Levneshtein Distance

Given two words, it’s defined as the number inserts and deletes of characters necessary to make one word same as the other. The Levenshtein Distance between two words w1 and w2 is calculated as follows

*distance(w1,w2) = length(w1) + length(w2) – 2 \* maxCommonSubSeqLength(w1,w2)  
where  
maxCommonSubSeqLength is the length of longest common sub sequence between w1 and w2.*

The distance is normalized, by dividing the distance with sum of the length of the two words as below, forcing the normalized distance to be always be between 0 and 1.0.

*normDistance(w1,w2) = distance(w1,w2) / (length(w1) + length(w2))*