

Quiz Topic: Tf-Idf Include Formulas

1. What does TF-IDF stand for?

- 1 Term Factor-Internal Document Frequency
- 2 Term Factor-Inner Document Frequency
- 3 Term Frequency-Inverse Data Frequency
- 4 Term Frequency-Inverse Document Frequency

Answer: Term Frequency-Inverse Document Frequency

2. Which of the following formulas correctly represents the Term Frequency (TF)?

- 1 $TF = (\text{Number of times term } t \text{ appears in a document}) / (\text{Total number of terms in the document})$
- 2 $TF = (\text{Number of times term } t \text{ appears in a document}) * (\text{Total number of terms in the document})$
- 3 $TF = (\text{Number of documents containing term } t) / (\text{Total number of documents})$
- 4 $TF = (\text{Total number of terms in the document}) / (\text{Number of times term } t \text{ appears in a document})$

Answer: $TF = (\text{Number of times term } t \text{ appears in a document}) / (\text{Total number of terms in the document})$

3. What is the formula for Inverse Document Frequency (IDF)?

- 1 $IDF = \log_e(\text{Total number of documents} / \text{Number of documents containing term } t)$
- 2 $IDF = \text{Total number of documents} / (\text{Number of documents containing term } t)$
- 3 $IDF = (\text{Number of documents containing term } t) / (\text{Total number of documents})$
- 4 $IDF = \log_e(\text{Number of documents containing term } t)$

Answer: $IDF = \log_e(\text{Total number of documents} / \text{Number of documents containing term } t)$

4. How is the TF-IDF score calculated?

- 1 $TF-IDF = TF / IDF$
- 2 $TF-IDF = \log(TF) + \log(IDF)$
- 3 $TF-IDF = TF * IDF$
- 4 $TF-IDF = TF + IDF$

*Answer: $TF-IDF = TF * IDF$*

5. If a term appears 10 times in a document of 100 terms, what is the Term Frequency (TF) of that term?

Answer: 0.1

6. An IDF value of 3 indicates what about the term's occurrence across documents?

- 1 The term is very common across most documents
- 2 The term is rare and appears in few documents
- 3 The term appears in every document
- 4 The term appears in a moderate number of documents

Answer: The term is rare and appears in few documents

7. What is the typical purpose of using TF-IDF in text analysis?

- 1 To measure document readability
- 2 To rank the importance of terms in a document relative to a corpus
- 3 To check for spelling errors
- 4 To translate text into different languages

Answer: To rank the importance of terms in a document relative to a corpus