

La minería de datos es una ciencia avanzada que puede ser difícil de hacer correctamente. Este curso demuestra el potencial de la minería de datos y muestra cómo descubrir patrones útiles y tendencias de los datos, consejos prácticos valiosos, adquirida durante años de experiencia en el mundo real, se centra en cómo construir adecuadamente modelos predictivos fiables e interpretar sus resultados con confianza. Los ejemplos se han extraído de varias industrias, incluyendo la puntuación de crédito, detección de fraudes, la biología, las inversiones y la venta cruzada.

1. Objetivos del curso

- Identificar proyectos con una alta probabilidad de éxito
- Traducir un problema de negocio en un conjunto estrechamente relacionado de tareas técnicas
- Transformar los datos en bruto para crear funciones de orden superior que revelan información
- Construir modelos no lineales utilizando técnicas de gran alcance, tales como árboles de decisión y redes neuronales
- Utilizar varias técnicas de remuestreo para evitar sobreajuste y predecir cómo los modelos también llevará a cabo en el uso real
- Interpretar y validar los resultados de la modelización tomar conciencia de los errores analíticos superiores y cómo evitarlos.

2. Contenidos

Resumen ejecutivo

- Introducción, resumen ejecutivo de data mining.
- SAS Enterprise Miner como plataforma de data mining

Algoritmos

- Aplicaciones de ModelApplication
- Procesos mining
- Detección
- Charts acumulativos

Validación de modelo

- Modelos
- Test de ejemplo
- overfit
- Validaciones crossvalidation

Algoritmos II

- Redes neuronales
- Modelos de regresión
- Árboles de decisión

Conjuntos

- Conjuntos
- Debilidades de un modelo único embolsado
- Académicos ejemplo : los árboles con bolsas de cinco frente a ocho nodos
- Ejemplo del mundo real : la puntuación de crédito