

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:	2019-20
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
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 más 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 y Arboles de Regresión.
- 5) Métodos de evaluación de la metodología estadística utilizada.
- 6) Técnicas de Exploración de Datos

Evaluación

El alumno será evaluado continuamente a lo largo del curso, a través de

prácticas. La nota media de éstas supondrá el 80% de la nota final, siendo obligatorio presentar todas las prácticas. Las intervenciones en clase con resolución de cuestiones planteadas por el profesor y la realización de ejercicios propuestos, representará el 20% de la nota final. En caso de que un alumno no haya superado la asignatura por éste método, o quiera mejorar su calificación se podrá presentar a un examen final y su nota será el 100% de la nota del examen.

Bibliografía

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.

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.

