Final Exam STAT 715

Dheeman Saha

5/4/2018

Question

Is it possible to build an optimal classifier in the sense of Theorem 5.5.1. Justify your answer.

Answer

The Theorem 5.5.1 states that:

A classification rule X^* that minimizes the error rate γ^* is given by

$$X^*_{opt} = \begin{cases} 1 \text{ if } y \in C_1, \\ 2 \; if \; y \in C_2, \end{cases}$$

where

$$C_1 = \Big\{ y \in R^p : \pi_1 \; f_1(y) \geq \pi_2 \; f_2(y)$$

and

$$C_2 = \Big\{ y \in R^p : \pi_1 \; f_1(y) \; < \; \pi_2 \; f_2(y)$$

Based on the above equations it can be stated that the best classification procedure can be determined where the error rate in minimum. The same logic for the classification rule also works for the maximum posterior probability i.e the *Bayes Rule* when compared with the other list of variables. The theorem also stated that the optimal classifier is the one that has the minimum error rate.

In our case we are trying to to classify 13 different pens where each pen has 22 different samples. Then each sample consists of a data set which has a vector size of 6200 (200 * 31).

Therefore, if we want to built an optimum based on the classification rule idea we can set up the our classifier in the following manner:

$$X^*_{opt} = \begin{cases} 1 \text{ if } y \in C_1, \\ 2 \; if \; y \in C_2, \\ 3 \text{ if } y \in C_3, \\ 4 \text{ if } y \in C_4, \\ . \\ . \\ . \\ 13 \text{ if } y \in C_{13}, \end{cases}$$

where

y is the sample number that belongs to that specific class

and

$$C_1 = \Big\{ y \in R^p : \pi_1 \ f_1(y) \geq \pi_2 \ f_2(y) \ and \ \pi_1 \ f_1(y) \geq \pi_3 \ f_3(y) \ .... \ \pi_1 \ f_1(y) \geq \pi_{13} \ f_{13}(y)$$

and

$$C_2 = \Big\{ y \in R^p : \pi_2 \ f_2(y) \geq \pi_1 \ f_1(y) \ and \ \pi_2 \ f_2(y) \geq \pi_3 \ f_3(y) \ .... \ \pi_2 \ f_2(y) \geq \pi_{13} \ f_{13}(y)$$

and

$$C_3 = \Big\{ y \in R^p : \pi_3 \ f_3(y) \geq \pi_1 \ f_1(y) \ and \ \pi_3 \ f_3(y) \geq \pi_2 \ f_2(y) \ .... \ \pi_3 \ f_3(y) \geq \pi_{13} \ f_{13}(y)$$

and

$$C_4 = \Big\{ y \in R^p : \pi_4 \ f_4(y) \geq \pi_1 \ f_1(y) \ and \ \pi_4 \ f_4(y) \geq \pi_4 \ f_4(y) \ .... \ \pi_4 \ f_4(y) \geq \pi_{13} \ f_{13}(y)$$

and

$$C_4 = \Big\{ y \in R^p : \pi_4 \ f_4(y) \geq \pi_1 \ f_1(y) \ and \ \pi_4 \ f_4(y) \geq \pi_2 \ f_2(y) \ .... \ \pi_4 \ f_4(y) \geq \pi_{13} \ f_{13}(y)$$

.

.

.

and

$$C_{13} = \Big\{ y \in R^p : \pi_{13} \ f_{13}(y) \geq \pi_{12} \ f_{12}(y) \ and \ \pi_{13} \ f_{13}(y) \geq \pi_{11} \ f_{11}(y) \ .... \ \pi_{13} \ f_{13}(y) \geq \pi_1 \ f_1(y)$$

In our case the prior probability $\pi_1 .....\pi_{13}$ are the score values s_n (as mentioned in the paper) are generated from the given data set. The $f_1 ......f_{13}$ are the Multivariate Normal Distribution Density Functions which generate the list of density values based on the prior scores. So, the value of $C_1 ....C_{13}$ reflect the value that is classifed from the outcome of $\pi_n f_n \geq \pi_m f_m$ where $n \neq m$. So the X^*_{opt} will be the one that generate the maximum posterior value.

Therefore, making use of the Theorem 5.5.1 we can generate the optimum classifer that can work with the logic we have executed in the paper.