

Optimization

K-Means

N Chakradhar¹ Havish²

¹EE16BTECH11022

²EE16BTECH11023

March 22, 2019

Introduction

- ▶ K-Means is an unsupervised learning algorithm used generally to cluster the given unlabelled data.
- ▶ Our given task was to classify the marks scored into 4 or 5 different classes where each class is the grade scored.
- ▶ Since we do not have pre-determined labels, we used K-Means

The algorithm

- ▶ K-Means algorithm calculates "K" centroids which represent each of the "K" classes
- ▶ The "K" centroids are initialised randomly(preferably points in the given dataset itself)
- ▶ The points in dataset are mapped to the closest centroid of the "K" centroids(distance is euclidean)
- ▶ Now after this mapping, new centroids are calculated and the same procedure is repeated iteratively until the movement in centroid locations falls below a certain small value(until convergence)