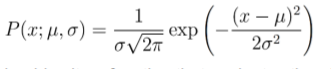
Implementation of the Naïve Bayes algorithm using python *(***gaussian distribution**)

**We will suppose that the features in each class follows a gaussian distribution. In other words, we have to calculate the mean and standard deviation of each feature in each class and then the class conditional probabilities of each feature in each class is calculated using the following equation**:



**we will use the iris data: 4 feutures and 3 classes**

*1. sepal length in cm  
2. sepal width in cm  
3. petal length in cm  
4. petal width in cm  
5. class:  
-- Iris Setosa -> class label 0  
-- Iris Versicolour -> class label 1  
-- Iris Virginica -> class label 2*

**We will code each class label to a number : for more flexibility in calcualtion w hek 😊: class 1 -> 0 / class 2-> 1/ class3->2**

**Steps:**

**1-encode the class label to numbers 0,1,2**

**2-function that split the data to 50% trainig set and 50% test set**

**3-group the trainig set under each class :**

**Ex: { 0: [p1,p2,p3] , 1: [p10,p11,p12] , 2: [p20,p21,p21] } /Pi:[f1,f2,f3,f4,ci]**

**4-write 2 function that claculate means and standard deviation**

**5-Create a function that take this :**

**{ 0: [p1,p2,p3] , 1: [p10,p11,p12] , 2: [p20,p21,p21] } Pi:[f1,f2,f3,f4,ci]**

**and return for each class the mean and standar devation of each feutures in each class :**

**ex: return like this: 0 ,1 and 2 are the classes, 0 :** μ1 is the the mean of f1 in class 0 and so on.

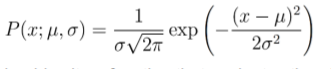
**INFO:**

**{ 0 : [ (***μ1,* σ1**), (***μ2,* σ2**), (***μ3,* σ3**), (***μ4,* σ4**), ] ,**

**1 : [ (***μ1,* σ1**), (***μ2,* σ2**), (***μ3,* σ3**), (***μ4,* σ4**), ] ,**

**2: [ (***μ1,* σ1**), (***μ2,* σ2**), (***μ3,* σ3**), (***μ4,* σ4**), ] , }**

**6-Create a function that take a ‘X ,** *μ ,*σ**’ and return the value of following eq:**



**7-Now make a fct that take INFO and a test pattern (ex : [f1,f2,f3,f4,ci]) and return the class probability of this test pattern in each class**

**Ex : return like this : for a test pattern :**

**{0:res 1 , 1: res 2, 2:res 3} , res *i* is the value of the equation above**

**8-Make a fnction PREDICT that take Info and a test pattern and use the step 7 to return the predicted label of this test pattern**

**9-Now make a function that take test set (more than one test pattern) and the info and return a lis of predicition of each test patttern in wichclass.**

**10-finaly get the accurency rate: function that take a predictions and the real test set from the dataset and get how many prediction was right and get the pourcentage**