HW 1 Text Classification and Naïve Bayes

for label in y:

index = np.where(classes == label)[0][0]

prior[index] += 1

prior /= sum(prior)

for feature, label in zip(x,y):

index = np.where(classes == label)[0][0]

likelihood[:, index] += feature

if self.smooth:

likelihood += 1

likelihood /= (np.sum(likelihood, axis=0) + self.smooth\_param \* n\_words)

else:

likelihood /= np.sum(likelihood, axis=0)



Accuracy on training set: 0.980000, on test set: 0.810000