Cairo University
Faculty of Engineering
Computer Engineering Department

CMPN450
Fall 2018

Pattern Recognition and Neural Networks.

Lab 3 – Bayesian Classifier

Given files "***data1.csv***", "***data2.csv***", "***test_data.csv***". The first two files, contain list of points and their corresponding classes. You will use the first file for the first problem, and the second file for the second problem. The test data file contains test points that are unlabeled (i.e. each point is not known to which class it belongs). The format of data files is shown as in Table 1.

Table 1 A sample of data1.csv

	Class	Feature 1	Feature 2
Point #1	1	0.271633	-2.93224
Point #2	1	7.020786	-1.98966
Point #3	1	2.901827	-0.91291

Requirement 1:

1.	Read data from the file **data1.csv**. How many features are there in this dataset? How many classes span the points?
2.	Given that the points of each class are normally distributed (i.e. the distribution is Gaussian). Find the **parameters** of the Gaussian distribution for **each class** (**μ,**Σ).
3.	Apply **Bayesian Classifier** to classify the test points found in ***test_data.csv***. *What parameters and probabilities do you need to compute first in order to apply the Bayesian classifier?*
4.	Compute the accuracy of your classifier by comparing against the actual classes found in "***test_data_true.csv*"**
4.	Plot the probability distribution of the three classes in one 3-D plot. *You should expect to see three Gaussian surfaces. How can you judge your plot is correct?*

Note: You are can use file "**lab3-1.py**". It contains starting code for this requirement.

Requirement 2:

1.	Read data from the file **data2.csv**. How many features are there in this dataset? How many classes span the points?
2.	Given that the points of each class are normally distributed (i.e. the distribution is Gaussian). Find the **parameters** of the Gaussian distribution for **each class** (**μ,**Σ).
3.	Compute the coefficients of the decision boundary equation. *What parameters are needed to compute the coefficients of the decision boundary equation?*
4.	Plot the decision boundary plane together with the probability distributions of the given classes in one 3-D plot.

Note: You are can use file "**lab3-2.py**". It contains starting code for this requirement.