



Universidad San Carlos de Guatemala

Escuela de Estudios de Postgrado de Ingeniería

Introducción a la minería de datos

Estudiante: Jorge David Ambrocio Ventura

Carnet: 999013655

Análisis Random Forest

Se realiza el análisis del conjunto de datos Migración db, mediante el algoritmo random forest, para la predicción de departamento del migrante, la predicción de si el migrante es mayor de edad o no; finalmente se predice el sexo de la persona en función de la zona en que vive, la edad que tenía al irse y el año en que se fue.

Random forest 01

Predicción de departamento mediante las variables sexo de la persona, edad que tenía cuando se fue y en qué año se fue.

Se predice tres nuevos datos,:

- hombre que se fue a 18 años y del año 2015, predice que es del departamento Huehuetenango.
- mujer que se fue a los 25 años, que se fue en el año 2010, predice que es del departamento Guatemala.
- mujer que se fue a los 45 años, que se fue en el año 2020, predice que es del departamento Guatemala.

Random forest 02

Predicción de si es mayor mediante las variables sexo de la persona, departamento al que pertenece y el año en que se fue.

Se predice dos nuevos datos:

- Hombre de Guatemala que se fue en el año 2015, predice que es mayor.
- Mujer de Quetzaltenango que se fue en el año 2006, predice que es menor.

Random Forest 03

Predicción del sexo mediante las variables edad que tenía cuando se fue, año en que se fue y la zona geográfica.

Se predice dos nuevos datos:

- Hombre de 1 año que se fue en el año 2015 y residía en la zona 7, predice que es hombre.
- Mujer de 20 años que se fue en el año 2006 y residía en la zona 17, predice que es mujer.