ASIGNATURA DE MÁSTER:

MINERÍA DE DATOS (MÁSTER EN ING. DE SISTEMAS Y DE CONTROL)

Curso 2012/2013

(Código:31104233)

1.PRESENTACIÓN

El objetivo de esta guía es orientar al alumno en el estudio de la asignatura. Se recomienda la lectura completa de la guía al comienzo del cuatrimestre para tener una idea de la temática de la asignatura y el plan de trabajo que se piensa seguir en su desarrollo.

La asignatura de “Minería de Datos” se imparte en el primer cuatrimestre del curso, consta de 6 créditos y tiene carácter optativo.

2.CONTEXTUALIZACIÓN

La asignatura de “Minería de Datos” pertenece a la materia de Tratamiento de Datos. La inclusión de esta asignatura en el plan de estudios persigue los siguientes objetivos generales:

1. Adquirir conceptos básicos sobre la teoría de la minería de datos.
2. Proporcionar herramientas y conocimientos necesarios para otras asignaturas que forman parte de este Máster.
3. Ayudar a adquirir las competencias genéricas y específicas propias de este máster.

Los dos primeros objetivos son propios de cualquier enseñanza tradicional de carácter técnico. En el tercer objetivo se menciona la adquisición de competencias propias de las enseñanzas impartidas en el Espacio Europeo de Educación Superior. En este sentido, la asignatura de “Minería de Datos” contribuye al desarrollo de distintas competencias genéricas y específicas de las planteadas en el plan de estudios en el que se enmarca. Entre ellas se deben destacar:

Competencias Genéricas:

Capacidad de gestión y planificación: A la hora de resolver un problema hay que ser capaz de establecer de forma adecuada los objetivos así como secuenciar y organizar el tiempo de realización.

Capacidad de análisis, síntesis y toma de decisiones: El proceso de análisis y síntesis de la información suministrada en el enunciado del problema a resolver es fundamental para poder aplicar adecuadamente los conocimientos adquiridos a la práctica. La toma de decisiones sobre que camino tomar para resolver un problema determinado, de entre todos los posibles, también es tarea importante.

Comunicación científica y tecnológica: Se irá introduciendo progresivamente la terminología común a la materia facilitando además el intercambio de contenidos entre las distintas ramas de la tecnología. Se hará uso de diverso material en lengua inglesa.

Uso de las herramientas y recursos de la Sociedad del Conocimiento: Manejo de las TIC. La mayor parte de la información (materiales), seguimiento y desarrollo del curso está disponible a través de una plataforma de gestión del conocimiento que el alumno utilizará con frecuencia. Por tanto el uso de las TIC, así como su manejo, es fundamental para poder realizar el

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4E0F06713DAC786F4068CA7AD7A04F72

aprendizaje de la asignatura.

Compromiso ético. El estudiante realizará las pruebas y ejercicios que el equipo docente propondrá a lo largo de la asignatura, comprometiéndose a la realización de estos trabajos sin plagios.

Competencias Específicas:

• Búsquedas bibliográficas.

• Presentación de resultados de investigación.

• Técnicas de extracción de información a partir de datos.

• Técnicas de reconocimiento de patrones.

• Técnicas de inferencia y aprendizaje a partir de datos.

• Sintetizar algoritmos de tratamiento de datos y señales.

• Analizar sistemas de tratamiento de datos y señales.

• Evaluar opciones de algoritmos de tratamiento de datos y señales.

Papel de la asignatura dentro del Plan de Estudios

Los contenidos de esta asignatura guardan estrecha relación con la asignatura de “Sistemas Inteligentes” que forma parte de la misma materia y son fundamentales para otras muchas del máster como por ejemplo el procesado de señales.

3.REQUISITOS PREVIOS RECOMENDABLES

No se requiere ningún nivel de conocimientos específico para abordar el estudio de la asignatura, salvo los propios para el ingreso al máster. No obstante es recomendable el conocimiento de análisis matemático, álgebra, estadística, estructura de datos y programación. Además, y dado que parte del material de estudio está en inglés, es necesario tener un nivel básico de esta lengua.

4.RESULTADOS DE APRENDIZAJE

El objetivo básico de la asignatura “Minería de datos” es adquirir conceptos básicos sobre la teoría y conceptos fundamentales utilizados en la minería de datos.

Los resultados de aprendizaje más significativos son los siguientes:

1. Conocer las distintas fases implicadas en un proyecto de minería de datos.

2. Ser capaz de caracterizar un conjunto de datos seleccionando y generando sus características más relevantes.

3. Conocer las distintas técnicas de clasificación, supervisada y no supervisada.

4. Saber seleccionar la técnica de clasificación más adecuada para resolver un problema concreto.

5. Ser capaz de interpretar y analizar los resultados obtenidos en la resolución del problema.

6. Realizar comparaciones entre diferentes aproximaciones para la resolución de un problema determinado.

7. Ser capaz de evaluar la bondad de la técnica utilizada para resolver el problema.

8. Realizar comparaciones entre diferentes aproximaciones para la resolución de un problema determinado.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4E0F06713D4C786F4068CA7AD7A04F72

5.CONTENIDOS DE LA ASIGNATURA

Los contenidos de este curso se organizan en cinco módulos. Se indica a continuación la temática que se trata en cada uno de estos módulos.

Módulo 1: Introducción a la Minería de Datos
Módulo 2: Preparación de Datos
Módulo 3: Técnicas de Clasificación Supervisadas
Módulo 4: Técnicas de Clasificación no Supervisadas
Módulo 5: Aplicaciones

Módulo 1: Introducción a la Minería de Datos
En este módulo se presenta el problema que se quiere abordar, la relación de este problema con otras disciplinas, los distintos ejemplos de aplicación donde se enmarca, el ciclo de diseño y los distintos tipos de modelos.

Módulo 2: Preparación de Datos
Como las técnicas que se estudian en el curso están fundamentalmente dirigidas a la manipulación de grandes bases de datos, es fundamental la reducción de la dimensionalidad de los datos. Para ello en este módulo se abordan distintas técnicas de preprocesamiento de datos tanto de selección como de generación de características.

Módulo 3: Técnicas de Clasificación Supervisadas
Las técnicas de clasificación supervisada parten de la hipótesis de que se conocen a priori las clases donde se deben clasificar los nuevos objetos. Este módulo se centra fundamentalmente en la implementación de Máquinas de Vectores Soporte como técnica eficiente a la hora de diseñar clasificadores supervisados.

Módulo 4: Técnicas de Clasificación no Supervisadas
En este módulo se presentan distintas técnicas de clustering. En primer lugar se detalla de forma general el problema de este tipo de clasificación y a continuación se explican, utilizando para ello distintos ejemplos los diferentes algoritmos. Además se analiza la complejidad de cada uno.

Módulo 5: Aplicaciones
Se presentarán diferentes ejemplos que utilizan las técnicas estudiadas en el curso. Estos ejemplos son aplicaciones reales que hoy en día se encuentran funcionando en el campo de la fusión nuclear.

6.EQUIPO DOCENTE

- SEBASTIAN DORMIDO CANTO
- NATIVIDAD DURO CARRALERO
- JOSE SANCHEZ MORENO
- RAQUEL DORMIDO CANTO

7.METODOLOGÍA

La metodología que se usará en la asignatura será la metodología a distancia propia de la UNED, basada en una educación que puede realizarse totalmente a distancia con el apoyo del uso de las tecnologías de la información y el conocimiento. El alumno cuenta inicialmente con esta guía de estudio que explica en detalle el plan de trabajo propuesto para la asignatura y proporciona orientaciones sobre el estudio y las actividades que se deben realizar. Además, en esta guía encontrará información sobre cómo está organizada la asignatura, cómo utilizarla y qué papel están llamados a desempeñar los materiales y medios que se van a utilizar para llevar a cabo el estudio de la asignatura, así como qué actividades y prácticas se pondrán en marcha, qué calendario deben seguir para realizarlas y como debe enviar los documentos y trabajos a realizar.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4E0F06713D4C786F4068CA7AD7A04F72

El alumno dispondrá además de una serie de presentaciones elaboradas por el equipo docente, sobre los distintos módulos que componen el temario de la asignatura. Además en el curso virtual se facilitarán otros materiales como: artículos y enlaces de interés.

Por otro lado, el alumno estará en todo momento apoyado por el curso virtual de la asignatura donde encontrará el apoyo del equipo docente para cualquier duda que se le presente.

La distribución del tiempo de estudio de la asignatura que se proporciona a continuación es orientativa, ya que no puede ser rigurosa pues depende del tipo de alumno que curse la asignatura.

1. Trabajo con contenidos teóricos, lectura de orientaciones, intercambio de información con el equipo docente, etc, puede suponer hasta un 15%.

2. Trabajo autónomo, donde se incluye el estudio de los contenidos teóricos, la realización de trabajos prácticos libres, u obligatorios, la realización de las pruebas de evaluación a distancia puede suponer el 85% restante.

8.BIBLIOGRAFÍA BÁSICA

Comentarios y anexos:

El material será facilitado por el equipo docente y estará disponible en el curso virtual.

9.BIBLIOGRAFÍA COMPLEMENTARIA

ISBN(13): 9780126858754
Título: PATTERN RECOGNITION (Second Edition)
Autor/es: Konstantinos Koutroumbas ; Sergios Theodoridis ;
Editorial: Elsevier Academic Press

Buscarlo en libreria virtual UNED

Buscarlo en bibliotecas UNED

Buscarlo en la Biblioteca de Educación

Buscarlo en Catálogo del Patrimonio Bibliográfico

ISBN(13): 9780471056690
Título: PATTERN CLASSIFICATION (2nd ed.)
Autor/es: Hart, Peter E. ; Stork, David G. ;
Editorial: WILEY

Buscarlo en libreria virtual UNED

Buscarlo en bibliotecas UNED

Buscarlo en la Biblioteca de Educación

Buscarlo en Catálogo del Patrimonio Bibliográfico

ISBN(13): 9780471154938

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4E0F06713D4C786F4068CA7AD7A04F72

Título: LEARNING FROM DATA :
Autor/es: Mulier, Filip ;
Editorial: JOHN WILEY AND SONS

Buscarlo en libreria virtual UNED

Buscarlo en bibliotecas UNED

Buscarlo en la Biblioteca de Educación

Buscarlo en Catálogo del Patrimonio Bibliográfico

ISBN(13): 9783540422976
Título: PATTERN RECOGNITION: CONCEPTS, METHODS AND APPLICATIONS
Autor/es: J.P. Marques De Sá ;
Editorial: SPRINGER VERLAG

Buscarlo en libreria virtual UNED

Buscarlo en bibliotecas UNED

Buscarlo en la Biblioteca de Educación

Buscarlo en Catálogo del Patrimonio Bibliográfico

Comentarios y anexos:

En la siguiente URL:

http://ai.stanford.edu/~nilsson/mlbook.html

se puede tener acceso al libro:

INTRODUCTION TO MACHINE LEARNING

Autores: Nilson, N. J.

Editorial: Stanford University

10.RECURSOS DE APOYO AL ESTUDIO

Los recursos que brinda la UNED al estudiante para apoyar su estudio son de distintos tipos, entre ellos cabe destacar:

1. Plan de trabajo y orientaciones para su desarrollo, accesible desde el Curso virtual.

2. Curso virtual: Su uso es ineludible para cualquier estudiante, tendrá las siguientes funciones:

a. Atender y resolver las dudas planteadas en los foros siguiendo el procedimiento que indique el Equipo docente .

b. Indicar la forma de acceso a diverso material multimedia de clases y video-tutoriales, que se consideren apropiados.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4E0F06713D4C786F4068CA7AD7A04F72

c. Establecer el calendario de actividades formativas.

d. Ser el medio para realizar pruebas de nivel y evaluación continua.

3. Bibliotecas: Además de los recursos anteriores, el uso de la Biblioteca, donde el estudiante podrá encontrar solución autónoma a distintas cuestiones, dada la gran cantidad de material existente en ellas.

4. Internet: Existen muchos recursos en Internet en los que el estudiante se puede basar para un mayor aprovechamiento del estudio. Con frecuencia se le remitirá a ellos.

11.TUTORIZACIÓN Y SEGUIMIENTO

El alumno podrá ponerse en contacto directo con el equipo docente, los lunes y martes de 12:00 a 14:00 en los despachos y teléfonos siguientes:

Raquel Dormido Canto
Teléfono: 913987192
Despacho: 6.01

Natividad Duro Carralero
Teléfono: 913987169
Despacho: 6.01

Sebastián Dormido Canto
Teléfono: 913987194
Despacho: 5.11

José Sánchez Moreno
Teléfono: 913987146
Despacho: 5.11

O a través del correo electrónico: mdatos@dia.uned.es

El departamento se encuentra ubicado en el edificio de la Escuela Técnica de Ingeniería Informática de la UNED situado en la C/ Juan del Rosal 16, 28040 Madrid. Indicaciones sobre cómo acceder a la Escuela se encuentran disponibles es:

UNED Inicio >> Tu Universidad>> Facultades y Escuelas >> ETS de Ingeniería Informática >> Como llegar

Las consultas sobre los contenidos o sobre el funcionamiento de la asignatura se plantearán preferentemente en el curso virtual, utilizando los foros públicos. Si el alumno no puede acceder a los cursos virtuales, o cuando necesite privacidad, se podrá poner en contacto con el equipo docente mediante correo electrónico.

Puesto que se dispone de un curso virtual, la participación en el mismo mediante el planteamiento de preguntas así como la participación en los debates que pueden surgir entorno a las mismas será de gran ayuda. No sólo se enriquece el que recibe la respuesta a su pregunta sino el que la responde dado que pone a prueba los conocimientos adquiridos hasta ese momento.

12.EVALUACIÓN DE LOS APRENDIZAJES

La evaluación de los aprendizajes se llevará a cabo mediante la realización uno o varios trabajos que serán detallados por el equipo docente en el curso virtual.

13.COLABORADORES DOCENTES

Véase equipo docente.

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/
4E0F06713D4C786F4068CA7AD7A04F72

Ámbito: GUI - La autenticidad, validez e integridad de este documento puede ser verificada mediante el "Código Seguro de Verificación (CSV)" en la dirección https://sede.uned.es/valida/

4E0F06713D4C786F4068CA7AD7A04F72