CPE 372/641 Natural Language Processing

**Individual Homework 3: Sentence Classification**

*Due March 28, 2020 via Piazza*

Let’s try implementing some sentence classification from scratch. Below you will find the rough code that uses naïve bayes classifier from TextBlob library (built on top of NLTK). (Rough code means that you might need to figure out how to make this work by adding something, like libraries, etc. Please try to find out yourself, it is part of the learning)

**What to do:**

1. Download everything your need for this simple program (NLTK, TextBlob).
2. Type the code below. Run and test the test corpus. Keep the results of 5 sentences in test corpus.
3. Expand the training corpus by add 5 more of class A sentences and 5 more of class B sentences (now you have to think what class A and class B mean) as well as adding 5 new sentences in the testing corpus
4. Try training and testing with your new test sentences.
5. Analyze the model accuracy (before and after expanding training corpus). Discuss.

**Deliverable: One pdf file that contains**

1. Source code that includes

* your expanded training and testing sentence
* the results of classifying your testing sentences (5 old and 5 newly added)

1. Short write-up that explains model accuracy and discussion

```

from textblob.classifiers import NaiveBayesClassifier as NBC

from textblob import TextBlob

training\_corpus = [

('I am exhausted of this work.', 'Class\_B'),

("I can't cooperate with this", 'Class\_B'),

('He is my badest enemy!', 'Class\_B'),

('My management is poor.', 'Class\_B'),

('I love this burger.', 'Class\_A'),

('This is an brilliant place!', 'Class\_A'),

('I feel very good about these dates.', 'Class\_A'),

('This is my best work.', 'Class\_A'),

("What an awesome view", 'Class\_A'),

('I do not like this dish', 'Class\_B')]

test\_corpus = [

("I am not feeling well today.", 'Class\_B'),

("I feel brilliant!", 'Class\_A'),

('Gary is a friend of mine.', 'Class\_A'),

("I can't believe I'm doing this.", 'Class\_B'),

('The date was good.', 'Class\_A'), ('I do not enjoy my job', 'Class\_B')]

model = NBC(training\_corpus)

print(model.classify("Their codes are amazing."))

>>> "Class\_A"

print(model.classify("I don't like their computer."))

>>> "Class\_B"

print(model.accuracy(test\_corpus))

>>> 0.83

```