

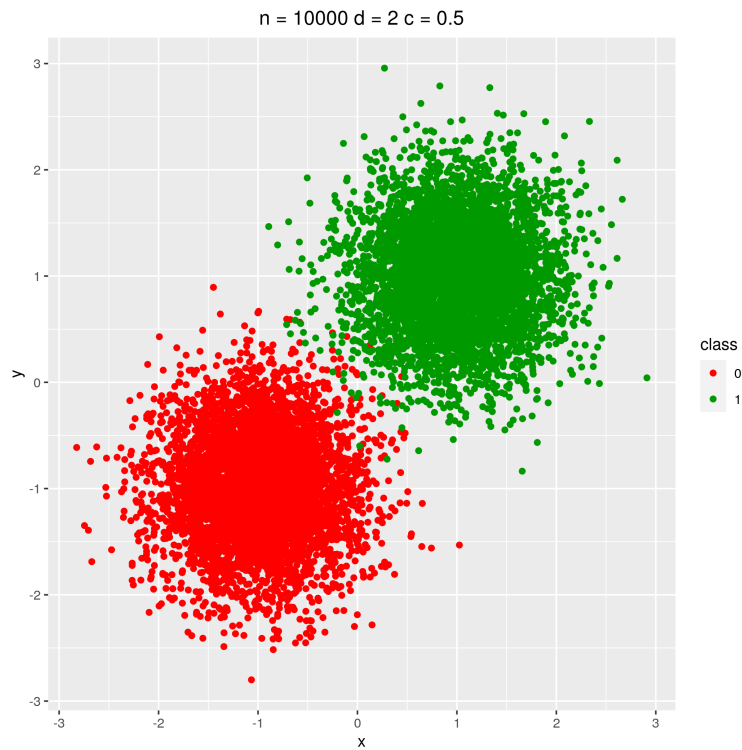
Práctica 1: Introducción a R

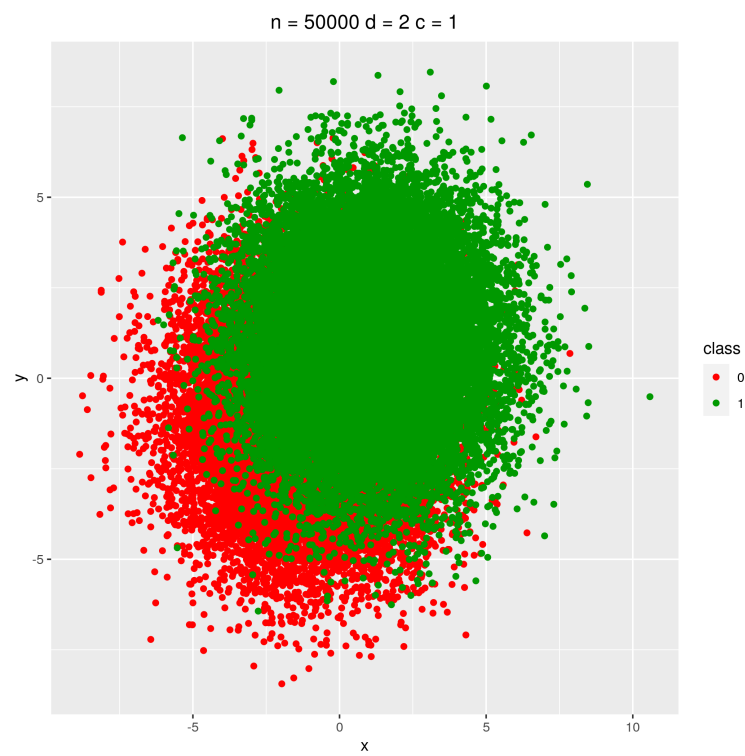
Alumno:Pablo Alonso

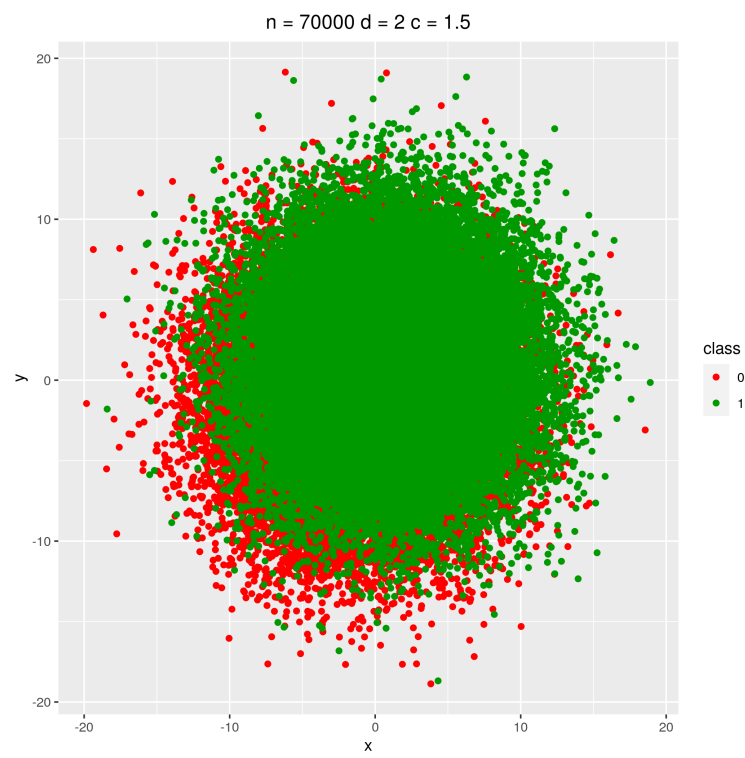
Para correr el script:

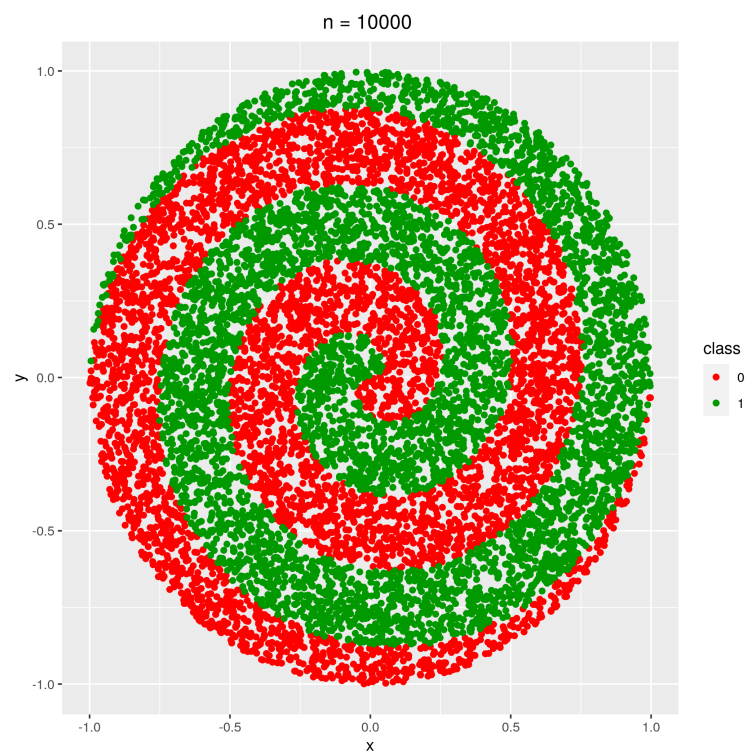
- `$chmod +x script.R`
- `$. /script.R`

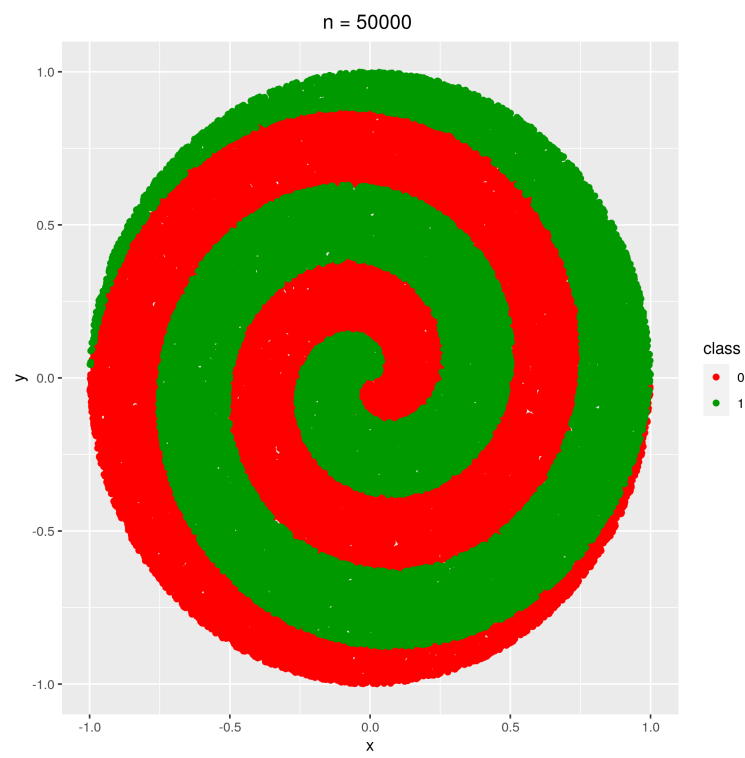
2)

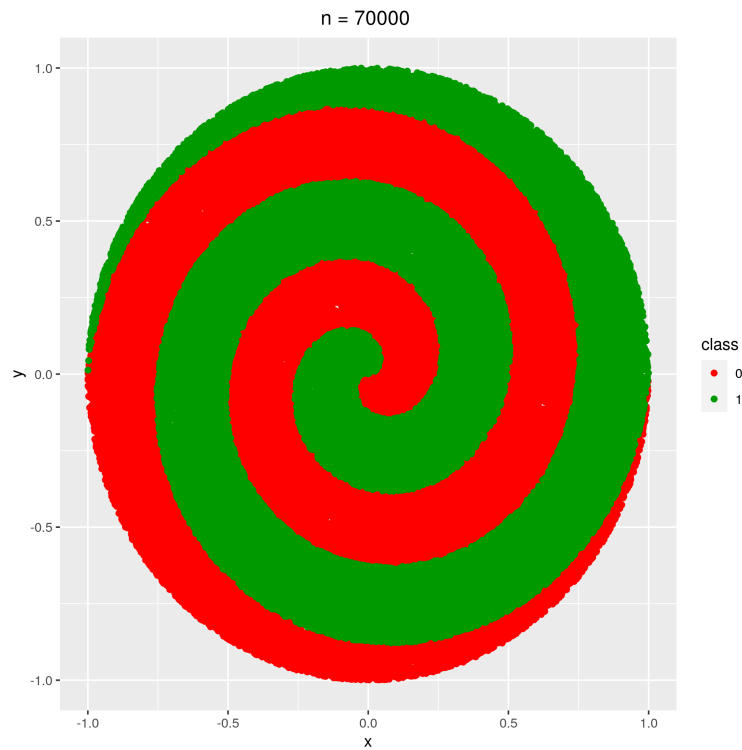












3)

Cross validation usa $4/5$ del training set para formar el modelo y $1/5$ para testear en cada iteración del 5 fold y ajustar los parámetros libres de cada problema (el k de KNN y el cp de decision tree), sin cross validation se usan el training set completo para formar el modelo y evaluar sobre test para cada valor de parámetro de una amplia gama que nosotros mismos definimos.

Finalmente se crean modelos usando los parámetros optimizados y se evalúa sobre test, teniendo mejores resultados con parámetros que fueron optimizados sobre un training set de mayor tamaño, es decir sin cross validation.

Por otro lado, el método local (knn) logra mejores resultados que el método global (decision tree) aunque también es más costoso porque debe computar todas las distancias euclídeas entre el training set y el punto que se quiere clasificar.