Fundamentos tecnológicos en el mundo Big Data

Curso Académico: (2022 / 2023) **Fecha de revisión: 30-08-2022**

Departamento asignado a la asignatura: Departamento de Informática

Coordinador/a: CARRETERO PEREZ, JESUS

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 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 de software y conocer las tecnologías hardware necesarias para la 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

Competencias Específicas

Aplicar conocimientos de programación los que basar la enseñanza de tecnologías y métodos avanzados para el tratamiento de grandes volúmenes de datos.
Conocer las principales tecnologías usadas para construir sistemas big data, así como las principales características de las mismas.
Emplear los resultados básicos de paralización y distribución de trabajos como fundamento para métodos avanzados de predicción y clasificación

Resultados de aprendizaje
-Utilización de técnicas de programación en el diseño y análisis de métodos de tratamiento de datos
- Habilidades para la aplicación de tecnología avanzada al tratamiento de datos
- Conocimiento y utilización de conceptos básicos de programación paralela y distribuida, diseño de algoritmos y sistemas de almacenamiento para el desarrollo de aplicaciones de tratamiento de datos

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1 Introduction
2 Computing platforms for Big Data
  * Computing architectures
  * Parallel computing
  * Advanced computing platforms
  * Big Data frameworks
3 Distributed Memory Systems:
  * Clusters
  * Clouds
4 Storage systems for big data
  * Distributed file systems
  * Parallel file systems
  * Storage technologies
5 Virtualization techniques
  * Virtualization fundamentals
  * Virtualization techniques
  * Server consolidation techniques
6 Parallel and distributed computing paradigms
  * Parallel Python
  * Web services

* Pyspark
* Fault tolerance

7 Use cases:
* Data analytics
* Data streaming

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Actividades formativas
* Clases teóricas
* 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

SISTEMA DE EVALUACIÓN

1.- Evaluación continua (70%)

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

2.- Examen final (30%)

* Resolución de problemas teórico-prácticos asociados a la asignatura
* Nota mínima 3 para hacer media con la evaluación continua.

Peso porcentual del Examen Final:	30
Peso porcentual del resto de la evaluación:	70

BIBLIOGRAFÍA BÁSICA

- Kai Wan and Min Chen Big-Data Analytics for Cloud, IoT and Cognitive Computing, Wiley, ISBN-13: 978-1119247029
- Vivek Kale Parallel Computing Architectures and APIs: IoT Big Data Stream Processing 1st Edición, Amazon, Edición Kindle, 2019