Context-Aware Synonym Replacement Using Learned Representations

Sebastian Lopez Diego Bonilla
Texas Tech University

CS4391 – Homework 1
Spring 2026

1 Introduction

This homework investigates whether pretrained semantic representations can support context-aware synonym replacement without relying on large language models, masked language models, or lexical databases. The objective is not fluent generation, but understanding how embeddings encode meaning and where they fail—particularly in metaphorical or poetic language.

We design a controlled pipeline that replaces a selected word in each sentence while attempting to preserve sentence-level meaning using embedding-based similarity.

2 System Overview

The system follows a modular pipeline:

- Sentence segmentation
- Target word selection (last non-stopword)
- Synonym candidate generation using word embeddings
- Heuristic grammatical filtering
- Context-based scoring using sentence embeddings

A replacement is accepted only if semantic similarity with the original sentence exceeds a predefined threshold.

3 Model Selection

3.1 Word Embeddings

Candidate synonyms are generated using the `glove-wiki-gigaword-100` pretrained embeddings. This model offers broad semantic coverage and efficient nearest-neighbor retrieval, making it suitable for exploratory experimentation.

3.2 Sentence Embeddings

Contextual compatibility is evaluated using the `all-MiniLM-L6-v2` Sentence-BERT model. This lightweight transformer produces sentence-level embeddings that allow cosine similarity comparisons between original and modified sentences.

4 Experimental Results

4.1 Filtering Effects

Early experiments with minimal filtering produced frequent grammatical errors, such as plural mismatches and noun–adjective shifts. Adding lightweight grammatical constraints significantly reduced these failures while preserving semantic flexibility.

4.2 Threshold Sensitivity

We tested similarity thresholds from 0.70 to 0.90. Lower thresholds allowed more substitutions but often damaged metaphorical meaning, while higher thresholds became overly restrictive. A threshold of **0.86** provided the best balance and was used for final results.

5 Examples

5.1 Literal Replacement

Original: "Two roads diverged in a yellow wood."
Modified: "Two roads diverged in a yellow timber."

The substitution preserves meaning and structure, confirmed by a high sentence similarity score.

5.2 Metaphorical Preservation

Original: "Hope is the thing with feathers."
Modified: (unchanged)

No candidate exceeded the similarity threshold, correctly preserving the metaphor.

6 Failure Analysis

Despite filtering and context-based scoring, several systematic failure modes persist. Below we illustrate each with concrete examples.

6.1 Polysemy and Sense Conflation

Example: Shakespeare — *players* $\rightarrow$ *teams*

Analysis: In Shakespeare's metaphor, "players" refers to theatrical actors. However, pretrained word embeddings conflate multiple senses of the word (theatrical, sports, music). As a result, "teams" appears as a close semantic neighbor due to sports-related co-occurrence, and sentence similarity remains high despite the loss of intended meaning.

Cause: Word embeddings average over all observed contexts and cannot perform word sense disambiguation.

6.2 Metaphorical Meaning Collapse

Example: Dickinson (low threshold) — *feathers* $\rightarrow$ *birds*

Analysis: Although "feathers" and "birds" are semantically related, the substitution destroys the metaphor. "Feathers" evokes lightness and hope abstractly, while "birds" introduces a concrete entity that alters the poem's intent.

Cause: Embedding similarity reflects association rather than figurative meaning.

6.3 Grammatical Number Drift

Example: Blake — *night* $\rightarrow$ *nights*

Analysis: Sentence embeddings score this replacement as highly similar, yet the grammatical shift subtly alters meaning. "The night" refers to a specific concept, while "the nights" generalizes it.

Cause: Sentence-level embeddings downweight grammatical distinctions such as number when computing semantic similarity.

6.4 Stylistic and Phonetic Insensitivity

Example: Haiku — *Splash* $\rightarrow$ *splashes*

Analysis: While semantic content is preserved, stylistic impact is reduced. "Splash!" is sharp and immediate, whereas "splashes" is softer and continuous. Capitalization and onomatopoeia contribute meaning that embeddings fail to capture.

Cause: Learned embeddings encode semantic content but ignore phonetic, rhythmic, and stylistic properties of language.

7 Use of AI in This Homework

AI tools were used as interactive collaborators to clarify concepts, accelerate experimentation, diagnose failures, and assist with report structuring. All architectural decisions and interpretations were critically evaluated by the authors.

7.1 AI-Assisted Workflow

Figure 1 illustrates the collaborative development process.

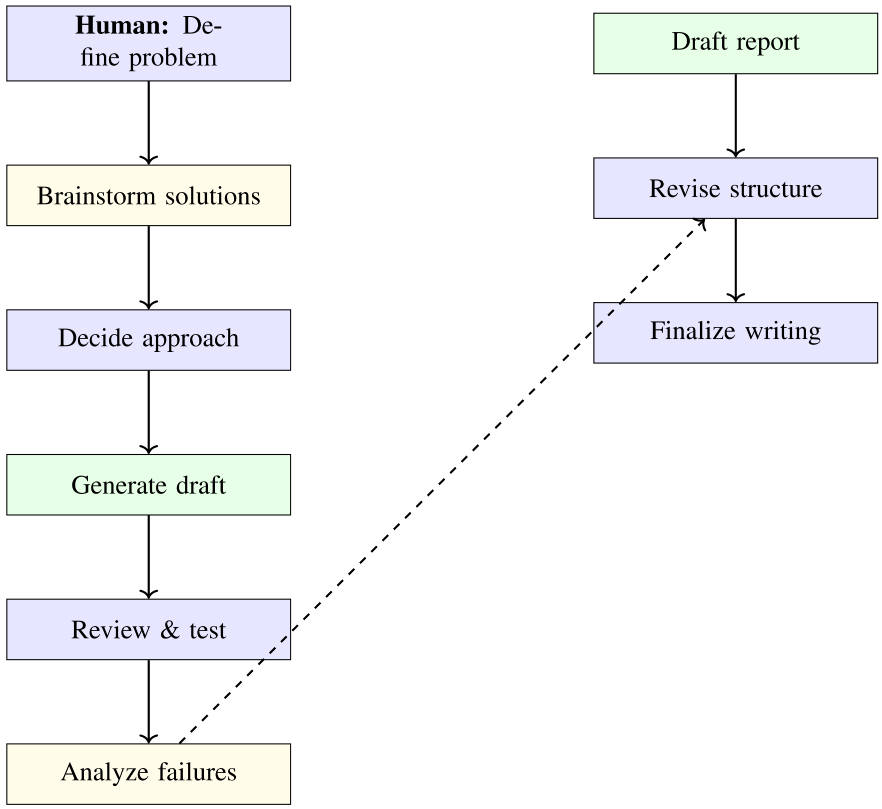

Figure 1: AI-assisted workflow for system development and report writing.

8 Conclusion

This work demonstrates that embedding-based similarity enables effective literal synonym replacement but fails systematically on metaphor, polysemy, and stylistic language. These limitations highlight the gap between distributional semantics and human language understanding and motivate more expressive models for context-sensitive text manipulation.