
PPGEE2249 – Aprendizado de Máquina
Prof. Daniel Guerreiro e Silva

Assignment 2

It is forbidden to use polynomial regression and Bayesian Classification libraries/toolboxes.

- 1) Assume you want to design a classifier which receives univariate data as input and must choose between class +1 or class -1. The data associated with class “+1” follow a gaussian density with mean 2 and unit variance. On the other hand, the data associated with class “-1” follow a uniform density between -2 and 2. The prior probabilities are $P(C_{+1}) = 0.6$ and $P(C_{-1}) = 0.4$.
 - a) Apply Bayes’ methodology to obtain the optimal classification model and show, using a diagram, the model’s architecture (with clear indication of the discriminant function(s)).
 - b) Calculate the model’s decision boundary. Analytically calculate the classifier error probability (the integrals can be numerically solved).
 - c) Recalculate the model’s decision boundary if the Maximum-Likelihood criterion is adopted, this time. Analytically calculate the classifier error probability (the integrals can be numerically solved).
 - d) Compare the performances of both methodologies.
- 2) **Do your own implementation** and train a univariate polynomial regression model for a given problem. Pick some dataset (suggestions: Kaggle, UCI Machine Learning Repository) and employ cross-validation to define the polynomial order that better suits your problem. Comment your results.