



Afi

Escuela
de Finanzas

Regresión Avanzada

Máster en Data Science y Big Data en Finanzas

Javier Nogales – PhD Matemáticas
Catedrático, Estadística e IO, UC3M

fcojavier.nogales@uc3m.es 2022

Ejercicio GLM + Regresión Avanzada

- El Ejercicio Final consistirá en dos partes: GLM (40 %) y Regresión Avanzada (60 %)
- Se pueden usar dos bases de datos distintas para cada parte, o bien solo una
- De hecho, se podría usar la base de datos utilizada en el Ejercicio de Análisis Multivariante, siempre que dicha base disponga de variables target adecuadas para este Ejercicio (ver detalles a continuación)
- En todo caso, siempre se aconseja utilizar bases de datos nuevas para ganar experiencia y habilidades

GLM: 40 %

- **Objetivo:** desarrollar modelos GLM (que no sean LM) que expliquen una variable target en función de predictores
- La variable target debe seguir una distribución de la familia exponencial que no sea la normal
- Proporcionar *interpretación* y *conclusiones*. Esta parte es muy importante
- Evaluar capacidad predictiva (dividiendo muestra en train y test), considerando intervalos de predicción. Esta parte no es tan importante
- Subir al campus virtual: datos utilizados (o link), notebook (Rmd o Jupyter), y versión compilada (html)

Advanced Regression: 60 %

- **Objetivo:** desarrollar herramientas de regresión avanzada (no GLM) para predecir una variable target en función de predictores
- La base de datos de esta parte tiene que ser de *alta dimensión*, pudiendo considerar predictores ruidosos o no relevantes
- Usar las herramientas vistas en clase para obtener un modelo final que explique la variable target (usando toda la muestra) y otro que obtenga predicciones de dicha variable (para muestra test) junto con intervalos de predicción
- Se valora originalidad: p.e. considerar interacciones entre variables, reducción de la dimensionalidad del problema, extracción de features, etc.
- Se permite coger inspiración en otros análisis, citando siempre la fuente
- Subir al campus virtual: datos utilizados (o link), notebook (Rmd o Jupyter), y versión compilada (html)

Evaluación

- GLM preparación datos: 0.5 puntos
- GLM desarrollo e interpretación de modelos: 2 puntos
- GLM predicción: 0.5 puntos
- GLM originalidad y report: 1 punto
- Regresión Avanzada preparación datos: 1 punto
- Regresión Avanzada parte explicativa: 2 puntos
- Regresión Avanzada parte predictiva: 2 puntos
- Regresión Avanzada originalidad y report: 1 punto



Afi

Escuela
de Finanzas

© 2015 Afi Escuela de Finanzas. Todos los derechos reservados.