

calculate_word_frequency_documentation

Mackenzie Ramalho

2024-04-26

Function 3: 'calculate_word_frequency()'

This calculates the frequency of each word in the text.

Parameters:

- 'text': A character vector containing the text for which word frequency is to be calculated.

Returns:

A named integer vector where each element represents the frequency of a word in the input text.

```
tokenize_text <- function(text) {  
  tokens <- strsplit(text, "\\s+")[1]  
  return(tokens)  
}  
  
# Function to calculate word frequency  
calculate_word_frequency <- function(text) {  
  tokens <- tokenize_text(text)  
  word_freq <- table(tokens)  
  return(word_freq)  
}  
  
#Example  
word_freq <- calculate_word_frequency("This is a sample sentence. This sentence is a sample.")  
  
word_freq
```

```
## tokens  
##      a      is    sample  sample. sentence sentence.      This  
##      2      2      1      1      1      1      2
```

Use Case:

This function is useful for understanding the distribution of words in a text document by providing the frequency of each word. It can be used for various text analysis tasks such as identifying important terms, detecting common phrases, and measuring document similarity.