Wei

N26041678

4A02C014@stust.dedu.tw

March 22 2016

In order to highlight the focus of the article ,the special and important words in the article need to be found mostly.

The article will promote these words to make the article's mainpoint prominent.

Compute all the words in the article and show the number of occurrences of most rankings.

By these words ,the analyzer know which words in the article should be analyzed.

operate process

1. distinguish : distinguish all the words in the article

2. classification : Divided into four categories by word count.

a. Word Count >15

b.15> Word Count>10

c. 10 > Word Count >5

d. 5>Word Count

3. search : Search Word repetition rate in the article.

4. statistics : count the repetition rate of words in the article.

5. sort : Show the number of occurrences of most rankings in every categorie.

**Reference**

1. Collins2011 Natural Language Processing(Almost) Machine Learning Research

2. Weiscbedel2006 White Paper on Natural Language Processing