

Internship Challenge

Ram Krishna Rahul, 8285460211 , rahul.iitd13@gmail.com

Steps I used to find the similarity between two resumes:

- This Assignment is made in Python language.
- Store both resumes in two Strings.
- Convert all character in to lower case character
- remove all special characters like (. : , ' ").
- Convert Strings in to list of words .
- Our list contains lots of word that will not decide the similarity. ex. Is, am that etc. Then I removed those noises from list.
- Now I make a dictionary in which i stored frequency of each word.
- Then Compare dictionary of both resumes and list out all matching words.
- I also make a list of power keyword which generally used in resume. For ex. Any programming language like java. If java word is used in both resume then java should decide the similarity rather than other words like “along, together” etc.
- Then I search power keyword in list of matching words and increase the weight of power keyword by increasing frequency of that word in both list by 4 times.
- Now I calculated the similarity in following manner:
 1. calculate total words in both resumes and add those numbers.
 $P = (p1+p2)$
p1 number of words in resume 1
p2 number of words in resume 2
 2. i have a list of matching word, i compare frequency of each matching word for both resume and choose lower frequency number. And divide that number by total number of words.

ex. Business is used in resume1 5 times and in resume2 8 times
then i choose 5 and divide 5 by total number of words in resume 1 and resume 2.
 3. Do this for all matching word and add them.
 4. Final answer is the similarity