Congratulations! You passed!

Grade received 100% **To pass** 80% or higher

Go to next item

Vector Space Models

Total points 10

1. Given a corpus A, encoded as $\begin{pmatrix}1\\2\\3\end{pmatrix}$ and corpus B encoded as $\begin{pmatrix}4\\7\\2\end{pmatrix}$, What is the euclidean distance between the two documents? **1 / 1 point**

- (●) 5.91608
- ( ) 35
- ( ) 2.43
- ( ) None of the above

Correct
Yes, this is correct.

2. Given the previous problem, a user now came up with a corpus C defined as $\begin{pmatrix}3\\1\\4\end{pmatrix}$ and you want to recommend a document that is similar to it. Would you recommend document A or document B? **1 / 1 point**

- (●) Document A
- ( ) Document B

Correct
That is correct

3. Which of the following is true about euclidean distance? **1 / 1 point**

- [x] When comparing similarity between two corpuses, it does not work well when the documents are of different sizes.

Correct
That is correct.

- [x] It is the norm of the difference between two vectors.

Correct
That is correct.

- [ ] It is a method that makes use of the angle between two vectors
- [ ] It is the norm squared of the difference between two vectors.

4. What is the range of a cosine similarity score, namely *s*, in the case of information retrieval where the vectors are positive? **1 / 1 point**

- [ ] $$ -1 \leq s \leq 1 $$
- [ ] $$-\infty \leq s \leq \infty $$
- [x] $$0 \leq s \leq 1 $$

Correct
That is correct.

- [ ] $$-1 \leq s \leq 0 $$

5. The cosine similarity score of corpus A = $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ and corpus B = $\begin{pmatrix} 2 \\ 8 \\ 1 \end{pmatrix}$ is equal to ? **1 / 1 point**

- (•) 0.08512565307587486
- ( ) 0
- ( ) 1.251903
- ( ) -0.3418283

Correct
This is correct.

6. We will define the following vectors, USA = $\begin{pmatrix} 5 \\ 6 \end{pmatrix}$, Washington = $\begin{pmatrix} 10 \\ 5 \end{pmatrix}$, Turkey = $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$, Ankara = $\begin{pmatrix} 9 \\ 1 \end{pmatrix}$, Russian = $\begin{pmatrix} 5 \\ 5 \end{pmatrix}$, and Japan = $\begin{pmatrix} 4 \\ 3 \end{pmatrix}$. Using only the following vectors, Ankara is the capital of what country? **1 / 1 point**

- ( ) Japan
- ( ) Russia
- ( ) Morocco
- (•) Turkey

Correct
Yes, you should compute (USA - Washington) + Ankara and then compare that vector to the country vectors to decide.

7. Please select all that apply. PCA is **1 / 1 point**

- [x] used to reduce the dimension of your data.

Correct
This is correct.

- [x] visualize word vectors

Correct

This is correct.

- [ ] make predictions
- [ ] label data

8. Please select all that apply. Which is correct about PCA? **1 / 1 point**

- [x] You can think of an eigenvector as an uncorrelated feature for your data.

Correct
That is correct.

- [x] The eigenvalues tell you the amount of information retained by each feature.

Correct
This is correct.

- [ ] If working with features in different scales, you do not have to mean normalize.
- [x] Computing the covariance matrix is critical when performing PCA

Correct
This is correct.

9. In which order do you perform the following operations when computing PCA? **1 / 1 point**

- (•) mean normalize, get Σ the covariance matrix, perform SVD, then dot product the data, namely X, with a subset of the columns of U to get the reconstruction of your data.
- ( ) mean normalize, perform SVD, get Σ the covariance matrix, then dot product the data, namely X, with a subset of the columns of U to get the reconstruction of your data.
- ( ) get Σ the covariance matrix, perform SVD, then dot product the data, namely X, with a subset of the columns of U to get the reconstruction of your data, mean normalize.
- ( ) get Σ the covariance matrix, mean normalize, perform SVD, then dot product the data, namely X, with a subset of the columns of U to get the reconstruction of your data.

Correct
This is correct.

10. Vector space models allow us to **1 / 1 point**

- [x] To represent words and documents as vectors.

Correct
This is correct.

- [x] build useful applications including and not limited to, information extraction, machine translation, and chatbots.

Correct
This is correct.

- [x] create representations that capture similar meaning.

Correct
This is correct.

- [ ] build faster training algorithms