

## HW 10

### Chong Sun

- (d) The linear discriminant scores of the point  $x_0 = [3.5 \ 1.75]'$  for groups Iris setosa, Iris versicolor and Iris virginica are given respectively by

$$\begin{aligned}d_1 &= 51.56, \\d_2 &= 23.77, \\d_3 &= -1.12.\end{aligned}$$

Thus we classify the point into group Iris setosa.

- (f) We have the APER= 0.047.

To find out the  $\hat{E}(\text{AER})$ , we use stratified 10-fold classification method described as follows.

- Randomly partitioned the original data set into 10 equal sized subsamples.
- Of the 10 subsamples, a single subsample is retained as the validation data for testing the model, and the remaining 9 subsamples are used as training data.
- The above process is then repeated 10 times (the folds), with each of the 10 subsamples used exactly once as the validation data.
- The 10 results from the folds can then be averaged to produce a single estimation.
- We then repeat the above cross-validation procedure 100 times, yielding 100 random partitions of the original sample. The 100 results are again averaged to produce the final estimate of error rate.

We have  $\hat{E}(\text{AER}) = 0.053$  as our final result.