PROBLEM --- Classifier with option of Rejection

In many pattern classification problems, one has the option either to assign the pattern to one of the C classes, or to reject it as being unrecognizable. If the cost for rejects is not too high, rejection may be a desirable action. Let:

$$\lambda_{ij} = \begin{cases} 0 & i = j \\ \lambda_r & i = c+1 \\ \lambda_s & otherwise \end{cases}$$

Where λ_r is the loss incurred for choosing the $(C + 1)^{th}$ action, rejection, and λ_s is the loss incurred for making a substitution error. Here, we assume the following values for the losses:

$$\lambda_r = 0.8, \lambda_s = 1$$

Modify the classifier that you designed in problem 3 to add the option of rejection.

PROBLEM --- Non-Parametric pdf Estimation & Related Classifiers

I.   Repeat Problem 3 (part I) with Parzen non-parametric estimate of pdfs. Study the effect of window size carefully, report the probability of classification error and correct classification rate. Consider two different windows: a. Rectangular and b. Gaussian. Compare the results for the windows.

II. Repeat problem 3 (part I) with k-nearest neighbor (k-NN) non-parametric estimate of pdfs. Study the effect of number of samples k. Report the probability of classification error and correct classification rate.

III. Design and Implement a k-nearest neighbor classifier. Report the correct classification rate for k= 1, 3, 5, 10.

IV. Design a minimum mean distance classifier. Report the correct classification rate.

PROBLEM --- Classifiers with using Scikit-Learn

I.   You have already implemented a number of different classifiers. Using pre-defined functions of scikit-learn package try to implement these classifiers:
   - KNN Classifier (KNeighborsClassifier)
   - Parzen non-parametric estimate of pdfs (RadiusNeighborsClassifier)
   - Gaussian Naive Bayes(GaussianNB)

6.2. Compare the classifiers of problems 3, 5 and 6 in terms of:
   a) Correct Classification Rate
   b) Confusion Matrix
   c) Confidence Matrix
   d) Required time for Training the algorithm
   e) Required time for testing the algorithm

Which classifier is your choice for given dataset? Explain why.