1. Read data

|  |
| --- |
| #load module data  import data  #create a dataset object and read data from file sample.txt  sample = data.DataSet()  sample.read('sample.txt')  #view the data  sample.x  array([[ 5., 1., 1., ..., 3., 1., 1.],  [ 5., 4., 4., ..., 3., 2., 1.],  [ 3., 1., 1., ..., 3., 1., 1.],  ...,  [ 5., 10., 10., ..., 8., 10., 2.],  [ 4., 8., 6., ..., 10., 6., 1.],  [ 4., 8., 8., ..., 10., 4., 1.]])  #view class labels  sample.y[:10]  ['2', '2', '2', '2', '2', '4', '2', '2', '2', '2']  #view dataset dimension  sample.dim()  (699, 9)  #view features  sample.label  ['Clump Thickness', 'Uniformity of Cell Size', 'Uniformity of Cell Shape', 'Marginal Adhesion', 'Single Epithelial Cell Size', 'Bare Nuclei', 'Bland Chromatin', 'Normal Nucleoli', 'Mitoses']  #view subject ids  sample.key[:5]  ['1000025', '1002945', '1015425', '1016277', '1017023']  #create train dataset and test dataset using 1:10 hold out  train,test = data.holdOut(sample,0.1) |

2. Run kNN

|  |
| --- |
| #load kNN module  import kNN  model1 = kNN.train(train,4)  kNN.save(model1,’test’)  kNN.test(test1,model1)  model2 = kNN.load(‘test’)  kNN.test(test1,model2) |