



Guía docente				
Datos Identificativos				2024/25
Asignatura (*)	Minería de Datos	Código	710G04030	
Titulación	Grao en Xestión Dixital de Información e Documentación			
Descritores				
Ciclo	Periodo	Curso	Tipo	Créditos
Grado	2º cuatrimestre	Tercero	Optativa	6
Idioma	CastellanoGallego			
Modalidad docente	Presencial			
Prerrequisitos				
Departamento	Matemáticas			
Coordinador/a	Gómez Rodríguez, Marcos	Correo electrónico	marcos.gomez.rodriguez@udc.es	
Profesorado	Gómez Rodríguez, Marcos	Correo electrónico	marcos.gomez.rodriguez@udc.es	
Web				
Descripción general	El objetivo fundamental de esta asignatura es que el alumno conozca los conceptos fundamentales y los principales modelos de la minería de datos, y su aplicación en el campo de las ciencias de la información y documentación.			

Competencias / Resultados del título	
Código	Competencias / Resultados del título

Resultados de aprendizaje			
Resultados de aprendizaje	Competencias / Resultados del título		
Adquisición de habilidades para la selección, tratamiento, gestión y análisis de la información mediante técnicas de minería de datos.			
Conocimiento y habilidades para la aplicación de las principales técnicas de clasificación.			
Adquirir conocimientos de análisis de datos computacional, incluyendo programas como el software estadístico R.			
Elegir las técnicas cuantitativas adecuadas a los objetivos para tareas de investigación, administración y gestión.			
Capacidad de análisis y de síntesis aplicada a la gestión y organización de la información.			
Conocimiento y habilidades para la aplicación de técnicas de regresión, detección de anomalías y series de tiempo.			
Adquisición de habilidades para la toma de decisiones a partir del análisis estadístico bases de datos complejas.			
Conocimiento y habilidades para la aplicación de las principales técnicas de clasificación.			
Adquisición de habilidades para la toma de decisiones a partir del análisis estadístico bases de datos complejas.			
Conocimiento y habilidades para la aplicación de técnicas de regresión, detección de anomalías y series de tiempo.			
Capacidad de análisis y de síntesis aplicada a la gestión y organización de la información.			
Elegir las técnicas cuantitativas adecuadas a los objetivos para tareas de investigación, administración y gestión.			
Adquirir conocimientos de análisis de datos computacional, incluyendo programas como el software estadístico R.			
Adquisición de habilidades para la selección, tratamiento, gestión y análisis de la información mediante técnicas de minería de datos.			
Adquisición de habilidades para la selección, tratamiento, gestión y análisis de la información mediante técnicas de minería de datos.			
Adquirir conocimientos de análisis de datos computacional, incluyendo programas como el software estadístico R.			
Elegir las técnicas cuantitativas adecuadas a los objetivos para tareas de investigación, administración y gestión.			
Capacidad de análisis y de síntesis aplicada a la gestión y organización de la información.			
Conocimiento y habilidades para la aplicación de técnicas de regresión, detección de anomalías y series de tiempo.			
Adquisición de habilidades para la toma de decisiones a partir del análisis estadístico bases de datos complejas.			
Conocimiento y habilidades para la aplicación de las principales técnicas de clasificación.			



Contenidos	
Tema	Subtema
Introducción a la minería de datos.	Conceptos preliminares. Tipos de problemas en minería de datos: descripción, clasificación, predicción, clustering, detección de anomalías, etc. Tipos de aprendizaje: supervisado y no supervisado.
Métodos de clasificación no supervisada o clúster	Conceptos básicos. Métodos de clasificación jerárquica. Métodos de agrupamiento por particiones. Casos prácticos.
Métodos de clasificación supervisada.	Conceptos básicos. Modelos principales de clasificación supervisada o reconocimiento de patrones. Validación de modelos de clasificación (¿cómo de bien predicen?). Casos prácticos.
Métodos avanzados de regresión.	Introducción. Modelos de regresión univariantes y multivariantes. Selección de variables relevantes. Validación de modelos de regresión (¿cómo de bien se ajusta a los datos?, ¿cómo de bien hace predicciones?). Casos prácticos.
Series de tiempo	Conceptos básicos. Análisis descriptivo de series de tiempo. Uso práctico de los modelos de series de tiempo. Casos prácticos.
Técnicas estadísticas para minería de textos y recuperación de la información.	Conceptos básicos. Casos prácticos de aplicación de la minería de textos.

Planificación				
Metodologías / pruebas	Competencias / Resultados	Horas lectivas (presenciales y virtuales)	Horas trabajo autónomo	Horas totales
Sesión magistral		19	0	19
Prácticas a través de TIC		13	0	13
Estudio de casos		7	7	14
Trabajos tutelados		1	101	102
Prueba objetiva		1	0	1
Atención personalizada		1	0	1

(*) Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos

Metodologías	
Metodologías	Descripción
Sesión magistral	Serán sesiones expositivas en las que se introducirán y describirán los diversos temas de la materia, mediante presentaciones (usando los adecuados medios audiovisuales) que incluirán teoría y ejemplos.
Prácticas a través de TIC	Se desarrollarán clases prácticas mediante software estadístico, en las que se introducirá su programación y aplicación a partir de casos reales y simulados.
Estudio de casos	Se aplicarán las técnicas estadísticas impartidas en la materia para la resolución de ejercicios y casos de estudio reales y simulados en el ámbito de la gestión digital de la información.



Trabajos tutelados	Se realizarán trabajos individuales y/o en grupo, tutelados por los docentes de la materia, en los que se abordará la resolución, mediante la aplicación de técnicas estadísticas y el software R, de ejercicios prácticos o de casos de estudio particulares relacionados con el ámbito de la comunicación y de las ciencias de la información. También se podrá realizar un estudio de revisión acerca de un tema concreto de la materia o del software utilizado. Los trabajos podrán ser propuestos por los docentes o por los propios alumnos (las propuestas se tendrán en cuenta o no siempre según el criterio del docente).
Prueba objetiva	Consistirá en una prueba tipo test sobre los contenidos impartidos en la materia.

Atención personalizada

Metodologías	Descripción
Prácticas a través de TIC	En las clases magistrales se fomentará en todo momento el debate entre los alumnos y entre los alumnos y el profesor. Para la resolución de problemas será importante atender personalmente a los alumnos ante las posibles dudas que puedan surgir.
Estudio de casos	Esta atención servirá también, por una parte, al profesor para detectar posibles problemas en la metodología utilizada para impartir la asignatura y, por otra, a los alumnos para consolidar conocimientos teóricos y para expresar sus inquietudes acerca de la asignatura. Será también fundamental la atención personalizada al estudiante durante las clases de prácticas TIC, sobre todo hasta que no se familiarice con el software estadístico a utilizar, además de en la resolución de casos de estudio.
Sesión magistral	
Trabajos tutelados	

Evaluación

Metodologías	Competencias / Resultados	Descripción	Calificación
Prácticas a través de TIC		Se valorará la asistencia y/o el desempeño de los alumnos con el software estadístico.	20
Prueba objetiva		Prueba tipo test consistente en un número de preguntas entre 10 y 20 con 3 respuestas posibles.	40
Trabajos tutelados		Se realizarán trabajos individuales y/o en grupo, tutelados por los docentes de la materia, en los que se abordará la resolución, mediante la aplicación de técnicas estadísticas y el software R, de ejercicios prácticos o de casos de estudio particulares relacionados con el ámbito de la comunicación y de las ciencias de la información. También se podrá realizar un estudio de revisión acerca de un tema concreto de la materia o del software utilizado. Los trabajos podrán ser propuestos por los docentes o por los propios alumnos (las propuestas serán tenidas en cuenta o no siempre según el criterio del docente).	40

Observaciones evaluación

<p>Primera oportunidad</p> <p>Se realizará una prueba de respuesta múltiple de 10 a 20 preguntas que representa el 40% de la nota. Por otra parte, la evaluación continua constará de la asistencia y/o entrega de prácticas relacionadas con el aprendizaje y aplicación del software estadístico R para la resolución de problemas en el campo de la gestión digital de la información (20% de la nota global), además de la entrega de uno o varios trabajos de aplicación de la estadística para la resolución de casos de estudio en documentación digital (alternativamente podrán ser trabajos de revisión o ampliación de la materia) que representa el 40% de la nota total.</p> <p>Segunda oportunidad</p> <p>En la evaluación de la segunda oportunidad se seguirá el mismo criterio que en la primera.</p> <p>Convocatoria adelantada</p> <p>Todas las observaciones previas son aplicables a los estudiantes que soliciten la convocatoria adelantada del examen.</p> <p>Calificación de no presentado</p> <p>En cualquiera de las dos oportunidades anuales figurará un NO PRESENTADO en aquellos casos en los que el alumnado no acuda al examen oficial de la materia.</p> <p>Estudiante con reconocimiento de dedicación a tiempo parcial y dispensa académica de exención de asistencia</p> <p>En el caso del alumnado con reconocimiento de dedicación a tiempo parcial y dispensa académica de exención de asistencia que decida no asistir a clases, éste será evaluado en las dos oportunidades como el resto del alumnado que se encuentra en una situación similar.</p> <p>La realización fraudulenta de las pruebas o actividades de evaluación implicará directamente la calificación de suspenso "0" en la materia en la convocatoria correspondiente, invalidando así cualquier calificación obtenida en todas las actividades de evaluación de cara a la convocatoria extraordinaria.</p>



Fuentes de información

Básica	<ul style="list-style-type: none"> - Silge, J. y Robinson, D. (2017). Text Mining with R: A Tidy Approach. O'Reilly - Cirillo, A. (2017). R Data Mining: Implement Data Mining Techniques Through Practical Use Cases and Real-world Datasets. . Packt Publishing. - Williams, G. (2011). Data mining with Rattle and R: The art of excavating data for knowledge discovery. . Springer Science &amp;amp; Business Media. - Jockers, M.L. (2014). Text Analysis with R for Students of Literature. Springer
Complementaria	

Recomendaciones

Asignaturas que se recomienda haber cursado previamente

Ciencia de Datos/710G04026
Fundamentos Estadísticos/710G04040

Asignaturas que se recomienda cursar simultáneamente

Asignaturas que continúan el temario

Otros comentarios

<p>Para ayudar a conseguir un entorno inmediato sostenible y cumplir con el objetivo de la acción número 5: ?Docencia e investigación saludable y sustentable ambiental y social? del "Plan de Acción Green Campus Ferrol":<p><p>1.- La entrega de los trabajos documentales que se realicen en esta materia: <p><p>1.1. Se solicitará en formato virtual y/o soporte informático. <p><p>1.2. Se realizará a través de Moodle, en formato digital sin necesidad de imprimirlos <p><p>1.3. De realizarse en papel: <p><p>- No se emplearán plásticos. <p><p>- Se realizarán impresiones a doble cara. <p><p>- Se empleará papel reciclado. <p><p>- Se evitará la impresión de borradores. <p><p>2.- Se debe hacer un uso sostenible de los recursos y la prevención de impactos negativos sobre el medio natural. <p><p>3.- Se debe tener en cuenta a importancia de los principios éticos relacionados con los valores de la sostenibilidad en los comportamientos personales y profesionales. <p><p>4.- Según se recoge en las distintas normativas de aplicación para la docencia universitaria se deberá incorporar la perspectiva de género en esta materia (se usará lenguaje no sexista, se utilizará bibliografía de autores de ambos sexos, se propiciará la intervención en clase de alumnos y alumnas?).<p><p>5.- Se trabajará para identificar y modificar prejuicios y actitud sexistas, y se influirá en el entorno para modificarlos y fomentar valores de respeto e igualdad. <p><p>6. Se deberán detectar situaciones de discriminación por razón de género y se propondrán acciones y medidas para corregirlas. <p><p>7. Se facilitará la plena integración del alumnado que por razón físicas, sensoriales, psíquicas o socioculturales, experimenten dificultades a un acceso idóneo, igualitario y provechoso a la vida universitaria.<p>

(*) La Guía Docente es el documento donde se visualiza la propuesta académica de la UDC. Este documento es público y no se puede modificar, salvo cosas excepcionales bajo la revisión del órgano competente de acuerdo a la normativa vigente que establece el proceso de elaboración de guías