

Project Documentation

Wordle Game Strategy

CS 410, Fall 2022

The video demonstration is hosted at Illinois Media Space:

https://mediaspace.illinois.edu/media/t/1_bon8om0s

Skip ahead to 13:40 for a demonstration on using this strategy tool to play Wordle.

1. Code Overview

- **POS_Wordle.ipynb**

- Python notebook that uses the NLTK package and perceptron part-of-speech (POS) tagger to tag the past Wordle game solution words and 12,972 allowable guess words. An interesting finding is that 62% of the possible words are tagged 'NN', which represents a singular noun in the Penn Treebank schema, whereas only 36% of allowable words have the 'NN' tag.
- Outputs `all_pos.csv` which is used for POS likelihoods in the Wordle game strategy tool.
- Users do not need to run this code to operate the Wordle game strategy tool.

- **TF_IDF_Wordle.ipynb**

- Python notebook that uses Google Book Ngram Viewer API to find the TF-IDF likelihood of the past Wordle game solution words and 12,972 allowable guess words. An interesting finding is that past Wordle game solutions have an average frequency likelihood that is 6x greater than the allowable words.
- Outputs `all_tf.csv` which is used for TF-IDF likelihoods in the Wordle game strategy tool.
- Users do not need to run this code to operate the Wordle game strategy tool.

- **Strategy_Wordle.ipynb**

- Python notebook that implements the Wordle game strategy.
- There are two main strategies:
 1. Select next guess word based on entropy
 2. Select next guess word based on the POS and TF-IDF likelihoods.
- Users should run all code blocks to prepare and load the Wordle strategy tool, but users can optionally skip the Evaluation and Results code blocks to save compute time.

2. Implementation Details

- This project was implemented using Python notebooks on Google Colab
- Packages used by the strategy notebook are NumPy and Pandas. In addition, the POS tagger notebook uses NLTK and the TF-IDF likelihood notebook uses the requests package to interface with the Google Books Ngram Viewer API.

3. Usage Instructions

- The notebooks are well organized into code blocks with commentary and examples throughout.
- For ease of use, it is recommended to copy the notebooks and data folder within a `Colab Notebooks` subfolder on Google Drive.
 - At the top of each notebook, there is a flag `GOOGLE_DRIVE = True` which a user can set to `False` if desired to run the code locally.
 - If using Google Drive, please update the `datapath` variable in this same code section, which by default is set to `/content/drive/My Drive/Colab Notebooks/UIUC/CS_410/Wordle/data/`
 - The challenge is selecting an appropriate corpus.
- Using the Wordle game strategy tool
 - Run these code blocks to load and prep the strategy tool:
 1. Prepare Environment
 2. Matrix Operations Tool
 3. Prepare Words
 4. Game Strategy
 - At the bottom of the Strategy_Wordle.ipynb notebook, there is a code block `Game Demo` where users can use the strategy tool in coordination with the official Wordle online game to aid in selecting guesses during live gameplay.
 - The video demonstration starting at 13:40 shows how to play Wordle using this strategy tool.

4. Evaluation

- The goal of this strategy tool was to improve upon the average human player score of 3.919, and that was accomplished using two different techniques.
- The code is flexible enough to be extended and improved, and the matrix operations make it relatively efficient.