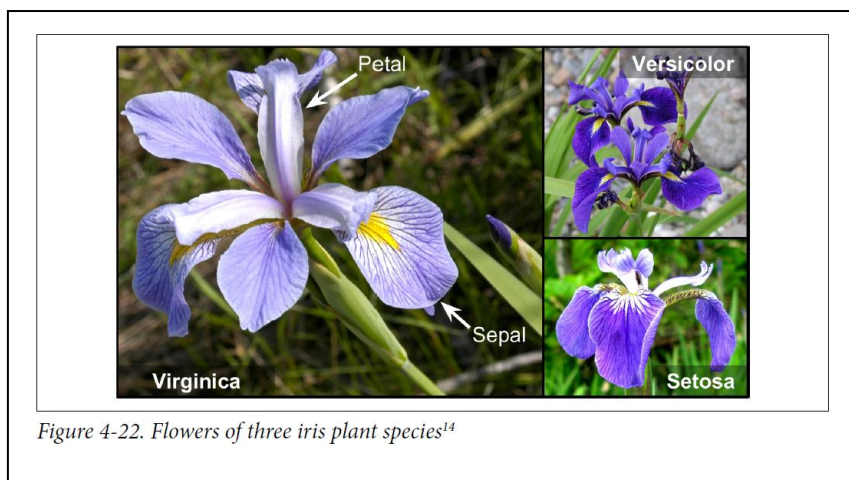| **COMPUTING SUBJECT:**	Machine Learning
TYPE:	WORK ASSIGNMENT
IDENTIFICATION:	Support Vector Machines Iris
COPYRIGHT:	*Michael Claudius*
DEGREE OF DIFFICULTY:	Medium
TIME CONSUMPTION:	1-3 hours
EXTENT:	< 150 lines
OBJECTIVE:	Basic understanding of binary classification. MNIST data set
COMMANDS:	

IDENTIFICATION: SVMIris/MICL

The Mission
To understand the idea behind Support Vector Machines and margin boundaries.

Precondition
You must have read the second part of Training Models in chapter 4.

The problem
Given a data set of 150 Iris flowers (X) and the label, the correct value (Y), you are to train a logistic regression classification both binary and multinomial with 3 classes. You are to use the Iris data set built in Sklearn.

Figure 4-22. Flowers of three iris plant species[14]

Useful links

https://matplotlib.org/3.1.0/tutorials/introductory/pyplot.html

Assignment 1: Download data set and program
As the data set is built in we don't need to download it.
For the program it is **easiest** to download a downsized version with only the necessary code for
SVM's from your teacher's home page. The files *"SVMIris.ipynb"* and "*SVMoon.ipynb*" are
encapsulated in .zip files.

Alternatively You could choose to from the Github to copy the Chapter 5 Jupyter program, "*05-
support-vector-machines.ipynb*", into the folder for your solutions (Machine Learning/Solutions)
folder. Rename it "*MySVMIris.ipynb*", before you cut away a lot of not–necessary code.

Assignment 2: Application program, adjusting the program
Start Jupyter and open the file. If you are using the the downloaded SVMIrisCh5.ipynb You don't
need to adjust the program, otherwise you will now have to delete and out-comment many lines/cells
like "moon-code" and some examples.

Now we can start to execute the cells.

Assignment 3: SVM classifier
Run the cells one by one and on the way discuss some of the topics and write down the answers to
the following questions:

   a. What is the idea behind SVM ?

   b. Can SVM also be used for regression ?

   c. Is it important to scale? Why ?

   d. Try to change some of the input parameters for predict- function.
       Do it in the following way:
       Add more cells and
            Use the petal data: [5, 2] as input parameters for predict- function
            Use the petal data: [5, 1.5] as input parameters for predict- function

   e. Just after the svm_clf.predict function (Cell[7])
       Add a new cell printing svm_clf.intercept_, svm_clf.coef_
       What one earth are these values ?

   f. How to declare binary LinearSVC ? (Show the code)

   g. How to declare a SVC with kernel ? (Show the code)

   h. Try to change the hyperparameter values for C and see the outcome.

   i. Go to the homepage for Sklearn.org and see the definition of LinearSVC and/or SVC
       constructor. There are really many parameters.
       Compare the parameters with the one used in the program code and explain 3 of them.

Tip:
https://scikit-
learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html

Assignment 4: Using septal data
Lets see how it looks using only septal data.
At the bottom of the program try to add cells using only SVC and change it to use septal-data.
So you copy and paste the cells and then change the data to be used and also the values of data to be predicted.
Tip: X = iris["data"][:, (0, 1)]
Tip: Use [5,2] and [7,2]
Remember to write down the probabilities.

Congratulation: You are now an expert-level 1 in SVM classification☺