

# BIL712 Machine Learning

Fall 2016

## Programming Assignment 2 – Naïve Bayes Classifier

**Due Date: December 16, 2016**

In this programming assignment, you are going to write a text categorization program using Naïve Bayes Classifier. In the given corpus ([ProgAssg02.rar](#)), there are 50 documents for each category. There are five categories: *ekonomi*, *kadin*, *otomobil*, *saglik*, *teknik*. Your program should use the first 40 documents of each category for training, and their last 10 documents for testing. This means that you have 200 documents for training and 50 documents for testing. After your program learns from training set, it should try to categorize each document in the test set. Your program should print the precision result for the testing set.

### You should hand-in:

- Use Java, C or C++ to implement your homework.
- Put the following files into a .rar file (named as `yourname.rar`) and you will send it use an online submission system.
  - Your source program files
  - Executable of your program and a message indicating how to run your executable
- You will use the online submission system (<https://submit.cs.hacettepe.edu.tr/>) to submit your projects. You should use your cs email accounts (with domain `cs.hacettepe.edu.tr`). You should write your student id with lowercase (i.e. `n12000000` or `o120000000`) to login into the system. If you have any problem with the online submission system, you may contact with the assistant Ali Osman Serhatoğlu ([aoserhatoglu@cs.hacettepe.edu.tr](mailto:aoserhatoglu@cs.hacettepe.edu.tr)).