Indian Institute of Information Technology Guwahati
CS640: Machine Learning

Lab 4: Bayesian Learning

Date: 09.02.2020 Total Marks: 20
Deadline: 12.02.2020

Implement the following questions in Matlab:

Q.1) Implement the Bayes minimum error classification on the given *breast cancer* data set (check the attached files). The training data set can be used to obtain relevant probability distribution to obtain a posteriori probability and evaluate the classifier using test data set. The file *Readme.txt* gives the information regarding the dataset. Each tuple contains multiple features, ie., feature vector $\mathbf{x} = (x_1, x_2, \ldots x_n)$ and

$$p(y|x_1, x_2, \ldots, x_n) = \frac{p(x_1|y)p(x_2|y)\ldots p(x_n|y)p(y)}{p(x_1)p(x_2)\ldots p(x_n)}$$

10 marks

Q.2) Implement Bayes minimum risk classifier using the same dataset. Use the following cost/loss while minimizing the risk:

$$\begin{aligned}
\lambda(\text{no-recurrence-events/recurrence-events}) &= 0.8 \\
\lambda(\text{recurrence-events/no-recurrence-events}) &= 0.2 \\
\lambda(\text{no-recurrence-events/no-recurrence-events}) &= 0 \\
\lambda(\text{recurrence-events/recurrence-events}) &= 0
\end{aligned}$$

Compare the result with bayes minimum error classifier and create confusion matrix for both classifiers to analyse the result.

10 marks