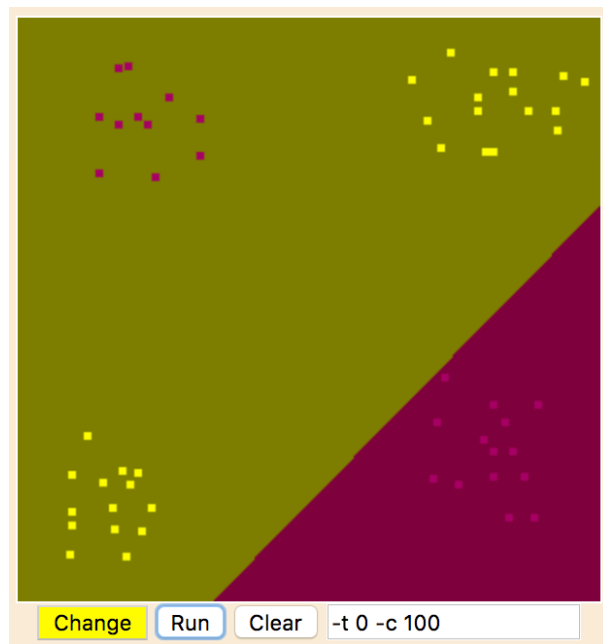


Machine Learning Assignment 5 report

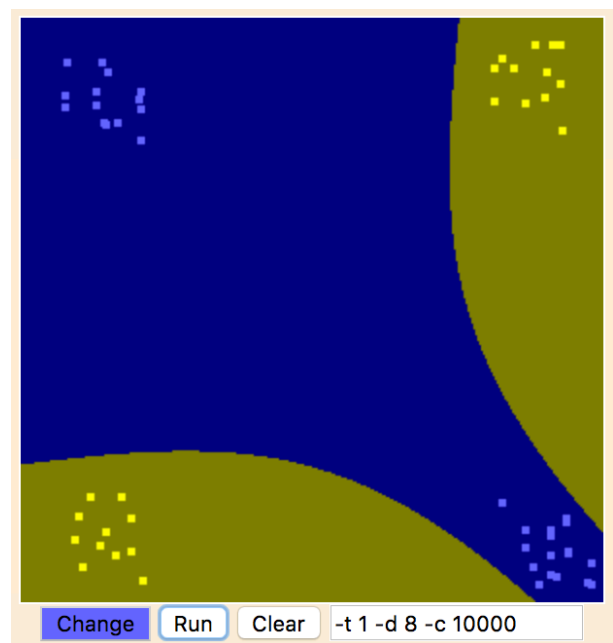
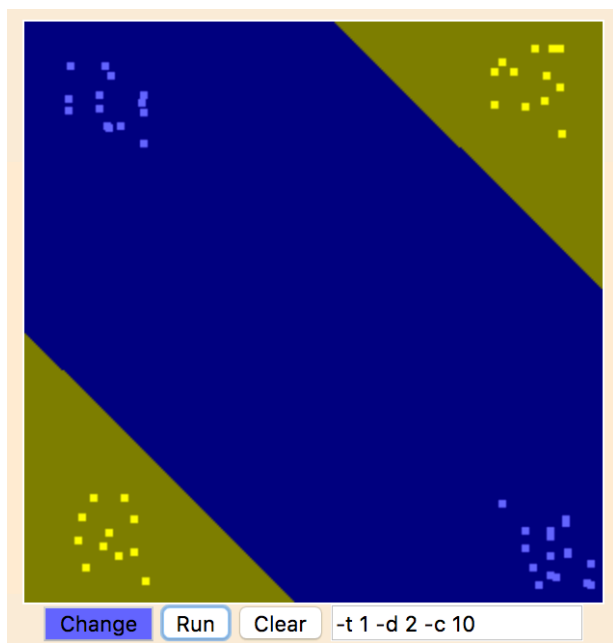
Massimo and Jiyan

Assignment 5.03

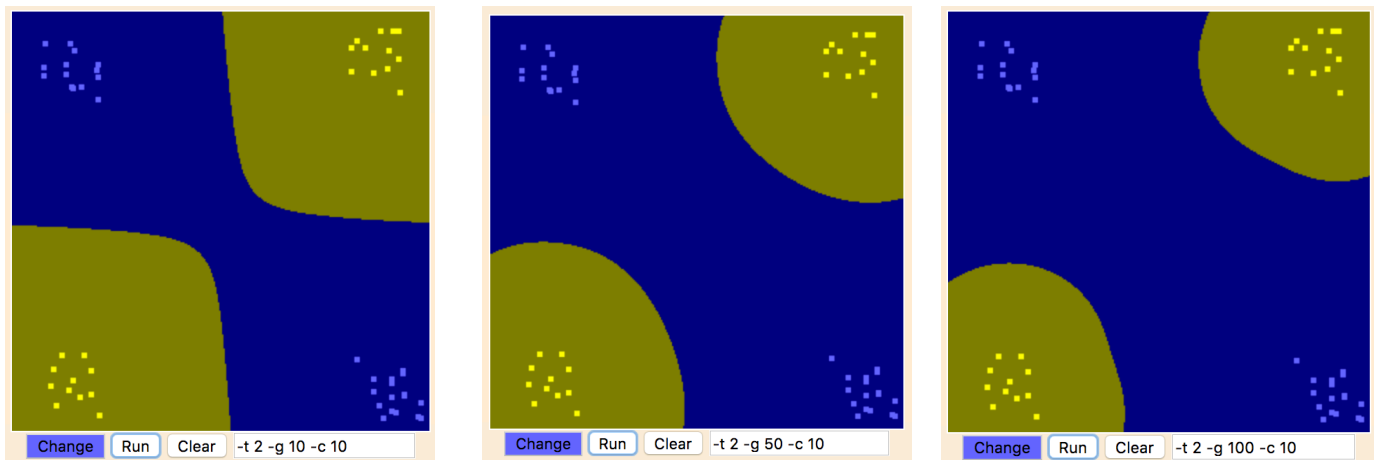
Part a - Linear kerne with C=100



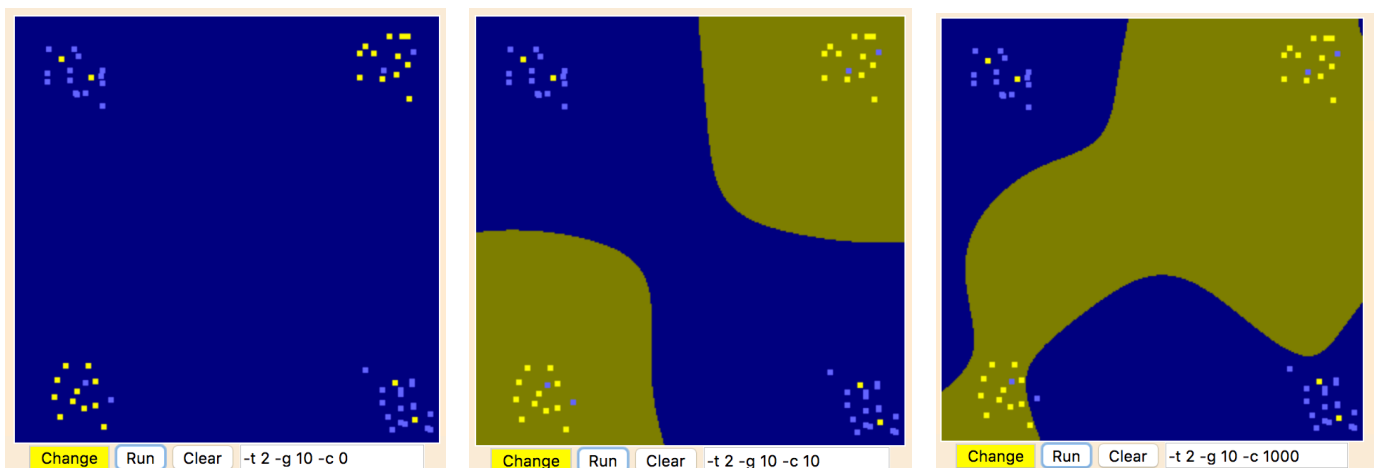
Part b - Polynomial kernel with degree 2 and 8. For degree 2 C=10 and for degree 8 C=10000



Part c - Gaussian kernel with sigma/beta = 10, 50, 100 and C=10



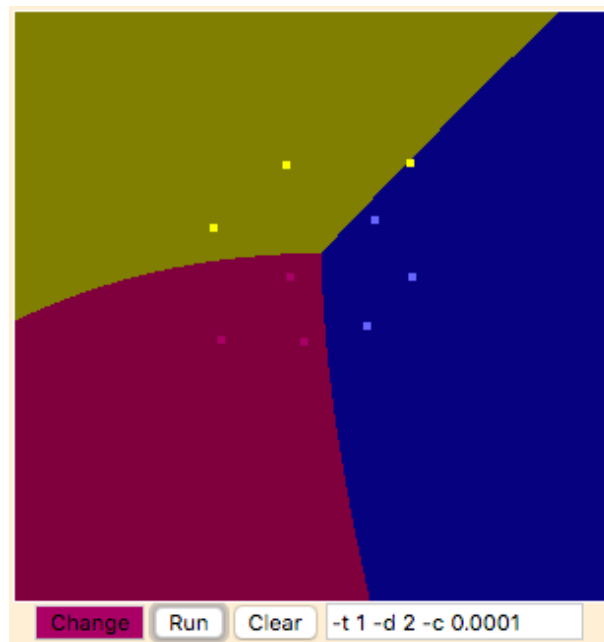
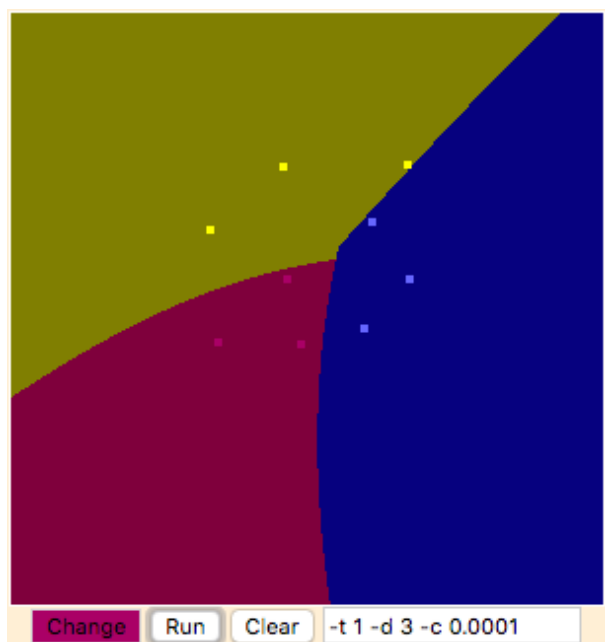
Part d - Gaussian kernel with noisy data, beta = 10 and C= 0, 10, 1000



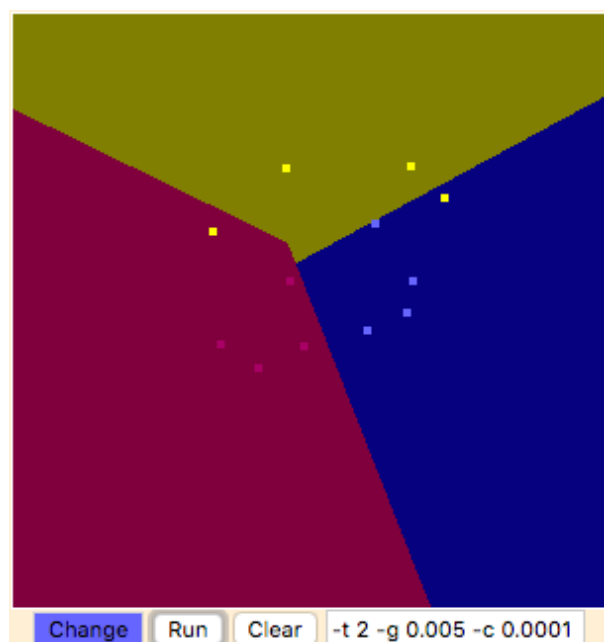
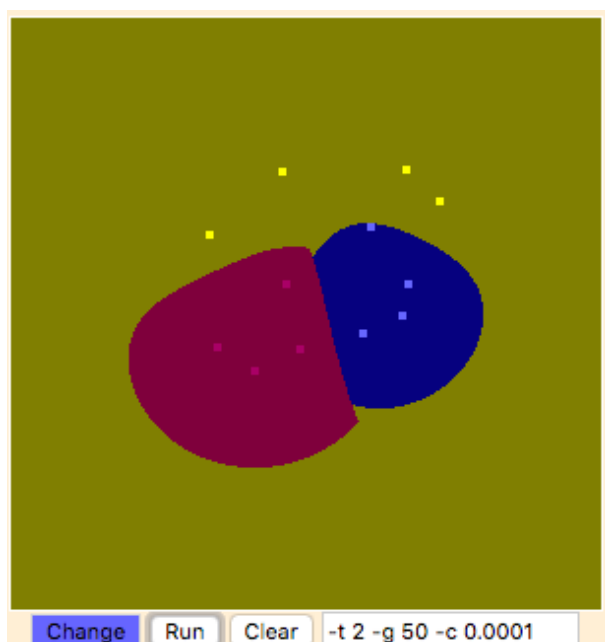
According to my observations the parameter C describes the amount of misclassification we are willing to allow when choosing an hyperplane. A small C means we are not willing to have misclassification and a large C means we don't mind. In fact we can see that from the graphs above where as the C increase we find possible hyperplanes classifying the data set however few points are misclassified.

Assignment 5.04

Part a - Data set separable with Fixed $C = 0.0001$ and polynomial kernel with degree = 3 but not with degree = 2.



Part b - Data set separable with fixed $C = 0.0001$ and radial basis function kernel with small σ but not with large σ .



Part c - Data set separable with fixed $C = 0.0001$ and radial basis function kernel but not with polynomial of degree 3.

