Problem

Suppose we have a list of word embeddings. How can we use this list to compare the similarity of two sentences?

1. Euclidean Distance - Add up all the semantic vectors for each sentence to get a sentence vector. Scale each sentence based on the number of words in the sentence. Compute the norm of their difference
2. Jaccard Similarity - For each word in a sentence, find the top 5 most similar words and append them to a list associated with each sentence. Compute the number of elements in intersection of each list / number of elements union of both lists
3. Cosine Similarity - Add up all the semantic vectors to get a sentence vector. Find the cosine angle between the two vectors

<https://studymachinelearning.com/jaccard-similarity-text-similarity-metric-in-nlp/>