Computación de altas prestaciones para Big Data en las Empresas

Curso Académico: (2019 / 2020) **Fecha de revisión: 29-04-2019**

Departamento asignado a la asignatura: Departamento de Informática

Coordinador/a: GARCIA BLAS, FRANCISCO JAVIER

Tipo: Obligatoria Créditos ECTS : 3.0

Curso : 1 Cuatrimestre : 1

OBJETIVOS

Competencias Básicas

* Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación
* Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio
* Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

Competencias Generales

* Aplicar los fundamentos teóricos de las técnicas de recogida, almacenamiento, tratamiento y presentación de información, especialmente para grandes volúmenes de datos, como base para el desarrollo y adaptación de dichas técnicas a problemas concretos
* Identificar diferentes técnicas para almacenar, replicar y distribuir grandes cantidades de datos, y diferenciarlas en función de sus características teóricas y prácticas
* Utilizar habilidades para el trabajo en equipo y para relacionarse con otros de forma autónoma

Competencias Específicas
* Aplicar conocimientos básicos de programación y bases de datos sobre los que basar la enseñanza de tecnologías y métodos avanzados para el tratamiento de grandes volúmenes de datos
* Identificar las oportunidades que las técnicas de tratamiento de datos pueden suponer para la mejora de la actividad de empresas y organizaciones
* Disponer de conocimientos básicos y fundamentales de arquitectura de computadores, arquitecturas de red y sistemas de almacenamiento
* Aplicar diferentes técnicas para almacenar, replicar y distribuir grandes cantidades de datos
* Identificar y seleccionar las herramientas software adecuadas para el tratamiento de grandes cantidades de datos
* Hacer uso eficiente de plataformas distribuidas para distribución de contenido y técnicas para el almacenamiento de su topología

Resultados de aprendizaje
* Manejo de los conceptos básicos de arquitectura y tecnología computacional y de redes.
* Capacidad de analizar y diseñar arquitecturas y tecnologías de altas prestaciones para grandes volúmenes de datos.
* Conocimiento de técnicas de diseño y desarrollo de aplicaciones de computación de altas prestaciones.
* Conocimientos para analizar y modelar las tecnologías más adecuadas para cada problema, adecuándolas a las especificaciones de los casos concretos

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1.- Introducción a procesamiento Big Data
2.- Paradigma Map-Reduce
3.- Sistemas de almacenamiento en entornos Big Data
    * HDFS como sistema de ficheros distribuido
    * Comandos básicos para la gestión de ficheros en HDFS
4.- Frameworks para la computación intensiva en datos
    * Introducción a Apache Hadoop

* Programación funcional en Scala
* Apache Spark
* Acceso y procesamiento sobre un gran volumen de datos
* Procesamiento de Streaming de datos

4.- Gestión de recursos computacionales
* Introducción a Apache Yarn
* Despliegue de aplicaciones Big Data en entornos corporativo
* Herramientas para monitorización de aplicaciones Big Data

5.- Casos de uso en la empresa

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Actividades formativas:
* Clase teórica
* Clases prácticas
* Prácticas de laboratorio
* Trabajo individual del estudiante

Metodologías docentes:
* Exposiciones en clase del profesor con soporte de medios informáticos y audiovisuales, en las que se desarrollan los conceptos principales de la materia y se proporciona la bibliografía para complementar el aprendizaje de los alumnos.
Lectura crítica de textos recomendados por el profesor de la asignatura:Artículos de prensa, informes, manuales y/o artículos académicos, bien para su posterior discusión en clase, bien para ampliar y consolidar los conocimientos de la asignatura.
Resolución de casos prácticos, problemas, etc.¿ planteados por el profesor de manera individual o en grupo
Elaboración de trabajos e informes de manera individual o en grupo

SISTEMA DE EVALUACIÓN

1.- Evaluación continua (50%)

* Participación en clase
* Trabajos individuales o en grupo realizados durante el curso

2.- Examen final (50%)

Será necesario obtener una calificación de al menos 4 puntos sobre 10 en cada una de las partes evaluadas.

Peso porcentual del Examen Final:	50
Peso porcentual del resto de la evaluación:	50

BIBLIOGRAFÍA BÁSICA

- Holden Karau, Andy Konwinski, Patrick Wendell & Matei Zaharia Learning Spark, O¿Reilly, 2015
- Martin Odersky, Lex Spoon, Bil Venners Programming in Scala, Artima.