Information Retrieval
Edit Distance
Lecture-13

Prepared By

Dr. Rasmita Dash & Dr. Rasmita Rautray

Associate Professor

Dept. of CSE

Content

- Levenshtein Edit distance

Example

1. Compute Levenshtein distance matrix for OSLO − SNOW
2. What are the Levenshtein editing operations that transform *cat* into *catcat*?

			s		n		o		w	
		0	1	1	2	2	3	3	4	4
o		$\frac{1}{1}$								
s		$\frac{2}{2}$								
l		$\frac{3}{3}$								
o		$\frac{4}{4}$								

		s		n		o		w	
	0	1	1	2	2	3	3	4	4
o	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{?}$						
s	$\frac{2}{2}$								
l	$\frac{3}{3}$								
o	$\frac{4}{4}$								

		s	n	o	w
	0	1 1	2 2	3 3	4 4
o	$\frac{1}{1}$	$\frac{1}{2}$ $\frac{2}{1}$			
s	$\frac{2}{2}$				
l	$\frac{3}{3}$				
o	$\frac{4}{4}$				

			s		n		o		w	
		$\frac{}{0}$	1	1	2	2	3	3	4	4
o	$\frac{1}{1}$		1 2	2 1	2 2	3 ?				
s	$\frac{2}{2}$									
l	$\frac{3}{3}$									
o	$\frac{4}{4}$									

		s	n	o	w
	0	1 1	2 2	3 3	4 4
o	1 / 1	1 2 / 2 1	2 3 / 2 2		
s	2 / 2				
l	3 / 3				
o	4 / 4				

		s		n		o		w	
	0	1	1	2	2	3	3	4	4
o	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{1}$	$\frac{2}{2}$	$\frac{3}{2}$	$\frac{2}{3}$	$\frac{4}{?}$		
s	$\frac{2}{2}$								
l	$\frac{3}{3}$								
o	$\frac{4}{4}$								

		s		n		o		w	
		0		1	1	2	2	3	3	4	4
o	1/1	1 2	2 1	2 3	2 2	2 4	3 2		
s	2/2								
l	3/3								
o	4/4								

		s		n		o		w	
	0	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**
o	**1** / **1**	**1** / 2	2 / **1**	**2** / 2	3 / **2**	**2** / 3	4 / **2**	4 / 3	5 / ?
s	**2** / **2**								
l	**3** / **3**								
o	**4** / **4**								

		s		n		o		w	
	0	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**
o	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{1}$	$\frac{2}{2}$	$\frac{3}{2}$	$\frac{2}{3}$	$\frac{4}{2}$	$\frac{4}{3}$	$\frac{5}{3}$
s	$\frac{2}{2}$	$\frac{1}{3}$	$\frac{2}{?}$						
l	$\frac{3}{3}$								
o	$\frac{4}{4}$								

		s		n		o		w	
	0	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**
o	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{1}$	$\frac{2}{2}$	$\frac{3}{2}$	$\frac{2}{3}$	$\frac{4}{2}$	$\frac{4}{3}$	$\frac{5}{3}$
s	$\frac{2}{2}$	$\frac{1}{3}$	$\frac{2}{1}$						
l	$\frac{3}{3}$								
o	$\frac{4}{4}$								

		s		n		o		w	
	$\frac{}{0}$	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**
o	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{1}$	$\frac{2}{2}$	$\frac{3}{2}$	$\frac{2}{3}$	$\frac{4}{2}$	$\frac{4}{3}$	$\frac{5}{3}$
s	$\frac{2}{2}$	$\frac{1}{3}$	$\frac{2}{1}$	$\frac{2}{2}$	$\frac{3}{?}$				
l	$\frac{3}{3}$								
o	$\frac{4}{4}$								

		s		n		o		w	
	0	1	1	2	2	3	3	4	4
o	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{1}$	$\frac{2}{2}$	$\frac{3}{2}$	$\frac{2}{3}$	$\frac{4}{2}$	$\frac{4}{3}$	$\frac{5}{3}$
s	$\frac{2}{2}$	$\frac{1}{3}$	$\frac{2}{1}$	$\frac{2}{2}$	$\frac{3}{2}$				
l	$\frac{3}{3}$								
o	$\frac{4}{4}$								

		s		n		o		w	
	0	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**
o	$\frac{1}{1}$	**1** / 2	2 / **1**	**2** / 2	3 / **2**	**2** / 3	4 / **2**	4 / **3**	5 / **3**
s	$\frac{2}{2}$	**1** / 3	2 / **1**	**2** / 2	3 / **2**	3 / 3	3 / ?		
l	$\frac{3}{3}$								
o	$\frac{4}{4}$								

		s		n		o		w	
	0	1	1	2	2	3	3	4	4
o	1/1	1/2	2/1	2/2	3/2	2/3	4/2	4/3	5/3
s	2/2	1/3	2/1	2/2	3/2	3/3	3/3		
l	3/3								
o	4/4								

		s	n	o	w
	0	1 1	2 2	3 3	4 4
o	$\frac{1}{1}$	1 2 / 2 1	2 3 / 2 2	2 4 / 3 2	4 5 / 3 3
s	$\frac{2}{2}$	1 2 / 3 1	2 3 / 2 2	3 3 / 3 3	3 4 / 4 ?
l	$\frac{3}{3}$				
o	$\frac{4}{4}$				

		s		n		o		w	
	0	1	1	2	2	3	3	4	4
o	1/1	1/2	2/1	2/2	3/2	2/3	4/2	4/3	5/3
s	2/2	1/3	2/1	2/2	3/2	3/3	3/3	3/4	4/3
l	3/3								
o	4/4								

		s		n		o		w	
	0	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**
o	**1**/**1**	**1**/2	2/**1**	**2**/**2**	3/2	**2**/3	4/**2**	4/**3**	5/**3**
s	**2**/**2**	**1**/3	2/**1**	**2**/**2**	3/2	**3**/**3**	3/**3**	**3**/4	4/**3**
l	**3**/**3**	3/4	**2**/**2**	**2**/3	3/**2**	**3**/**3**	4/**3**	**4**/**4**	**4**/**4**
o	**4**/**4**	4/5	**3**/**3**	**3**/4	3/**3**	**2**/4	4/**2**	4/**3**	5/**3**

		s	n	o	w
	0	**1** **1**	**2** **2**	**3** **3**	**4** **4**
o	**1** / **1**	**1** 2 / 2 **1**	**2** 3 / **2** **2**	**2** 4 / 3 **2**	4 5 / **3** **3**
s	**2** / **2**	**1** 2 / 3 **1**	**2** 3 / **2** **2**	**3** **3** / **3** **3**	**3** 4 / **4** **3**
l	**3** / **3**	3 **2** / 4 **2**	**2** 3 / 3 **2**	**3** 4 / **3** **3**	**4** **4** / **4** **4**
o	**4** / **4**	4 **3** / 5 **3**	**3** **3** / 4 **3**	**2** 4 / 4 **2**	4 5 / **3** **3**

		s	n	o	w
	0	**1** 1	**2** 2	**3** 3	**4** 4
o	**1** / **1**	1 2 / 2 1	2 3 / 2 2	2 4 / 3 2	4 5 / 3 3
s	**2** / **2**	1 2 / 3 1	2 3 / 2 2	3 3 / 3 3	3 4 / 4 3
l	**3** / **3**	3 2 / 4 2	2 3 / 3 2	3 4 / 3 3	4 4 / 4 4
o	**4** / **4**	4 3 / 5 3	3 3 / 4 3	2 4 / 4 2	4 5 / 3 3

cost	operation	input	output
1	delete	o	*
0	(copy)	s	s
1	replace	l	n
0	(copy)	o	o
1	insert	*	w

Thank You