

Hide menu

Lecture: Autocorrect and Minimum Edit Distance

Video: Intro to Course 2
1 min

Video: Week Introduction
55 sec

Video: Overview
1 min

Reading: Overview
3 min

Video: Autocorrect
2 min

Reading: Autocorrect
4 min

Video: Building the model
4 min

Reading: Building the model
3 min

Lab: Lecture notebook: Building the vocabulary
1 h

Video: Building the model II
3 min

Reading: Building the model II
4 min

Lab: Lecture notebook: Candidates from edits
1 h

Video: Minimum edit distance
3 min

Reading: Minimum edit distance
5 min

Video: Minimum edit distance algorithm
5 min

Reading: Minimum edit distance algorithm
3 min

Video: Minimum edit distance algorithm II
4 min

Reading: Minimum edit distance algorithm II
5 min

Video: Minimum edit distance algorithm III
3 min

Reading: Minimum edit distance III
4 min

Video: Week Conclusion
51 sec

Ungraded App Item: [IMPORTANT] Have questions, issues or ideas? Join our Community!
10 min

Week 1 Minimum edit distance algorithm II

Minimum edit distance algorithm II

To populate the following table:

Source: play → Target: stay
Cost: insert: 1, delete: 1, replace: 2

$$D[i, j] = \min \begin{cases} D[i-1, j] + del_cost \\ D[i, j-1] + ins_cost \\ D[i-1, j-1] + \begin{cases} rep_cost; & \text{if } src[i] \neq tar[j] \\ 0; & \text{if } src[i] = tar[j] \end{cases} \end{cases}$$

		0	1	2	3	4
	#	#	s	t	a	y
0	#	0	1	2	3	4
1	p	1	2			
2	l	2				
3	a	3				
4	y	4				

There are three equations:

- D[i,j] = D[i-1, j] + del_cost:** this indicates you want to populate the current cell (i,j) by using the cost in the cell found directly above.
- D[i,j] = D[i, j-1] + ins_cost:** this indicates you want to populate the current cell (i,j) by using the cost in the cell found directly to its left.
- D[i,j] = D[i-1, j-1] + rep_cost:** the rep cost can be 2 or 0 depending if you are going to actually replace it or not.

At every time step you check the three possible paths where you can come from and you select the least expensive one. Once you are done, you get the following:

Source: play → Target: stay
Cost: insert: 1, delete: 1, replace: 2

		0	1	2	3	4
	#	#	s	t	a	y
0	#	0	1	2	3	4
1	p	1	2	3	4	5
2	l	2	3	4	5	6
3	a	3	4	5	4	5
4	y	4	5	6	5	4

Mark as completed

Like Dislike Report an issue

