

# Assignment No #5 IDS

Name :- Adnan Safdar  
Reg No :- SP20-BLS-116

Q:-1

$S_1 \Rightarrow$  "Sunshine state enjoy sunshine"  
 $S_2 \Rightarrow$  "brown fox jump high, brown fox run"  
 $S_3 \Rightarrow$  "Sunshine state fox run fast"

BOG:-

$S_1 \Rightarrow$  Sunshine state enjoy.

$S_2 \Rightarrow$  brown, fox, jump, high, run.

$S_3 \Rightarrow$  Sunshine, state, fox, run, fast.

	Sunshine	state	enjoy	brown	fox	Jump
$S_1$	2	1	1	0	0	0
$S_2$	0	0	0	2	2	1
$S_3$	1	1	0	0	1	0

	high	run	fast
$S_1$	0	0	0
$S_2$	1	1	0
$S_3$	0	1	1

Term Frequency

$S_1$

Sunshine  $\Rightarrow 2/4$

state  $\Rightarrow 1/4$

enjoy  $\Rightarrow 1/4$



S<sub>2r</sub>  
 Brown  $\Rightarrow 2/7$   
 Fox  $\Rightarrow 2/7$   
 Jump  $\Rightarrow 1/7$   
 High  $\Rightarrow 1/7$   
 Run  $\Rightarrow 1/7$

S<sub>3r</sub>

Sunshine  $\Rightarrow 1/5$   
 state  $\Rightarrow 1/5$   
 Fox  $\Rightarrow 1/5$   
 Run  $\Rightarrow 1/5$   
 Fast  $\Rightarrow 1/5$

Inverse Document Frequency

Sunshine  $\Rightarrow \log(3/2) \Rightarrow 0.176$   
 state  $\Rightarrow \log(3/2) \Rightarrow 0.176$   
 Enjoy  $\Rightarrow \log(3/1) \Rightarrow 0.477$   
 brown  $\Rightarrow \log(3/2) \Rightarrow 0.477$   
 Fox  $\Rightarrow \log(3/2) \Rightarrow 0.176$   
 Jump  $\Rightarrow \log(3/1) \Rightarrow 0.477$   
 High  $\Rightarrow \log(3/1) \Rightarrow 0.477$   
 Run  $\Rightarrow \log(3/2) \Rightarrow 0.176$   
 Fast  $\Rightarrow \log(3/1) \Rightarrow 0.477$

TF-IDF

S<sub>1</sub>

Sunshine  $2/4 \times 0.176 \Rightarrow 0.088$   
 state  $1/4 \times 0.176 \Rightarrow 0.044$   
 Enjoy  $1/4 \times 0.477 \Rightarrow 0.1192$



S<sub>2r</sub>

Brown  $\rightarrow \frac{2}{7} \times 0.477 \Rightarrow 0.136$

Fox  $\rightarrow \frac{2}{7} \times 0.176 \Rightarrow 0.051$

Jump  $\rightarrow \frac{1}{7} \times 0.477 \Rightarrow 0.068$

high  $\rightarrow \frac{1}{7} \times 0.477 \Rightarrow 0.068$

Run  $\rightarrow \frac{1}{7} \times 0.176 \Rightarrow 0.025$

S<sub>3r</sub>

Sunshine  $\Rightarrow \frac{1}{5} \times 0.176 \Rightarrow 0.0352$

state  $\Rightarrow \frac{1}{5} \times 0.176 \Rightarrow 0.0352$

Fox  $\Rightarrow \frac{1}{5} \times 0.176 \Rightarrow 0.0352$

Run  $\Rightarrow \frac{1}{5} \times 0.176 \Rightarrow 0.0352$

Fast  $\Rightarrow \frac{1}{5} \times 0.477 \Rightarrow 0.0954$

TF-IDF Table

	<u>d<sub>1</sub></u>	<u>d<sub>2</sub></u>	<u>d<sub>3</sub></u>
Sunshine	0.088	0	0.0352
state	0.044	0	0.0352
enjoy	0.11925	0	0
brown	0	0.136	0
Fox	0	0.051	0
Jump	0	0.068	0.0352
high	0	0.068	0
Run	0	0.025	0.0352
Fast	0	0	0.0954

Q:-2

Similarity b/w  $S_2$  and  $S_3$  ~

$$\frac{\vec{S_2} \cdot \vec{S_3}}{|\vec{S_2}| \cdot |\vec{S_3}|}$$

Document vector ~

$$S_2 = [2, 1, 1, 0, 0, 0, 0, 0, 0]$$

$$S_3 = [1, 1, 0, 0, 1, 0, 0, 1, 1]$$

$$\vec{S_2} \cdot \vec{S_3} \Rightarrow (2 \times 1) + (1 \times 1) + (1 \times 0) + (0 \times 0) + (0 \times 1) + (0 \times 0) + (0 \times 0) + (0 \times 1) + (0 \times 1)$$

$$\Rightarrow 2 + 1$$

$$\Rightarrow 3$$

$$|\vec{S_2}| \Rightarrow 2 \times 2 + 1 \times 1 + 1 \times 1$$

$$\Rightarrow 6$$

$$|\vec{S_3}| \Rightarrow 1 \times 1 + 1 \times 1 + 1 \times 1 + 1 \times 1 + 1 \times 1$$

$$\Rightarrow 5$$

Now

$$\frac{3}{(6)(5)}$$

$$\frac{3}{30} = \frac{1}{10}$$

$$= \frac{1}{10}$$

$$\cos(S_2, S_3) \Rightarrow 0.1$$