

Assignment 2: Bayesian Decision Theory (I)

Submission: Thursday August 25th
Groups of maximum 2 students

Prof. Fabio A. González
Machine Learning - 2011-II
Maestría en Ing. de Sistemas y Computación

1. Download the dataset from the course website. The dataset is a text file with a number of data samples, one per line. Each line has the following structure:

$$x_i \ y_i \ C_i,$$

where $(x_i, y_i) \in \mathbb{R}^2$ and $C_i \in \{0, 1, 2\}$.

2. Use a portion of the dataset (80% of the samples) to estimate the parameters of a bivariate Gaussian distribution for each class.
3. Write a program that calculates the discriminant function for each class, taking into account the possibility of rejection with a cost λ and cost 1 for misclassification ([Alp04] Eq. (3.10)).
4. Draw the discriminant functions showing the boundary for each class and, implicitly, the rejection area.
5. Classify the rest of the dataset that was not used for training, using a classifier based on the discriminant functions. Evaluate the results.

References

- [Alp04] Alpaydin, E. 2004 Introduction to Machine Learning (Adaptive Computation and Machine Learning). The MIT Press.