



Aprendizagem Automática Avançada (2º Ciclo de Informática)

SVM Exercises

The following exercises are based on the examples available at the following resource: <https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/>.

Problem 1

Run the examples provided in the resource linked above, in the case of applying a linear SVM for the iris dataset. Familiarize yourself with the code.

Problem 2

Analyze the influence of the **C** hyperparameter on the decision boundary.

Problem 3

Apply a polynomial kernel SVM to the iris dataset.

Discover how the solutions change as the degree values differs. Afterwards, analyze the impact of the **gamma** hyperparameter on the decision boundary.

Problem 4 *

If you remove a few support vectors and re-trained the SVM, what should happen to the decision boundary?

Verify your original assumptions by removing a few support vectors from the dataset, training a new SVM and comparing the obtained solution to the previous one.

Problem 5 *

Optimize the **C** and **gamma** hyperparameters using grid search with cross validation of 5 folds. How does the performance compare to base values? **Tip:** remember to split the dataset into a train and test split, so you can test on an isolated section of the data.

Problem 6 *

Do the practice problem from the resource by:

- finding the additional feature to have a hyper-plane that separates the classes;

- visualizing in 3D the separation of the data (since the actual point positions are not provided, use approximate coordinates for the blue star and red dot points);
- use the SVM classifier from scikit-learn, with different kernels, to solve the problem and compare it with the solution you had previously obtained.