

Ejercicio 4. Optimización de parámetros del modelo de árbol de decisión en RapidMiner

Referencia: Blog de Bala Despande

En este ejercicio se establece una competencia entre los equipos, en la cual cada uno procurará obtener un modelo de árbol de decisión ajustado (sintonizando los parámetros) para obtener los mejores resultados posibles en la predicción.

En la tarea “UT4-TA3 Ej4” se provee el dataset a ser utilizado: “Indian Liver Patient Dataset (ILPD).csv”, y el proceso de RapidMiner con los operadores actualizados.

1. Abrir y analizar cuidadosamente el tutorial “[How to choose optimal decision tree model parameters in Rapidminer](https://goo.gl/TZtY4D)” de Bala Despande, disponible en <https://goo.gl/TZtY4D>
2. Importar el dataset provisto.
3. Realizar las tareas de preparación de datos indicadas en el tutorial.
4. Analizar el funcionamiento del operador “Optimize”
5. Seleccionar los parámetros que se deseen optimizar, rangos, criterios, etc.
6. Ejecutar el proceso y luego obtener los resultados de rendimiento del modelo.
7. Listar parámetros seleccionados y valores obtenidos, así como los resultados de rendimiento. Subir a la tarea un archivo comprimido conteniendo un documento con los resultados y el proceso RapidMiner realizado.