Homework 2:
Bayes Optimal Classifiers

Instructions:

1. All questions in this assignment require coding in python. Submit a jupyter notebook with both your code and your answers to the questions.
2. You may discuss this assignment with other students in the class, but you must submit your own answers to the questions below.
3. Include an honor pledge with your submission.
4. Submit online.
5. This homework is worth 100 points and the point totals for each question are shown in parentheses.

2. (30) The authors in [1] describe a marketing campaign by a bank in Portugal. Use the data (bank-full.csv) from this campaign with only the predictor variables age, balance, and duration and the response variable, y, to create Bayes optimal decision rules each of the following nine conditions:

   (a) Assume Gaussian class conditional likelihoods with unequal variance-covariance matrices with each of the following additional assumptions applied singularly to each decision rule in this class:

      i. Equal class priors and equal costs for misclassification;
      ii. The prior for not selecting the new bank service is 0.9 and misclassification costs are equal; and
      iii. The prior for not selecting the new bank service is 0.9 and the cost of classifying a customer as not a new service candidate when they are is 15 times the cost of classifying a customer as a new service customer

   (b) Assume Gaussian class conditional likelihoods with equal variance-covariance matrices with each of the following additional assumptions applied singularly to each decision rule in this class:

      i. Equal class priors and equal costs for misclassification;
      ii. The prior for not selecting the new bank service is 0.9 and misclassification costs are equal; and
      iii. The prior for not selecting the new bank service is 0.9 and the cost of classifying a customer as not a new service candidate when they are is 15 times the cost of classifying a customer as a new service customer

(c) Assume multivariate t distributions for the class conditional likelihoods with equal shape (precision) matrices with each of the following additional assumptions applied singularly to each decision rule in this class:

i. Equal class priors and equal costs for misclassification;

ii. The prior for not selecting the new bank service is 0.9 and misclassification costs are equal; and

iii. The prior for not selecting the new bank service is 0.9 and the cost of classifying a customer as not a new service candidate when they are is 15 times the cost of classifying a customer as a new service customer

Show your resulting decision rules for each of these cases.

References

[1] S. Moro, R. Laureano and P. Cortez. "'Using Data Mining for Bank Direct Marketing: An Application of the CRISP-DM Methodology." In P. Novais et al. (eds.), *Proceedings of the European Simulation and Modelling Conference* - ESM'2011, pp. 117—121, Guimares, Portugal, October, 2011.