

Curso Académico: (2023 / 2024)

Fecha de revisión: 26-04-2023

Departamento asignado a la asignatura: Departamento de Estadística

Coordinador/a: AUSIN OLIVERA, MARIA CONCEPCION

Tipo: Optativa Créditos ECTS : 6.0

Curso : 4 Cuatrimestre :

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Métodos de Regresión

Análisis Multivariante

No se recomienda cursar esta asignatura si ya se ha cursado Aprendizaje Estadístico

OBJETIVOS

1. Comprender la importancia de transformar grandes volúmenes de datos en información relevante para la toma de decisiones y desarrollo de negocio en organizaciones, empresas y particulares.
2. Aprender las técnicas básicas de preprocesamiento y visualización de datos. Conocer métodos para trabajar con datos faltantes y atípicos. Adquirir el manejo de técnicas de reducción de la dimensión.
3. Conocer los principales métodos de aprendizaje supervisado en regresión y su utilidad en problemas de predicción. Distinguir entre modelos lineales y no lineales y comprender la importancia de los métodos de selección de modelos.
4. Familiarizarse con los procedimientos habituales de aprendizaje supervisado para clasificación. Entender los clasificadores más habituales y sus limitaciones. Conocer métodos avanzados en clasificación y sus beneficios en la empresa.
5. Ser capaz de identificar las técnicas adecuadas del Big Data en problemas reales en la empresa: clasificación de clientes, scoring, gestión de riesgos, detección de fraude, predicción de bancarrota, etc.

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1. Introducción.
2. Recopilación, muestreo y preprocesamiento de datos.
 - 2.1. Tipos de datos
 - 2.2. Muestreo
 - 2.3. Herramientas de visualización de datos
 - 2.4. Valores perdidos
 - 2.5. Detección y tratamiento de atípicos.
 - 2.6. Transformaciones de datos.
 - 2.7. Reducción de la dimensión.
 - 2.8. Aplicación: Gestión de riesgos en el Mercado de valores.
3. Aprendizaje supervisado: Regresión.
 - 3.1. Regresión lineal y polinómica.
 - 3.2. Validación cruzada.
 - 3.3. Selección de modelos y métodos de regularización (ridge y lasso).
 - 3.4. Modelos no lineales, splines y modelos aditivos generalizados.
 - 3.5. Aplicación: Predicción de credit-scoring.
4. Aprendizaje supervisado: Clasificación.
 - 4.1. Clasificadores de Bayes
 - 4.2. Regresión logística.
 - 4.3. K-vecinos más cercanos.
 - 4.4. Random forest
 - 4.5. Máquinas de vector soporte.
 - 4.6. Boosting.

- 4.7. Aplicación: Riesgo de créditos
- 4.8. Aplicación: detección de fraude
- 4.9. Aplicación: Predicción de bancarrota.

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Teoría (2 ECTS). Clases teóricas con material de apoyo disponible en la Web. Prácticas (4 ECTS) Clases de resolución de problemas. Prácticas computacionales en aulas informáticas. Trabajos a realizar en grupo. Tutorías semanales para asesorar a los estudiantes individualmente o en grupos.

SISTEMA DE EVALUACIÓN

Examen final (60%). Presentaciones en clase, vídeos explicativos y proyecto (40%)

Peso porcentual del Examen Final: 60

Peso porcentual del resto de la evaluación: 40

BIBLIOGRAFÍA BÁSICA

- Daniel Peña Análisis de datos multivariantes., McGraw-Hill, 2002