25.10.24
Friday.

(I) Syntax

(II) Semantics

word level

1. word sense disambiguation – same word in different sense
2. Semantic role labelling
3. Semantic parser
4. Text entailment

sentence level

He has strong interest in NLP

I deposited money to get interest

(III) Pragmatics / Discourse

Semantic role labelling – Agent, patient, Source, destination, Instrument

eg: John drove Mary from Delhi to Chennai in his car

Semantic parser – How many cities are there in US

↓

answer (A, count (B, city (B), loc (B, C), country (C, name (USA))), A)

↓

country id

Text entailment - Eyeing a huge market potential, currently led by google, yahoo. Took over search company minutes in last one year.

↳ yahoo bought overture

(3) Pragmatics / Discourse / Anaphora resolution -

John put the carrot on the plate & ate it.

↓

carrot or plate ?

Ellipses - wise men talk because they have something to say, fools because they have to say something.

↓

talk is implied here

machine will answer that

Statistical NLP :-

1. Cosine Similarity

eg: (i) Bear is funny and dangerous

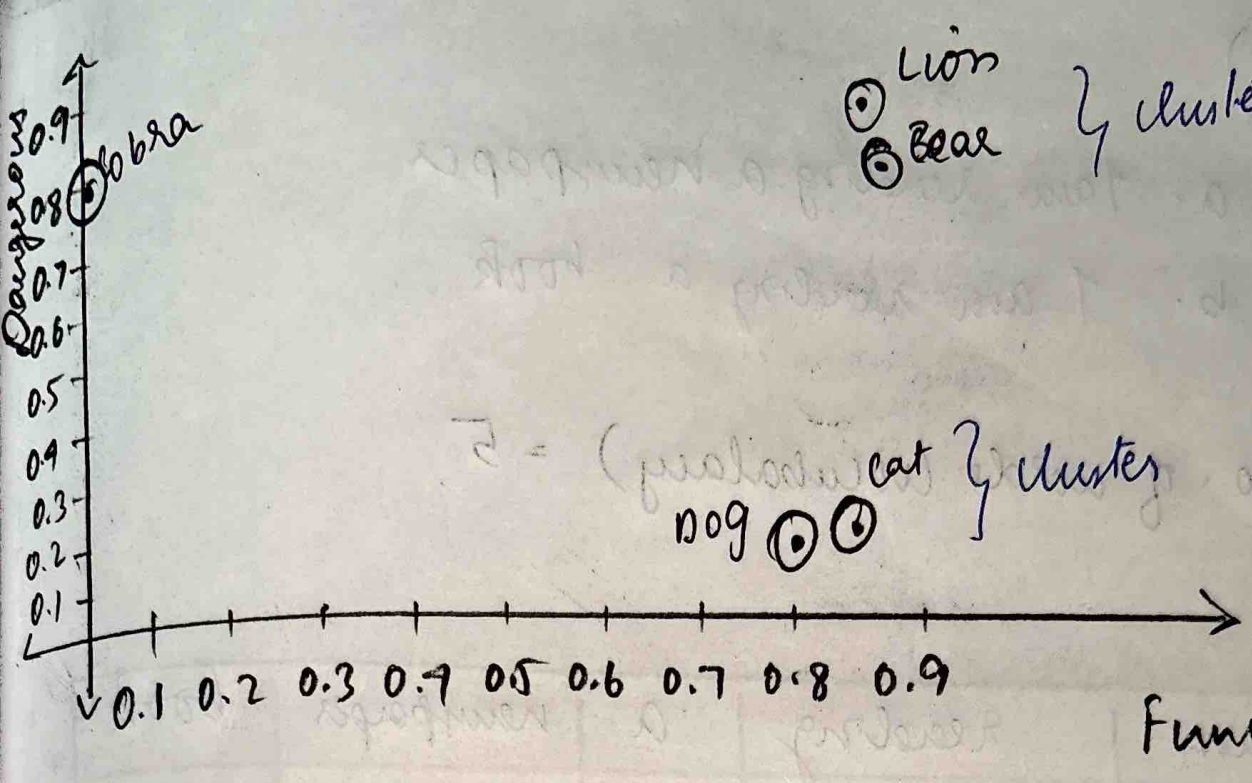
(ii) cat is funny and not dangerous

(iii) cobra is not funny and dangerous

(iv) Lion is funny & dangerous

(v) Dog is funny and not dangerous

	Funny	Dangerous.
Bear	0.9	0.85
cat	0.85	0.15
cobra	0.0	0.8
Lion	0.85	0.9
Dog	0.8	0.15



$$\cos(a,b) = \frac{\sum a_i b_i}{\sqrt{\sum a^2}\sqrt{\sum b^2}}$$

a, b are animals

$$\cos(\text{lion}, \text{dog}) = \frac{(0.85)(0.8) + (0.9)(0.15)}{\sqrt{(0.85)^2 + (0.91)^2} * \sqrt{(0.8)^2 + (0.15)^2}} = 0.809$$

$\cos(\text{cobra}, \text{dog}) = 0.074$

$\cos(\text{lion}, \text{bear}) = 0.998$

$\cos(\text{cobra}, \text{lion}) =$

(Document)

a. I am reading a newspaper

b. I am reading a book

Total no. of words (vocabulary) = 5

V	I am	reading	a	newspaper	book
(a)	1	1	1	1	0
(b)	1	1	1	0	1

$$\cos(a,b) = \frac{a \cdot b}{|a||b|} = \frac{1.1 + 1.1 + 1.1 + 1.0 + 0.}{\sqrt{4} * \sqrt{4}}$$

$$= 0.75$$

c) you are playing tennis

$\cos(a, c) = 0.$