

# LING 575K HW5

Due 11PM on May 6, 2021

In this assignment, you will

- Develop understanding of a feed-forward neural language model
- Implement components of data processing and text generation
- Implement key pieces of the model architecture

All files referenced herein may be found in `/dropbox/20-21/575k/hw5/` on patas.

## 1 Understanding the Feed-Forward Language Model

**Q1:**

**Q2: `tanh`** The model uses the hyperbolic tangent (`tanh`) activation function, defined as:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- Show that  $\tanh(x) = 2\sigma(2x) - 1$ , where  $\sigma(x)$  is the sigmoid function.
- Show that  $\frac{d}{dx} \tanh(x) = 1 - \tanh^2(x)$ .

## 2 Implementing the Feed-Forward Language Model

**Q1: Data processing** The basic ingredient of a language model is a dataset of next-token predictions. In `data.py`, you will find a basic dataset class `SSTLanguageModelingDataset`. In its `from_file` method, it iterates through the lines in a file, and calls a helper function to generate example pairs.

- Implement the method `examples_from_characters`. Read the docstring closely for desired behavior.

**Q2: Implementing `tanh`** In `ops.py`, you will find a skeleton Operation for `tanh`. Using your written answer above as a guide, implement the forward and backward methods for this op.

**Q3: Implementing the Language Model** In `model.py`, you will find the main model class `FeedForwardLanguageModel`, with its initialization method written. Implement the `.forward` method, using its docstring as a guide. [Hint: `ops.concat`, which we provide, will be necessary.]

**Q4: Generating the next character**

### 3 Running the Language Model

### 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 .
- `hw5.tar.gz` containing:
  - `run_hw5.sh`. This should contain the code for activating the conda environment and your run commands for XX above. You can use `run_hw2.sh` from the previous assignment as a template.
  - `data.py`
  - `model.py`
  - `ops.py`
  - `run.py`