**Práctica Clasificación**

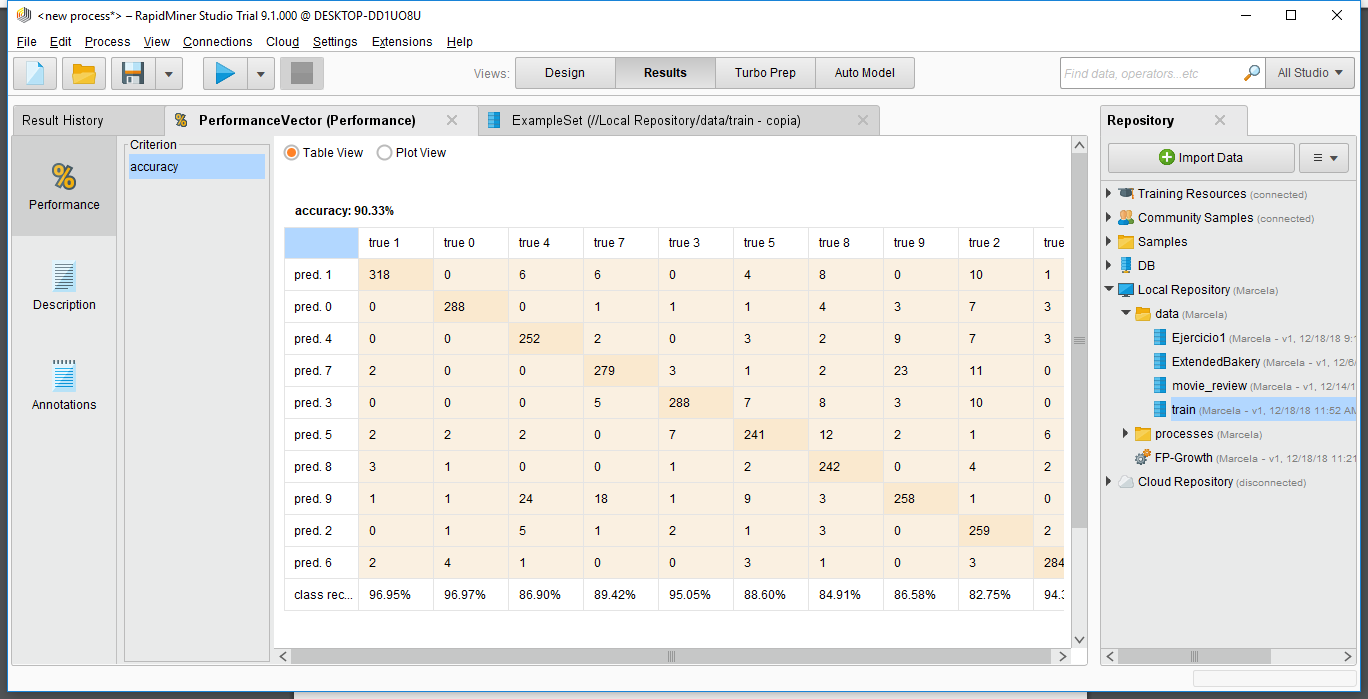
Presentar un análisis de la matriz de confusión obtenida luego de realizar la práctica, cómo interpreta los valores de recall y de precisión más altos y más bajos?

Trabajar con las primeras 10 000 filas del dataset train.csv que se encuentra disponible en: <https://www.kaggle.com/c/digit-recognizer/data>

Formatear las columnas, para indicar cual es la clase (primera columna), el rol de la primera columna será label y el tipo de la columna a polynominal.

Dentro del proceso:

1. Cargar el dataset
2. Normalizar las columnas: Normalize
3. Una vez normalizada las, procederemos a crear un training set y un testing set, para esto utilizaremos el bloque llamado Split It, el cual nos permite dividir nuestra base de datos en dos.
4. Y vemos en los parámetros y seleccionar Edit Enumeration.
5. Agregaremos una entrada, le pondremos 70% y 30% para el testing (0,7 y 0,3)
6. Una vez agregado nuestro Split Data, procederemos a agregar al operador y también a la métrica de evaluación.
7. El operador se llama K-NN, de vecinos cercanos (realizar la prueba con 1 y 4 vecinos cercanos).
8. Añadir el bloque Apply Model, el cual nos permite poder agregar este modelo en datos no antes vistos.
9. Conectamos la salida mod con la entrada mod y la salida par, que corresponde a nuestro testing con nuestra entrada de Apply Model.
10. Este operador lo que nos generará es una nueva base de datos, la misma que teníamos anteriormente con una columna extra, que es la columna predecida para cada una de estas filas.
11. Por último, agregaremos el bloque llamado Performance, el cual nos permitirá poder medir el performance de nuestro clasificador.
12. Conectamos la salida lab con la entrada lab y la salida per a la entrada res.
13. Una vez que tenemos todo conectado, podemos proceder a ejecutar el proceso.



Y la matriz de confusión:

