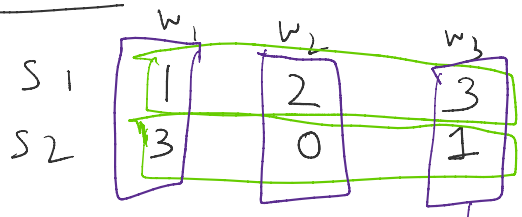
$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \cdot \|\vec{b}\|}$$

↓

$\|\vec{a}\| \cdot \|\vec{b}\|$

$(0,1)$

Input Matrix



$\left(\frac{1.2+3.0}{\sqrt{10}\ \sqrt{4}}\right)$

		S_1	S_2
battery =	W_1	1	3
drain =	W_2	2	0
	W_3	3	1

$$\frac{1.3+2.0+3.1}{\sqrt{14}\cdot\sqrt{10}} = \frac{6}{2\sqrt{35}} = \frac{3}{\sqrt{35}}$$

	S_1	S_2	S_3	S_4	S_5
S_1	1	$\frac{3}{\sqrt{35}}$			
S_2	$\frac{3}{\sqrt{35}}$	1			
S_3					
S_4					
S_5					

3×3 cosine similarity

---

① Language models

↳ "n-gram generator"

② Word2 Vec

↳ ① SVD, PCA

↳ ② Neural Networks

"Word Embeddings"

Word = [—, —, —, —, —]

↳ Because cat is hungry, give food.

↳ Because dog is hungry, give "bone"

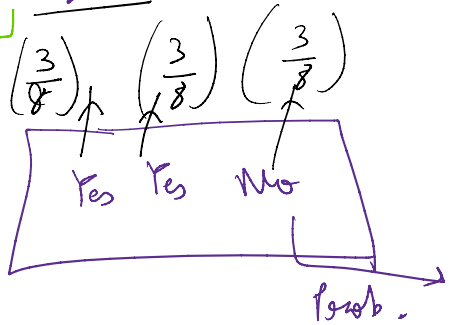
→ Bigram model ← Trigram model

Every piece of text of finite length gets some prob. associated with it.

"I ate a cherry" ← Pr(↑)

"Eye eight a cherry" ← (Pr ↓)

→ Yes No Maybe No Yes Yes No Maybe



V = 3 → {Yes, No, Maybe}

P(Yes Yes No)

= P(w1 = Yes, w2 = Yes, w3 = No)

= P(w1 = Yes) · P(w2 = Yes | w1 = Yes) · P(w3 = No | (w1 = Yes, w2 = Yes))

(3/8) (1/3) P(w3 = No | w2 = Yes) (2/3)

→ [A class happening online on 23rd June in Roorkee] ←

P(w9 = Roorkee | (w1 = A, w2 = class, ---, w8 = in))

≈ P(w9 = Roorkee | w8 = in)

+ Unigram model

+ Unigram model

① "Bigram model" →

② (Trigram model) →

= "n-gram models"

+

Language models

"State of Art Models"

"n- gram model as random text generator"

Yes — Yes — No — Maybe —

← Bigram model.

Yes:
- $\frac{2}{3}$ prob → No
- $\frac{1}{3}$ prob → Yes

No:
- $\frac{2}{3}$ prob → Maybe
- $\frac{1}{3}$ prob → Yes

Maybe:
- $\frac{1}{2}$ → No
- $\frac{1}{2}$ → "STOP"