9 Support Vector Machine - Soft Margin Classification

Exercise 9.1
Consider the following training dataset :

Class	x_1	x_2
+1	1	1
+1	2	2
+1	2	0
-1	0	0
-1	1	0
-1	0	1

1. Write formally the optimization problem when using the soft margin classification with logistic loss.
2. With the Jupyter notebook, code the gradient descent algorithm to compute the optimal separating line.
3. Plot the training set and the separating line. Comment the solution you obtain.
4. Add the new training samples $(\text{Class} = -1, x_1 = 0, x_2 = 2)$ and $(\text{Class} = +1, x_1 = 1, x_2 = -1)$. Train again the classifier and discuss the role of the parameter C.
5. Save the Jupyter notebook as a pdf file.
   You must install Pandoc as described on `http://pandoc.org/installing.html`

Exercise 9.2
The Jalon website contains the notebook file "plot_iris.ipynb".
This file is coming from `http://scikit-learn.org/`

1. Run the file and explain carefully each command.
2. Modify the file to keep only the classifier "lin_svc = svm.LinearSVC(...)". The other classifiers are ignored in the rest of the exercise.
3. Create your own Python function (based on the logistic relaxation) to compute the multiclass classifier equivalent to "lin_svc". You must use again the program you have coded in the previous exercise. The classifier will use the One-vs-All multiclass approach.
4. Comment the differences between "svm.LinearSVC(...)" and your own classifier. The comments should be based on numerical results and graphical illustrations.
5. Discuss the results of your classifier in fonction of C.
6. The classifier can make some classification errors. Define formally what can be called a classification error. Compute numerically all the classification errors.
7. Save the Jupyter notebook as a pdf file.