Exp. 4: Using any data apply the concept of: logistic regression

Logistic regression aims to solve **classification** problems. It does this by predicting **categorical outcomes**, unlike linear regression that predicts a continuous outcome.

In the simplest case there are two outcomes, which is called binomial, an example of which is predicting if a tumor is malignant or benign. Other cases have more than two outcomes to classify, in this case it is called multinomial. A common example for multinomial logistic regression would be predicting the class of an iris flower between 3 different species.

Here we will be using basic logistic regression to predict a binomial variable. This means it has only two possible outcomes.

In Python we have modules that will do the work for us. Start by importing the **NumPy** module.

import numpy

Store the independent variables in X.

Store the dependent variable in y.

Below is a sample dataset:

#X represents the size of a tumor in centimeters.

X = numpy.array([3.78, 2.44, 2.09, 0.14, 1.72, 1.65, 4.92, 4.37, 4.96, 4.52, 3.69, 5.88]).reshape(-1,1)

#Note: X has to be reshaped into a column from a row for the LogisticRegression() function to work.

#y represents whether or not the tumor is cancerous (0 for "No", 1 for "Yes").

y = numpy.array([0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1])

We will use a method from the **sklearn module**, so we will have to import that module as well

from sklearn import linear\_model

From the sklearn module we will use the LogisticRegression() method to create a logistic regression object.

This object has a method called fit() that takes the independent and dependent values as parameters and fills the regression object with data that describes the relationship:

logr = linear\_model.LogisticRegression()

logr.fit(X,y)

Now we have a logistic regression object that is ready to whether a tumor is cancerous based on the tumor size:

#predict if tumor is cancerous where the size is 3.46mm:

predicted = logr.predict(numpy.array([3.46]).reshape(-1,1))

Example[Get your own Python Server](https://www.w3schools.com/spaces/)

See the whole example in action:

import numpy  
from sklearn import linear\_model  
  
#Reshaped for Logistic function.  
X = numpy.array([3.78, 2.44, 2.09, 0.14, 1.72, 1.65, 4.92, 4.37, 4.96, 4.52, 3.69, 5.88]).reshape(-1,1)  
y = numpy.array([0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1])  
  
logr = linear\_model.LogisticRegression()  
logr.fit(X,y)  
  
#predict if tumor is cancerous where the size is 3.46mm:  
predicted = logr.predict(numpy.array([3.46]).reshape(-1,1))  
print(predicted)

Result

[0]

We have predicted that a tumor with a size of 3.46mm will not be cancerous.