**Bigram Grammar Exercise (Option 2)**

Abuchi Godswill Okeke

School of Computer and Information Sciences, University of the Cumberlands

MSAI532 - M51 Natural Language Processing

Dr. Toni Farley

March 14, 2025

**Result Discussion**

In the week 10 assignment 8, I worked on Option 2, where I set up an experiment to evaluate the robustness of a part-of-speech tagging algorithm using spaCy. In this experiment, I processed the grammatically correct sentence “**Learning NLP helps me excel as a software engineer at Walmart**” to obtain POS tags and then compared these results with those generated from a modified version containing intentional spelling errors (“**Lerning NLP helps me excel as a softwre engneer at Walmart**”). The analysis revealed that while the algorithm accurately tagged the original sentence, it largely maintained the correct POS assignments in the modified sentence as well, relying on contextual cues to overcome the minor orthographic deviations. However, the presence of misspellings did appear to slightly diminish the confidence of some tag assignments, suggesting that although the model is resilient to small errors, substantial deviations might further compromise its performance (Jurafsky & Martin, 2017).

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* Python Script\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

*import* spacy

*#requirements*

*#!pip install spacy*

*#!python -m spacy download en\_core\_web\_sm*

nlp = spacy.load("en\_core\_web\_sm")

*# 1. Original sentence (grammatically correct) related to the course and Walmart:*

sentence = "Learning NLP helps me excel as a software engineer at Walmart."

doc = nlp(sentence)

print("Original sentence POS tags:")

*for* token *in* doc:

print(f"{token.text:20s} {token.pos\_}")

*# 2. Introduce 3 spelling errors: "Learning" -> "Lerning", "software" -> "softwre", "engineer" -> "engneer"*

sentence\_misspelled = "Lerning NLP helps me excel as a softwre engneer at Walmart."

doc\_misspelled = nlp(sentence\_misspelled)

print("\nModified sentence (with spelling errors) POS tags:")

*for* token *in* doc\_misspelled:

print(f"{token.text:20s} {token.pos\_}")

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* End Script\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**Reference**

Jurafsky, D., & Martin, J. H. (2007). Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition (Draft, October 30, 2007). Pearson.