

GUÍA DOCENTE

1. DATOS BÁSICOS

Asignatura	Extracción de datos empresariales y modelos predictivos
Titulación	Máster en Tecnología e Inteligencia de Datos Empresariales
Escuela/ Facultad	Facultad de Ingeniería y Tecnología Empresarial
ECTS	6
Carácter	Obligatoria
Idioma/s	Español
Modalidad	Presencial
Semestre	Primero
Docente coordinador	Sara Rodríguez Pastoriza

2. PRESENTACIÓN

La asignatura de Extracción de Datos Empresariales y Modelos Predictivos tiene como objetivo proporcionar a los estudiantes una comprensión integral de los procesos y técnicas fundamentales en el ámbito del análisis de datos empresarial. Inicialmente, se enfoca en el desarrollo de habilidades para la comprensión y preparación de datos, incluyendo la detección y tratamiento de valores faltantes, atípicos y la codificación de variables. Posteriormente, se introduce a los estudiantes en los principios de la minería de datos, explorando metodologías más conocidas como KDD, CRISP-DM y SEMMA, junto con el estudio de herramientas de software para su aplicación práctica. Luego, se profundiza en técnicas de modelado predictivo y aprendizaje automático, abarcando desde algoritmos supervisados como regresión lineal y SVM, hasta algoritmos no supervisados como K-Means y PCA, con énfasis en su aplicación en entornos empresariales. Además, se analizan herramientas avanzadas como la automatización de modelos predictivos y técnicas de procesamiento analítico en línea, junto con la exploración de series temporales y datos en formato texto, demostrando su relevancia en diversos sectores, particularmente en el empresarial.

3. COMPETENCIAS Y RESULTADOS DE APRENDIZAJE

Competencias	Código	Descripción
Básicas	CB06	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.
	CB07	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.
	CB08	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.
	CB09	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.
Generales	CG01	Utilizar las técnicas que permiten obtener y analizar la información, evaluar su relevancia y validez, sintetizarla y adaptarla al contexto empresarial.
	CG02	Utilizar el potencial de las tecnologías de vanguardia para contribuir a la mejora de la competitividad de la empresa.
	CG03	Tratar situaciones complejas e impredecibles de forma sistemática, creativa y con juicio crítico, tomando decisiones con información incompleta y asumiendo riesgos.
	CG04	Comunicarse de forma correcta, tanto oralmente como por escrito, en cualquier ámbito y con el apoyo de la tecnología más actual.
Transversales	CT04:	Aplicar los conocimientos adquiridos a la práctica profesional y al aprendizaje experiencial.
	CT05:	Demostrar habilidades y actitudes para el trabajo autónomo y el trabajo en equipo.
	CT06:	Utilizar conocimientos, habilidades, destrezas y actitudes para comunicarse en entornos digitales.
	CT07:	Seleccionar y utilizar herramientas de las tecnologías de la información y comunicación en el ámbito de su campo de estudio y profesional.
Específicas	CE06:	El titulado o titulada podrá aplicar modelos y técnicas de extracción y preparación de datos para el análisis predictivo en el ámbito empresarial.

Código	Descripción
RA01	Conocer los distintos tipos de fuentes de datos y su adecuación para la preparación de documentos, informes y presentaciones.
RA02	Ser capaz de localizar información precisa, relevante y oportuna en los medios digitales, reconociendo las fuentes fiables y cuáles no son fiables desde el punto de vista del rigor científico y académico.
RA03	Ser capaz de revisar distintos tipos de documentos de forma rápida y eficaz para poder localizar la información clave o más relevante sobre un determinado tema.
RA04	Ser capaz de analizar y revisar textos y argumentaciones aplicando el pensamiento crítico.
RA05	Ser capaz de redactar trabajos e informes claros, concisos y creíbles, con rigor académico y profesional.
RA06	Incorporar de manera correcta las ideas y aportaciones de otros autores en sus trabajos, evitando el plagio mediante el uso de métodos de citación correctos.
RA07	Elaborar correctamente referencias bibliográficas formateadas y organizadas, utilizando el sistema APA.
RA08	Ser capaz de adaptar la presentación de datos al público objetivo, utilizando de forma eficaz las distintas herramientas y tipos de gráficos.

4. CONTENIDOS

- **Unidad I.** Comprensión y preparación de datos.
 - 1.1. Identificación de variables.
 - 1.2. Detección de valores atípicos y datos faltantes.
 - 1.3. Codificación de variables.
 - 1.4. Creación y transformación de variables.
- **Unidad II.** Introducción a la Minería de Datos.
 - 2.1. Comparación entre KDD, CRISP-DM y SEMMA.
 - 2.2. Definición de minería de datos.
 - 2.3. Importancia del análisis predictivo.
 - 2.4. Herramientas de software.
- **Unidad III.** Modelado predictivo y métricas de evaluación de modelos.
 - 3.1. Técnicas de modelado predictivo.
 - 3.2. Algoritmos de aprendizaje supervisado y no supervisado.
 - 3.3. Evaluación y métricas del modelo predictivo.
 - 3.4. Aplicación en la empresa y demás áreas de interés.
- **Unidad IV.** Herramientas avanzadas.
 - 4.1. Automatización de modelos predictivos.
 - 4.2. Herramientas de procesamiento analítico en línea OLAP.
 - 4.3. Exploración de series temporales y datos textuales y modelo Bayesiano.
 - 4.4. Tópicos especiales de minería de datos y aplicaciones futuras.
 - 4.5. Proyecto final de minería de datos

5. METODOLOGÍAS DE ENSEÑANZA-APRENDIZAJE

La UIE desarrolla un modelo académico innovador centrado en el sujeto que aprende, combinando diferentes corrientes filosóficas de Enseñanza-Aprendizaje (E-A), una amplia variedad de actividades de aprendizaje, en especial aquellas donde el estudiante asume un rol activo en la construcción del conocimiento, el acompañamiento permanente y el uso intensivo de las tecnologías, como herramienta facilitadora del proceso, conformando un ecosistema de aprendizaje único e innovador.

La formación se desarrolla en la modalidad presencial, incluyendo la modalidad virtual síncrona, con un campus virtual vanguardista, que proporciona flexibilidad y personalización, en un modelo ubicuo de aprendizaje (U-Learning).

Por otra parte, en armonía con los principios fundacionales y corporativos de responsabilidad social, en la UIE además de promover la participación de toda su comunidad universitaria en actividades de voluntariado y servicio social, incorpora la actividad formativa “Aprendizaje Servicio (ApS)” y la habilita como parte de las estrategias de enseñanza-aprendizaje.

Código	Actividad	Tipo	Modalidad E-A	Modo
MD01	Primer Contacto y Motivación	I	Introductoria	PR
MD02	Presentación, Plan de Trabajo y Compromiso	I		
MD03	Clase Magistral	T	Expositiva y Participativa	PR
MD04	Charlas Invitadas de Expertos	T		
MD07	Actividad en el-Campus Virtual UIE	T/P		
MD08	Estudio de Contenidos	T	Autónoma	NP
MD09	Elaboración de Proyectos y Trabajos	T/P		
MD13	Exposiciones	T/P	Guiada	PR
MD16	Uso de Herramientas de Software	P		
MD17	Prácticas de Laboratorio	P		
MD20	Tutorías	T/P	Personalizada (Individual / Grupal)	PR
MD21	Contrato de Aprendizaje	I/T/P		
MD24	Análisis y Síntesis de Material Documental	T		
MD25	Seguimiento y Finalización	C	Autoevaluación continua	NP

I: Informativa T: Teórica P: Práctica C: Complementaria

PR: Presencial NP: No presencial

6. ACTIVIDADES FORMATIVAS

A continuación, se identifican los tipos de actividades formativas que se realizarán:

Actividades formativas			
Código	Nombre	Modalidad	Tipo
AF01	Introductoria	PR	Motivacional / Informativa
AF02	Expositiva y Participativa	PR	Teórica
AF03	Guiada	PR	Teórica / Práctica
AF04	Personalizada (Individual / Grupal)	PR	Teórica / Práctica
AF05	Autónoma	NP	Teórica / Práctica
AF06	Aprendizaje Servicio	PR	Aprendizaje Servicio
AF07	Autoevaluación continua	NP	Evaluación de la Calidad

PR: Presencial NP: No presencial

7. EVALUACIÓN

El modelo incluye además el proceso de evaluación continua como parte esencial de la verificación de las competencias adquiridas. Para la UIE y en armonía con la mejora del proceso de enseñanza-aprendizaje propuesto para el EEES, el sistema de evaluación que se ha denominado Revisión de los Resultados de Aprendizaje (RRA) se desarrolla como un proceso más humanizado, alejado de los sistemas tradicionales en donde los estudiantes se juega su suerte en exámenes (convocatorias), en ocasiones con pesos porcentuales elevados y definitivos, con la consiguiente generación de estrés, frustración y en ocasiones la deserción.

El sistema RRA de la UIE es de carácter continuo, compartido y progresivo, permitiendo un seguimiento del aprendizaje a lo largo del todo el período, haciendo de ello un proceso natural al que los estudiantes acuden sin emociones negativas y conscientes de la necesidad de conocer su propio progreso.

Código	Actividad de Evaluación	Instrumento	Cantidad	Ponderación %	Tipo	Modo
AE01	Pruebas Parciales	Prueba	2	30	Discreta	O/E
AE03	Proyectos	Proyecto (20%)	1	20	Discreta	O/E
AE04	Exposición	Exposición (15%)	2	15	Discreta	O
AE05	Participación en el Campus Virtual	Cuestionarios (2% c/u%)	5	15	Discreta	E
		Foros (5%)	5		Discreta Pass/Fail	E
AE06	Participación, Actividades diarias y Voluntariado	Tutorías (2%)	3	20	Discreta Pass/Fail	O
		Gestión de Calidad (2%)	3		Discreta Pass/Fail	E
		Prácticum (1%)	25 h		Discreta Pass/Fail	O/E
		Prácticas de Laboratorio (5 % c/u)	3			
AE08	Aprendizaje Servicio					
AE10	Recuperar Parciales (20%)					
				100		

O: Oral

E: Escrito

CD: Carpeta Digital

8. BIBLIOGRAFÍA

- Pérez Castillo, JN. (2020). Introducción a la ciencia de datos en R: un enfoque práctico. [https://www.google.es/books/edition/Introducci%C3%B3n_a_la_ciencia_de_datos_en_R/VXmiEAAAQBAJ?hl=es&gbpv=1].
- Salas Eljatib, C. (2021). Análisis de datos con el programa estadístico R: Una introducción aplicada. [https://www.google.es/books/edition/An%C3%A1lisis_de_datos_con_el_programa_estad/15dOEAAAQBAJ?hl=es&gbpv=1].
- Maurits Kaptein, E.H. (2022). Statistics for Data Scientists: An Introduction to Probability, Statistics, and Data Analysis
- Biecek, P. Burzykowski, T. (2021). Explanatory model analysis: explore, explain, and examine predictive models.
- Kuhn, M. Johnson, K. (2019). Feature Engineering and Selection: A Practical Approach for Predictive Models. <https://doi.org/10.1201/9781315108230>.
- Khder, M. Abu-Alsondos, IA. (2021). The impact of implementing data mining in business intelligence.
- Chatfield, C. Xing H. (2019). The Analysis of Time of Series. Seventh Edition.
- S De, S Dey, S Bhattacharyya, S Bhatia. (2022). Advanced Data Mining Tools and Methods for Social Computing.

9. TUTORÍAS

MD20 Tutoría (2%): Los estudiantes deben asistir como mínimo de tres tutorías personalizadas a lo largo del semestre. Es una actividad del tipo todo o nada ("Pass- Fail"), es decir se deben completar las tres tutorías.

10. ENCUESTAS DE CALIDAD

MD25 Gestión de la calidad (2%): Los estudiantes deben rellenar a lo largo del semestre cuatro formularios referidos a la gestión de la calidad de la UIE. Es una actividad del tipo todo o nada ("Pass- Fail"), es decir se deben completar los cuatro formularios en los plazos previstos en el plan de actividades de la asignatura. La actividad tiene como objetivo valorar oportunamente el desarrollo del proceso de enseñanza-aprendizaje y la competencia transversal referida al pensamiento crítico y autocrítico.