Nombre: León Emiliano García Pérez

Matrícula: A00573074

Repeating the Tutorial with de Iris Dataset

In [23]:
```
# Define where you are running the code: colab or local
RunInColab          = False     # (False: no  | True: yes)

# If running in colab:
if RunInColab:
    # Mount your google drive in google colab
    from google.colab import drive
    drive.mount('/content/drive')

    # Find location
    #!pwd
    #!ls
    #!ls "/content/drive/My Drive/Colab Notebooks/MachineLearningWithPython/"

    # Define path del proyecto
    Ruta            = ""

else:
    # Define path del proyecto
    Ruta            = 'datasets'
```

In [24]:
```
# Import the packages that we will be using

import pandas as pd

# Dataset url

url = '/iris/iris.csv'

# Load the dataset

datos = pd.read_csv(Ruta+url)
datos.columns = ["PetalWidth", "PetalLength", "SepalWidth", "SepalLength", "Type"]
```

In [25]:
```
# Print the dataset

print(datos)
```

```
     PetalWidth  PetalLength  SepalWidth  SepalLength           Type
0           4.9          3.0         1.4          0.2    Iris-setosa
1           4.7          3.2         1.3          0.2    Iris-setosa
2           4.6          3.1         1.5          0.2    Iris-setosa
3           5.0          3.6         1.4          0.2    Iris-setosa
4           5.4          3.9         1.7          0.4    Iris-setosa
..          ...          ...         ...          ...            ...
144         6.7          3.0         5.2          2.3 Iris-virginica
145         6.3          2.5         5.0          1.9 Iris-virginica
146         6.5          3.0         5.2          2.0 Iris-virginica
147         6.2          3.4         5.4          2.3 Iris-virginica
148         5.9          3.0         5.1          1.8 Iris-virginica

[149 rows x 5 columns]
```

In [26]:
```
# Print the number of rows

print("Rows: ", datos.shape[0])
```

```
Rows:  149
```

In [27]:
```
# Print the number of columns

print("Columns: ", datos.shape[1])
```

```
Columns:  5
```

Activity: work with the iris dataset

1. Load the iris.csv file in your computer and understand the dataset
2. How many observations (rows) are in total?
3. How many variables (columns) are in total? What do they represent?
4. How many observations are for each type of flower?
5. What is the type of data for each variable?
6. What are the units of each variable?

In [28]:
```
Ruta            = 'datasets'

newUrl = '/iris/iris.csv'

newDatos = pd.read_csv(Ruta+newUrl) #Load the csv into a DataSet

print("Rows: ", newDatos.shape[0]) #Number of observations

print("Columns: ", newDatos.shape[1]) #Number of variables - They represent the dif

newDatos.columns = ["PetalWidth", "PetalLength", "SepalWidth", "SepalLength", "Type

#Obersavtions of each type of flower
print("Iris-setosa: ", newDatos['Type'].value_counts()['Iris-setosa']) #Count of th
print("Iris-versicolor: ", newDatos['Type'].value_counts()['Iris-versicolor']) #Cou
print("Iris-virginica: ", newDatos['Type'].value_counts()['Iris-virginica']) #Count

print(newDatos) #TypeOfData


print ("What are the units of each variable? Centimeters") #What are the units of e
```

```
Rows:  149
Columns:  5
Iris-setosa:  49
Iris-versicolor:  50
Iris-virginica:  50
     PetalWidth  PetalLength  SepalWidth  SepalLength            Type
0           4.9          3.0         1.4          0.2     Iris-setosa
1           4.7          3.2         1.3          0.2     Iris-setosa
2           4.6          3.1         1.5          0.2     Iris-setosa
3           5.0          3.6         1.4          0.2     Iris-setosa
4           5.4          3.9         1.7          0.4     Iris-setosa
..          ...          ...         ...          ...             ...
144         6.7          3.0         5.2          2.3  Iris-virginica
145         6.3          2.5         5.0          1.9  Iris-virginica
146         6.5          3.0         5.2          2.0  Iris-virginica
147         6.2          3.4         5.4          2.3  Iris-virginica
148         5.9          3.0         5.1          1.8  Iris-virginica

[149 rows x 5 columns]
What are the units of each variable? Centimeters
```

In [ ]: