

Actividad 3 (FORO)

En esta actividad usted deberá aportar con 1 respuesta principal a la o las preguntas enunciadas por el profesor y comentar y/o intervenir 2 respuestas de sus compañeros (en otros hilos de conversación).

Además, junto a la respuesta principal, usted deberá subir un informe del trabajo realizado, el cual también **será considerado en la evaluación del foro**. Se permite 1 solo archivo adjunto (tamaño máximo 800MB).

Considere los datos 'hr.csv', que contiene información de 14.999 empleados de una multinacional de EE.UU., generada con el objeto de estudiar las razones que explican la rotación de personal. Las variables que se encuentran en el conjunto de datos se describen en la siguiente tabla.

Variable	Descripción
Satisfaction_level	Nivel de satisfacción porcentual del empleado.
Last_evaluation	Puntaje porcentual obtenido en la última evaluación laboral.
Number_project	Número de proyectos en los que participa el empleado.
Average montly hours	Promedio de horas mensuales que trabajó el empleado.
Time spend company	Tiempo (en meses) de trabajo del empleado en la compañía.
Work_accident	Indica 1 si el empleado tuvo accidentes laborales y 0 si no.
Left	Indica 1 si el empleado dejó la compañía y 0 si no.
Promotion last 5 years	Indica 1 si el empleado fue ascendido en los últimos 5 años y 0 si no.
Area	Indica el área en la que se desempeña el empleador.
Salary	Indica el sueldo del empleado (medio, bajo o alto).

Realice las siguientes actividades:

- 1) Cargue el conjunto de datos en la sesión de trabajo de R usando la función `read.table`. Defina la variable `left` como factor, utilizando la función `factor()`.
- 2) Seleccione de manera aleatoria 2/3 de los datos para crear sus datos de entrenamiento y guarde el tercio restante para objeto de validación. Utilice la semilla 1, mediante la función `set.seed(1)`.
- 3) Construya un clasificador de la variable `left` del tipo Bagging, usando la función `bagging()` de la librería `adabag`. Utilice la semilla 1, mediante la función `set.seed(1)`. Utilice 10 árboles, mediante la opción `mfinal=10`.
- 4) Construya un clasificador de la variable `left` del tipo Boosting, usando la función `bagging()` de la librería `adabag`. Utilice la semilla 1, mediante la función `set.seed(1)`. Utilice 10 árboles, mediante la opción `mfinal=10`.
- 5) Construya un clasificador de la variable `left` del tipo Random Forest, usando la función `randomForest()` de la librería `randomForest`. Utilice la semilla 1, mediante la función `set.seed(1)`. Utilice 100 árboles, mediante la opción `ntree=100`.
- 6) Utilizando los datos de validación, calcule:

- a. La sensibilidad de cada procedimiento de clasificación.
- b. La especificidad de cada procedimiento de clasificación.
- c. La precisión de cada procedimiento de clasificación.

Interprete cada uno de los resultados obtenidos y compare los clasificadores compartiendo sus resultados en el foro. Utilice la función `predict()` para realizar las predicciones y puede utilizar la función `mmetric()`, de la librería `rminer` para calcular estos indicadores.
