

Estadística avanzada

22/07/2023.

Universidad Internacional de Valencia.

Prof. Yérali Gandica.

AG1: Regresiones (20%)

El trabajo consiste en realizar la regresión más apropiada para tu variable dependiente en función de las variables independientes.

Instrucciones :

- Antes de la selección de variables, tu data debe tener entre 4 y 10 variables independientes. Si hay más puedes hacer ya una selección basado en el mismo estudio.
- La data debe tener al menos 300 puntos de registro, no hay máximo.
- He creado el foro: Base de datos para Actividad AG1. Ahí debes escribir tu data antes de empezar tu actividad.
- Deben entregar el informe en pdf, entre 5 y 7 páginas más un script con las instrucciones de R, de manera que yo las pueda ejecutar.
- La data la deben guardar en un repositorio, por ejemplo github (al cual yo debería tener acceso). Aprender a hacer ésto es parte de la evaluación.
- Dentro de las actividades está una primera parte describiendo los pasos que deben tomar en cuenta para el análisis.
- Por supuesto, el informe debe contener una breve descripción de la base de datos.
- Errores en el informe (como errores ortográficos) quitan puntos. Se va a evaluar como si fuera un informe en su lugar de trabajo.
- Dentro del análisis deben tomar en cuenta el material visto en clase: Diferentes tipos de regresiones: simples, multidimensionales, polinómicas, logísticas, etc. Colinealidad y penalización Ridge y/o Lasso.
- Si la variable dependiente no es binaria, pueden usar una de las variables binarias para intentar una regresión logística en función de otras variables. Ésto da 3 puntos extra.
- Una vez entregada la actividad yo podría preguntar a cada alumno y/o hacer tutorías a parte para chequear que la actividad fue bien hecha por esa persona. Si la persona no entiende algo de lo que hizo no tiene puntos.
- Cualquier consulta la deben hacer desde hoy sábado 22 al viernes 29. Después estaré de viaje.

Buen trabajo y mucho éxito!! _____