

Q530 Homework 3: K Nearest Neighbors

Due date: 26 September, midnight

1. Using MATLAB to implement KNN algorithms.
2. To test your implementation, you will need to randomly generate two data sets, each with 300 2-dimensional data points. Each data set is generated by 3 Gaussian distributions with certain means. You need to maintain the same means of 3 Gaussians but assign 2 sets of variances to the two data sets.
3. Writing a report to describe your implementation. You will also need to provide three query examples that the KNN method would not work based on your training data set. Show those examples visually by plotting both the training data and your query data. Based on your observation, you will need to describe how to improve the KNN.
4. (bonus) implementing the advanced version of KNN to fix the problem you observe above.