

Assignment-2

Jay Patel

December 7, 2016

1 Machine learning algorithms implementation

1.1 Supervised Learning

1. Kth Nearest Neighbor Considering one county for each entry, the algorithm is been implemented based on counties, where the distance is differentiated using the Euclidean formula for measuring distance between two objects having many attributed. The accuracy turns out to be around 80 percent. It takes K neighbours based on the distance formula and makes prediction for the query been made. The training the dataset is loaded and then the testing data set is used to predict the value of the queries in them. It is reflecting the democrat column values on the testing data set and not the value of accuracy of the prediction of the algorithm.
2. Perceptron Given the dataset remains the same throughout , the algorithm is been implemented accordingly. The weight vector is been randomly initialized for all of the attributes of the training data set, inorder to model the hyper plane. Introducing the errors during the iterations we make sure the hyperplane yields as realist and accurate predictions. The prediction in this implementation are not accurate because the some how the a proper hyper plane is not formed is which may lead to an accurate prediction. The procedure of the algorithm is the similar where the train data set is loaded and test with the values of the testing data set.

1.2 Clustering

1. K means The algorithm would define K centroids for each cluster and the centroids are to be placed as far as possible from each other. Then recalculate the values of the centroids and repeat the process untill the centroids do not move any longer.