Student: Stephanie Drake

Experiment Notebooks: File Description

This document provides a guide on how to navigate the five notebooks, each containing a separate experiment. Below is a brief description of the contents and purpose of each notebook:

(Please note: All models, including those with both tuned and fixed embeddings, are included in the notebooks. However, in the paper, I only included the best-performing models: the Tree-LSTM and LSTM with tuned embeddings, as well as the Tree-LSTM with node-level supervision and tuned embeddings.)

Key for notebooks

Research question	Name of file
Research question 1	**Experiment_1_wordorder.ipynb**
Research question 2	**Experiment_1_wordorder.ipynb**
Research question 3	**Experiment_2_sentencelength**
Research question 4	**Experiment_3_supervise**
Research question 5	**Experiment_4_sentiment**

Notebook 1: stephanie_drake.ipynb

Purpose: The original notebook provided by the course.

- **Contents:**
    - Data preprocessing steps.
    - Model training setup
    - Decisions regarding iterations
    - Evaluation metrics and results.
- **Key Outputs:** Loss and accuracy plots

Notebook 2: Experiment_1_wordorder.ipynb

- **Purpose:** To investigate the influence of word order, specifically by comparing the accuracy of DeepCBOW to LSTM models.

- **Contents:**
    - Extension of Notebook 1 (e.g., same model architecture, dataset).
    - Runs each model over 3 seeds using chosen iterations
- **Key Outputs:** mean accuracy, standard deviation

Notebook 3: Experiment_2_sentencelength

- **Purpose:** investigation into the accuracy of models across different sentence length
- **Contents:**
    - Extension of notebook 2
    - Evaluation of each model across sentence length
        - Two functions added to create bins (sentence length) and to evaluate bins
- **Key Outputs:** Line graph showing the performance of each model across sentence lengths. Found at the end of the notebook.

Notebook 4: Experiment_3_supervise

- **Purpose:** Supervising the sentiment at nodes to improve model performance
- **Contents:**
    - Includes a new function to parse subtrees with their sentiment.
    - Creates new training data to train the model on
    - Predict the sentiment classes using the original test data
- **Key Outputs:** code specific to this experiment is under the heading 'Experiment 3' of the notebook

---

Notebook 5: Experiment_4_sentiment

- **Purpose:** To evaluate sentiment class accuracy across depth and phrase length of n-ary tree model.
- **Contents:**
    - Error logging at nodes – change to Tree-lstm model

    - Evaluate node logging function
    - Generation of graphs for sentiment classes across depths and phrase lengths
- **Key Outputs:** All code added for the experiment is under heading ‘Experiment 4’ Visualizations produced at bottom of notebook

---