Introducción a la Ciencia de los Datos

Tarea 8

Nota: entregar las respuestas el próximo miércoles al correo de Daniela ([daniela.pena@alumnos.cide.edu](mailto:daniela.pena@alumnos.cide.edu)) en un archivo: nombre\_apellido\_tarea8. py

(i) A partir de la base de datos 'breast\_cancer' disponible en sklearn obtén los accuracy scores con kNN para valores de k entre 1 y 10, usando tanto al training set como al testing set como observaciones a evaluar (puntos en cuestión)

(ii) Almacena a estos valores en dos listas diferentes y haz es un mismo plot las gráficas de los valores de k -en el eje horizontal- contra las listas con los accuracy scores (training y testing) correspondientes. No se te olvide poner las etiquetas en los ejes y una leyenda para distinguir a cada tipo de score.

(iii) Como interpretas el resultado que observas en estas dos gráficas