Dewey Vo

CS 584

September 26, 2024

# In-Class Activity 3: Edit Distance Calculation

## Activity: Calculate Edit Distance (By Hand)

In this activity, you will manually compute the Levenshtein distance between

pairs of words. Each edit operation (insertion, deletion, substitution) has a

cost specified by the edit distance. Follow the steps carefully to calculate the

minimum number of edits required to transform one word into another.

## Instructions

For each word pair below, calculate the Levenshtein distance using the formula

above, where each edit operation (insertion, deletion, substitution) has a cost

of 1.

Word Pairs

• Word Pair 1: apple → applesauce

1. apple + s
2. apples + a
3. applesa + u
4. applesau + c
5. applesauc + e
6. applesauce

Only 5 distance is needed here

• Word Pair 2: kitten → sitting

1. S + itten (replace)
2. Sitt + I + n (replace)
3. Sittin + g (add)
4. sitting

Only 3 distance is needed here

• Word Pair 3: flaw → lawn

1. F – law (remove)
2. Law + n (add)
3. Lawn

Only 2 distance is needed here

• Word Pair 4: intention → execution

1. E + ntention (replace)
2. Ex + tention (replace)
3. Exe + ention (replace)
4. Exec + ntion (replace)
5. Execu + tion (replace)
6. exectuion

Only 5 distance is needed here

## Guidelines for Calculation

• Write down the steps for each transformation (e.g., insert, delete, or sub-

stitute characters) and count the total number of operations.

• Compare your answer with your partner to verify the results.

# Activity 2: Create Your Own Edit Distance (Group Activity)

In this activity, you will form groups of two. Each group will design its own custom edit distance by assigning different costs to the operations (insertion, deletion, substitution). After designing your custom distance, you will exchange your distance rules with another group and ask them to compute the edit distances for a set of word pairs based on your custom distance. Instructions

1. Form Groups: Pair up with another student to form a group of 2.
2. Design Your Edit Distance:
   1. Assign different costs to the three operations (insertion, deletion, substitution). For example, you could assign:

– Insertion = 1

– Deletion = 2

– Substitution = 3

Be creative with your cost assignment! The only rule is that each

operation must have a non-zero cost.

3. Select Word Pairs:

• Choose at least 3 word pairs (they can be as simple or complex as

you’d like) for the other group to compute using your custom edit

distance. For example:

– Word Pair 1: match → scratch

– Word Pair 2: mechanical → chemical

– Word Pair 3: potassium → magnesium

4. Exchange Distances and Word Pairs:

• Once you’ve finalized your custom distance and word pairs, exchange

them with another group. You will receive their custom edit distance

rules and word pairs as well.

5. Calculate the Edit Distance:

• Using the other group’s custom edit distance, calculate the number

of operations required to transform each word into the other. Be sure

to factor in the different costs for insertion, deletion, and substitution

based on the rules given by the other group.

6. Discuss and Compare:

3

• After both groups have completed their calculations, discuss your re-

sults with each other. Compare how the different costs influenced the

calculations and whether certain operations were more ”expensive”

under one system versus another.

Example: Custom Edit Distance

• Group A’s Edit Distance:

– Insertion = 2

– Deletion = 1

– Substitution = 3

• Group A’s Word Pairs:

– book → cook

– cat → bat

– blue → blues

• Group B will compute the edit distance for these word pairs using Group

A’s custom rules.

—

3 Deliverables

At the end of the class, you should submit the following in a word doc, or photo

of CLEARLY handwritten work on Canvas, under ’In-Class Activity 2’ in the

Assignments section:

• Submit your solutions for Activity 1.

• Submit your custom edit distance rules and word pairs for Activity 2,

along with your calculations based on the rules provided by the other

group