

① History of NLP.

Sat & Sun

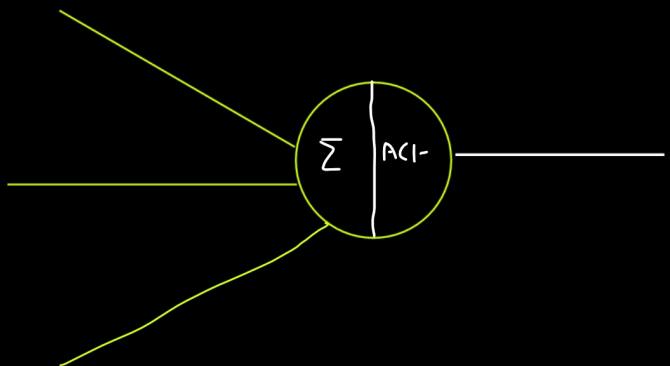
② Text Data Handling

- ↳ Preprocess
- ↳ Encoding
- ↳ Embedding (word2vec)

Deep Learning

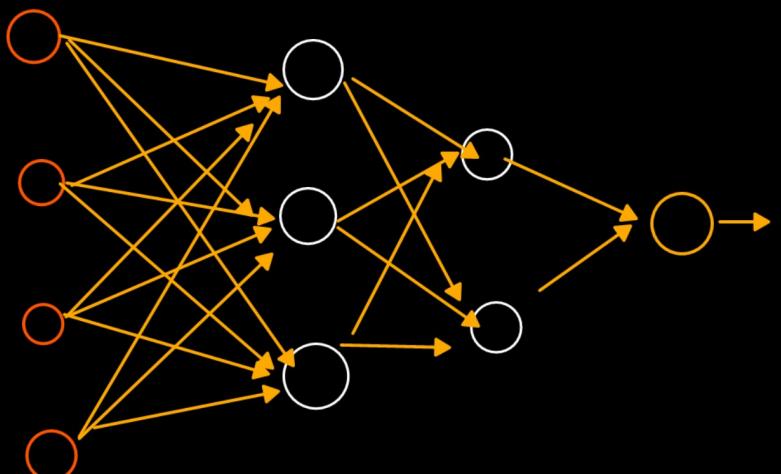
- ① ANN
- ② CNN
- ③ RNN
- ④ GAN / Autoencoder
- ⑤ RF

Perceptron

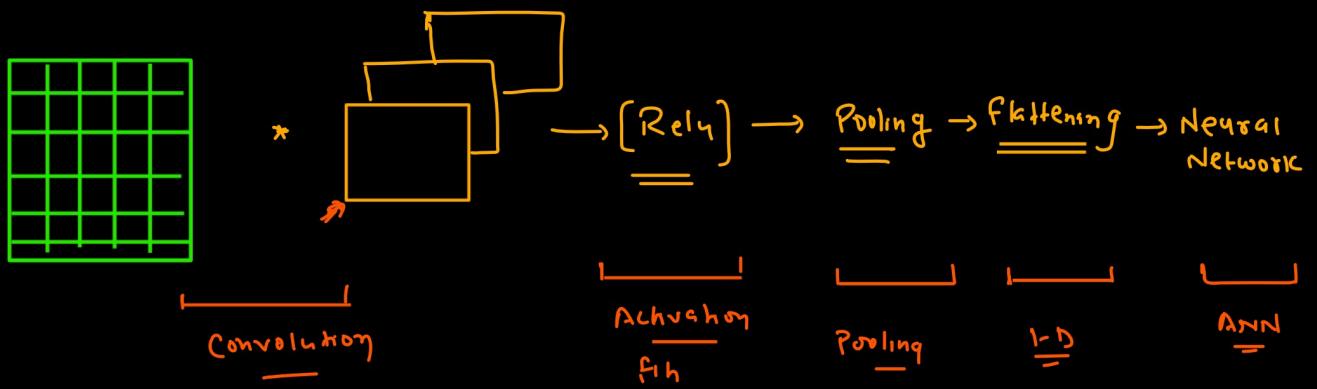


MLP

ANN

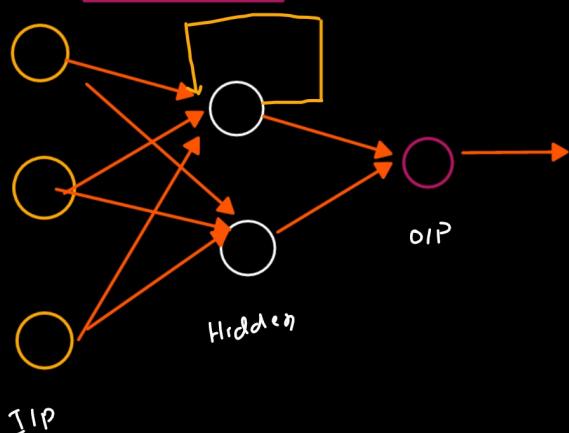


② CNN Pixel [0-255]



③ RNN (Recurrent Neural Network)

⇒ Process the Data based on Time stamp



② LSTM

③ GRU

RNN, LSTM, GRU

Seq to Seq mapping

① One to many technique

② many to one

→ ③ many to many

- ① Synchronize (Same Length)

- ② ASynchronize (Different length) ←

④ Encoder- Decoder (2014) Google AI

⑤ Attention Mechanism

⑥ Transformer (self-Attention)

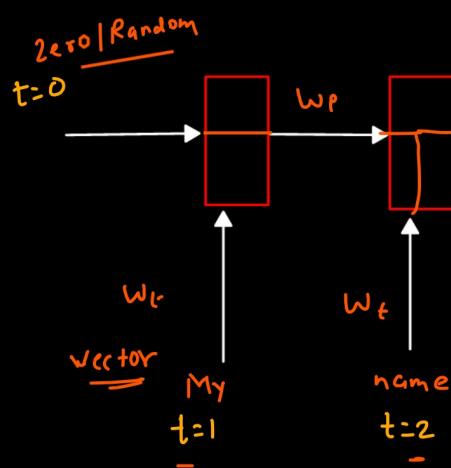
⑦ Transfer Learning, Finetuning

⑧ Few LLM

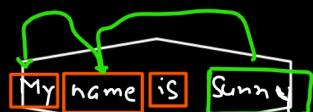
⑨ Chatgpt-

1 RNN

1980 1990



OIP

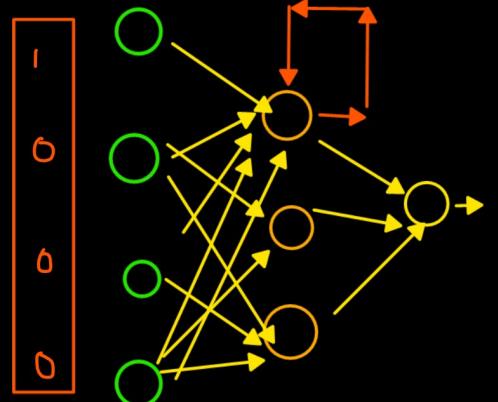


↓

+x1-

Sequence

$\begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$
 $\begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}$
 $\begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$

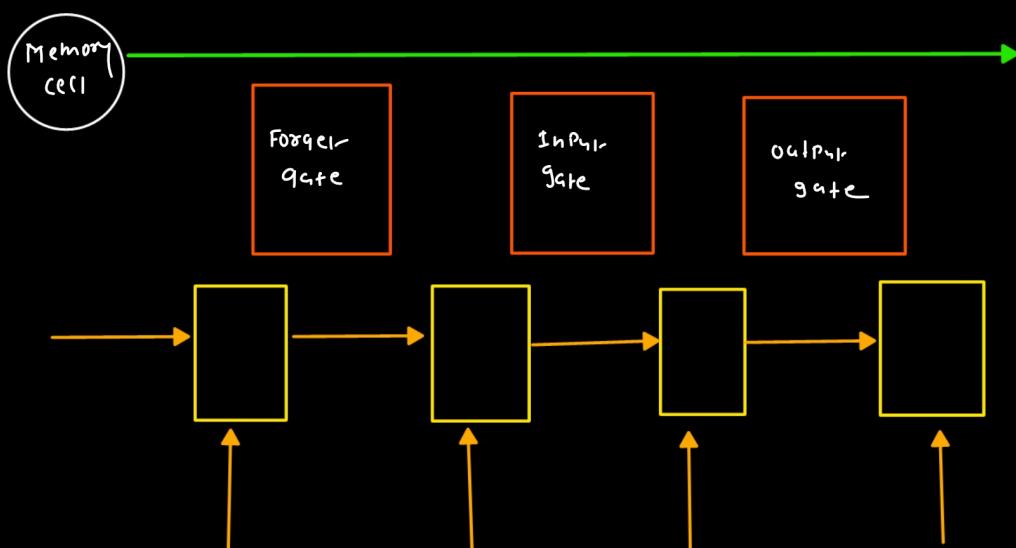


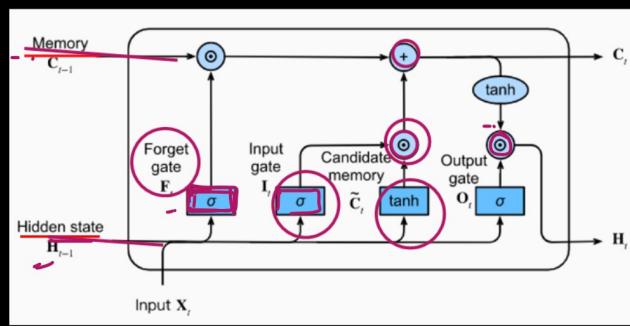
Problem \Rightarrow Longer sequence

{ my sentence is getting short
it was not able to sustain the context }
 { vanishing gradient }

2 LSTM (Long short-term memory)

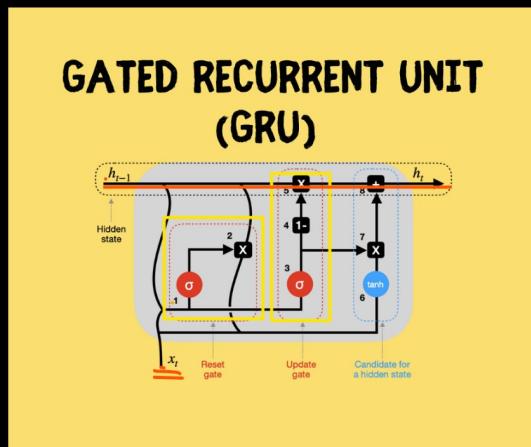
1997





3 GRU

2014



Seq to Seq (Mapping)

1 ONE to MANY

<Caption>

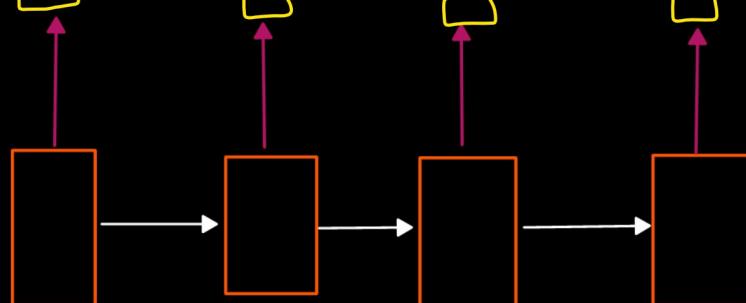
$t=1 \text{ OIP}$

Ex: Image Captioning ($\text{NN} \rightarrow \text{Image}$ Processing)
($\text{RNN} \rightarrow \text{Caption}$)

$t=2 \text{ OIP}$

$t=3 \text{ OIP}$

$t=4 \text{ OIP}$

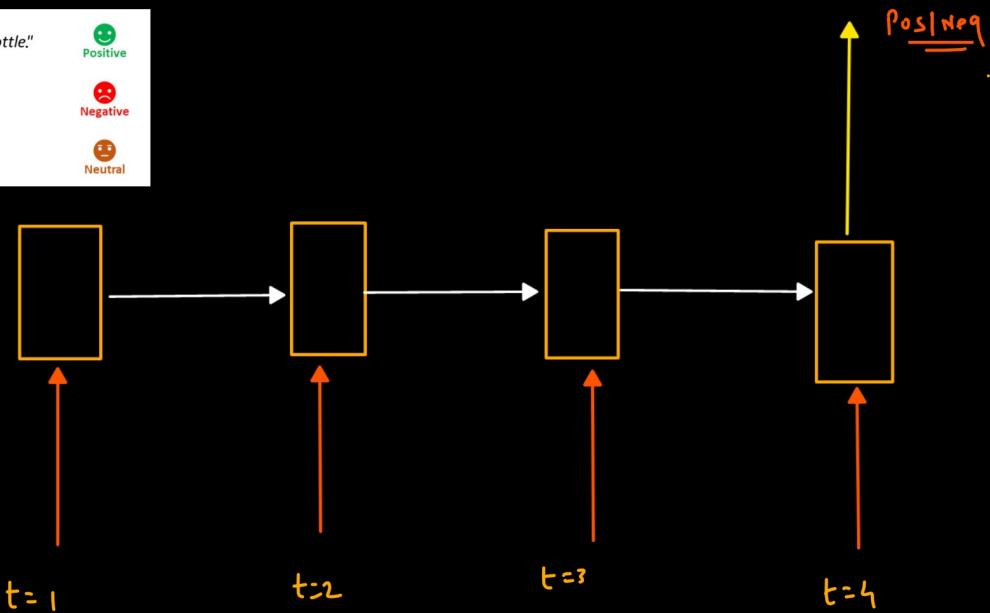
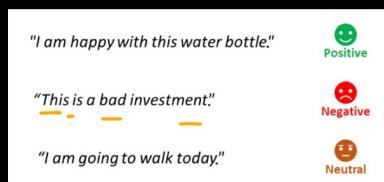


(RNN, LSTM, GRU)

2

Many to ONE

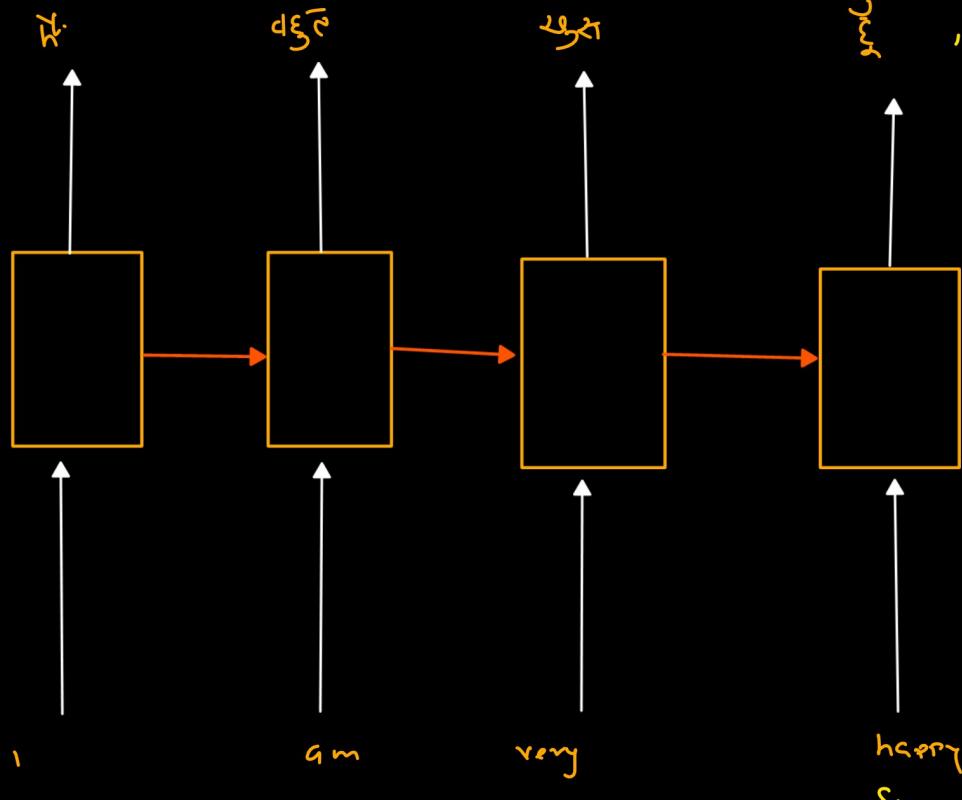
Example \approx Sentiment Analysis



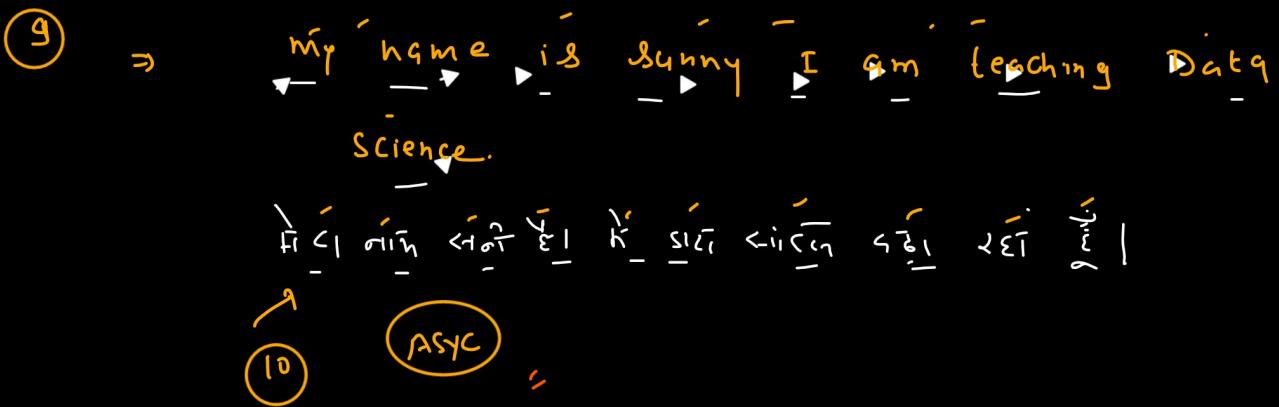
2 Posttagging // 3 NER //

- Ex: Machine translation

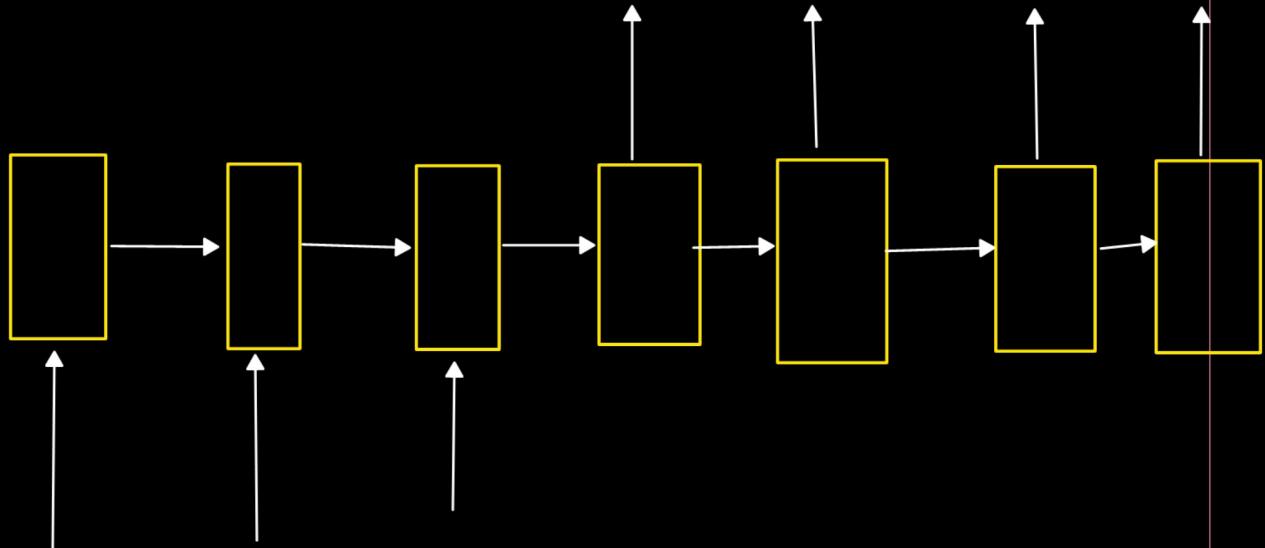
3 MANY to MANY (Same Length) (Synchronized) - RNN / LSTM , GRU



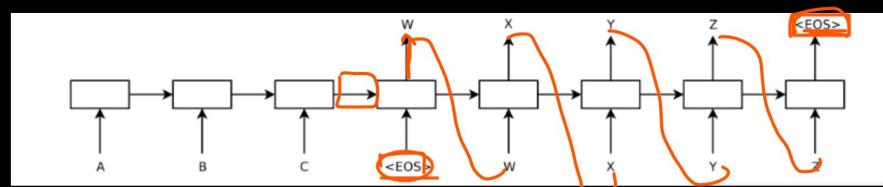
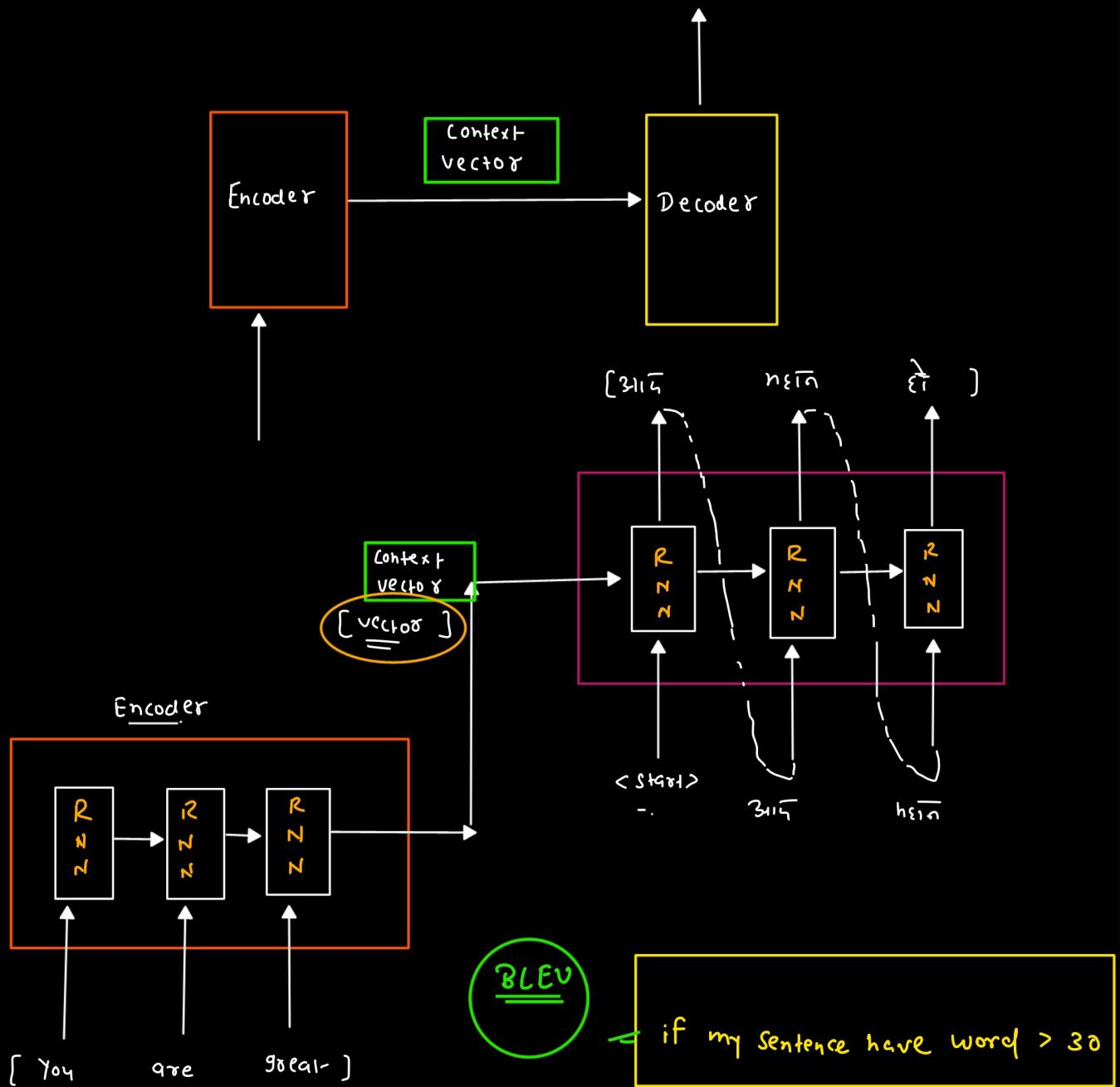
$\rightarrow \left\{ \begin{array}{l} \text{I} \rightarrow \text{हूँ} \\ \text{am} \rightarrow \text{अ} \\ \text{very} \rightarrow \text{उत्तम} \\ \text{happy} \rightarrow \text{सुन्न} \end{array} \right. \Rightarrow \text{Hindi} \Rightarrow \text{हिन्दी}$



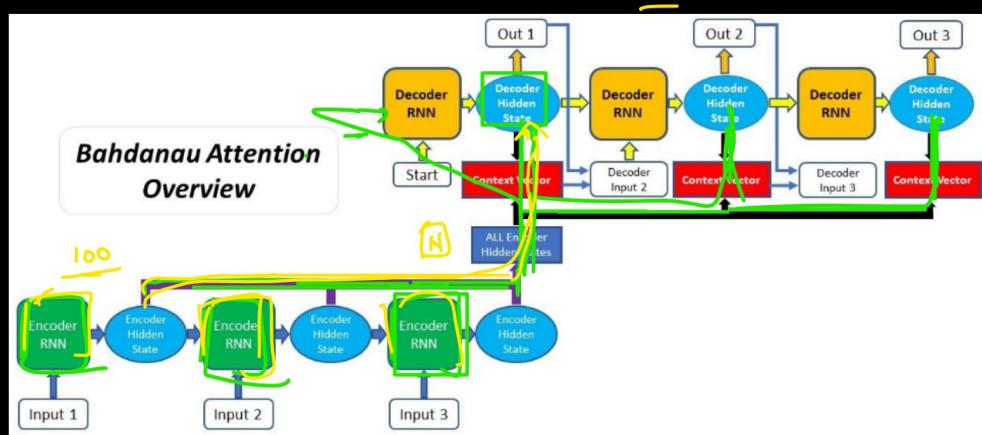
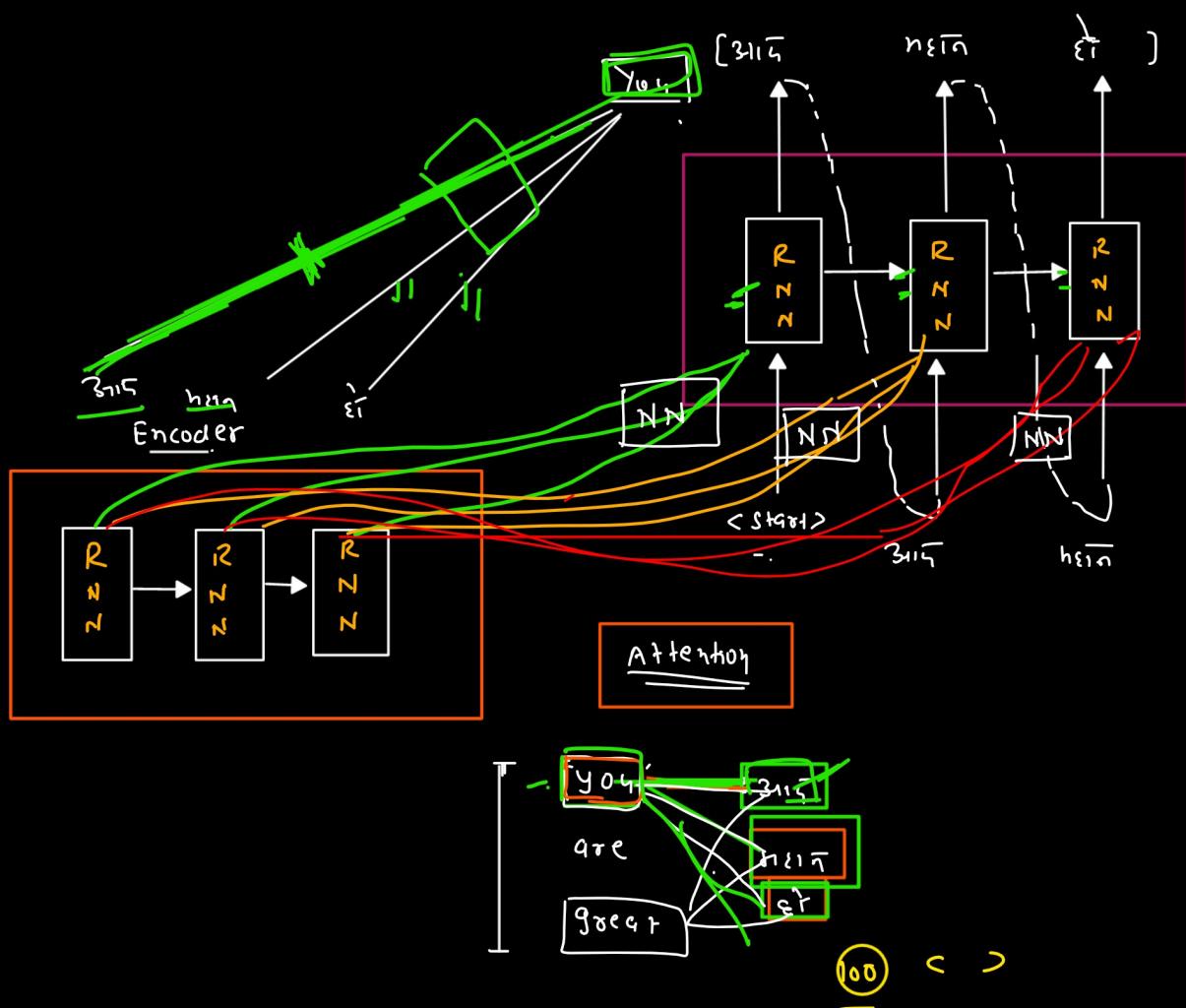
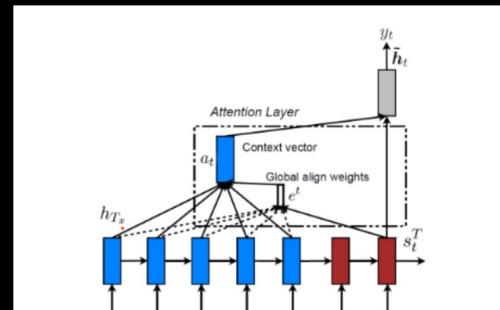
Encoder-Decoder (seq to seq)

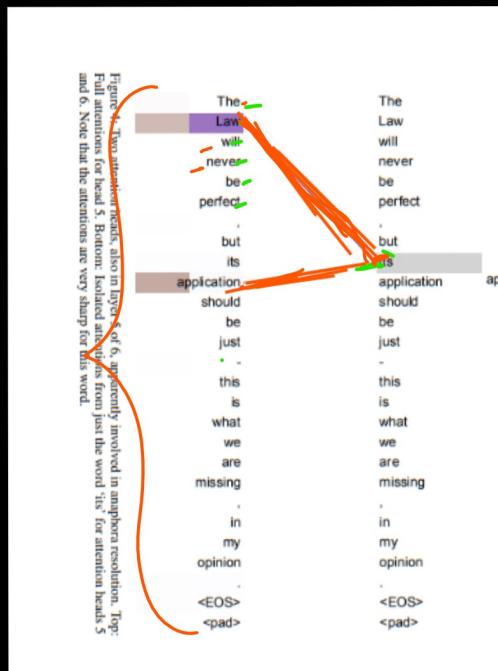


4 Encoder & Decoder

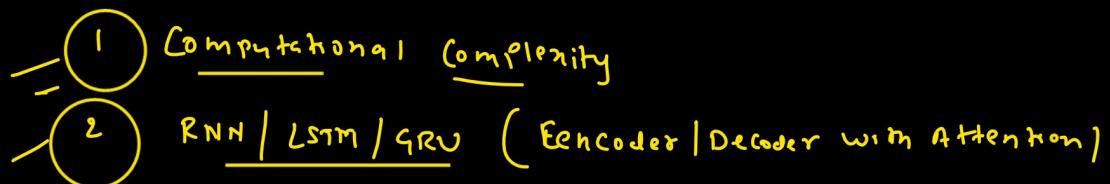


② Attention





Problem with Self Attention

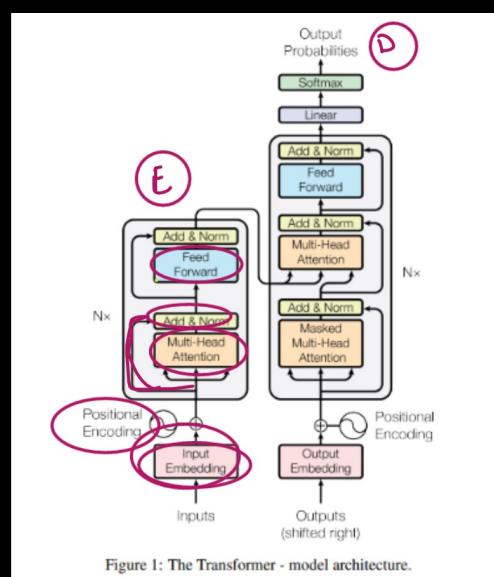


<https://arxiv.org/pdf/1706.03762.pdf>

5 Transformer \Rightarrow RNN | LSTM | GRU ~~X~~

= Computational complexity X

LM
Encoder Decoder



- 1 Parallel input -
 2 Positional encoding -
 3 Query, Key, Value (Self attention)
 4 Multihead attention -
 5 Normalization
 6 Skip connection -

~~RNN LSTM GRU~~

~~Computational complexity~~

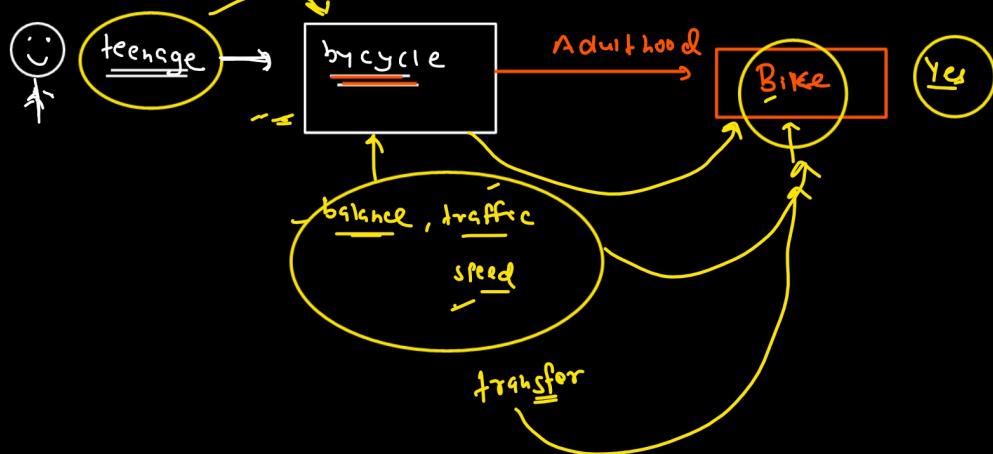


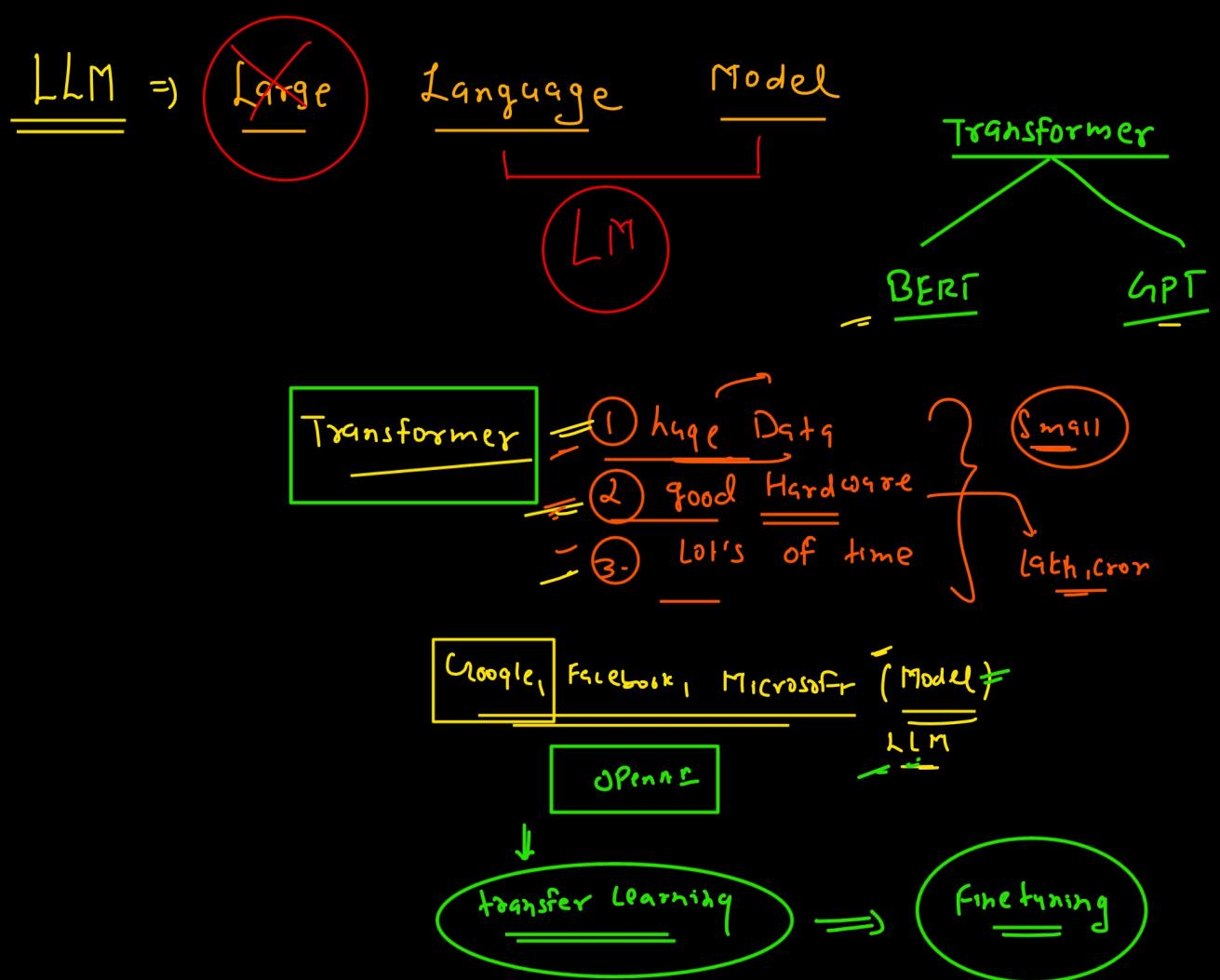
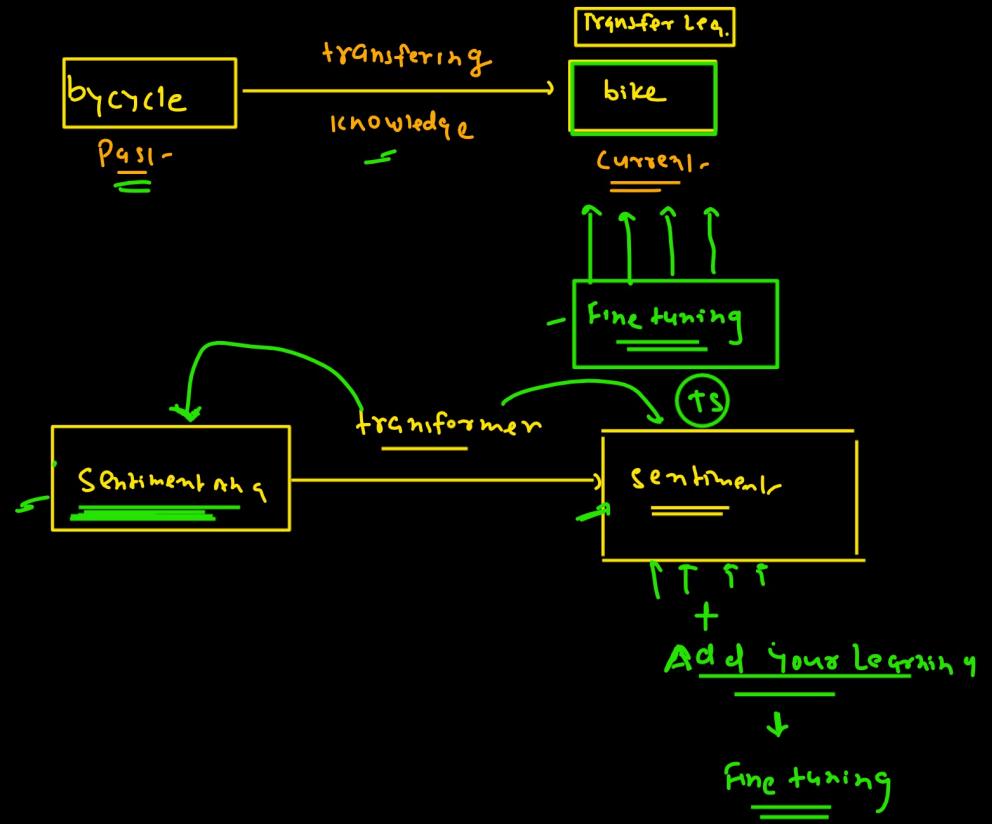
Transfer learning
 fine tuning
 Language modelling

Transformer

Transfer Learning / Finetuning

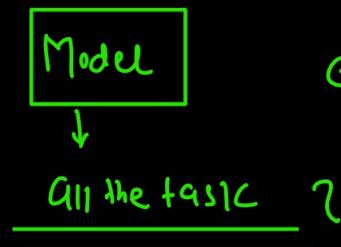
Transfer Learning \Rightarrow transferring the learning from previous task to the next task.





Language modelling →

Huge



- 1 text summarization
- 2 MT
- 3 PoS, NER
- 4 Sentimental
- 5 text generation

- unsupervised
- 1 Language Modelling
 - 2 Supervised Fine tuning

I am Sushant My Des. is OS I am living in

bangalore



I taught to my \Rightarrow Language

by predicting next word

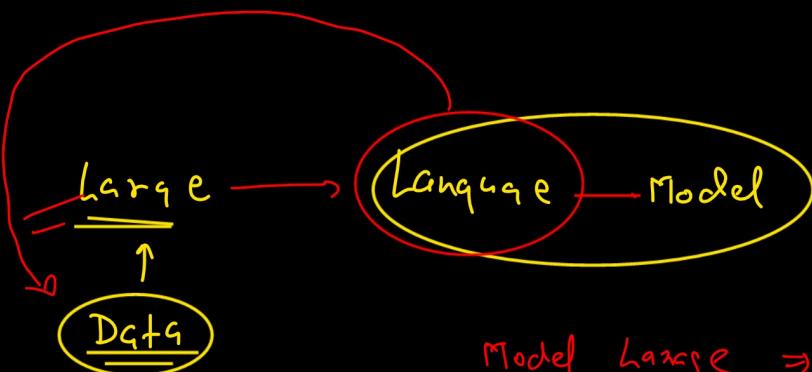
Base

- 1 Unsupervised Learning (LM)
- 2 Supervised fine tuning (train your model on specific task)

Changer

RLHIZ

huge
GSTB



Model Large \Rightarrow huge