



viu

Universidad
Internacional
de Valencia

Guía de Asignatura

ASIGNATURA: *Minería de datos*

Título: *Máster Universitario en Big Data y Ciencia de Datos*

Materia: *Minería de datos*

Créditos: *6 ECTS*

Código: *05MBID*

Índice

1. Organización general	3
1.1. Datos de la asignatura.....	3
1.2. Introducción a la asignatura.....	3
1.3. Competencias y resultados de aprendizaje.....	3
2. Contenidos/temario	5
3. Metodología	5
4. Actividades formativas.....	5
5. Evaluación.....	8
5.1. Sistema de evaluación	8
5.2. Sistema de calificación	8
6. Bibliografía	10

1. Organización general

1.1. Datos de la asignatura

TITULACIÓN	<i>Máster Universitario en Big Data y Ciencia de Datos</i>
ASIGNATURA	<i>Minería de datos</i>
CÓDIGO - NOMBRE ASIGNATURA	<i>05MBID_Minería de datos</i>
Carácter	Obligatorio
Cuatrimestre	Primero
Idioma en que se imparte	Castellano
Requisitos previos	No existen
Dedicación al estudio por ECTS	25 horas

1.2. Introducción a la asignatura

El objetivo de este curso es introducir al alumnado los conceptos básicos de la minería de datos. Estas técnicas permiten describir y modelar conjuntos de datos complejos y procedimientos automáticos encuadrados bajo la categoría de “minería de datos”; también conocida como Knowledge Discovery in Databases (KDD).

Estos procedimientos sirven para extraer información relevante, generalmente implícita y desconocida de los datos. Este mecanismo está diseñado para detectar patrones y regularidades en dichos datos y que puedan ser usados para predecir respuestas ante nuevas situaciones. Además de conseguir detectar los campos clave que determinan las características del problema, realizar segmentaciones de forma automática, lanzar pronósticos, detectar excepciones, proponer escenarios; en definitiva, anticipar el futuro de forma objetiva y basada en datos reales.

Los objetivos propios de la asignatura Minería de Datos son:

- Introducir al alumno al análisis inteligente de datos en entornos Big Data.
- Que el alumno conozca y sea capaz de aplicar un proceso completo de extracción de conocimiento de una base de datos, con sus diferentes etapas.
- Que el alumno conozca las técnicas provenientes tanto de la estadística como de la Inteligencia Artificial para el análisis de datos, que conozca las diferentes utilidades de estas en cuanto a la adecuación de los diferentes problemas a tratar y que distinga claramente los diferentes resultados (conocimiento) que se podrían alcanzar.

1.3. Competencias y resultados de aprendizaje

COMPETENCIAS GENERALES

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 ESPECÍFICAS DE LA ASIGNATURA

CE2 - Utilizar técnicas y herramientas de programación especializada en analítica y procesamiento de datos en entornos de Big Data

CE4 - Resolver problemas reales de clasificación, modelización e interpretación de un conjunto de datos en el contexto de Big Data

CE6 - Utilizar conjuntamente métodos estadísticos avanzados, recursos informáticos y de diagnóstico estratégico para la elaboración de soluciones Big Data de alto valor añadido

CE7 - Aplicar los principales métodos de minería de datos para la extracción de características y la contextualización de información para la toma de decisiones

CE8 - Utilizar técnicas de aprendizaje automático en la detección de potenciales dependencias entre un conjunto de variables

CE9 - Interpretar los resultados obtenidos del análisis de datos a los que se han aplicado técnicas de aprendizaje automático

CE10 - Aplicar técnicas de generación de visualizaciones de datos para el análisis y comunicación de resultados

CE11 - Elaborar modelos de decisión estratégica basados en técnicas de clasificación y tratado de datos en función de los requisitos y necesidades del usuario final

CE15 - Identificar la solución Big Data óptima para un problema, en términos de eficiencia, eficacia e interpretación de resultados

CE16 - Diseñar estrategias de visualización de resultados y presentación de conclusiones obtenidos en el proceso de análisis de datos en un contexto de toma de decisiones.

2. Contenidos/temario

- Proyectos de minería de datos
- Proceso de KDD
- Técnicas de pre-procesamiento de datos
- Métodos de clasificación
- Sistemas de recomendación
- Análisis de datos con cubos y modelos de minería
- Evaluación y selección de modelos: matriz de confusión, métricas, costes. Curvas ROC.

3. Metodología

La metodología de la Universidad Internacional de Valencia (VIU) se caracteriza por una apuesta decidida en un modelo de carácter e-presencial. Así, siguiendo lo estipulado en el calendario de actividades docentes del Título, se impartirán en directo un conjunto de sesiones, que, además, quedarán grabadas para su posterior visionado por parte de aquellos estudiantes que lo necesitasen. En todo caso, se recomienda acudir, en la medida de lo posible, a dichas sesiones, facilitando así el intercambio de experiencias y dudas con el docente.

En lo que se refiere a las metodologías específicas de enseñanza-aprendizaje, serán aplicadas por el docente en función de los contenidos de la asignatura y de las necesidades pedagógicas de los estudiantes. De manera general, se impartirán contenidos teóricos y, en el ámbito de las clases prácticas se podrá realizar la resolución de problemas, el estudio de casos y/o la simulación.

Por otro lado, la Universidad y sus docentes ofrecen un acompañamiento continuo al estudiante, poniendo a su disposición foros de dudas y tutorías para resolver las consultas de carácter académico que el estudiante pueda tener. Es importante señalar que resulta fundamental el trabajo autónomo del estudiante para lograr una adecuada consecución de los objetivos formativos previstos para la asignatura.

4. Actividades formativas

Durante el desarrollo de cada una de las asignaturas se programan una serie de actividades de aprendizaje que ayudan a los estudiantes a consolidar los conocimientos trabajados. A continuación, se relacionan las actividades que forman parte de la asignatura:

1. Clases expositivas

Sesiones dedicadas al desarrollo de los contenidos mediante una metodología de lección magistral. El profesor expone los contenidos de forma que el alumno pueda participar en dicho espacio para interactuar y realizar cuestiones.

2. Sesiones con expertos en el aula

Participación de expertos en la materia dedicadas a ofrecer sus experiencias profesionales en el ámbito de la materia de estudio, ya sea en diseño de proyectos o en cuestiones técnicas.

3. Observación y evaluación de recursos didácticos audiovisuales

Los estudiantes visualizan recursos didácticos audiovisuales como complemento a las sesiones del profesor. El docente dispone al alumnado los recursos para después evaluar con una prueba.

4. Estudio y seguimiento de material interactivo

Los estudiantes disponen de un documento sobre los contenidos de la asignatura en un formato multimedia que complementan el texto base.

5. Clases prácticas

El profesor diseña una serie de sesiones prácticas aplicando diferentes metodologías como el estudio de casos, resolución de problemas, simulación, trabajo cooperativo, diseño de proyectos y trabajo práctico sobre el uso de herramientas propias del Big Data y la Ciencia de Datos. En dichas sesiones el alumno puede recibir comentarios sobre los ejercicios propuestos.

6. Prácticas observacionales

El profesor propone una serie de recursos audiovisuales complementarios para que los alumnos puedan observarlos y así adquirir conocimientos que puedan ser aplicados en otras actividades de carácter práctico.

7. Actividades de seguimiento de la asignatura

El profesor puede proponer en las sesiones programadas una serie de actividades como la exposición de trabajos sobre un tema y tratar de comentar sobre la participación de los alumnos. También proponer el desarrollo de resúmenes, mapas conceptuales, one minute paper, test de autoevaluación..., o, también, un espacio de reflexión sobre el aprendizaje y el análisis crítico.

8. Tutorías

Los estudiantes pueden solicitar tutorías por correo para la orientación sobre dudas de la asignatura.

9. Lectura, análisis y estudio del manual de la asignatura

Trabajo autónomo del estudiante sobre el estudio del texto base.

10. Lectura, análisis y estudio de material complementario

Trabajo autónomo del estudiante sobre el estudio de otros materiales dispuestos por el profesor.

11. Desarrollo de actividades del portafolio

Trabajo autónomo del estudiante sobre el desarrollo de actividades prácticas evaluables.

12. Trabajo cooperativo

Los estudiantes trabajan conjuntamente para la resolución de actividades propuestas en clase, del portafolio, con el objetivo de fomentar la interacción y el trabajo en grupo.

13. Prueba objetiva final

Como parte de la evaluación de cada una de las asignaturas (a excepción de las prácticas y el Trabajo fin de título), se realiza una prueba (examen final). Esta prueba se realiza en tiempo real (con los medios de control antifraude especificados) y tiene como objetivo evidenciar el nivel de adquisición de conocimientos y desarrollo de competencias por parte de los estudiantes. Esta actividad, por su definición, tiene carácter síncrono.

5. Evaluación

5.1. Sistema de evaluación

El Modelo de Evaluación de estudiantes en la Universidad se sustenta en los principios del Espacio Europeo de Educación Superior (EEES), y está adaptado a la estructura de formación virtual propia de esta Universidad. De este modo, se dirige a la evaluación de competencias.

Sistema de Evaluación	Ponderación
Portafolio (*)	40% - 60%
Actividades de evaluación continua: (5% - 10%) Informe sobre resolución de problemas, diseño de proyectos o estudio de casos: (10% - 20%) Informe sobre laboratorio informático virtual: (20% - 30%) Participación en grupos de debate (**): (10% - 10%)	
Sistema de Evaluación	Ponderación
Prueba final (*)	40% - 60%
<i>Examen con preguntas teóricas y prácticas</i>	

(*) Es requisito indispensable para superar la asignatura aprobar cada apartado (portafolio y prueba final) con un mínimo de 5 para ponderar las calificaciones.

() Sólo disponibles en primera convocatoria.**

Los enunciados y especificaciones propias de las distintas actividades serán aportados por el docente, a través del Campus Virtual, a lo largo de la impartición de la asignatura.

Atendiendo a la Normativa de Evaluación de la Universidad, se tendrá en cuenta que la utilización de **contenido de autoría ajena** al propio estudiante debe ser citada adecuadamente en los trabajos entregados. Los casos de plagio serán sancionados con suspenso (0) de la actividad en la que se detecte. Asimismo, el uso de **medios fraudulentos durante las pruebas de evaluación** implicará un suspenso (0) y podrá implicar la apertura de un expediente disciplinario.

5.2. Sistema de calificación

La calificación de la asignatura se establecerá en los siguientes cálculos y términos:

Nivel de aprendizaje	Calificación numérica	Calificación cualitativa
----------------------	-----------------------	--------------------------

Muy competente	9,0 - 10	Sobresaliente
Competente	7,0 - 8,9	Notable
Aceptable	5,0 -6,9	Aprobado
Aún no competente	0,0 -4,9	Suspenso

Sin detrimento de lo anterior, el estudiante dispondrá de una **rúbrica simplificada** en el aula que mostrará los aspectos que valorará el docente, como así también los **niveles de desempeño que tendrá en cuenta para calificar las actividades vinculadas a cada resultado de aprendizaje**.

La mención de «**Matrícula de Honor**» podrá ser otorgada a estudiantes que hayan obtenido una calificación igual o superior a 9.5. Su número no podrá exceder del cinco por ciento de los estudiantes matriculados en una materia en el correspondiente curso académico, salvo que el número de estudiantes matriculados sea inferior a 20, en cuyo caso se podrá conceder una sola «Matrícula de Honor».

6. Bibliografía

Bibliografía de referencia

- Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. (1996). The KDD Process for Extracting Useful Knowledge from Volumes of Data. Communications of the ACM, November 1996/ Vol 39, Nº 11, 27–34.
- Siegel, E. (2013). Analítica predictiva. Predecir el futuro utilizando Big Data. Anaya Multimedia-Anaya Interactiva.
- Mayer-Schönberger, V., Cukier, K. (2013). Big data. La revolución de los datos masivos. Turner

Bibliografía complementaria

- Agrawal, D., Das, S., Abadi, A. E. (2010). Big Data and Cloud Computing: New Wine or Just New Bottles?. Proc. of the VLDB 2010, Vol. 3, No. 2.
- Mayer-Schönberger, V., Cukier, K. (2013). Big data. La revolución de los datos masivos. Turner. Olivas, J. A. (2011). Búsqueda eficaz de información en la Web, Edulp, La Plata, Argentina. <http://hdl.handle.net/10915/18401>.
- Piatetsky-Shapiro, G., Frawley, W. (1991). Knowledge Discovery in Databases. AAAI/MIT Press, Cambridge MA.
- Hastie, Trevor and Tibshirani, Robert and Friedman, Jerome (2009), The elements of statistical learning: data mining, inference, and prediction, Springer Science & Business Media

Videos docentes

- Jupyter Notebook for data Science: Publishing, Packt. Jupyter Notebook for Data Science. PACKT Publishing, 2018, video.alexanderstreet.com/watch/jupyter-notebook-for-data-science. Accessed 17 May 2021.
- Minería de datos por José Ángel Olivas Varela