
FICHA DE ASIGNATURA

Título: Procesamiento de datos masivos

Descripción:

Analizar masivamente datos requiere de nuevos modelos de procesamiento Big Data que superen las limitaciones de las tecnologías tradicionales. Estos análisis suelen ejecutarse sobre grandes infraestructuras escalables y distribuidas que se pueden desplegar en la nube. El objetivo principal de esta materia es dar a conocer al alumnado las técnicas de búsqueda, extracción y procesamiento distribuido de datos masivos a través de los principales frameworks Big Data para que se puedan desplegar de forma escalable en la nube.

Carácter: Obligatoria

Créditos ECTS: 6

Contextualización:

El programa de esta asignatura proporciona al alumnado conocimientos para desarrollar sistemas de procesamiento masivo de datos que se puedan desplegar en infraestructuras distribuidas en la nube.

Modalidad: Online

Temario:

1. Introducción a tecnologías y servicios cloud en el análisis masivo de datos: infraestructuras, tipos, desarrollo y despliegue de aplicaciones.
2. Desarrollo de aplicaciones escalables
3. Tipos de procesamiento Big Data para modelar la lógica de negocio: batch, streaming, arquitectura Lambda, y Kappa.
4. Modelo de procesamiento MapReduce
5. Herramientas y lenguajes de alto nivel para el procesamiento masivo de datos.
6. Aplicación de soluciones en la nube para procesamiento masivo de datos.
7. Diseño de una solución Big Data

Competencias:

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

CE3: Aplicar diferentes modelos de almacenamiento de datos masivos, así como sistemas de bases de datos, para su procesamiento en infraestructuras distribuidas.

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	12	0%
Clases prácticas sobre laboratorio informático	12	0%
Tutorías online	10	30%
Trabajo autónomo	116	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:

Beltrán, M., Sevillano, F. (2014), *Cloud computing, tecnología y negocio*. Paraninfo, Madrid.

Gates, A. (2011), *Programming Pig*. O'Reilly Media, Sebastopol, CA.

Joyanes, L. (2012), *Computación en la nube*. Marcombo, barcelona, 2ª ed.

Lam, C. (2011), *Hadoop in action*. Manning, Nueva York.

Martín, N., Valverde, R. (2016), *Internet de las cosas*. Unidad Editorial, Madrid.

Miner, D., Shook, A. (2012), *MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems*. O'Reilly Media, Sebastopol CA.

Perera, S. (2013), *MapReduce Patterns: Hadoop Essentials How-to*. Packt Publishing, Birmingham.

Perera, S., Gunarathne, T. (2013), *Hadoop MapReduce Cookbook*. Packt Publishing, Birmingham.

Tannir, K. (2014), *Optimizing Hadoop for MapReduce*. Packt Publishing, Birmingham.

Velte, T., Velte, A., Eisenpeter, C. (2009), *Cloud computing. A practical approach*. McGraw-Hill, Nueva York.

White, T. (2015), *Hadoop: The Definitive Guide (5ed.)*, O'Reilly Media, Sebastopol CA, 4ª ed.