

[| 일반 소스 공유 | >](#)

## 유사도 함수

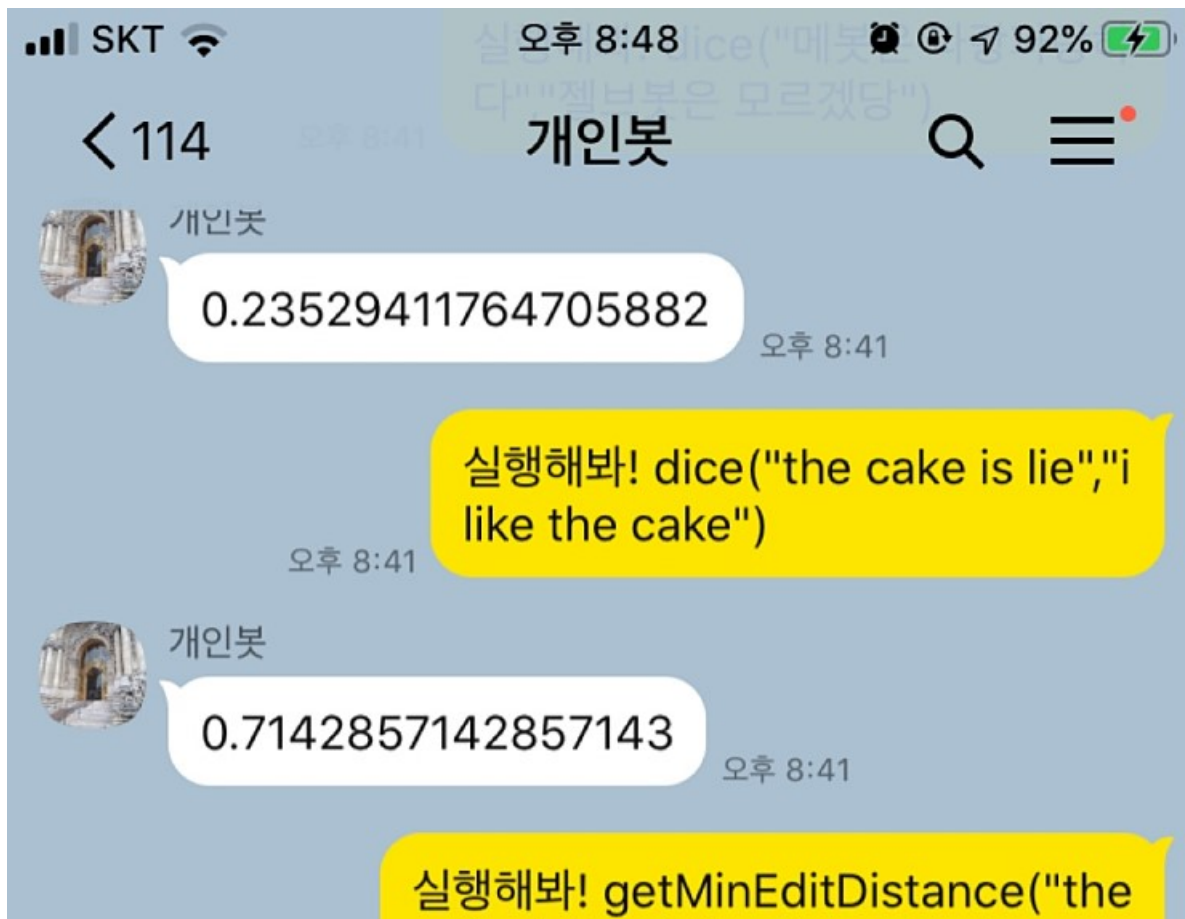


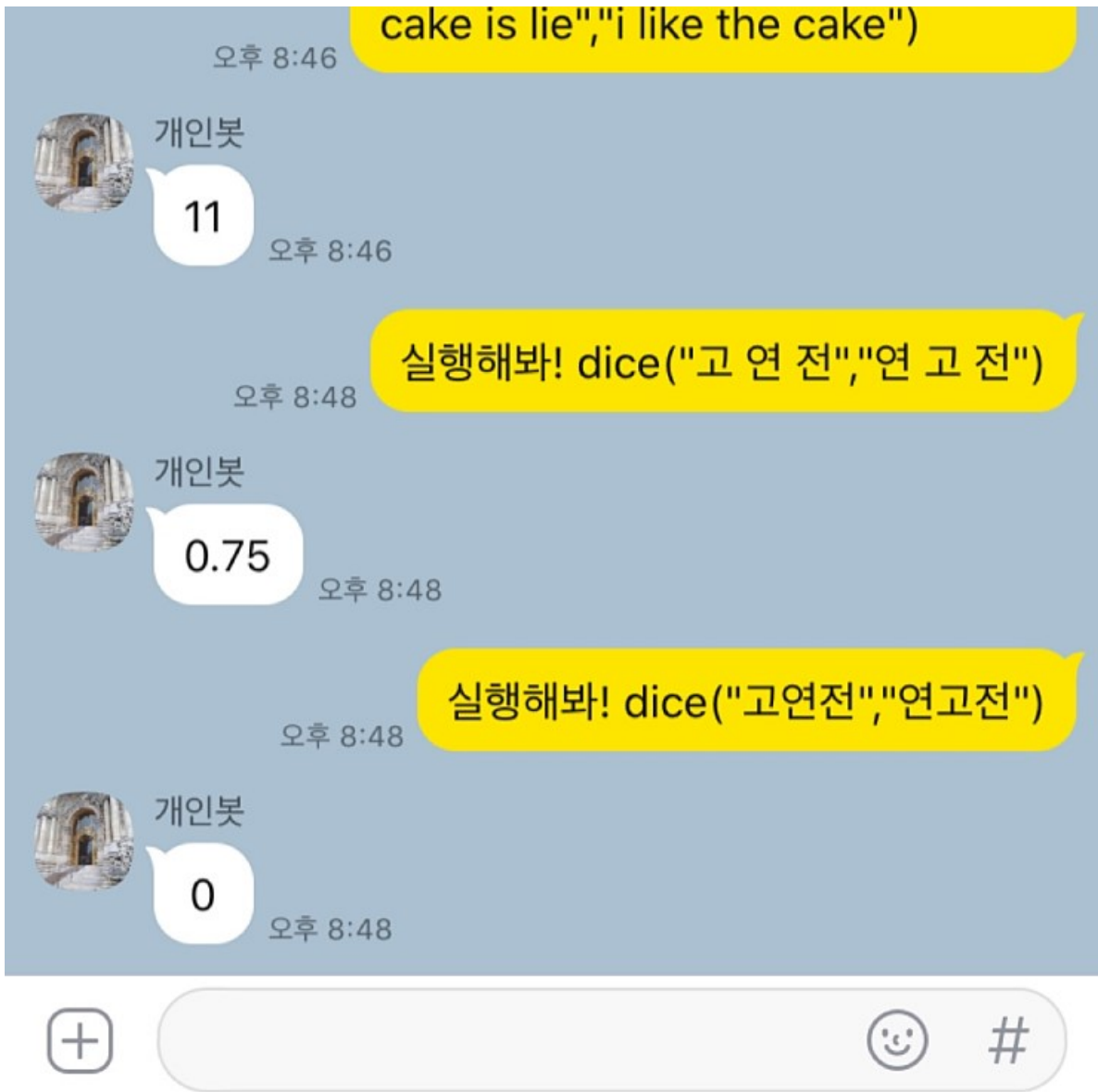
kyleok 챗봇 고수 1:1 채팅

2019.10.28. 20:54 조회 249


💬 댓글 21 URL 복사 ⋮

```
function dice_coefficient(string1, string2) {  
  var intersection = 0;  
  var length1 = string1.length - 1;  
  var length2 = string2.length - 1;  
  if(length1 < 1 || length2 < 1) return 0;  
  var bigrams2 = [];  
  for(var i = 0; i < length2; i++) {  
    bigrams2.push(string2.substr(i,2));  
  }  
  for(var i = 0; i < length1; i++) {  
    var bigram1 = string1.substr(i, 2);  
    for(var j = 0; j < length2; j++) {  
      if(bigram1 == bigrams2[j]) {  
        intersection++;  
        bigrams2[j] = null;  
        break;  
      }  
    }  
  }  
  return (2.0 * intersection) / (length1 + length2);  
}
```







옛날에는 편집거리쓰고 문자열 길이로 나눠서 유사도를 퍼센트르 띄우고 했었는데 말그대로 편집거리만 보여주지 우리가 생각하는 유사도랑은 다르더군요. 저는 개인적으로 이걸 배열같은거에 연동해서 입력어가 배열에 없는데 유사 있다면 제안하는 방식으로 쓰니당. 이 친구의 장점은 문장처럼 여러 단어가 있을때 더 두드러집니당.


 **kyleok**님의 게시글 더보기 >


 좋아요 0  댓글 21


 공유 | 신고

댓글 등록순 최신순 

댓글알림 

 **Sjy0916**  
11...??  
2019.10.28. 20:55 답글쓰기

 **kyleok** 작성자  
편집거리임니당  
2019.10.28. 20:55 답글쓰기

 **Sjy0916**  
**kyleok** 아 최소편집거리인가요...? 그럼 질문 하나만 해도 될까요?  
2019.10.28. 20:56 답글쓰기

 **kyleok** 작성자  
**Sjy0916** 네에

2019.10.28. 20:56 답글쓰기



Sjy0916

kyleok 최소편집거리를 이용해 유사도를 측정하고 싶고 0-100 퍼센트의 비율로 표현하고 싶은데, 편집 횟수를 비율로 어떻게 변환할까요?

2019.10.28. 20:58 답글쓰기



kyleok 작성자

Sjy0916 스트링1,스트링2 둘중에 긴놈의 글자수로 나누면 유사도가 0~1로 나오겠져

2019.10.28. 20:59 답글쓰기



kyleok 작성자

Sjy0916 저 경우 공백포함 15자 스트링에 11편집거리니까 11/15 \*100(%)겠쥬

2019.10.28. 21:00 답글쓰기



Sjy0916

kyleok 아 ㄷㄷ 그렇네요.. 발상부터 다른 고인물...

2019.10.28. 21:01 답글쓰기



청정수123

Sjy0916 그쵸 최소편집거리가 아무리 길어도 max(str1.length, str2.length)는 못넘으니깐..

2019.10.28. 21:56 답글쓰기



AlphaDo

kyleok 저는 두 스트링의 길이의 평균으로 나눴어용

2019.10.28. 22:43 답글쓰기



kyleok 작성자

AlphaDo 그러면 아주 재수없는 경우 확률이 1을 넘길수도 있어서...!

2019.10.29. 00:55 답글쓰기



AlphaDo

kyleok 아아앙

2019.10.29. 00:56 답글쓰기



kyleok 작성자

AlphaDo 본인 등판하셨으니 여쭙봐야징  
핑퐁소스 base64하라는 얘기는 무슨뜻이에요?  
강 api키 넣으면 안되나여

2019.10.29. 00:59 답글쓰기



AlphaDo

kyleok api 키를 base64로 인코딩...  
아니 근데 저보다 고이셨으면서 왜그래요 ㅠ ㅠ

2019.10.29. 01:00 답글쓰기



AlphaDo

kyleok 아마 거기 사이트에서 base64 인코딩을 헤드에 담아 보내라거 했던거 같은데 안되면 그냥 넣어보셔요

2019.10.29. 01:03 답글쓰기



kyleok 작성자

AlphaDo 예? 저 자스라는 단어를 처음본지 한달됐어여

2019.10.29. 01:04 답글쓰기



AlphaDo

kyleok 그렇다면 다른 언어를 해보셨군ㅇ...알고리즘 수준이

2019.10.29. 01:04 답글쓰기



kyleok 작성자

AlphaDo 훌륭한 개발자는 검색을 잘합니다(?)

2019.10.29. 01:05 답글쓰기



kyleok 작성자

아 한글의 경우 자음모음 찢어놓고하면 더 인식을 좋아지니 쓰실분은 참고하시길

2019.10.28. 20:57 답글 쓰기



jomin398

역시 ㅋㅋㅋ 제작하셨네요  
저번에 단톡방에 아이디어 드렸었죠

2019.10.28. 21:34 답글 쓰기



AlphaDo

LCS!

2019.10.28. 21:36 답글 쓰기

Hibot

댓글을 남겨보세요



등록

글쓰기

답글

목록

▲ TOP

## '| 일반 소스 공유 |' 게시판 글

이 게시판 새글 구독하기 ☐

|                                |          |             |
|--------------------------------|----------|-------------|
| 아이폰키알을 대체하기위해 하나 만들었는데... [8]  | 커피먹고쿠우   | 2019.10.29. |
| 노다가 실검 (Jsoup) ☹️ [22]         | Anseojin | 2019.10.28. |
| 유사도 함수 ☹️ [21]                 | kyleok   | 2019.10.28. |
| 제이썬 인사봇 [2]                    | Anseojin | 2019.10.28. |
| [Rhino] 개선된 세이프 이발(실행) 함수 [13] | OnTheWay | 2019.10.27. |

1 2 3

전체보기

## 이 카페 인기글

가르치기 리로드..

Milk2

♡0 💬9

공식 카카오톡 유희

Lunar

♡0 💬9



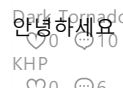
링크를 포함한 글 홍보시 링크 분석(?)하기



틱택토 (카카오링크 적용)



태양, 달 정보 구현 완료



안녕하세요

[ 치\*\*\*, 한\*\* ] 님 강제 탈퇴

AlphaDo  
♡1 💬2

eval 질문  
5.612005  
♡1 💬2  
Kiri  
♡0 💬5

path질문  
젤리파덕  
♡0 💬2