

FICHA DE ASIGNATURA

Título: Fundamentos de la tecnología Big Data

Descripción:

Esta asignatura introduce el Big Data y el mercado de profesionales especializados en sus distintas tecnologías. La asignatura profundiza en las fases iniciales del proceso de gestión de la información masiva, que son la captura y el almacenamiento de la información. La elección de los modelos correctos en estas dos fases determinará en gran manera la eficiencia del proceso de análisis. De ahí la importancia que se concede a las técnicas para identificar posibles fuentes de datos y para el almacenamiento de la información.

Carácter: Obligatoria

Créditos ECTS: 3

Contextualización:

El programa de esta asignatura proporciona al alumno una visión global del campo de Big Data, así como le introduce la familia de herramientas que componen el nuevo paradigma de almacenamiento y gestión de datos masivos.

Modalidad: Online

Temario:

- Introducción al Big Data. Cadenas de valor y áreas del Big Data: modelado, ingesta, almacenamiento, procesado, análisis y visualización. Definiciones relacionadas con Big Data: Data Science, Data Analytics, Data Mining, Data Science, Machine Learning, Business Intelligence. Perfiles profesionales relacionados con Big Data: Chief Data Officer, Data Manager, Data Scientist.
- Fuentes de datos en entornos Big data: Open Data, Internet of Things, Web semántica. Diferencias respecto a las tecnologías de datos tradicionales. Tipos de datos, flujo de datos, volumen de datos, calidad de datos, privacidad de datos.
- Estructuras de datos y tecnologías para selección de datos útiles.
- Beneficios y riesgos inherentes a la aplicación de técnicas de procesamiento masivo de datos.
- Criterios de calidad de datos en Big Data.
- Técnicas de rastreo, procesamiento, indexación y recuperación de información estructurada y no estructurada.
- Principales estrategias de scraping y crawling.

Competencias:

CE1: Conocer los fundamentos de la ingeniería de datos (modelado, ingesta, almacenamiento, procesado, análisis y visualización), las técnicas de rastreo, procesamiento, indexación y recuperación de información.

CE2: Desarrollar capacidades de programación especializada en analítica y procesamiento de datos en entornos de Big Data.

CE4: Resolver problemas reales en la clasificación, modelización e interpretación de un conjunto de datos haciendo uso de los recursos técnicos disponibles y apropiados en cada caso en particular en el contexto de Big Data.

CE15: Evaluar las diferentes soluciones Big Data frente a un problema y seleccionar en base a criterios de eficiencia y otros, las técnicas óptimas para cada problema, así como ser capaz de ejecutar la solución de forma adecuada e interpretar los resultados obtenidos.

CE16: Comunicar con claridad, a los grupos decisores usuarios de la información, las conclusiones obtenidas en el proceso de análisis de datos.

Actividades Formativas:

Actividad Formativa	Horas	Presencialidad
Clases expositivas	6	0%
Clases prácticas sobre laboratorio informático	6	0%
Tutorías online	5	30%
Trabajo autónomo	58	0%

Metodologías docentes:

- Lección magistral
- Estudio de casos
- Resolución de problemas
- Simulaciones
- Laboratorio informático virtual
- Seguimiento

Sistema de Evaluación:

Sistemas de evaluación	Ponderación mínima	Ponderación máxima
Evaluación de portafolio. Informe sobre resolución de problemas o estudio de casos	20%	30%
Evaluación del portafolio. Informe sobre Diseño de Proyectos	10%	20%
Participación en grupos de debate	10%	10%
Evaluación de la prueba	40%	60%

Normativa específica:

No procede

Bibliografía:

- Abu-Mostafa, Y.S., Magdon-Ismail, M., Lin, H.T (2012) Learning from Data. Online: AML Book

- Baesens, B. (2014) Analytics in a Big Data World: The Essential Guide to Data Science and its Applications. New Jersey: John Wiley & Sons, Inc.
- Davenport, T. H. (2014). Big Data at Work: Dispelling the Myths, Uncovering the Opportunities. USA: Harvard Business School Publishing Corporation
- Foreman, J.W. (2014) Data Smart: Using Data Science to Transform Information into Insight. Indiana: John Wiley & Sons, Inc.
- Maheshwari, A. (2014). Data Analytics Made Accessible (2017 edition) Online: Amazon.
- Simon, P. (2013). Too Big to Ignore. The Business Case for Big Data. New Jersey: John Wiley & Sons, Inc.
- White, T. (2015) Hadoop: The Definitive Guide (4th Edition). USA: O'Really Media, Inc.