



Grado: GRADO EN INGENIERÍA INFORMÁTICA		Curso: Optativas generales 3º y 4º (2C)
Asignatura: 803347 - Minería de datos y el paradigma Big Data	Abrev: MIN	6 ECTS
Asignatura en Inglés:		
Materia: Complementos de gestión y procesamiento de la información		24 ECTS
Otras asignaturas en la misma materia:		
Análisis de redes sociales		6 ECTS
Bases de Datos noSQL		6 ECTS
Gestión de la información en la web		6 ECTS
Módulo: Optativo		
Departamento: Sistemas Informáticos y Computación		Coordinador: Valero Espada, Miguel Ángel

Descripción de contenidos mínimos:

Programación lógica y bases de datos deductivas.
Repositorios y minería de datos.
Gestión de la información en la web.
Técnicas de programación para problemas de optimización y planificación.
Modelado y visualización de gráficos.
Procesamiento de señales multimedia.

Programa detallado:

El término Big Data describe un paradigma computacional en las que las colecciones de datos son extremadamente grandes y complejas, lo cual ha propiciado que al aparición de nuevas técnicas de tratamiento y análisis para la extracción de información valiosa de las mismas. La asignatura estudia las principales tecnologías para el tratamiento de grandes cantidades de datos; especialmente enfocada a los algoritmos distribuidos y en la nube.

OBJETIVOS

El objetivo es familiarizar con un amplio espectro de técnicas de manipulación y análisis de grandes cantidades de datos. Desde la clasificación, ordenación y minería.

TEMARIO

Data Warehouse
-Introducción a los almacenes de datos y cubos OLAP
-ETL (extract, transfer and load)
Bases de datos orientadas a documentos
- Evolución de las bases de datos
- Tecnologías de bases de datos NoSQL: BigTable, Hadoop y MongoDB
Bases de datos como sistemas de archivos distribuidos. Introducción a MapReduce.
El paradigma Big Data y la “inteligencia web” •
-Minería de textos: clasificación, búsqueda y ranking
-Algoritmos de web: bag of words, pagerank, adsense, ...
-Búsqueda y clasificación de objetos (imágenes, huellas dactilares,...).
-Introducción a sistemas de recomendación.
Minería y visualización de datos:
-Algoritmos de aprendizaje supervisado.
-Adquisición de datos: web crawling, graph and stream APIs
-Minería de grafos: algoritmos de extracción de comunidades, ranking, social network mining
-Visualización de datos

Programa detallado en inglés:

Introduction to data warehouse
OLAP cubes
ETL (extract, transfer and load)
NoSQL databases
The bigdata paradigm
MapReduce and introduction to cloud computing
Data mining: clustering and machine learning
Text mining
Data acquisition: web crawling
Graph mining
Data visualization

Competencias de la asignatura:

Fecha: ____ de ____ de ____

Firma del Director del Departamento:

**Generales:**

CG18-Conocimiento y aplicación de las herramientas necesarias para el almacenamiento, procesamiento y acceso a los Sistemas de información, incluidos los basados en web.

Específicas:

No tiene

Básicas y Transversales:

CT1-Capacidad de comunicación oral y escrita, en inglés y español utilizando los medios audiovisuales habituales, y para trabajar en equipos multidisciplinares y en contextos internacionales.

CT3-Capacidad para gestionar adecuadamente la información disponible integrando creativamente conocimientos y aplicándolos a la resolución de problemas informáticos utilizando el método científico.

Resultados de aprendizaje:

No tiene

Evaluación:

Todas las pruebas realizadas en cada asignatura serán comunes a todos los grupos de la misma.

Al tener las materias optativas muy diversas características la calificación de las mismas podrá ser muy variada, por lo que los rangos se dejan muy abiertos:

- Exámenes sobre la materia: 0-60%
- Otras actividades: 100-40%

En el apartado “Otras actividades” se podrá valorar la participación activa en el proceso de aprendizaje, la realización de prácticas y ejercicios y la realización de otras actividades dirigidas. La realización de las prácticas de laboratorio y del resto de las actividades evaluables será obligatoria.

Antes del comienzo de cada curso escolar se concretarán en las fichas docentes los porcentajes exactos que se utilizarán durante ese curso para la evaluación de la materia, siendo comunes estos criterios para todos los grupos de una misma asignatura.

La calificación reflejará los resultados de aprendizaje de las diferentes competencias que se adquieren en el módulo o materia.

Evaluación detallada:

20% asistencia y participación en clase
40% prácticas obligatorias
30% examen tipo test
10% presentación en clase

Exámenes:

- | | |
|---|--------------------------------------|
| <input checked="" type="checkbox"/> En Aula | <input type="checkbox"/> En Lab |
| <input type="checkbox"/> Final Feb | <input type="checkbox"/> Parcial Feb |
| <input checked="" type="checkbox"/> Final Jun | <input type="checkbox"/> Parcial Jun |
| <input checked="" type="checkbox"/> Final Sep | <input type="checkbox"/> Sin Examen |

Actividades formativas:

Las actividades formativas que se van a realizar para esta materia se dividen en tres grupos:

Actividades presenciales: 30-40% de la dedicación del alumno. Estas actividades podrán incluir:

Clases teóricas magistrales.

Clases de problemas.

Laboratorios.

Seminarios.

Actividades dirigidas: 10-15% de la dedicación del alumno. Estas actividades podrán incluir:

Trabajos dirigidos.

Tutorías dirigidas.

Trabajo personal: 50-55% de la dedicación del alumno. Estas actividades podrán incluir:

Trabajo personal no dirigido: Estudio, preparación de exámenes, realización de ejercicios.

Realización de exámenes.

Actividades docentes:

Reparto de créditos:

Teoría: 2,00

Problemas: 1,00

Laboratorios: 3,00

Otras actividades:

No tiene

Fecha: ____ de ____ de ____

Firma del Director del Departamento:



UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE INFORMATICA

Bibliografía:

Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More, Matthew A. Russell

Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques , Morgan Kaufmann

Cursos online

Intro to Artificial Intelligence

<https://www.udacity.com/course/cs271>

La "inteligencia web" y los "grandes datos"

<https://www.coursera.org/course/bigdata>

Introducción a la ciencia de datos

<https://www.coursera.org/course/datasci>

Software

<http://www.knime.org/>

<http://gephi.org/>

<http://community.pentaho.com/projects/mondrian/>

<http://www.mongodb.org/>

<http://www.cs.waikato.ac.nz/~ml/weka/>

Ficha docente guardada por última vez el 25/06/2015 12:42:00 por el usuario: Vic. Estudios

Fecha: ____ de ____ de ____

Firma del Director del Departamento:



UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE INFORMÁTICA

Fecha: ____ de _____ de ____

Firma del Director del Departamento: