MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science
6.806/6.864 Advanced Natural Language Processing
Fall 2017

Assignment 0, Due: 9am on Thursday Sep 14. Upload pdf to Stellar.

Goal: The aim of this assignment is three fold:

1. Install and try Machine Learning libraries that we will use in the class
2. Familiarize yourself with the class policies
3. Acquaint yourself with the submission portal through a demo submission

Setup: Given specified Python libraries, install them on your local machine and try them on toy examples.

Software: All installations, as stated, will be for Python.

Problem 1 [1pt] : Installations

1. **Scipy stack and Numpy:** You can install numpy for Python via `https://www.scipy.org/install.html`.
2. **Scikit-learn:** Install scikit-learn for Python via `http://scikit-learn.org/stable/install.html`. Run test_1.py. This trains a news categorization model using a Multinomial Naive Bayes classifier. *Run the same model with 50% of the training data and state the performance. This modified code will be a part of your submission.*
3. **PyTorch:** You can install PyTorch for Python via `http://pytorch.org/`. (Select **None** for CUDA on your local machines). Run test_2.py, this is a modification of `http://pytorch.org/tutorials/intermediate/char_rnn_classification_tutorial.html`. This runs a neural network model that classifies names according to their culture of origin. *Run the model and observe the performance. Increase the dimensions of the hidden layer (n_hidden) from 128 to 200 and add a Linear layer on top of the output just before it gets fed to the softmax layer. State the new performance. This modified code will be a part of your submission.*

Problem 2 [1pt] : Policies

The following questions will help you familiarize with the class policies. Answers to these will be discussed in class and are present on stellar.

1. **Late submission policy:**
   (a) Will there be any partial credit for submitting an assignment 1 day late ?
   (b) How will your grades be re-arranged if you have an S^3 note for submitting a particular assignment late?
2. **Lecture notes:**
   (a) Is there any text book specified for the course ?

(b) Will there be exact notes provided as per the flow of each class ?

Problem 3: Submission Details

1. Create and submit a directory with the report and the code.
2. **Answers:** report should be called Answers.pdf and should contain all the answers to the questions.
3. **Code:** There should be a main.py present in your solution directory. On running the command "main.py –test_1" it should run test_1.py. On running the command "main.py –test_2" it should run test_2.py.
4. Please don't submit the assignment pdf as part of the solutions.