# LING 575K HW4

Due 11PM on Apr 30, 2021

In this assignment, you will

- Develop your understanding of feed-forward networks for text classification
- Implement the Deep Averaging Network
- Understand a more sophisticated optimizer than vanilla SGD
- Explore some regularization techniques

## 1 Implementing the Deep Averaging Network + Loss

**Q1: Implement DeepAveragingNetwork.forward**

## 2 Implementing Adagrad

## 3 Running the Deep Averaging Network

In `run.py`, you will find a basic training loop for building a DAN and training it on the Stanford Sentiment Treebank. There are several command-line arguments specifying different arguments. The default arguments for various hyper-parameters are:

- Hidden dimension: 100
- Embedding dimension: 100
- Batch size: 32
- Number of epochs: 8
- Word dropout: 0.0 [i.e. no word dropout is applied]
- $L_2$ regularization: 0.0 [i.e. no regularization is applied]

Each run will print to stdout (and, therefore, the output file you specify in your condor job script) the training and dev set loss for each epoch, and then the final model's accuracy on the dev set.

**Q1: Run 1, default arguments** Run the main training loop by calling `run.py` with all of the default arguments. Record the outputs of this run (per epoch train/dev loss, final dev accuracy) here:

In 2-3 sentences, describe any trends that you see in the training and dev set losses over the course of training, and any differences between the two. What do these trends suggest to you?

## 4 Testing your code

In the dropbox folder for this assignment, we will include a file `test_all.py` with a few very simple unit tests for the methods that you need to implement. You can verify that your code passes the tests by running `pytest` from your code's directory, with the course's conda environment activated.

## Submission Instructions

In your submission, include the following:

- readme.(txt|pdf) that includes your answers to §1 as well as Q3 of §2.
- `hw3.tar.gz` containing:
  - run_hw3.sh. This should contain the code for activating the conda environment and your two commands for Q3 above. You can use run_hw2.sh from the previous assignment as a template.
  - word2vec.py
  - ops.py