

Documentación para el archivo **Proyecto7.py**, el código es un archivo de Python que funciona tanto en la versión 2.7 como 3.5, para su correcto funcionamiento es necesario tener instaladas las librerías de **matplotlib** y **numpy**.

El archivo consta de 3 funciones invocables y 3 de uso interno en conjunto ejecutan el algoritmo de K Means, las funciones, en el orden deseado de ejecución:

- **getDataFromFile( fileName , delimiter=" " ):**

- *fileName* es un string que contiene el *path* hacia el archivo con la información
- *delimiter* es el carácter usado para separar los datos, opcional

La función devuelve una matriz con los datos dentro del archivo, el archivo debe tener 2 columnas de números separadas por un " " (o cualquier otro carácter que se debe de pasar a la función), la matriz devuelta es de tipo *numpy* y sus dimensiones son Nx2

- **kMeansInitCentroids( X , K, choice=False ):**

- *X* es una matriz numpy de Nx2 (obtenida de la función anterior)
- *K* es un valor numérico y nos dice cuantos *clusters* serán inicializados
- *choice* es un valor booleano que nos dice que algoritmo sera usado, opcional

La función devuelve una matriz de Kx2 , donde cada renglón es un centroide inicial para esos datos, si *choice* es *True* los centroides regresados son elegidos de manera aleatoria de la matriz X, de otra forma los puntos son generados aleatoriamente considerando x's y y's criticas dentro de la matriz X.

- **runkMeans( X , centroids , max\_iters, plot=False ):**

- *X* es una matriz numpy de Nx2 (obtenida de la primer funcion)
- *centroids* es una matriz de Kx2 que contiene los centroides
- *max\_iters* es un contador que nos dice cuantas veces se debe ejecutar el algoritmo k means(movimientos de centroides)
- *plot* es un valor booleano que nos dice si debemos graficar o no los datos y los centroides

Devuelve una matriz de Kx2 que contiene los centroides mas óptimos partiendo de los *centroids* dados, el algoritmo invoca *max\_iters* veces las funciones *findClosestCentroids* y *computeCentroids* para calcular unos nuevos centroides. En caso de querer graficar invoca la función *graficarDatos*.