17-18

PROGRAMA DE DOCTORADO EN
SISTEMAS INTELIGENTES

GUÍA DE ESTUDIO PÚBLICA

MINERÍA DE DATOS

CÓDIGO 31101061

UNED

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/
3CA30A38A2151C6BC3F590F5F78C1715

17-18

MINERÍA DE DATOS

CÓDIGO 31101061

ÍNDICE

PRESENTACIÓN Y CONTEXTUALIZACIÓN

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

EQUIPO DOCENTE

HORARIO DE ATENCIÓN AL ESTUDIANTE

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

RESULTADOS DE APRENDIZAJE

CONTENIDOS

METODOLOGÍA

SISTEMA DE EVALUACIÓN

BIBLIOGRAFÍA BÁSICA

BIBLIOGRAFÍA COMPLEMENTARIA

RECURSOS DE APOYO Y WEBGRAFÍA

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/

3CA30A38A2151C6BC3F590F5F78C1715

Nombre de la asignatura	MINERÍA DE DATOS
Código	31101061
Curso académico	2017/2018
Títulos en que se imparte	PROGRAMA DE DOCTORADO EN SISTEMAS INTELIGENTES (máster seleccionado) / MÁSTER UNIVERSITARIO EN I.A. AVANZADA: FUNDAMENTOS,MÉTODOS Y APLICACIONES / MÁSTER UNIVERSITARIO EN LENGUAJES Y SISTEMAS INFORMÁTICOS
Tipo	
Nº ETCS	0
Horas	0.0
Periodo	SEMESTRE
Idiomas en que se imparte	

PRESENTACIÓN Y CONTEXTUALIZACIÓN

El presente curso pretende dar una visión panorámica de la teoría y conceptos fundamentales utilizados en Minería de Datos (MD), del conjunto de tareas abordadas por esta disciplina y del repertorio de técnicas y métodos existentes que permiten resolver cada una de estas tareas.

Ficha técnica:

•Tipo: Optativa

•Duración: Anual

•Créditos Totales y Horas: 6 / 150

•Horas de estudio teórico: 55

•Horas de trabajo práctico: 50

•Horas de actividades complementarias: 45

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

El alumno debe haber cursado las asignaturas de Fundamentos Matemáticos de la Informática y Estadística impartidas en el primer ciclo de la titulación de Informática de la UNED o asignaturas equivalentes en otras universidades.

En particular, debe haber adquirido competencias básicas en el manejo algebraico de matrices, cálculo de determinantes , inversión de matrices y diagonalización de éstas. Debe ser capaz de calcular con soltura derivadas parciales e integrales de funciones multivariantes (Análisis Matemático). Finalmente, debe conocer conceptos básicos de Estadística como las propiedades de la distribución gaussiana multivariante o los tests estadísticos de contraste de hipótesis.

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/
3CA30A38A2151C6BC2F590F5F78C1715

EQUIPO DOCENTE

Nombre y Apellidos	LUIS MANUEL SARRO BARO
Correo Electrónico	lsb@dia.uned.es
Teléfono	91398-8715
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	INTELIGENCIA ARTIFICIAL

Nombre y Apellidos	JOSE LUIS AZNARTE MELLADO
Correo Electrónico	jlaznarte@dia.uned.es
Teléfono	91398-9688
Facultad	ESCUELA TÉCN.SUP INGENIERÍA INFORMÁTICA
Departamento	INTELIGENCIA ARTIFICIAL

HORARIO DE ATENCIÓN AL ESTUDIANTE

La tutorización de los alumnos se llevará a cabo exclusivamente a través de la plataforma de e-learning Alf.
Los horarios de los profesores son:
Luis M. Sarro Baro
Guardia: Lunes, de 10:30 a 14:00
Horario de Atención al Estudiante: Miércoles y Viernes de 10:00 a 14:00 h
José Luis Aznarte
Guardia:Lunes, de 16:00 a 20:00

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

RESULTADOS DE APRENDIZAJE

Destrezas y competencias
•Conocer las relaciones existentes de la MD con otras disciplinas.
•Conocer las distintas fases implicadas en un proyecto de minería de datos y las relaciones existentes entre ellas.
•Conocer y saber aplicar las distintas técnicas existentes en MD para realizar preparación de datos.
•Distinguir entre tarea, técnica y método en MD.
•Saber relacionar las distintas tareas propias de MD con las técnicas que permiten resolverlas.
•Conocer que tipo de tarea es capaz de abordar cada técnica de MD.
•Conocer varios tipos de algoritmos o métodos para cada técnica de MD.
•Dominar, tanto desde un punto de vista teórico como práctico, los distintas técnicas/algoritmos utilizados en MD.

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/
3CA30A38A2151C6BC3F590F5F78C1715

•Aplicar técnicas de evaluación adecuadas en función del tipo de modelo a evaluar.
•Practicar con algunas de las herramientas software de minería de datos.
•Afrontar la solución de un proyecto de MD siempre desde un punto de vista metodológico o ingenieril, nunca como un arte.
•Conocer y aplicar las metodologías de MD dedicadas a la creación y seguimiento de un proyecto de minería de datos.
•Saber responder a la pregunta de: ¿Cuándo implantar un proyecto de minería de datos en una organización?
•Conocer las repercusiones de la MD en distintos campos: social, legal y ético.
•Conocer los retos que plantea la MD actualmente y las tendencias futuras.

CONTENIDOS

METODOLOGÍA

La general del Máster adaptada a las directrices del EEES, de acuerdo con el documento del IUED. Junto a las actividades y enlaces con fuentes de información externas, existe material didáctico propio preparado por el equipo docente. La asignatura no tiene clases presenciales. Los contenidos teóricos se impartirán a distancia, de acuerdo con las normas y estructuras de soporte telemático de la enseñanza en la UNED.
En particular, en la asignatura se abordarán de manera secuencial las diversas fases del proceso de descubrimiento de conocimiento desde el punto de vista algorítmico, de manera que es conveniente seguir los contenidos de manera igualmente secuencial. Algunos temas vienen acompañados de una o varias actividades cuya memoria servirá de base para la evaluación. Recomendamos leer primero los contenidos teóricos de cada tema (y especificos de cada actividad) antes de abordar las actividades.
No es necesario memorizar expresamente los contenidos del temario (no hay examen presencial de la asignatura), pero el quipo docente hará especial énfasis en la comprensión de los contenidos mostrada en las actividades. Éstas están diseñadas de manera que el/la estudiante debe realizar una tarea importante de contextualización y análisis. Si el/la estudiante se limita a generar resultados sin demostrar la comprensión de los conceptos en la discusión de dichos resultados se considerará que la práctica es insuficiente.

SISTEMA DE EVALUACIÓN

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/
3CA30A38A2151C6BC3F590F5F78C1715

BIBLIOGRAFÍA BÁSICA

ISBN(13):9788420540917
Título:INTRODUCCIÓN A LA MINERÍA DE DATOS (1ª)
Autor/es:Ferri Ramírez, César ; Ramírez Quintana, Mª José ; Hernández Orallo, José ;
Editorial:PEARSON

Muy importante: Existen textos más actualizados y modernos que el que se recomienda en la bibliografía básica (ver la bibliografía complementaria). Hemos mantenido éste por razones históricas y por estar en castellano, pero el estudiante tiene libertad total para escoger el texto que más se adecúe a su nivel o a sus preferencia de aprendizaje. Puede consultar al equipo docente si necesita asesoramiento sobre qué libro escoger.

El material docente del presente curso está compuesto por el texto base indicado en la bibliografía básica, por textos alternativos indicados en la bibliografía general de consulta, por los artículos referenciados en las actividades y en los epígrafes "Orientaciones", pertenecientes al desglose que se hace más adelante de cada tema por separado y, finalmente, por aquellas herramientas software indicadas en algunas de las actividades a realizar.

El texto base será el hilo conductor para el estudio de los contenidos de este curso. No obstante, dado el carácter introductorio de dicho texto, existirán algunas cuestiones que será necesario ampliar mediante la lectura de bibliografía alternativa.

Tratándose de un master orientado a la investigación, las actividades de aprendizaje se pueden estructurar tanto desde un punto de vista teórico como práctico. En el primer caso, girarán en torno al estado del arte en cada una de las materias del curso y, en el segundo caso, lo harán en relación con la búsqueda de soluciones de distintos subproblemas propios del campo de la MD.

BIBLIOGRAFÍA COMPLEMENTARIA

Materiales y recursos de apoyo

Además de la bibliografía indicada anteriormente, los materiales de apoyo para la realización de las prácticas serán los siguientes:

De manera general, las prácticas se realizarán con el programa R. Si alguien desea hacerlo con python, podrá hacerlo.

Los ficheros con los datos de trabajo serán proporcionados por el equipo docente a través de la plataforma alF o formarán parte de la distribución del software empleado. Si no se indica que la actividad correspondiente haya de ser realizada con un conjunto de datos particular, el alumno podrá elegir un fichero de casos del repositorio de la Universidad de California Irvine http://kdd.ics.uci.edu/.

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/
3CA30A38A2151C6BC3F590F5F78C1715

La plataforma de e-Learning aLF, proporcionará el adecuado interfaz de interacción entre el alumno y sus profesores. Esta plataforma colaborativa permite impartir y recibir formación, gestionar y compartir documentos, crear y participar en comunidades temáticas, así como realizar proyectos online.
Se ofrecerán las herramientas necesarias para que, tanto el equipo docente como el alumnado, encuentren la manera de compaginar tanto el trabajo individual como el aprendizaje cooperativo.

BIbliografia general de consulta

•Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, An Introduction to Statistical Learning, Springer, 2013-2016
•C.M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006
•J. Han, M.Kamber, Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, 2001.
•H. Witten, E. Frank, Data mining: Practical Machine Learning Tools and Techniques (Second Edition). Morgan Kaufmann Publishers, 2005.
•B.Pyle, Data Preparation for Data Mining. Morgan Kaufmann Publishers, 1999

Además, véase la incluida en la descripción de las actividades.

RECURSOS DE APOYO Y WEBGRAFÍA

Ver la sección Comentarios y anexos de Bibliografía complementaria.

IGUALDAD DE GÉNERO

En coherencia con el valor asumido de la igualdad de género, todas las denominaciones que en esta Guía hacen referencia a órganos de gobierno unipersonales, de representación, o miembros de la comunidad universitaria y se efectúan en género masculino, cuando no hayan sido sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el sexo del titular que los desempeñe.

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/
3CA30A38A2151C6BC3F590F5F78C1715