Universidad de San Carlos de Guatemala

Escuela de estudios de postgrado

Maestría en ingeniería para la industria con especialización

en ciencias de la computación

Introducción a la minería de datos.

No carne: 999014035

Descripción

Imagen que contiene Texto

Descripción generada automáticamente

La variable objetivo, DEPARTAMENTO, se convierte en un factor, una codificación esencial para problemas de clasificación.

Los datos se reorganizan aleatoriamente (sample) para evitar sesgos en la partición.

Texto

Descripción generada automáticamente con confianza media

Se dividen los datos en un 80% para entrenamiento y 20% para prueba, utilizando índices seleccionados aleatoriamente.

Entrenamiento del modelo:

Variable objetivo: DEPARTAMENTO (departamento asociado al migrante).

Predictoras:

PEI3: Sexo del migrante.

PEI4: Edad del migrante.

PEI5: Año de emigración.

Hiperparámetros:

ntree = 40: Número de árboles en el bosque.

mtry = 3: Número de variables consideradas al dividir un nodo.

Texto

Descripción generada automáticamente con confianza media

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Se predice los datos para una persona con genero masculino, con 50 años de edad y que migro en 2004 y la predicción del modelo dio como resultado el número 13 de departamento que es el departamento de Huehuetenango según el diccionario de datos.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

**Repositorio**  
<https://github.com/Mruiz-99/apriori>

En la carepta Tarea extra – comodín

**Tarea por reemplazar**

Tarea 3