

Assignment 2: Bayesian Decision Theory and Parametric Estimation

Submission: Friday May 3rd
Groups of maximum 3 students

Prof. Fabio A. González
Machine Learning - 2019-I
Maestría en Ing. de Sistemas y Computación

Load the Iris dataset from sklearn and create a dataset using only the Petal Width and Sepal Width attributes. Plot the resulting dataset assigning a different color to each class.

1. (2.0)
 - (a) Assume that each class is generated by a bivariate Gaussian distribution with the a covariance matrix $\Sigma = I\sigma$, where σ is a scalar, shared by all the classes. This is, the distribution for each class has a different mean but the same covariance matrix. Calculate the parameters of the probability distribution functions for the three classes.
 - (b) Write a different Python functions that calculate the discriminant function for each class.
 - (c) Draw a plot, where the regions corresponding to the different classes are shown with different colors. A region corresponding to a class is the set of points where the particular class discriminant function is maximum (decision regions, [Alp14] Sect. 3.4).
 - (d) The boundary between class regions must be a line. Calculate the equation of these lines clearly explaining the deduction process. Draw the lines along with the regions.
 - (e) What happens with the boundary lines if we change the prior probabilities of the classes? Illustrate with a graphical example.
2. (1.0) Repeat steps (a) to (c) from previous item, but this time:
 - (a) Write again Python functions that calculate the discriminant function for each class, but now, taking into account the possibility of rejection with a cost λ and cost 1 for misclassification ([?] Eq. (3.10)). Look for values of λ that produce a rejection region easily distinguishable from the other regions.
3. (2.0) Repeat the previous item, but this time:
 - (a) The covariance matrix could be an arbitrary matrix (not diagonal) and different for each class.
 - (b) Use only a portion of the dataset (80% of the samples) to estimate the parameters of the probability distribution functions of each class. Print the values of the parameters for each class.
 - (c) Classify the rest of the dataset that was not used for estimation (20%), using a classifier based on the discriminant functions. Evaluate the results using a confusion matrix.

The assignment must be submitted as a Jupyter notebook through the following Dropbox file request, before midnight of the deadline date. The file must be named as `ml-assign2-unalusername1-unalusername2-unalusername3.ipynb`, where `unalusername` is the user name assigned by the university (include the usernames of all the members of the group).

References

- [Alp14] [Alp14] Alpaydin, E. Introduction to Machine Learning, 3Ed. The MIT Press, 2014
- [DHS00] Duda, R. O., Hart, P. E., and Stork, D. G. 2000 Pattern Classification (2nd Edition).
Wiley-Interscience.