



Universidad Tecnológica Centroamericana

Facultad de Ingeniería

Sistemas Inteligentes

Tarea # 5

Docente: Dr. Kenny Davila

Presentado por:

11541261

Carlos Rivera

11641381

Oswaldo Varela

17 de diciembre del 2021

Parte 1: Regresión Lineal

1. ¿Cuáles son los valores de MSE y pesos del modelo de regresión por mínimos cuadrados usando tanto los datos originales como también los normalizados?

Datos No Normalizados

Metodo	Normalizado	MSE	rating
Min Squares	Si	0.08554974516304864	0.7641349479746832

fav_genre	cast	advertising	length
0.14861188775785772	1.5761372651388774	0.1512885005197766	-0.0007136140042831307

Datos Normalizados

Metodo	Normalizado	MSE	rating
Min Squares	Si	0.08554974516304864	0.7641349479746832

fav_genre	cast	advertising	length
0.044670271042315746	0.48863597703178685	0.18650020436844972	-0.025797317434800696

2. ¿Cuáles son los valores de MSE y pesos del modelo de regresión de Lasso usando tanto los datos originales como también los normalizados?

Datos No Normalizados

Metodo	Normalizado	MSE	rating
Lasso	No	0.4705226387582382	0.13503828469000162

fav_genre	cast	advertising	length
0	0	0	-0.0015018383171048924

Datos Normalizados

Metodo	Normalizado	MSE	rating
Lasso	Si	1	0

fav_genre	cast	advertising	length
0	0	0	0

3. ¿Qué valores Alpha utilizó en cada dataset para obtener los resultados reportados para Lasso?
4. ¿Cuál de los dos modelos funciona mejor a su criterio y por qué?
R// En este caso el que mejor funciona para estos modelos sería el método de regresión mínimo cuadrado ya que el mse tiene menos peso que en la regresión por lasso
5. En base a datos no normalizados, ¿Cómo ordenaría la importancia de cada uno de los 5 atributos en la decisión final?
R// cast, rating, Advertising, fav_genre, length
6. En base a datos normalizados, ¿Cómo ordenaría la importancia de cada uno de los 5 atributos en la decisión final? ¿hay cambios con respecto al punto anterior? ¿Por qué sí o porque no?
R// rating, cast, advertising, fav_genre, length,
Si hay cambios,
Esto se debe a que en los datos normalizados el peso de los atributos cambia

Parte 2: Regresión Logística

1. ¿Considera que la regresión Logística funcionó bien para este problema?
R// Al observar las métricas se podría decir que la regresión logística funcionó bien con los datos dados.
2. Reporte los pesos aprendidos para cada atributo por la regresión logística. ¿Qué atributos tienen mayor peso absoluto y que atributos tienen menor peso absoluto?
R//

24_7	alto_millaje	cambiar_llanta
0.35981890057660587	0.33600814133388246	1.4840428791393143
carro_rentado	carro_viejo	conductor_experimental
1.3712870920548874	-0.9352834549517844	-0.9318223128188399
conductor_joven	dispuesto_pagar	
-0.1217933090462349	1.677762998134567	
historia_accidentes	maneja_mucho	
-0.8654238381261847	0.8599110074843545	

Atributos con mayor peso

cambiar_llanta, carro_rentado y dispuesto_pagar

Atributos con menor peso

24_7, alto_millaje y carro_viejo

3. Sorpresas. ¿Son los pesos aprendidos consistentes con lo que usted esperaba ver?

R// Se esperaba que algunos atributos tuvieran un mayor peso de lo obtenido, como conductor experimentado e historial_accidentes

Parte 3: Random Forest