Similarity Scores

You are provided with four documents, numbered 1 to 4, each with a single sentence of text. Determine the identifier of the document D which is the most similar to the first document, as computed according to the TF-IDF scores.

1. I'd like an apple.

2. An apple a day keeps the doctor away.

3. Never compare an apple to an orange.

4. I prefer scikit-learn to orange.

Output the integer D (which may be either 2 or 3 or 4), leaving no leading or trailing spaces.

You may either compute the answer manually and submit it in plain-text mode, or submit a program which computes the answer, in a language of your choice.