Guía Docente

DATOS DE IDENTIFICACIÓN

Titulación:	Grado en Análisis de Negocios/Business Analytics
Rama de Conocimiento:	Ciencias Sociales y Jurídicas
Facultad/Escuela:	Derecho, Empresa y Gobierno
Asignatura:	Minería de Datos I
Tipo:	Obligatoria
Créditos ECTS:	6
Curso:	2
Código:	5320
Periodo docente:	Tercer semestre
Materia:	Minería de Datos / Data Mining
Módulo:	Formación Disciplinar
Tipo de enseñanza:	Presencial
Idioma:	Castellano
Total de horas de dedicación del alumno:	150

Equipo Docente	Correo Electrónico
Pablo Hidalgo García	pablo.hidalgo@ufv.es
Julio Emilio Sandubete Galán	je.sandubete@ufv.es
María Carmen Lafuente Ibáñez	mcarmen.lafuente@ufv.es

DESCRIPCIÓN DE LA ASIGNATURA

La asignatura de Minería de Datos I inicia al alumno en el conocimiento y en la aplicación de técnicas y herramientas de predicción y de clasificación a una cantidad pequeña o grande de datos para extraer información clave. El curso ofrece formación básica en el análisis de datos multivarientes con el uso del programa R para la

toma de decisiones empresariales.

OBJETIVO

Formar profesionales capaces de comprender la información contenida en los datos almacenados por las organizaciones, o la información contenida en los ficheros de texto, Internet, etc. para encontrar patrones o reglas que una vez validados puedan ser incorporados en los sistemas de análisis de información de las empresas.

Los fines especficos de la asignatura son:

Conocer las técnicas que utiliza la minería de datos para obtener unos resultados.

Saber aplicar técnicas de exploración de datos y de detección de anomalías.

Conocer técnicas de estadística y algoritmos de inteligencia artificial

Comprender la información contenida en una base de datos con el uso de técnicas estadísticas y algoritmos de inteligencia artificial
Conocer y saber usar el programa R para el análisis de datos y su procesamiento mediante técnicas de minería de datos

Conocer y saber usar el programa R para el análisis de datos y su procesamiento mediante técnicas de minería de datos

CONOCIMIENTOS PREVIOS

Se recomienda haber cursado la asignatura de “Introducción a la Estadística y la Probabilidad” en el primer curso,

así como un nivel de comprensión de lectura medio en el idioma inglés.

CONTENIDOS

1. INTRODUCCIÓN A LA MINERÍA DE DATOS Introducción data mining.Tipos de algoritmos en minería de datosHerramientas de minería de datos
2. PROGRAMACIÓN EN R Y ESTADÍSTICA PARA MINERÍA DE DATOS Entorno y paneles de trabajoCRAN y sistema de librerías de R.Tipos y estructuras de datos.Importación y manipulación de bases de datos.Funciones básicas.Estadísticas descriptiva e inferencial.
3. PREPARACIÓN DE DATOS Tratamiento de datos faltantesDetección de anomalíasAnálisis exploratorio de datosLa importancia de visualización de datos
4. MÉTODOS DE REGRESIÓN Regresión lineal.Análisis de los residuos.
5. MÉTODOS BASADOS EN ÁRBOLES Árboles de decisión y de regresión.
6. EVALUACIÓN DE MODELOS División en entrenamiento, validadción y test.Métricas para evaluar modelos

ACTIVIDADES FORMATIVAS

La asignatura de Minería de Datos I tiene un enfoque eminentemente práctico.

La materia se desarrolla en dos tipos de sesiones:
Sesiones teórico-prácticas, en las que, por un lado, se llevará a cabo la exposición de los contenidos del programa mediante el uso de clases magistrales por parte del profesor, así como con el uso de metodologías ágiles que favorecen el proceso de enseñanza-aprendizaje; y por otro lado, se realizarán ejercicios, problemas y casos sencillos relacionados con los conocimientos impartidos en la misma sesión.
Sesiones prácticas que se desarrollarán en el aula, individualmente o en grupos de alumnos, con el programa R dónde se analizarán y procesarán datos obtenidos de distintas fuentes.

Por último, se organizarán grupos de alumnos que deberán desarrollar un trabajo que recoja todo el proceso de minería de datos: selección del conjunto de datos, análisis de las propiedades de los datos recopilados, selección y aplicación de la técnica de minería de datos (construcción de un modelo de predicción y/o de clasificación), extracción del conocimiento y validación del modelo y de que las conclusiones son satisfactorias. Para la realización de este trabajo se utilizará una metodología mixta que integra el Aprendizaje Basado en Proyectos con el Aprendizaje Basado en Problemas. Con esta metodología se trata de que el alumno construya conocimiento y que genere nueva información para resolver problemas que muestren su capacidad de razonar y aplicar conocimiento, así como mostrar su capacidad de proponer soluciones a una problemática real.
El alumno podrá hacer uso de tutorías tanto individuales como grupales para obtener unos mejores resultados del aprendizaje.
Los trabajos en equipo exigen una supervisión continuada por parte del profesor por lo que los alumnos deberán concertar tutorías a lo largo de todo el semestre para garantizar el cumplimiento de los tiempos de realización de cada parte del proceso, así como para orientar y resolver dudas sobre el trabajo a desarrollar.

DISTRIBUCIÓN DE LOS TIEMPOS DE TRABAJO

ACTIVIDAD PRESENCIAL	TRABAJO AUTÓNOMO/ACTIVIDAD NO PRESENCIAL
60 horas	90 horas

COMPETENCIAS

Competencias básicas

Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio

Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio

Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética

Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado

Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía

Competencias generales

Capacidad de organización, sistematización y planificación en la identificación de problemas, pautas y modelos en el contexto del big data

Compromiso ético en la sociedad de la información

Capacidad para el cumplimiento de objetivos, resolución de problemas y toma de decisiones en un entorno de datos masivos tanto cuantitativos como cualitativos

Capacidad de diseñar e implementar proyectos e informes, utilizando con naturalidad los canales digitales

Capacidad de liderazgo y de trabajar en equipo en la sociedad de la información

Capacidad de pensamiento crítico, autocrítico, analítico y reflexivo

Capacidad de aprendizaje autónomo en la sociedad de la información

Competencias específicas

Saber identificar y resolver problemas reales de la empresa, a través del análisis avanzado de datos y de la elección de técnicas adecuadas para la toma de decisiones

Saber manejar herramientas cuantitativas e informáticas para la toma de decisiones

Conocer los fundamentos de la estadística multivariante (data mining) así como su aplicación en el mundo del big data

RESULTADOS DE APRENDIZAJE

Conoce y selecciona métodos y algoritmos de la minería de datos para identificar y resolver problemas reales de la empresa.

Utiliza con soltura el programa R en el análisis de datos para la toma de decisiones empresariales

Conoce y aplica técnicas estadísticas para el análisis multivariente de datos.

Sabe extraer, comprender y analizar la información contenida en grandes volúmenes de datos

SISTEMA DE EVALUACIÓN DEL APRENDIZAJE

El sistema de evaluación del aprendizaje será mediante evaluación continua, examen teórico-práctico, realización de ejercicios, realización y ejecución de problemas y prácticas, valoración de trabajos y de la participación en clase. La asistencia a clases expositivas y prácticas es obligatoria para someterse a este sistema. Todos los alumnos deberán realizar un examen final independientemente de la calificación obtenida en exámenes parciales, evaluación continua, pruebas, etc. Para hacer media, la calificación obtenida en dicha prueba final debe ser de al menos un 5. En ningún caso, la superación de la asignatura podrá reducirse a la aprobación de un examen.
SISTEMA DE EVALUACIÓN PARA ALUMNOS DE PRIMERA MATRÍCULA

1. Elaboración de trabajos grupales: 15%
2. Elaboración de trabajos transdisciplinar: 15%

3. Exámenes parciales (MOOC, certificado R): 10%
4. Actividades diarias, ejercicios y trabajos individuales: 5%
5. Participación en las clase y actividades presenciales en clase: 5%
6. Examen final: 50%

ALUMNOS CON DISPENSA ACADÉMICA

1. Trabajos presentados a requerimiento del profesor:30%
2. Examen final (será el mismo que los alumnos de asistencia normal):70%

ALUMNOS DE SEGUNDA Y SUCESIVAS MATRÍCULAS
En este caso los alumnos podrán optar entre cualquiera de los dos sistemas anteriores, previa comunicación al docente al inicio del semestre.
CONVOCATORIA EXTRAORDINARIA
Se aplicará el mismo criterio que en los casos anteriores. El tiempo fijado para los trabajos prácticos está programado para ejecutarse en ese tiempo, por lo que las fechas de entrega se han de cumplir. El calendario de entrega de estos trabajos se proporcionará al comienzo del curso. Las conductas de plagio, así como el uso de medios ilegítimos en las pruebas de evaluación, serán sancionados conforme a los establecido en la Normativa de Evaluación y la Normativa de Convivencia de la universidad

BIBLIOGRAFÍA Y OTROS RECURSOS

Básica

Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani An Introduction to Statistical Learning 2021

Lind, Douglas A. Estadística aplicada a los negocios y la economía 15
México, D.F.: McGraw-Hill Interamericana, 2012

Complementaria

Joseph H. Hair ... [et al.]. Análisis multivariante / 5ª ed. Madrid :Prentice Hall,D.L. 1999.