivos que se pretenden conseguir a través de este curso de Big Data para Ingenierías son los siguientes:

Conocer las tecnologías disponibles para realizar estrategias de Big Data para Ingenierías.
Realizar un desarrollo con Spark y Hadoop.
Analizar datos con Pig, Hive e Impala.

A QUIÉN VA DIRIGIDO

Este curso online está dirigido a todas aquellas personas interesadas en especializarse en dicha materia correspondiente a la Familia Profesional de Sistemas, Redes y Comunicaciones y más concretamente del Área de Conocimiento Sistemas y Telemática y que quieran especializarse en Big Data para Ingenierías.

PARA QUÉ TE PREPARA

Este curso online le prepara para especializarse en Big Data para Ingenierías dentro de la Familia Profesional de Sistemas, Redes y Comunicaciones y más concretamente del área de conocimiento Sistemas y Telemática.

SALIDAS LABORALES



Una vez finalizada la formación en el presente curso, habrás adquirido los conocimientos y habilidades que aumentarán tus expectativas laborales en los siguientes sectores:

- Sistemas, Redes y Comunicaciones.
- Sistemas y Telemática.

Curso de Big Data para Ingenierías



DURACIÓN
200 horas



MODALIDAD
Online

CENTRO DE FORMACIÓN:

Euroinnova International
Online Education



EUROINNOVA
INTERNATIONAL ONLINE EDUCATION

TITULACIÓN

TITULACIÓN expedida por EUROINNOVA INTERNATIONAL ONLINE EDUCATION, miembro de la AEEN (Asociación Española de Escuelas de Negocios) y CLADEA (Consejo Latinoamericano de Escuelas de Administración)



EUROINNOVA
INTERNATIONAL ONLINE EDUCATION

Titulación Expedida por
**Euroinnova International
Online Education**

QUALIFICA2



Titulación Avalada para el
**Desarrollo de las Competencias
Profesionales R.D. 1224/2009**

Una vez finalizado el curso, el alumno recibirá por parte de Euroinnova International Online Education vía correo postal, la titulación que acredita el haber superado con éxito todas las pruebas de conocimientos propuestas en el mismo.

Esta titulación incluirá el nombre del curso/master, la duración del mismo, el nombre y DNI del alumno, el nivel de aprovechamiento que acredita que el alumno superó las pruebas propuestas, las firmas del profesor y Director del centro, y los sellos de la instituciones

que avalan la formación recibida (Euroinnova Internaional Online Education y la Comisión Internacional para la Formación a Distancia de la UNESCO).

MATERIALES DIDÁCTICOS

- Manual teórico: Big Data para Ingenierías - EF
- Paquete SCORM: Big Data para Ingenierías - EF



** Envío de material didáctico solamente en España.*

FORMAS DE PAGO Y FINANCIACIÓN

- Contrareembolso.
- Tarjeta de crédito.
- PayPal.
- Bizum.
- Amazon Pay.
- PayU.



Matricúlate en cómodos Plazos Sin Intereses + Envío Gratis.
Fracciona tu pago con la garantía de



Fracciona el pago de tu curso en cómodos plazos y sin intereses.

Calcula tus plazos con el simulador de cuotas:

 **LLÁMANOS GRATIS AL 900 831 200**

FINANCIACIÓN Y BECAS

EUROINNOVA continúa ampliando su programa de becas para acercar y posibilitar el aprendizaje continuo al máximo número de personas. Con el fin de adaptarnos a las necesidades de todos los perfiles que componen nuestro alumnado.

En EUROINNOVA Ofrecemos a nuestros alumnos facilidades económicas y financieras para la realización del pago de matrículas, todo ello **100% sin intereses**

20% BECA Desempleados

15% BECA Emprende

20% BECA Antiguos Alumnos

15% BECA Amigo

Llama gratis al 900 831 200 e infórmate de nuestras facilidades de pago.

LÍDERES EN FORMACIÓN ONLINE

7 Razones para confiar en Euroinnova

1 NUESTRA EXPERIENCIA

- ✓ Más de **20 años de experiencia**.
- ✓ Más de **300.000 alumnos** ya se han formado en nuestras aulas virtuales
- ✓ Alumnos de los 5 continentes.
- ✓ **25%** de alumnos internacionales.
- ✓ **97%** de satisfacción.
- ✓ **100% lo recomiendan**.
- ✓ **Más de la mitad** ha vuelto a estudiar en Euroinnova

Las cifras nos avalan

 **4,7** ★★★★★
2.625 opiniones **4,7** ★★★★★
12.842 opiniones **8.582**
suscriptores **5.856**
suscriptores

2 NUESTRO EQUIPO

En la actualidad, Euroinnova cuenta con un equipo humano formado por **más de 300 profesionales**. Nuestro personal se encuentra sólidamente enmarcado en una estructura que facilita la mayor calidad en la atención al alumnado.

3 NUESTRA METODOLOGÍA



100% ONLINE

Estudia cuando y desde donde quieras. Accede al campus virtual desde cualquier dispositivo.



APRENDIZAJE SIGNIFICATIVO

Con esta estrategia pretendemos que los nuevos conocimientos se incorporen de forma sustantiva en la estructura cognitiva del alumno.



EQUIPO DOCENTE ESPECIALIZADO

Euroinnova cuenta con un equipo de profesionales que harán de tu estudio una experiencia de alta calidad educativa



NO ESTARÁS SOLO

Acompañamiento por parte del equipo de tutorización durante toda tu experiencia como estudiante.

4 CALIDAD AENOR

- ✓ Somos Agencia de Colaboración N° 9900000169 autorizada por el Ministerio de Empleo y Seguridad Social.
- ✓ Se llevan a cabo auditorías externas anuales que garantizan la máxima calidad AENOR.
- ✓ Nuestros procesos de enseñanza están certificados por AENOR por la ISO 9001



5 CONFIANZA

Contamos con el sello de Confianza Online y colaboramos con la Universidades más prestigiosas, Administraciones Públicas y Empresas Software a nivel Nacional e Internacional.



6 BOLSA DE EMPLEO Y PRÁCTICAS

Disponemos de Bolsa de Empleo propia con diferentes ofertas de trabajo, y facilitamos la realización de prácticas de empresa a nuestro alumnado.

Somos agencia de colaboración N° 9900000169 autorizada por el Ministerio de Empleo y Seguridad Social.



7 SOMOS DISTRIBUIDORES DE FORMACIÓN

Como parte de su infraestructura y como muestra de su constante expansión, Euroinnova incluye dentro de su organización una editorial y una imprenta digital industrial.



ACREDITACIONES Y RECONOCIMIENTOS



Programa Formativo

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. - Securización 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-Valué 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?

Euroinnova

International Online Education

Esta es tu Escuela



¿Te ha parecido interesante esta formación? Si aún tienes dudas, nuestro **equipo de asesoramiento académico** estará encantado de resolverlas. Pregúntanos sobre nuestro método de formación, nuestros profesores, las becas o incluso simplemente conócenos.

Solicita información sin compromiso.

[Llamadme gratis](#)

[¡Matricularme ya!](#)