

GUÍA DOCENTE
EXPERIENCIA PILOTO DE APLICACIÓN DEL SISTEMA EUROPEO DE CRÉDITOS (ECTS)
UNIVERSIDADES ANDALUZAS

TITULACIÓN: Ingeniería Técnica en Informática de Gestión

DATOS BÁSICOS DE LA ASIGNATURA

NOMBRE: DATA WAREHOUSE Y DATA MINING

CÓDIGO: 935

AÑO DE PLAN DE ESTUDIO: 2004

TIPO (troncal/obligatoria/optativa): OPTATIVA

Créditos totales (LRU/ECTS):
4.5

Créditos LRU/ECTS teóricos: 2.

Créditos LRU/ECTS prácticos: 2.5

CURSO: 3º

CUATRIMESTRE(S): 2º

CICLO: 1º

EQUIPO DOCENTE

Responsable / Coordinador de la asignatura:

NOMBRE: FRANCISCO A. GÓMEZ VELA

CENTRO/DEPARTAMENTO: ESCUELA POLITECNICA SUPERIOR

ÁREA: LENGUAJES Y SISTEMAS INFORMATICOS

CATEGORÍA: PROFESOR SUSTITUTO

LA ASIGNATURA EN EL PROGRAMA FORMATIVO

1. DESCRIPTOR.

Comprender los conceptos fundamentales de las técnicas de minería de datos y de data warehouse.
Conocer y utilizar distintos algoritmos de minería de datos

2. UBICACIÓN EN EL PROGRAMA FORMATIVO.

2.1. PRERREQUISITOS:

Ninguno.

2.2. CONTEXTO DENTRO DE LA TITULACIÓN:

La asignatura introduce al alumnado en los conceptos de minería de datos y de datawarehouse

2.3. RECOMENDACIONES:

No hay.

3. LA ASIGNATURA EN LA ADQUISICIÓN DE COMPETENCIAS.

3.1. COMPETENCIAS TRANSVERSALES/GENÉRICAS:

Conocimiento generales básicos.

Solidez en los conocimientos básicos de la profesión.

Habilidades elementales en informática.

Resolución de problemas.

Capacidad de aprender.

Habilidad para trabajar de forma autónoma.

3.2. COMPETENCIAS ESPECÍFICAS:

- **Cognitivas (Saber):** Enseñar al alumno unos conocimientos sobre los conceptos fundamentales de la minería de datos y de datawarehouse: algoritmos de datamining, extracción de conocimientos de datos
- **Procedimentales/Instrumentales (Saber hacer):** Enseñar al alumno los conceptos necesarios para solucionar problemas relacionados con la análisis de datos.
- **Actitudinales (Ser):** Curiosidad técnica, habilidad analítica para solucionar problemas relacionados con la extracción de conocimientos de datos. Adaptación tecnológica a nuevas técnicas de data mining.

4. OBJETIVOS.

- A) Introducir al alumno en las técnicas de análisis inteligente de datos
- B) Conocer herramienta para la extracción de información de datos

DISTRIBUCIÓN DEL TRABAJO PRESENCIAL.

	Gran Grupo	Grupo de Docencia	Actividades dirigidas (seminarios)
Nº de grupos	1	2	2
Nº de horas	12	12	4
Nº de sesiones	12	8	2

5. METODOLOGÍA.

NÚMERO TOTAL DE HORAS DE TRABAJO DEL ALUMNO: 117

PRIMER SEMESTRE: 0 horas de trabajo

Nº de Horas: 0

- Enseñanzas básicas (Gran Grupo): ...
- Enseñanzas prácticas y de desarrollo (Grupo de Docencia): ...
- Actividades académicas dirigidas (Seminarios-Grupo de Trabajo): ...
- Tutorías especializadas (presenciales o virtuales): ...
 - A) Colectivas: ...
 - B) Individuales: ...
- Trabajo personal autónomo: ...
 - A) Horas de estudio de enseñanzas básicas: ...
 - B) Horas de estudio-preparación de las enseñanzas básicas y de desarrollo: ...
 - C) Horas de trabajo personal o en grupo derivadas de las actividades académicas dirigidas: ...
- Otras actividades (visitas, excursiones, etc.): ...
- Realización de pruebas de evaluación y/o exámenes: ...
 - C) Pruebas de evaluación y/o exámenes escritos: ...
 - D) Pruebas de evaluación y/o exámenes orales (control del Trabajo Personal): ...

SEGUNDO SEMESTRE: 117 horas de trabajo

Nº de Horas: 117

1. Enseñanzas básicas (Gran Grupo)* ⁽¹⁾: 12
2. Enseñanzas prácticas y de desarrollo (Grupo de Docencia)*: 12
3. Actividades académicas dirigidas (Seminarios-Grupo de Trabajo)*: 4
Compute aquí la parte de las actividades dirigidas realizada en presencia del profesor mediante los seminarios
4. Tutorías especializadas (presenciales o virtuales):
 1. Colectivas:
 2. Individuales: 1
5. Trabajo personal autónomo*: 84
 - Horas de estudio de enseñanzas básicas*: 24
 - Horas de estudio-preparación de las enseñanzas prácticas y de desarrollo*: 36
 - Horas de trabajo personal o en grupo derivadas de las actividades académicas dirigidas*: 24
6. Otras actividades (visitas, excursiones, etc.)
7. Realización de pruebas de evaluación y/o exámenes*:
 - A) Prueba de evaluación y/o exámenes escritos: 4
 - B) Pruebas de evaluación y/o exámenes orales (control del Trabajo Personal):

6. TÉCNICAS DOCENTES. (Señale con una X las técnicas que va a utilizar en el desarrollo de su asignatura. Puede señalar más de una).		
Sesiones académicas teóricas: X	Exposición y debate: X	Tutorías especializadas: X
Sesiones académicas prácticas: X	Visitas y excursiones:	Controles de lecturas obligatorias:
Otras (especificar): DESARROLLO Y JUSTIFICACIÓN: 		
7. BLOQUES TEMÁTICOS. (Dividir el temario en grandes bloques temáticos; no hay número mínimo ni máximo). 1. PRESENTACIÓN E INTRODUCCIÓN AL ENTORNO WEKA. 2. INTRODUCCIÓN AL ANÁLISIS INTELIGENTE DE DATOS. 3. TÉCNICAS DE PREPROCESADO 4. TÉCNICAS DE SELECCIÓN DE ATRIBUTOS. 5. TÉCNICAS DE CLUSTERING / TÉCNICAS DE REGLAS DE ASOCIACIÓN. 6. TÉCNICAS DE CLASIFICACIÓN/REGRESIÓN 7. TÉCNICAS DE VALIDACIÓN. 8. ESTRUCTURA DEL ANÁLISIS EXPERIMENTAL DE DATOS.		

8. BIBLIOGRAFÍA.

8.1 GENERAL:

- I.H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2005
- José Hernández Orallo, M.José Ramírez Quintana, César Ferri Ramírez, Introducción a la minería de datos, Pearson, 2004.

8.2 ESPECÍFICA: (con remisiones concretas en lo posible)

- Tema 1. I.H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, capítulo 9.
José Hernández Orallo, M.José Ramírez Quintana, César Ferri Ramírez, Introducción a la minería de dato, capítulo 1
- Tema 2. I.H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, capítulo 1, 2.
José Hernández Orallo, M.José Ramírez Quintana, César Ferri Ramírez, Introducción a la minería de dato, capítulo 2
- Tema 3. I.H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, capítulo 7.
José Hernández Orallo, M.José Ramírez Quintana, César Ferri Ramírez, Introducción a la minería de dato, capítulo 4
- Tema 4. I.H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, capítulo 7
- Tema 5. I.H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, capítulo 6.
José Hernández Orallo, M.José Ramírez Quintana, César Ferri Ramírez, Introducción a la minería de dato, capítulo 11
- Tema 6. I.H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, capítulo 6.
- Tema 7. I.H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, capítulo 5
- Tema 8. I.H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, capítulo 12

9. TÉCNICAS DE EVALUACIÓN.

La evaluación se basará principalmente en los conocimientos adquiridos tanto en clase de teoría como en aula de informática. La participación también será evaluada.

Para la realización de cualquier prueba evaluable no se permitirá el uso o consulta de documentación, salvo indicación expresa del profesor en convocatoria oficial.

Criterios de evaluación y calificación: (referidos a las competencias trabajadas durante el curso)

La nota oscilará entre 0 y 10 puntos, los cuales se acumularán en función de los porcentajes descritos a continuación:

- Enseñanza básica: 30%
- Actividades prácticas y de desarrollo: 45%
- Actividades académicas dirigidas: 25%

Más información en el documento "Criterios de Evaluación"

Nota: Título II. Capítulo II. Artículo 14.2 y 14.3 de la Normativa de Régimen Académico y de Evaluación del Alumnado (aprobada en Consejo de Gobierno de la UPO el 18 de julio de 2006): "En la realización de trabajos, el **plagio** y la utilización de material no original, incluido aquél obtenido a través de Internet, sin indicación expresa de su procedencia y, si es el caso, permiso de su autor, podrá ser considerada causa de calificación de **suspenso** de la asignatura, sin perjuicio de que pueda derivar en **sanción académica**."

Corresponderá a la Dirección del Departamento responsable de la asignatura, oídos el profesorado responsable de la misma, los estudiantes afectados y cualquier otra instancia académica requerida por la Dirección del Departamento, decidir sobre la posibilidad de solicitar la apertura del correspondiente **expediente sancionador**".

10. ORGANIZACIÓN DOCENTE SEMANAL. (Sólo hay que indicar el número de horas que a ese tipo de sesión va a dedicar el estudiante cada semana)

SEMANA	Enseñanzas básicas (Gran Grupo) Nº de horas	Enseñanzas básicas y de desarrollo (Grupo de Docencia) Nº de horas	Actividades académicas dirigidas (Seminarios-Grupos de Trabajo) Nº de horas	Visita y excursiones Nº de horas	Tutorías especializadas Nº de horas	Control de lecturas obligatorias Nº de horas	Exámenes	Temas del temario a tratar
Segundo Cuatrimestre								
SEMANA 1 (13-17 feb)	1	1.5						Tema 1, APD1
SEMANA 2 (20-24 feb)	1							Tema 2
SEMANA 3 (27 feb - 2 mar)	1							Tema 3
SEMANA 4 (5-9 mar)	1	1.5						Tema 4 (4.1) , APD2
SEMANA 5 (12-16 mar)	1	1.5						Tema 4 (4.2) , APD3
SEMANA 6 (19-23 mar)	1							Tema 5
SEMANA 7 (26 mar - 30 mar)	1	1.5						Tema 6 (6.1 – 6.2), APD4,
SEMANA 8 (2-6 abr)								
SEMANA 9 (9-13 abr)	1		2					Tema 6 (6.3), AAD1
SEMANA 10 (16-20 abr)	1	1.5						Tema 7 (7.1 -7.2), APD5
SEMANA 11 (23 abr-27 abr)								
SEMANA 12 (30 abr- 4 may)	1							Tema 7 (7.3) Tema 8 (8.1)
SEMANA 13 (7 may- 11 may)	1	1.5						Tema 8 (8.2 – 8.3), APD6
SEMANA 14 (14 may-18 may)			2					AAD2
SEMANA 15 (21-25 may)	1	1.5						Tema 9, APD7
SEMANA 16 (28 may-1 jun)		1.5						APD8
SEMANA 17 (4-8 jun)								
SEMANA 18 (18 jun – 22 jul)							4	
Evaluaciones finales								

11. TEMARIO DESARROLLADO. (Con indicación de las competencias que se van a trabajar en cada tema).
competencia

1. PRESENTACIÓN E INTRODUCCIÓN AL ENTORNO WEKA.
 1. Presentación de la asignatura.
 2. Introducción al entorno WEKA.
 3. Formato de los datos en el entorno WEKA.
2. INTRODUCCIÓN AL ANÁLISIS INTELIGENTE DE DATOS.
 1. Introducción al proceso KDD (Knowledge Discovery from Data)
 2. Información preliminar de WEKA.
3. TÉCNICAS DE PREPROCESADO.
 1. Técnicas de filtrado.
 2. Técnicas de selección de ejemplos.
4. TÉCNICAS DE SELECCIÓN DE ATRIBUTOS.
 1. Medidas de evaluación.
 2. Algoritmos de búsquedas.
5. TÉCNICAS DE CLUSTERING / TÉCNICAS DE REGLAS DE ASOCIACIÓN.
6. TÉCNICAS DE CLASIFICACIÓN/REGRESIÓN I.
 1. IBk.
 2. J48.
 3. M5P.
7. TÉCNICAS DE CLASIFICACIÓN/REGRESIÓN II.
 1. Naïve Bayes.
 2. Regresión Lineal.
 3. Perceptron.
8. TÉCNICAS DE VALIDACIÓN.
 1. Hold-out.
 2. Leave-one-out.
 3. K-fold cross-validation.
9. ESTRUCTURA DEL ANÁLISIS EXPERIMENTAL DE DATOS

ACTIVIDADES PRÁCTICAS Y DE DESARROLLO

- | | |
|-------|---|
| APD1. | Introducción a WEKA |
| APD2. | Preprocesado de datos |
| APD3. | Selección de Atributos |
| APD4. | Clustering y Reglas de asociación |
| APD5. | Clasificación y regresión I |
| APD6. | Clasificación y regresión II |
| APD7. | Técnicas de validación |
| APD8. | Estructura del análisis experimental de datos |

12. MECANISMOS DE CONTROL Y SEGUIMIENTO. (Al margen de los contemplados a nivel general para toda la Experiencia Piloto, se recogerán aquí los mecanismos concretos que los docentes propongan para el seguimiento de cada asignatura).