Assignment 2: Kernels and Differentiable Programming

Submission: Monday June 30th 3 students per group

Prof. Fabio A. González

We will implement a kernel version of logistic regression. The goal is to train a logistic regression model on a feature space F. Specifically, the discriminant function of the model is given by:

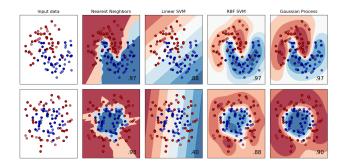
$$f(x) = P(C = 1|x) = \sigma(w\Phi(x)),$$

where $\Phi: X \to F$ is a mapping function associated with a kernel function $k: X \times X \to \mathbb{R}$ and σ is the logistic function.

Assume that the weight vector w is expressed as a linear combination of the training samples:

$$w = \sum_{i=1}^{\ell} \alpha_i \Phi(x_i)$$

- 1. Write a expression of the discriminant function expressed in terms of the kernel and the coefficients α_i .
- 2. Formulate the problem of learning the parameters of the model as an optimization problem that looks for the parameters α_i that minimize a cross entropy loss function.
- 3. Use Keras to implement a function that receives a training data set and a kernel function and finds a vector α that minimizes the loss function using gradient descent.
- 4. Test your algorithm using different kernels (linear, polynomial, Gaussian, etc.) on synthetic 2D datasets from sklearn (https://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html). Plot the decision regions and discuss the results:



5. Extend the model to deal with multi-class classification problems. Evaluate the model on the MNIST dataset.

The assignment must be submitted as a <u>Jupyter notebook</u> through the following <u>Dropbox file request</u>, 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). Do not submit additional files. Make sure that the notebook renders correctly and is free of errors before submitting.