Name :- NAMAT ULLAH
Reg no :- FAM-BCS-403
Assignment no :- # 5

Question 1 :- S_1 = " sunshine state enjoy sunshine"
S_2 = " brown fox jump high brown fox run"
S_3 = "sun shine real fox run fast"

⇒ for Bow -

	sunshine	state	enjoy	brown	fox	jump	high	run	fast
S_1	2	1	1	0	0	0	0	0	0
S_2	0	0	0	2	2	1	1	1	0
S_3	1	1	0	0	1	0	0	1	1

TF ⇒

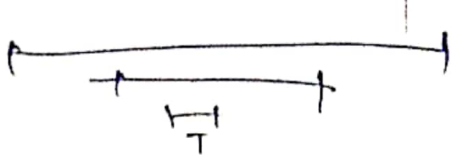
	sunshine	state	enjoy	brown	fox	jump	run	fast	high
S_1	2/4	1/4	1/4	0	0	0	0	0	0
S_2	0	0	0	2/7	2/7	1/7	1/7	0	1/7
S_3	1/5	1/5	0	0	1/5	0	1/5	1/5	0

IDF ⇒

- Sunshine = $\log(3/2) = 0.176$
- State = $\log(3/2) = 0.176$
- enjoy = $\log(3/1) = 0.471$
- brown = $\log(3/1) = 0.477$
- fox = $\log(3/2) = 0.176$
- jump = $\log(3/1) = 0.477$
- hight = $\log(3/1) = 0.477$
- run = $\log(3/2) = 0.176$
- fast = $\log(3/1) = 0.477$

⇒ TF - IDF ⇒ $(tf\_idf)_{t,d} = tf_{t,d} * idf_t$

	Sunshine	State	enjoy	brown	fox	jump	high	run	fast
S1	0.088	0.044	0.119	0	0	0	0	0	0
S2	0	0	0	0.135	0.05	0.068	0.068	0.025	0
S3	0.035	0.035	0	0	0.035	0	0	0.035	0.095



Name: NAmat ullah
Reg no: FA19-BCS-403
Assignmt no:- #5

Question no 2:- (Vectors) — $\frac{S_1 \cdot S_3}{|S_1| \times |S_3|}$

$S_1 = [2, 1, 1, 0, 0, 0, 0, 0, 0]$
$S_2 = [0, 0, 0, 2, 2, 1, 1, 1, 0]$
$S_3 = [1, 1, 0, 0, 1, 0, 0, 1, 1]$

→ Cosine-similarity b/w S_1 and S_3 :-

$S_1 * S_3 = (2*1) + (1*1) + (1*0) + (0*0) + (0*1) +$
$(0*0) + (0*0) + (0*1) + (0*1) \Rightarrow$

$\Rightarrow 2 + 1 + 0 + 0 + 0 + 0 + 0 + 0 + 0 \Rightarrow 3$ ✓

$|S_1| = (2^2 + 1^2 + 1^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2)^{1/2}$

$\Rightarrow (4+1+1)^{1/2} = (6)^{1/2} = 2.45$

$|S_3| = (1^2 + 1^2 + 0 + 0 + 1^2 + 0 + 0 + 1^2 + 1^2)^{1/2}$

$\Rightarrow (1+1+1+1+1)^{1/2} = (5)^{1/2} = 2.24$

$\cos(S_1, S_3) = \frac{S_1 \cdot S_3}{|S_1|\,|S_3|} = \frac{3}{(2.45)(2.24)} \approx 0.55$

answer ✓