

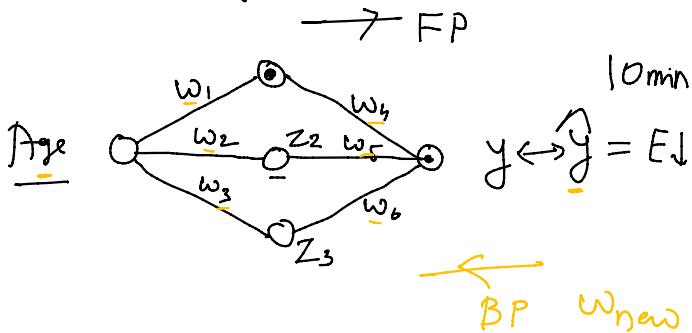
## Agenda

1. Revisit ANN & RNN, LSTM
2. Word2Vec
3. Implement LSTM with W2V
4. Collecting face images using  
OpenCV.

→ —

ANN  
→ FP  $z_1 = f(Age \times w + b)$

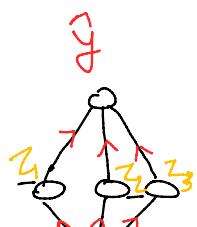
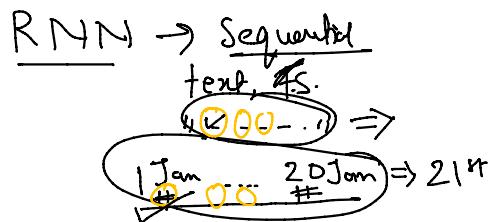
$$f(z) = f(z_1 \times w_4 + z_2 \times w_5 + z_3 \times w_6 + b) = y$$

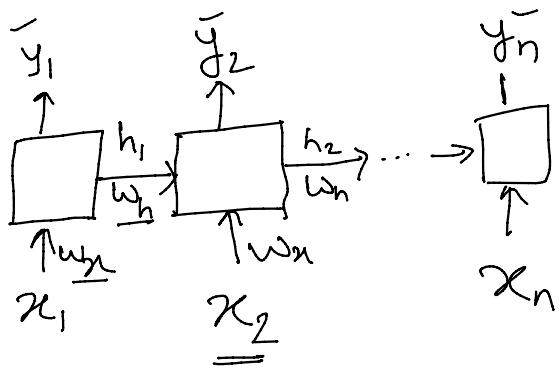
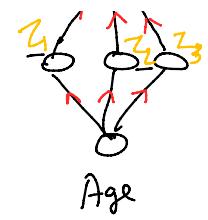


model = Sequential()  $\Rightarrow$  ANN

layer = Dense (units=#, activation="")  
model.add(layer)  $\rightarrow$  SGD, MB ...  
model.compile (optimizer, metric, loss)  
model.fit(xt, yt, epochs = , -, )  $\rightarrow$  ac. mae

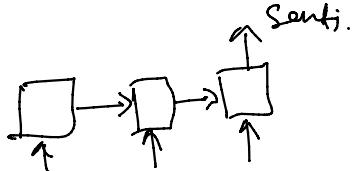
→ —





1. Many to one (Seq. to vec.)

$[- - - - -] \Rightarrow +ve / -ve$   
 $0 / 1$

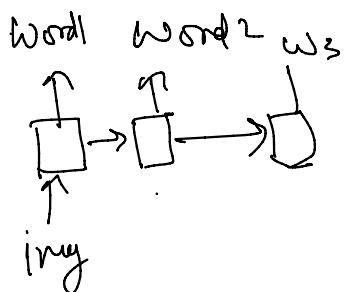


2. One to many (vec. to seq.)

image caption

$\Rightarrow [- - - - -]$

vect.

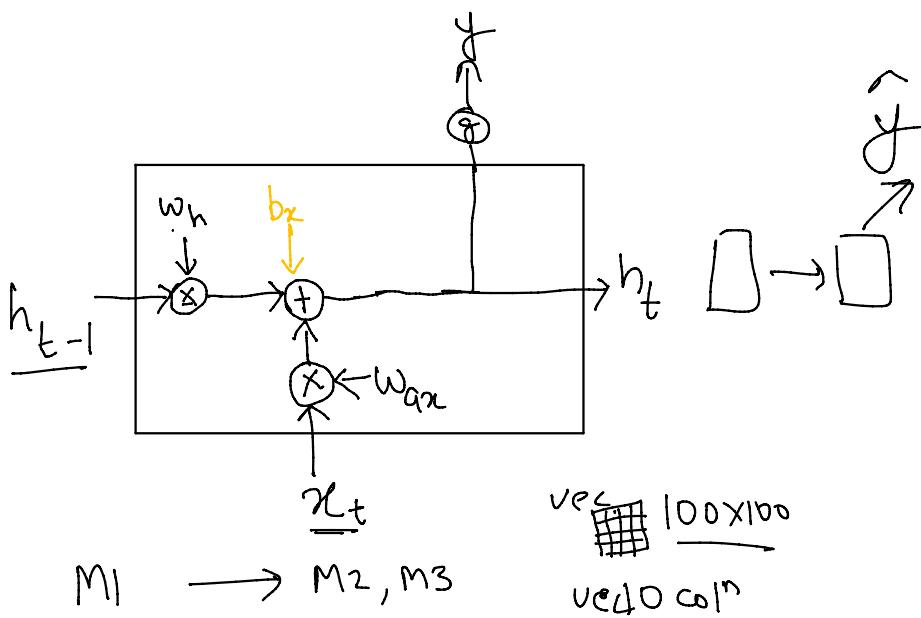
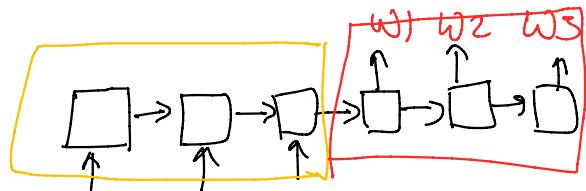


3. Seq. to Seq. (many to many)

### 3. Seq to Seq (many to many)

machine translation

Zulu [ - - - - ]  $\Rightarrow$  Eng [ - - - - ]



[cont + raw + plot + game]  $\rightarrow$   $\boxed{\text{Vec}}$

Seq

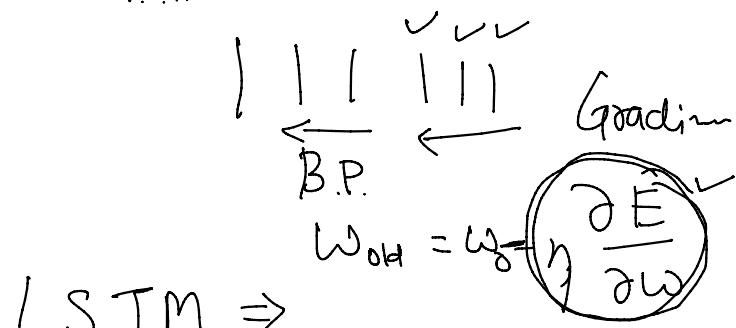
$h_t \rightarrow \boxed{\text{RNN}}$   $\rightarrow$   $h_t$  Short term many

I didn't like the product . . .

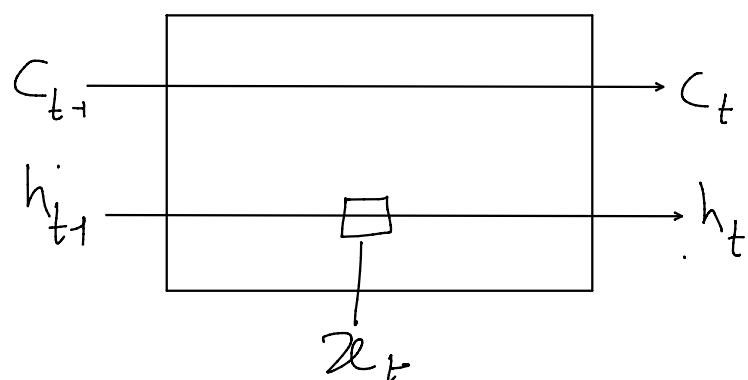
But after using it for 2 good prod.

~ The 1st half of the more said

$\Rightarrow$  The 1st half of the more good  
but 2nd half was quite boring



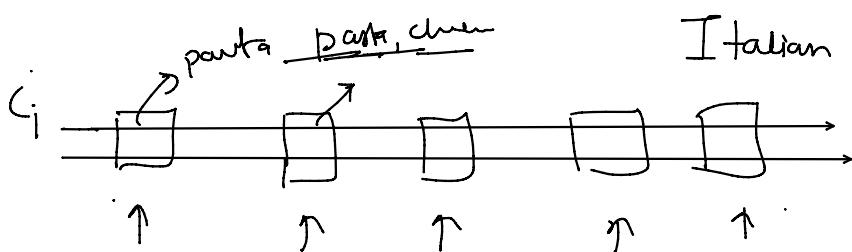
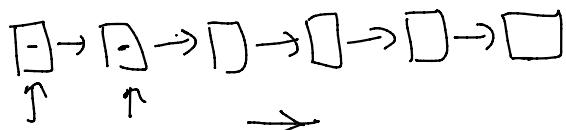
long short term memory



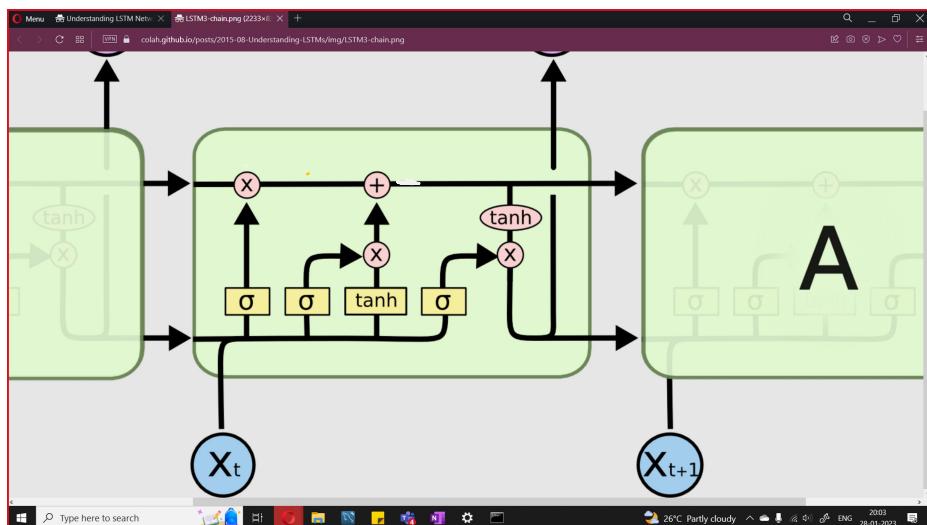
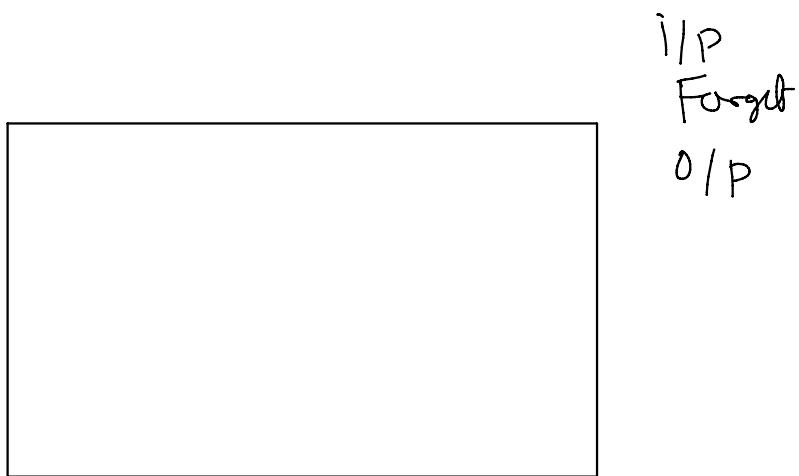
— — — — now

Ahmed eats pasta & cheese, --  
-- his fav. cuisine is italian.

RNN



I P T R P T  
para



LSTM (→)

W2V  $\Rightarrow$  CBOW, Skipgram

BOW, OHE, TF-IDF,

~~✓~~ I like this book

~~✓~~ I don't like this book

OHE

I like this

I: [1 0 0]

like: [0 1 0]

this: [0 0 1]

→ I like this book don't

A horizontal number line starting at 0 and ending at 1. There are four tick marks on the line, including the endpoints. A red arrow points to the right above the line. A red box encloses the number 1.

Word  $\Rightarrow$  vector

The diagram illustrates the probability distribution of word lengths for two different texts. The top row, labeled 'book', shows a distribution with four words of length 1, one word of length 2, and one word of length 3. The bottom row, labeled 'magazine', shows a distribution with one word of length 1, one word of length 2, and four words of length 3. The probabilities for each word length are indicated in the boxes.

Word Length	book	magazine
1	0.5, 0.5, 0.5, 0.5	0.6, 0.4
2	0.6, 0.33	0.09, 0.09
3	0.33, 0.33	0.41, 0.41, 0.41, 0.41

3 billion 

Man Woman King Queen Apple

## Genes

Royalties 0.00

Age | 0.00

P

190

King - Man + Queen = Woman

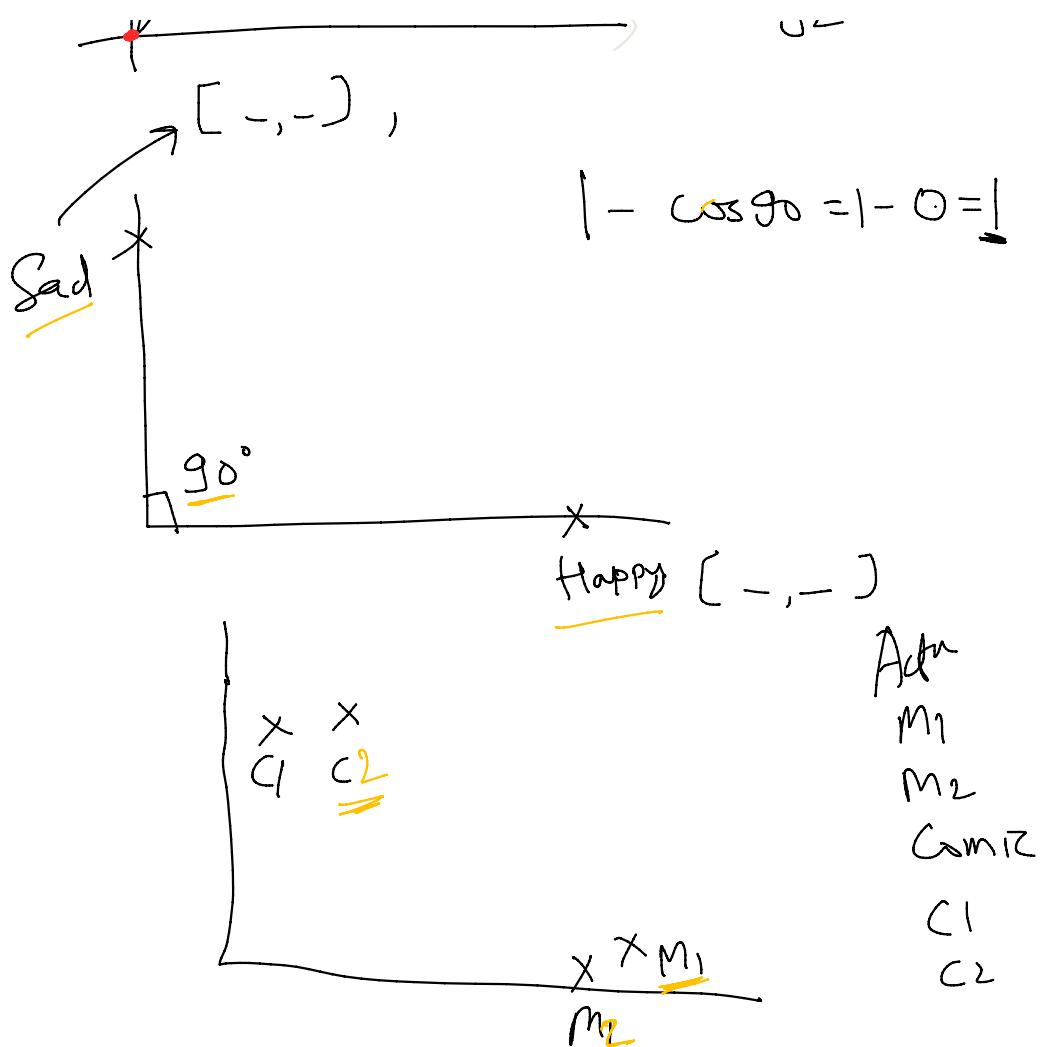
→  
Mom

[-, -]  
woman

$$\text{Dist} = 1 - \cos \theta$$

1 - 60s 45°

$$1 - \frac{1}{\sqrt{2}} = 0.2 \dots$$



Corpus  $\rightarrow$  Vocab size

3 B

tent data = Word 2 Vec  $\Rightarrow$  Sup  $\Rightarrow$   $\frac{1}{1} \Rightarrow \frac{0}{1} =$  CBOW

Digital Regenesys offers courses in Data Sci. ←

Window size = 5 / odd no. ✓ 300

i/p

Digital, Regenrys, courses, in  
Regenyr, offers, in, Data

O/P Transfer Learning  
3B, 300  
offers  
Courses  
in

122V 300

Digital  
Req  
Courses  
in

# Neuron  
= window size

Offer  
0.9  
0.2  
0.1  
0.66  
0.03  
.

$i/L$

$W_j$

$H/L$

$O/L$

Diagram illustrating an RNN/LSTM sequence processing model. The input sequence is 1, 2, 3, 4. The output sequence shows hidden states (hs) and cell states (cs) for each input. The cell state for input 4 is highlighted in an oval.

