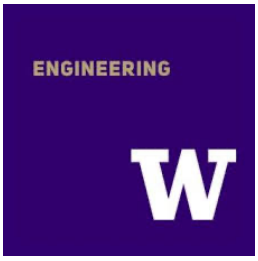
University of Washington – College of Engineering



Mini Project 1_Part 1

Subject:

From Transformers to GPT

Student:

Jens Jung

Professor:

Dr. Karthik Mohan

Date of submission:

28-Jan-26

(6) Analysis and Questions

Part A: Analysis

1. Which models got it right?

- SentenceTransformer and OpenAI consistently got it right. In the "Roses are red" test, they correctly prioritized the Flowers category although the presence of "trucks." In the Sentiment Tests they correctly identified the sentiment as Negative.
- GloVe failed both tests (classified the flower/truck sentence as Colors and misclassified terrible and upsetting as Positive).

2. Why fail?

- GloVe failed because it uses word embeddings that are static and the implementation uses averages. It treats a sentence as kind of a bag of words, losing the context of how words interact. In the sentiment test, words like terrible might have high co-occurrence with movie in training data, but without a structural understanding of negation or specific emotional weight the simple average was pulled toward a neutral or slightly positive baseline.
- Transformer based models (ST and OpenAI) succeeded because they use attention Mechanisms. They don't just look at word counts they actually look at the relationship between words. They recognize that upsetting is the core descriptor of the sentence's intent, giving it more weight in the final vector representation.

3. What does this reveal about word order?

- GloVe is invariant to word order: In the Part C experiment, GloVe produced the *exact same* confidence score (2.5466) for both cinema-hotel-restaurant and hotel-restaurant-cinema. Since it simply averages the vectors, the sequence doesn't matter.
- Transformers are sensitive to word order: The Sentence Transformer and OpenAI models produced different confidence scores when the words were shuffled. This is because they use positional encodings. The model knows which word came first, allowing it to capture the nuanced shift in focus or emphasis when a sentence is restructured.

Part B: Comparison

Feature	GloVe	Sentence Transformer	OpenAI
Accuracy	Low	High	High
Speed	Fast	Fast	Slow
Word Order	No	Yes	Yes (highly)