

Temario

Diplomado de Modelado Predictivo y Machine Learning

Hamdi Raissi

José Ruetter

Temas Básicos

1. Estadística descriptiva para exploración previa de datos
2. Estadística inferencial
 - a. Test de Hipótesis
 - b. Análisis de asociación de variables
3. Series de tiempo
 - a. Introducción de las series de tiempo
 - b. Suavizado exponencial

Temas Avanzados

Modelos predictivos aplicados durante el dictado de clases

- Regresión Lineal Múltiple
- Regresión Logística
- Árboles de decisión
- Redes Neurales
- Vectores auto soportados

Pasos para el modelado predictivo aplicado a los modelos propuestos

1. Entendiendo los modelos predictivos
 - a. Objetivo de los modelos predictivos
 - b. Elementos de los modelos predictivos
 - c. Pasos básicos para modelar
 - d. Aplicaciones de los modelos predictivos
 - e. Demostración de los modelos predictivos en marketing
 - f. Demostración de los modelos predictivos en banca
2. Retos de los modelos predictivos
 - a. Retos de los modelos predictivos con la data
 - b. Retos en términos de análisis
 - c. Muestreo en los modelos predictivos
 - d. Tratamiento del sesgo en los modelos
3. Ajuste de los modelos predictivos
 - a. Entendiendo el sobre muestreo

- b. Ajustando el sesgo de la muestra en los modelos
- 4. Preparando la data para el modelado
 - a. Las razones de porque se producen datos faltantes o errada
 - b. Estudio de un caso completo de ajuste de los datos faltantes
 - c. Métodos de imputación de datos
 - i. Imputación de datos categóricos
 - ii. Problemas causados por la imputación de datos categóricos
 - iii. Solucionando los problemas de imputación de datos categóricos
 - iv. Colapsando categorías
 - 1. Colapsando con umbrales
 - 2. Utilizando Greenacres
 - 3. Utilizando el peso de la evidencia
 - v. Imputación de datos no categóricos
- 5. Reduciendo el problema dimensional de las variables
 - a. El problema de la redundancia
 - b. Conglomeración de variables
 - i. Entendiendo componentes principales
 - ii. Conglomeración de variables por métodos divisivos
 - iii. Seleccionando una variable representativa de cada conglomerado de variables
- 6. Revisando la interacción de las variables
 - a. Visualizando relaciones no lineales
 - b. Suavizando variables
 - i. Estimación Nadaraya Watson
 - ii. Distribuciones no paramétricas
 - c. Remediación de relaciones no lineales
- 7. Midiendo el ajuste del modelo
 - a. El ajuste versus la complejidad
 - b. Métricas de evaluación de modelos
 - i. Métricas de clasificación
 - ii. Métricas de ranking
 - c. Validación cruzada y muestras de validación posterior
 - d. Bootstrapping
- 8. Modelos de pronóstico temporal
 - a. Modelos SARIMA
 - b. Modelos GARCH y evaluación de la VaR y expected shortfall