Curso Académico: (2022 / 2023)**Fecha de revisión: 20/05/2022 12:22:52**

Departamento asignado a la asignatura: Departamento de Informática**Coordinador/a: CALLE GOMEZ, FRANCISCO JAVIER****Tipo: Obligatoria Créditos ECTS : 6.0****Curso : 1 Cuatrimestre : 1**

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

- Bases de Datos Estructuradas
- Lenguajes de Datos Algebraicos
- SQL
- Bases de Datos OLAP y Data Warehouse

Habilidades de programación
(conveniente fundamentos de Javascript)

OBJETIVOS

Competencias básicas y generales

- 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 sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades
- 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.
- Capacidad para comprender y aplicar métodos y técnicas del ámbito de la Ingeniería Informática en los mercados financieros
- Capacidad de concebir, diseñar o crear, poner en práctica y adoptar un proceso sustancial de desarrollo o creación de software para mercados financieros

Competencias específicas

- Analizar y evaluar la viabilidad de implantación de un sistema de gestión de datos dependiendo de las necesidades
- Analizar y comprender las principales herramientas para el manejo de grandes cantidades de datos, su almacenamiento, acceso y revisión

Los resultados del aprendizaje vienen determinados tanto por los contenidos de la asignatura como por los de la materia en los que se engloba, que son los Sistemas de soporte a la decisión en el sector financiero

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

Bloque I: Fundamentación Teórica.

Tema 1: Introducción: Marco social y tecnológico

- La sociedad de las TI
- Papel actual de la información y los datos
- Paradigmas de almacenamiento
- Caracterización del concepto Big Data

Tema 2: Aproximación a Big Data

- BBDD Transaccionales y Analíticas
- Organizaciones físicas adaptadas al proceso
- Arquitecturas. Sistemas Distribuidos y CAP.
- Almacenes ROLAP. Operativa analítica en SQL.

Tema 3: Integración, transformación y Limpieza

- Integración de fuentes
- Transformación y Limpieza
- Google Refine
- SPARQL

Bloque II: Implementando Big Data

Tema 4: Operatividad de BigData

- El paradigma Map-Reduce
- Aspectos legales y éticos: Privacidad y Seguridad

Tema 5: Back-End para BigData I: MongoDB

- Introducción a Mongo DB.
- Operatividad Básica en MongoDB
- Agregación en MongoDB. Pipeline y Map-Reduce.
- Replicación y Distribución en MongoDB

Tema 6: Back-End para BigData II: Cassandra

- Fundamentos de Cassandra
- Diseño sobre Cassandra

Tema 7: Back-End para BigData III: Hadoop

- El ecosistema HADOOP y su instalación
- SandBox
- Funcionalidad de HADOOP
- Map-Reduce en HADOOP

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Clase teórica: Presentaciones teóricas acompañadas de material electrónico, como presentaciones digitales.

Clases teórico prácticas: Combinación de clases teóricas acompañadas de la resolución de ejercicios prácticos.

Prácticas de laboratorio: Prácticas a desarrollar en laboratorios específicos para las distintas asignaturas.

- Nº Horas totales: 23 Presencialidad: 100%

Tutorías: Tutorías de carácter presencial y/o a distancia (videoconferencia): Nº Horas totales: 7 Presencialidad: 100%

Actividades de e-learning: tutoriales, lecturas recomendadas, documentación: Nº Horas totales: 50 Presencialidad:0%

Trabajo individual del estudiante: Actividades individuales del alumno que complementan al resto de actividades

(tanto presenciales como no presenciales), así como la preparación de exámenes: Nº Horas totales: 70

Presencialidad: 0%

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

-Exposición y discusión en clase, bajo la moderación del profesor de temas relacionados con el contenido de la materia, así como de casos prácticos

-Elaboración de trabajos e informes de manera individual o en grupo

-Actividades específicas de e-learning, relacionadas con el carácter semi-presencial del título,

actividades de auto-corrección, participación en foros, y cualquier otro mecanismo de enseñanza on-line

SISTEMA DE EVALUACIÓN

Peso porcentual del Examen/Prueba Final: 20

Peso porcentual del resto de la evaluación: 80

La evaluación se llevará a cabo a través de entregas de trabajos a lo largo del curso (80%) y una prueba de evaluación de contenidos teóricos (20%).

BIBLIOGRAFÍA BÁSICA

- Apache¿ Hadoop® <http://hadoop.apache.org/>, Apache¿ Hadoop®, 2016

- MongoDB <http://www.mongodb.org>, MongoDB, 2016

RECURSOS ELECTRÓNICOS BÁSICOS

- Google Refine . Tutorial: <a href="https://www.irekia.euskadi.net/assets/attachments/2421/Tutorial4_IntroduccionGoogleR

- MongoDB . Documentación MongBD: <http://docs.mongodb.org/manual/>

- Sandbox . Documentación Sandbox: <http://hortonworks.com/products/hortonworks-sandbox/>