

Ejercicio 3 - Evaluación

11. **Ejecuta el proceso**, y en los resultados, observa que RM ha generado un nuevo atributo, “prediction(DeportePrimario”.
12. **Observa** las estadísticas correspondientes, en detalles y con gráficos
 - a. ¿cómo son los valores de la predicción?
 - b. ¿cómo son los valores de confianza en cada caso?

Ejercicio 4 – Deploy

El Maestro ahora cuenta con las predicciones para sus clientes actuales.

Podrá eventualmente tomar cada una de ellas y hablar con sus clientes.

Desde RM es posible extraer esta información de diferentes maneras. Si se trata de un conjunto relativamente pequeño, podríamos simplemente copiarlo en una planilla electrónica.

1. Crea una nueva planilla electrónica
2. En la vista de Data de los resultados de predicción, selecciona todos los ejemplos (“select all – ctrl+A”) y cópialos (“ctrl-C”)
3. Pégalos en la planilla. Puedes ahora manejar cómodamente los datos y trabajar con ellos.
 - a. Ej: contar cuántas ocurrencias de cada deporte se han predicho
 - b. Dado un identificador (en este caso sólo tenemos el número de tupla generado por RM), devolver todo el ejemplo, con la predicción y los valores de todos los atributos correspondientes, para así informar al cliente
4. Eventualmente sería conveniente generar un identificador único de ejemplo / cliente, ya en el dataset de entrada de scoring, para luego tener una referencia de identificación inmutable (recuerda que se han eliminado varios ejemplos del dataset)

Ejercicio 5 – Análisis de la performance del modelo

Utilizando las técnicas de validación que ya conoces (prueba con varias)

1. Separa el conjunto de entrenamiento en train / test
2. Ejecuta el modelo y registra los resultados
3. ¿qué consideraciones te merece?
4. Intenta mejorar los resultados, aplicando técnicas de preparación de los datos
5. ¿tienen los atributos las características más apropiadas para este algoritmo?