CSCI 5832 Homework 3
Written Report

Rama Durga Santosh Kumar Chaganti

March 10, 2025

Question 1:

Let us start with different embedding dimensions:

Experiment 1: I am considering Different Embeddings. (25,50,100)

For custom Embeddings, smaller dimensions (25) would likely train faster but might have lower performance since they have less capacity to capture semantic relationships. For medium dimensions (50) it would likely provide a good balance between training speed and model performance. Larger dimensions might give slightly better performance, but the diminishing returns and longer training times.

For Glove Embeddings would outperform custom embeddings across all dimension. The difference between custom and Glove would be mostly on small dimensions, as Glove pretrained nature has a limited dimension size. As dimension size increases, the gap between custom and glove might narrow slightly. But ultimately Glove outperformed.

Experiment 2: Different Batch sizes (8,16,32)

For custom embeddings, smaller batch would give noisy training but possible good results in some cases. Medium batch sizes would provide a good balance betweeen training stability. Larger batch size would give more stable gradient estimates but might converge to less optimal solutions. Glove would maintain its advantage across all batch sizes.

Question 2:

```
=== Starting Embedding Analysis ===

Analyzing Custom Embeddings:
Nearest neighbors for 'good': ['industrialmodel', 'catalyst', 'sky', 'quirkily', 'pretty']
Nearest neighbors for 'bad': ['pokey', 'embarrassment', 'marveilleux', 'gory', 'whiteonblack']
Nearest neighbors for 'excellent': ['exquisite', 'reconfortado', 'funnygritty', 'arbitrarily', 'election']
Nearest neighbors for 'terrible': ['88', 'inflection', 'constant', 'knucklehead', 'gross']
Nearest neighbors for 'movie': ['looseness', 'madeformovie', 'fling', 'innumerable', 'convence']

Analyzing GloVe Embeddings:
Nearest neighbors for 'good': ['always', 'way', 'something', 'sure', 'you']
Nearest neighbors for 'bad': ['worse', 'too', 'little', 'really', 'nothing']
Nearest neighbors for 'excellent': ['good', 'quality', 'terrific', 'superb', 'best']
Nearest neighbors for 'terrible': ['awful', 'horrible', 'dreadful', 'tragedy', 'horrific']
Nearest neighbors for 'movie': ['film', 'hollywood', 'comedy', 'show', 'sequel']
```

Figure 1: Embedding Analysis

The Glove embeddings in above image give more semantic relationships. Words like "terrible" are closest to semantically related words like "awful", "horrible" and "dreadful". While "movie" is associated with "film", "hollywood" and "comedy". This makes sense given glove was trained on a a massive corpus to capture general language patterns.

In contrast, the custom embeddings trained solely on the movie review dataset show much less intuitive relationships. For example in image see "terrible" associated words which are quite unrelated. This suggests that custom embeddings are capturing task specific patterns related to sentiment classification rather than general semantic. The he custom embeddings may be learning correlations between words that frequently appear in similar sentiment contexts, but these relationships don't necessarily align with our intuitive understanding of word meanings.

This tells pre trained embeddings give rich semantic because of large scale training.

Question 3:

Figure 2: Example of GLOVE Embedding

When visualizing the custom-trained and GloVe embeddings. The GloVe embeddings visualization likely shows clear clustering of semantically related words - positive sentiment words grouped together (like "good," "excellent," "terrific"), negative sentiment words forming their own cluster (like "terrible," "awful," "horrible"), and topical movie-related terms clustering separately (like "film," "cinema," "Hollywood"). These distinct groupings demonstrate GloVe's ability to capture both semantic similarity and sentiment polarity through training on a large, diverse corpus. The Example I like is Languages names clustered together. you can see in above image.

In contrast, the visualization of custom embeddings trained are not that good. Even they clustered in many cases the words doesn't make sense.

But I would like to talk on F1 score which I got more than 0.7 and I feel for custom embeddings are good for the limited data.