Teammate 1 NAME: Shenfan Feng netid: sfeng37
Teammate 2 NAME: Yiming Wang netid:ywang626
Instructor: Barbara Di Eugenio
Course: CS 421

Final project

Part one

(a) Length of the Essay: Counting Sentences

- Methodology: To determine the essay's length, SpaCy's language model (en_core_web_sm) is utilized for its ability to effectively segment text into sentences. This is achieved by loading the essay text into SpaCy's NLP pipeline and using its doc.sents attribute to extract and count sentences.
- Technical Implementation: The number of sentences (num_sentences) is compared against predetermined thresholds for low and high scores (10 and 20 sentences, respectively). This count helps scale the score linearly between these thresholds to provide a nuanced assessment of essay length.

(b) Spelling Mistakes: Spell Checking

- Methodology: The SpellChecker library is used for identifying spelling errors. This library provides a list of words it considers misspelled by comparing each word against a large English dictionary.
- Technical Implementation: Words from the essay are split and passed to the spell.unknown() method which checks their correctness. The count of these unrecognized words forms the basis of the spelling score. Note: The scoring logic here deviates from typical expectations, as more spelling errors uncharacteristically result in a higher score.

(c) Syntax/Grammar: POS Patterns and Agreement

c.i Subject-Verb Agreement:

- Methodology: Utilizes SpaCy's dependency parsing capabilities to identify verbs and their subjects within sentences. It checks agreement in number (singular vs plural).
- Technical Implementation: For each verb identified by the POS tag VERB, its related subjects (children with nsubj dependency) are examined. The agreement is validated based on POS tags (NN, NNP for singular and NNS, NNPS for plural) and corresponding verb forms (VBZ for singular, VBP for plural). Errors in agreement reduce the score.

c.ii Verb Tense/Usage:

- Methodology: Also leveraging SpaCy's parsing features, this component focuses on verb structures involving auxiliary verbs and main verb completeness.
- Technical Implementation: Each sentence is checked for the presence of auxiliary verbs (MD) and main verbs. The script ensures that auxiliary verbs are paired with appropriate

main verbs or vice versa. Incorrect or missing verb pairings contribute to lower scores based on the prevalence of such errors relative to the total sentence count.

Part 2:
c.iii Syntactic well-formedness
Tried:
1. determine the type of sentence, e.g., whether it is a question or declarative sentence
2. determine if the question begins with an auxiliary.
3. determine if the main clause is punctuated “.” / “!” as an ending.
4. determine if the sentence begins with a verb, then it’s an error
5. determine if the main clause does not begin with a noun, proper noun, preposition, adverb, etc. then it is a mistake.
6. determine the subordinating conjunction
7. determine whether a verb occurs when the modal appears.
8. determine whether a singular noun is preceded by a determiner.
9. determine if there is no determiner before a proper noun
10. determine whether a sentence has a main verb

Not sure:
1. In judging prepositional phrases, I considered two examples. When the prepositional phrase is “at night”, there is no determiner in front of the noun, while when the prepositional phrase is “in the morning”, the noun needs a determiner in front of it.

(d) Semantics (meaning) & Pragmatics (coherence)
d.i Does the essay address the topic?
Tried:
1. using spacy and vector, calculate cosine_similarity
2. divide scores based on similarity results

Not sure:
1. After the test results showed the scores, I found that some sentences were rated low even though their similarity was high, while others were rated the opposite.

d.ii Essay coherence
Tried:
1. Calculated using similarity and combined with standard deviation to classify scores.

Not sure:

1. As mentioned in the pdf, I would have a hard time using coherence alone to judge the quality of the essay

What learned:
1. For the division of sentences, the structures needed, such as nouns and determiners, modals and verbs, prepositional phrases, etc., they all need to be carefully considered when writing sentences
2. Cosine similarity is very relevant for both meaning and coherence.

Future work:
1. Need more time to test sentence by sentence. This would have made it more likely to avoid the possibility of misclassification.
2. There is still a portion of overlapping misclassification in the results. For example, if some of the results actually get a score higher than 4, the grade is considered high, but the actual quality is low
3. Regarding sentences that are not questions, there is actually still the possibility of starting with a verb, such as an imperative. “Run.”