

Auxiliar Machine Learning - Semestre Otoño 2019

Junio, 2019

Pregunta 1

Un importante centro médico lo ha contactado para realizar un modelo predictivo para detectar si un paciente padece cáncer. Usted decide utilizar algoritmos de clasificación de machine learning. Para lo cual procede como sigue:

- a) Entrene dos algoritmos de clasificación (*Decision Tree* y *K-nearest neighbors*) usando el 70 % de los datos para ello.
- b) Compare los algoritmos utilizando el 30 % restante de los datos, empleando métricas como *Precision*, *Recall* y *F1-score*.

Pregunta 2

Dado su excelente trabajo, un famoso banco lo ha contactado con el propósito de segmentar a sus clientes en función de dimensiones como el balance, límite de crédito y pagos.

- a) Entrene dos algoritmos de clustering (*k-mean* y *clustering jerárquico*) con el 70 % de los clientes. Mencione cómo determinó el número de grupos a formar. ¿Le hacen sentido?
- b) Establezca en qué cluster se encontrarán el 30 %.
- c) Mencione las características que tienen los clusters que formó y qué pueden significar en palabras de alguien que conoce el negocio, pero no sabe de estadística.