

Guía docente de Asignatura–Máster en Minería de Datos e Inteligencia de Negocios

Datos generales de la asignatura

Asignatura:	Técnicas y Metodología de la Minería de Datos (SEMMA) - 606538
Curso académico:	2020-21
Carácter:	Obligatoria
Semestre:	1
Créditos ECTS	
Presenciales:	1,8
No presenciales:	4,2
Total	6,0
Actividades docentes	
Clases teóricas:	30%
Clases prácticas:	70%
Total:	100%
Departamentos responsables:	Departamento de Estadística y Ciencia de los Datos
Profesor coordinador:	Aida Calviño Martínez
Profesores:	Aida Calviño Martínez

Datos específicos de la asignatura

Breve descriptor:	<p>Esta asignatura se contempla como una introducción al análisis de datos integrado utilizando software estadístico amigable (Clementine, Miner SAS). Se describen técnicas estadísticas habituales con el objetivo de que se conozca el entorno informático. Se hace especial hincapié en la presentación y resolución de problemas prácticos a los que se les va aplicando diferentes técnicas, entendiéndolas como un procedimiento sistemático de trabajo. Exploración y depuración de datos, resumen de la información y predicción de resultados.</p> <p>El método de trabajo consistirá en utilizar dos grandes bases de datos y trabajar con ellas de forma integrada. Será necesario repetir más de una vez los diferentes pasos que configuran la filosofía SEMMA con el objeto de que al final nos quedemos satisfechos con los resultados de los análisis, confirmando se cumplan todas las hipótesis asociadas a los modelos y técnicas empleadas.</p> <p>Además se buscará que el alumno se ayude a través de diagramas de flujo que describan adecuadamente el proceso del tratamiento de los datos. Por esta razón utilizaremos el interfaz gráfico presente en el software comercial mas generalizado.</p> <p>Las técnicas a las que se dará prioridad en esta asignatura son los métodos de selección de variables, la regresión lineal, la regresión logística y los arboles de regresión.</p>
--------------------------	--

Competencias

Generales:	El alumno debe saber depurar la información, conocer, aplicar, comprobar e interpretar los resultados que se deriven de los análisis realizados. Debe dominar el software de aplicación de la metodología SEMMA. Realizar diagramas informativos que expliquen la secuencia de técnicas a utilizar.
-------------------	---

Contenidos

- 1) Introducción al Enterprise Miner de SAS.
- 2) Estudio de casos.
- 3) Métodos de selección de variables.
- 4) Resolución de los casos a través de la metodología SEMMA usando regresión lineal, regresión logística, KNN y árboles de regresión/clasificación.
- 5) Métodos de evaluación de la metodología predictiva utilizada.

Evaluación

El alumno será evaluado continuamente a lo largo del curso. Para la nota final se tendrá en cuenta la resolución de dos prácticas realizadas fuera del aula, y cuyo peso supondrá el 80% (siendo obligatorio presentar todas las prácticas) y las intervenciones en clase con resolución de cuestiones planteadas por el profesor y la realización de ejercicios propuestos, cuyo peso representará el 20% restante. En caso de que un alumno no haya superado la asignatura por este método, haya faltado a más de un tercio de las clases, o quiera mejorar su calificación se podrá presentar a un examen final y su nota será el 100% de la nota del examen.

El sistema de evaluación será el mismo en cualquiera de los tres escenarios docentes que se plantean (docencia presencial, semi-presencial o en línea).

Bibliografía

Hastie, Trevor; Tibshirani, Robert y Friedman, Jerome (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.

Kuhn, Max y Johnson, Kjell (2016). Applied Predictive Modelling. Springer.

Gareth, James et. al (2015). An introduction to statistical learning : with applications in R. Springer.

Santin González, Daniel y Pérez, López, César (2007). Minería de datos, técnicas y herramientas. Tomson.

S. Sarma, Kattamury (2007). Predictive Modeling with SAS Enterprise Miner: Practical Solutions for Business Applications. SAS Institute.

Cerrito B. Patricia (2006). Introduction to Data Mining Using SAS Enterprise Miner. SAS Institute.

Otra información relevante

En los tres escenarios posibles las clases se llevarán a cabo de manera síncrona (presencialmente o a partir de algunas de las herramientas de docencia virtual disponibles).