

GUÍA DE ESTUDIO PÚBLICA



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

CÓDIGO 31104233

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	

Nombre de la asignatura	MINERÍA DE DATOS (MÁSTER EN ING. DE SISTEMAS Y DE CONTROL)
Código	31104233
Curso académico	2023/2024
Título en que se imparte	MÁSTER UNIVERSITARIO EN INGENIERÍA DE SISTEMAS Y DE CONTROL
Tipo	CONTENIDOS
Nº ETCS	6
Horas	150.0
Periodo	SEMESTRE 1
Idiomas en que se imparte	CASTELLANO

PRESENTACIÓN Y CONTEXTUALIZACIÓN

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 para conseguir los objetivos propuestos en la asignatura.

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

En esta asignatura se estudiarán técnicas de minería de datos que, posteriormente, se utilizarán en otras asignaturas del Máster o podrán ser futura fuente de trabajos de investigación. Los temas que se van a tratar en la asignatura son los siguientes:

- Introducción a la Minería de Datos
- Técnicas de Clasificación Supervisadas
- Técnicas de Clasificación no Supervisadas
- Aplicaciones.

La asignatura de Minería de Datos pertenece a la materia de Tratamiento de Datos.

Sobre la contribución de la asignatura al perfil profesional está directamente ligada con los perfiles que trabajan en entornos de inteligencia de negocio, ciencia de datos, análisis de datos y otras áreas relacionadas. En concreto, la asignatura cubre en parte conocimientos propios de un perfil de científico de datos, que requiere de un nivel alto de conocimientos matemáticos y de programación para conocer los detalles de implementación de los métodos y algoritmos de minería de datos y machine learning.

La importancia de la minería de datos en un máster denominado "Ingeniería de Sistemas y de Control" no requiere de mayor justificación. Todas las disciplinas científicas e ingenieriles que se agrupan bajo esta denominación, y que son abordadas en gran medida por las asignaturas del máster, utilizan datos en mayor o menor cantidad que deben ser tratados, manipulados y estudiados para poder obtener un mejor conocimiento de los sistemas que se desean estudiar. Por tanto, la citada asignatura guarda una gran relación con muchas otras asignaturas del máster, donde el estudio de las técnicas de minería de datos son fundamentales para poder tratar mejor la información de los sistemas de los cuales hacen uso. Con las técnicas que se presentan en la asignatura el alumno aprenderá a clasificar los

datos proporcionados por el proceso o sistema tanto de forma supervisada como no supervisada, lo que significa que podrá hacerlo tanto si conoce bien el proceso al que se está enfrentando como si no. Se puede indicar por tanto que 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 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”.

REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

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.

EQUIPO DOCENTE

Nombre y Apellidos
Correo Electrónico
Teléfono
Facultad
Departamento

NATIVIDAD DURO CARRALERO (Coordinador de asignatura)
nduro@dia.uned.es

Nombre y Apellidos
Correo Electrónico
Teléfono
Facultad
Departamento

SEBASTIAN DORMIDO CANTO
sebas@dia.uned.es

Nombre y Apellidos
Correo Electrónico
Teléfono
Facultad
Departamento

RAQUEL DORMIDO CANTO
raquel@dia.uned.es

Nombre y Apellidos
Correo Electrónico
Teléfono
Facultad
Departamento

JOSE SANCHEZ MORENO
jsanchez@dia.uned.es

HORARIO DE ATENCIÓN AL ESTUDIANTE

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:

Sebastián Dormido Canto

Teléfono: 913987194

Despacho: 5.11

José Sánchez Moreno

Teléfono: 913987146

Despacho: 5.11

Raquel Dormido Canto

Teléfono: 913987192

Despacho: 6.01

Natividad Duro Carralero

Teléfono: 913987169

Despacho: 6.01

Para cualquier pregunta, aclaración puede en ponerse en contacto con el equipo docente utilizando la dirección de correo: 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 por correo electrónico.

COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

Competencias Básicas:

CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación

CB7 - 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

CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

CB9 - 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

CB10 - 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.

Competencias Generales:

CG01 - Adquirir capacidad de iniciativa y motivación; planificación y organización; y manejo adecuado del tiempo.

CG02 - Ser capaz de seleccionar y manejar adecuadamente los conocimientos, recursos y estrategias cognitivas de nivel superior apropiados para el afrontamiento y resolución de diverso tipo de tareas/problemas con distinto nivel de complejidad y novedad: análisis y síntesis.

CG03 - Ser capaz de aplicar los conocimientos a la práctica y resolver problemas en entornos nuevos o poco conocidos.

CG04 - Ser capaz de desarrollar pensamiento creativo, razonamiento crítico y tomar decisiones

CG05 - Ser capaz de seguir, monitorizar y evaluar el trabajo propio o de otros, aplicando medidas de mejora e innovación.

CG06 - Ser capaz de comunicarse y expresarse, tanto oralmente como por escrito, en castellano y otras lenguas, con especial énfasis en inglés

CG07 - Desarrollar capacidades en comunicación y expresión matemática, científica y tecnológica

CG08 - Ser capaz de utilizar las herramientas y recursos de la Sociedad del Conocimiento: manejo de las TIC, búsqueda de información relevante, gestión y organización de la información, recolección de datos, el manejo de bases de datos y su presentación.

Competencias Específicas:

CE01 - Abordar el tratamiento de procesos industriales, aeronáuticos o navales de distinta tecnología (mecánicos, electrónicos, sociales, ...) recurriendo a diferentes soluciones.

CE02 - Montar sistemas de control sobre procesos reales, incluyendo sensores, actuadores, fusión de datos, comunicaciones, microcontroladores, etc.

CE03 - Ser capaz de realizar búsquedas bibliográficas y de documentación técnica para la resolución de problemas

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.

CONTENIDOS

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.

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.

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.

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.

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.

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.

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.

SISTEMA DE EVALUACIÓN

TIPO DE PRUEBA PRESENCIAL

Tipo de examen

No hay prueba presencial

CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad

No

Descripción

La asignatura se evalúa de la siguiente manera:

- un 90% corresponderá a la realización de un ejercicio práctico.
- un 10% corresponderá a la evaluación continua y final del estudiante mediante videoconferencia con el equipo docente, para garantizar la autoría del trabajo realizado por el alumno.

En el curso virtual estarán especificadas las distintas opciones para la elección del trabajo práctico que será de carácter obligatorio. Cada trabajo a realizar será acordado con el equipo docente a través del email y/o videoconferencia, por lo que el equipo docente tendrá conocimiento desde el inicio del trabajo que va a realizar el alumno y tendrá supervisión sobre el mismo, según el alumno lo vaya requiriendo.

El trabajo a realizar debe barrer toda la temática de la asignatura. Es decir, el trabajo debe partir de unos datos en bruto, que será procesados con las distintas técnicas estudiadas, y, posteriormente, clasificados con varias de las técnicas presentadas. Por último, el trabajo deberá presentar una interfaz o menú de interacción con el usuario donde se puedan seleccionar datos y técnicas y donde se muestren resultados. Todos los trabajos además de la parte de programación deberán ir acompañados de una memoria donde se indique de forma detallada el trabajo realizado y se comparen y analicen los resultados obtenidos.

Criterios de evaluación

La nota final de la asignatura irá en relación con el trabajo presentado por el alumno. El trabajo deberá cumplir una serie de requisitos que el alumno conocerá de antemano y en función de cómo sean desarrollados será llevada a cabo la calificación final.

Ponderación de la prueba presencial y/o los trabajos en la nota final	El trabajo final representa el 90% de la calificación final de la asignatura
---	--

Fecha aproximada de entrega	15/02/2022
-----------------------------	------------

Comentarios y observaciones

El trabajo del alumno es guiado desde el comienzo por el equipo docente, por lo que el profesor conoce de primera mano la evolución del aprendizaje del alumno

Cada alumno deberá ponerse en contacto con el equipo docente de la asignatura por correo electrónico a través de la dirección mdatos@dia.uned.es para fijar el tipo de trabajo a realizar para superar la asignatura.

PRUEBAS DE EVALUACIÓN CONTINUA (PEC)

¿Hay PEC?	No
-----------	----

Descripción

Se considera que al tratarse de una asignatura de máster, con un número de alumnos no excesivo, y de los cuales se puede llevar un seguimiento del trabajo final que realizan desde el inicio hasta el fin, no son necesarias pruebas de evaluación continua adicionales a realizar por parte del alumno.

Criterios de evaluación

Dado que no se disponen pruebas no ha calificación en base a ellas

Ponderación de la PEC en la nota final

Fecha aproximada de entrega

Comentarios y observaciones

OTRAS ACTIVIDADES EVALUABLES

¿Hay otra/s actividad/es evaluable/s? Si,no presencial

Descripción

Videoconferencia con el alumno para garantizar la autoría del trabajo y el seguimiento del mismo por parte del equipo docente.

Criterios de evaluación

Ponderación en la nota final 10% de la calificación final

Fecha aproximada de entrega 15/02/2022

Comentarios y observaciones

¿CÓMO SE OBTIENE LA NOTA FINAL?

La nota final coincide con la calificación obtenida en el trabajo realizado (90%) más la calificación obtenida en la videoconferencia de autoría y seguimiento (10%).

El alumno podrá entregar dicho trabajo en la convocatoria ordinaria de febrero (entrega 15 de febrero) o en la convocatoria extraordinaria de septiembre (entrega 15 de septiembre).

El alumno informará al equipo docente en cuanto sepa la convocatoria en la que se va a presentar.

BIBLIOGRAFÍA BÁSICA

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

BIBLIOGRAFÍA COMPLEMENTARIA

ISBN(13):9780126858754

Título:PATTERN RECOGNITION (Second Edition)

Autor/es:Konstantinos Koutroumbas ; Sergios Theodoridis ;

Editorial:Elsevier Academic Press

ISBN(13):9780471056690

Título:PATTERN CLASSIFICATION (2nd ed.)

Autor/es:Hart, Peter E. ; Stork, David G. ;

Editorial:WILEY

ISBN(13):9780471154938

Título:LEARNING FROM DATA :

Autor/es:Mulier, Filip ;

Editorial:JOHN WILEY AND SONS

ISBN(13):9783540422976

Título:PATTERN RECOGNITION: CONCEPTS, METHODS AND APPLICATIONS

Autor/es:J.P. Marques De Sá ;

Editorial:SPRINGER VERLAG

En la siguiente URL:

<https://ai.stanford.edu/~nilsson/mlbook.html>

se puede tener acceso al libro:

INTRODUCTION TO MACHINE LEARNING

Autores: Nilson, N. J.

Editorial: Stanford University

RECURSOS DE APOYO Y WEBGRAFÍA

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.
 - 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.

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 se hayan sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el

sexo del titular que los desempeñe.