**KECE407: Pattern Recognition**

**School of Electrical Engineering, KOREA UNIVERSITY**

**(Homework #1) (due via Blackboard, 5 pm, Wed 13 Oct 2021)**

**Report containing the code, results, discussions**

In this assignment, you will write a program to learn a naive Bayes classifier and use it to predict class labels of test data. The learned classifier should be tested on test instances with unknown class labels, and the predicted class labels for the test instances should be printed as output.

Please answer the following questions.

1. **(Data visualization)** Load the attached data using Matlab or Python and plot each class in a different color. In this case, the first feature is called length, and the second feature is called weight.
2. **(Data statistic)** Calculate the mean and covariance of each class and write them down.
3. **(Make predictions)**
   1. Load the test file and calculate the likelihood for each class using the mean and covariance calculated in Problem 2.
   2. Classify which class each sample of the test file belongs to and give the explanations.
   3. Perform the above process only for the first feature (length).
   4. Perform the above process only for the second feature (weight).
   5. Explain why length or weight is the better feature for classification.

T.A email: kdko@ korea.ac.kr