

Irs lab1

```
from sklearn.feature_extraction.text import TfidfVectorizer  
  
from sklearn.metrics.pairwise import cosine_similarity  
  
def compute_similarity(documents):  
    """
```

Computes the similarity between documents using the Vector Space Model and Cosine Similarity.

Args:

documents: A list of strings, where each string is a document.

Returns:

A list of similarity scores between each document pair.

"""

```
# 1. Text Preprocessing and Vectorization
```

```
vectorizer = TfidfVectorizer()
```

```
vectors = vectorizer.fit_transform(documents)
```

```
# 2. Calculate Cosine Similarity
```

```
similarity_matrix = cosine_similarity(vectors)
```

```
# 3. Return Similarity Scores
```

```
return similarity_matrix
```

```
# Example Usage
```

```
documents = [
```

```
"This is the first document.",  
"This document is the second document.",  
"And this is the third one.",  
"Is this the first document?"  
]  
  
similarity_scores = compute_similarity(documents)  
print(similarity_scores)
```

Created with H.ai - Your AI Learning Assistant