

Práctica 5 Introducción al Aprendizaje Automático

Jose Joaquín Rodríguez García y Araceli Teruel Domenech

Universidad Politécnica de Valencia

Master en Big Data Analytics

Octubre 2017

1. Primera Tarea

1.1. Enunciado

Estudiad el código MyKMeans.py para ver la implementación de lo que se ha explicado en clase. Probad dicha clase ejecutando el programa Python testing-k-means.py. Variad el número de clusters para ver la evolución en este ejemplo sintético. Probad el modo de inicialización de los clusters en base a Katsavounidis y observad las diferencias. ¿Qué tipo de inicialización es mejor?

1.2. Resolución

Tras hacer pruebas con la inicialización de los cluster en base a Katsavounidis y con una inicialización random y variando entre 15 y 25 el número de clusters, podemos observar que conseguimos mejores resultados inicializando los cluster en base a Katsavounidis obteniendo como mejor resultado

2. Segunda Tarea

2.1. Enunciado

Utilizad la implementación de MyKMeans.py para discretizar imágenes. Para ello primero debéis descargar el ejemplo de cuantificar imágenes de la siguiente URL: <http://scikit->

learn.org/stable/auto_examples/cluster/plot_color_quantization.html En este programa hay dos ejemplos de discretización de imágenes con objeto de reducir su tamaño. Bajando el número de colores distintos se consigue disminuir la cantidad de memoria necesaria para almacenar una imagen. También se consigue mediante interpolación para reducir la cantidad de píxeles, pero ese no es nuestro objetivo aquí. Se pide que añadáis un tercer ejemplo utilizando objetos de la clase `MyKMeans.py` y probad el efecto que tiene el tipo de inicialización `random` y la tipo `Katsavounidis`.

2.2. Resolución