

NLP Assignment 1

Report

Kaustav Vats | 2016048

$$D(i,j) = \min \begin{cases} D(i-1,j) + 1 \\ D(i,j-1) + 1 \\ D(i-1,j-1) + \begin{cases} 2; & \text{if } S_1(i) \neq S_2(j) \\ 0; & \text{if } S_1(i) = S_2(j) \end{cases} \end{cases}$$

Input 1:

How are you?

How r You

Output:

Edits performed on the String:

Replace: you -> You

Replace: are -> r

Min Edit Distance: 4

Assumptions-

1. Strings are case sensitive.
2. Punctuations/Special Character ($[^\\w]$) are not part of edit distance. (They can be removed before calculating edit distance).
3. Either words or two single line sentences would be passed as an input.
4. Total number of words in both the sentence wont exceed max possible limit for creating an int array in java.
5. Input would be given without any extra trailing space after a sentence or word.