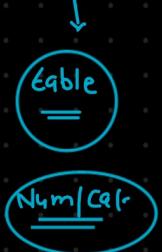
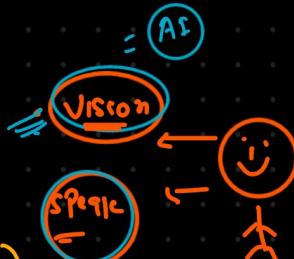
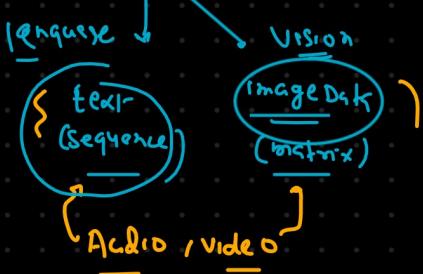


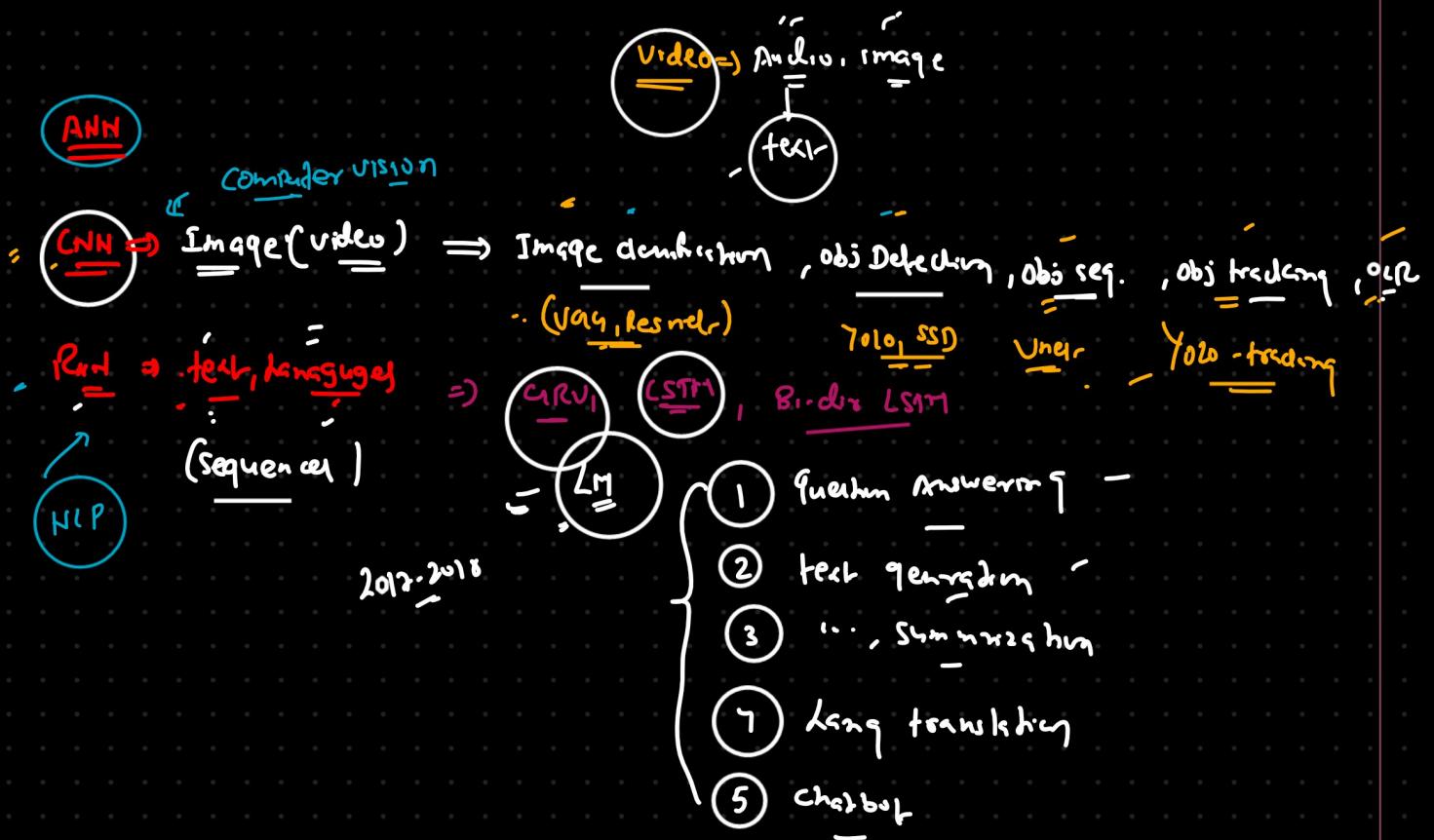
Data

Structure

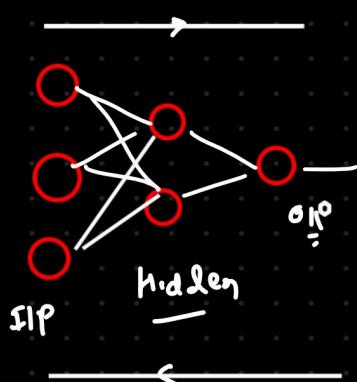


Unstructure





ANN
(Artificial Neural Network)



1. Wegdai
2. Bias
3. Activation fun
4. Loss fun
5. Optimizer
6. FP | BP

CNN
(Convolution NN)

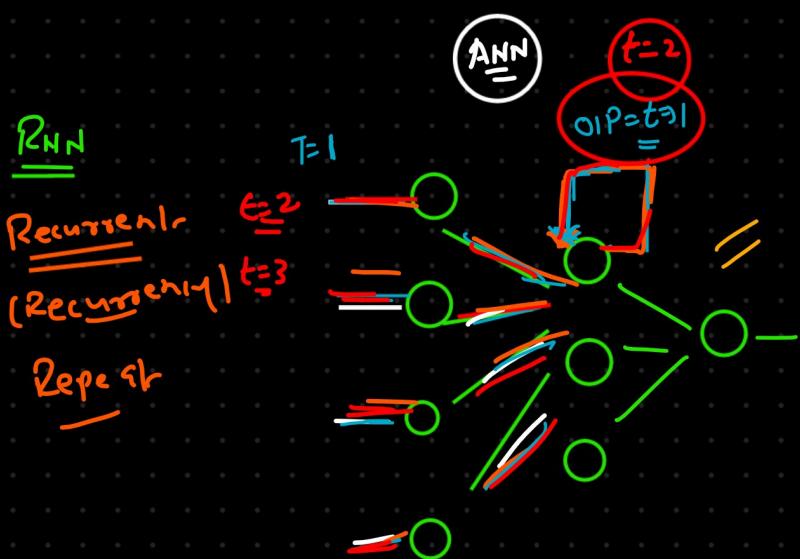
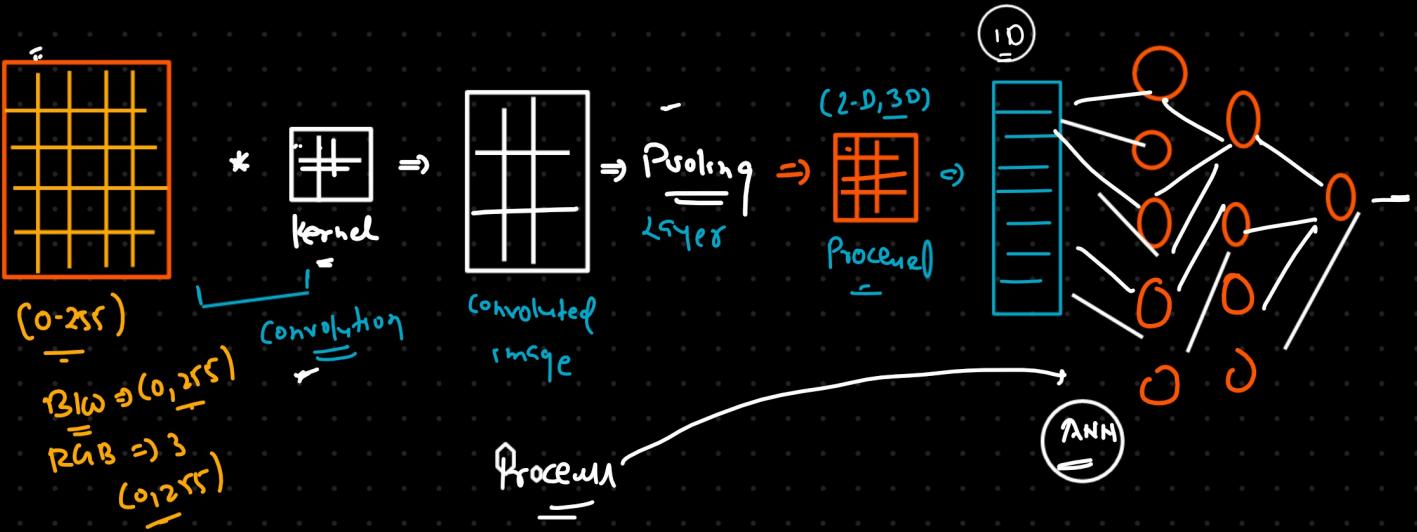
1. Convolution
 2. Padding
 3. flattening
 4. fully connection
- \Rightarrow Img processing

RNN
(Recurrent NN)

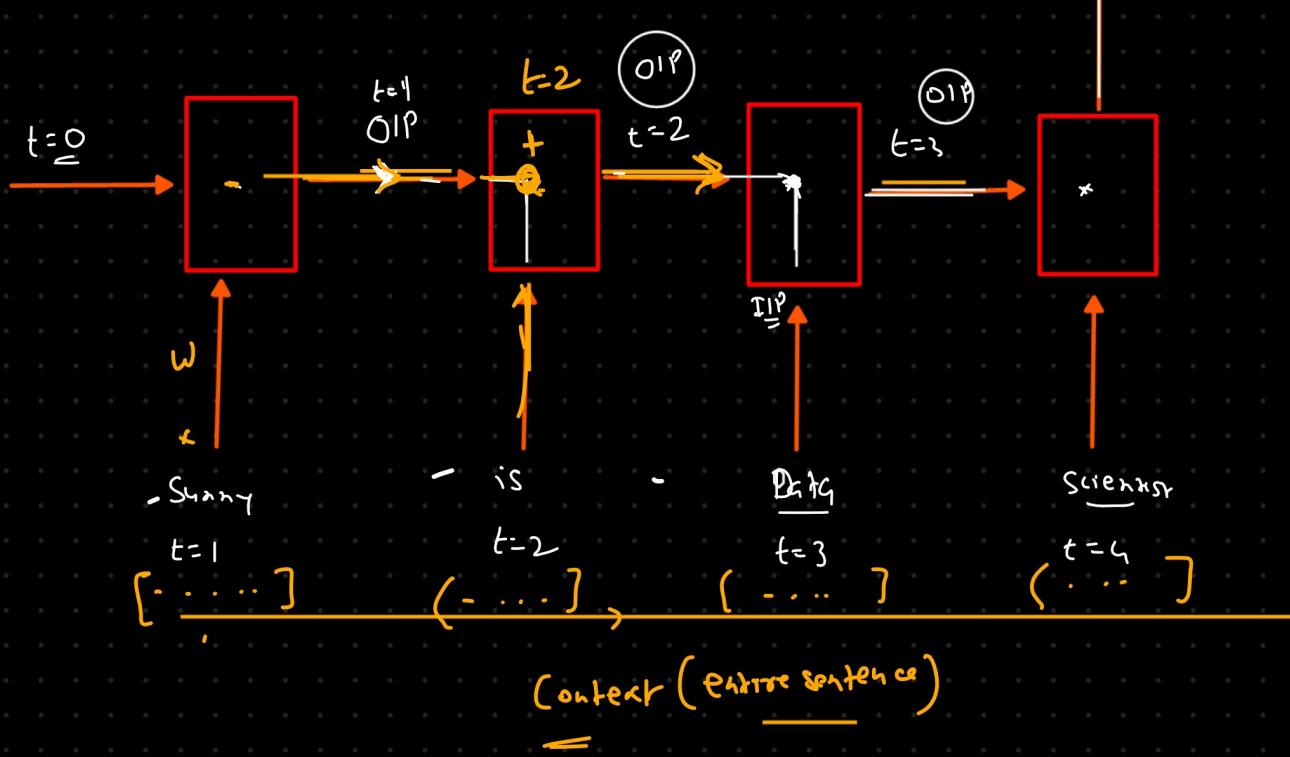
1. feedback loop
2. ANN

Sequence (text)

1. many to one
2. one to many
3. many to many



= Sunny is Data Scientist . Yet



1

One to many

2

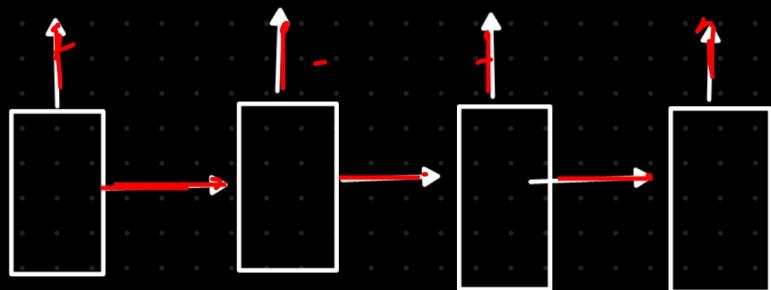
many to many

3

many to one

1

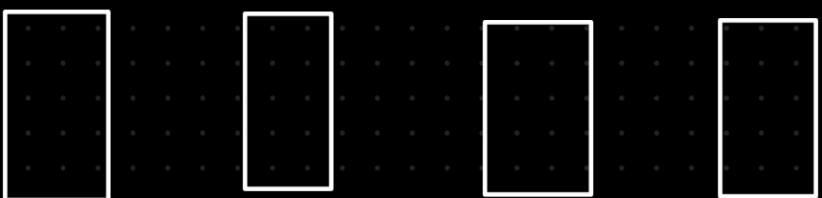
One to many
mapping
(loop to op)



Ex:- Image captioning

NN train → optimize

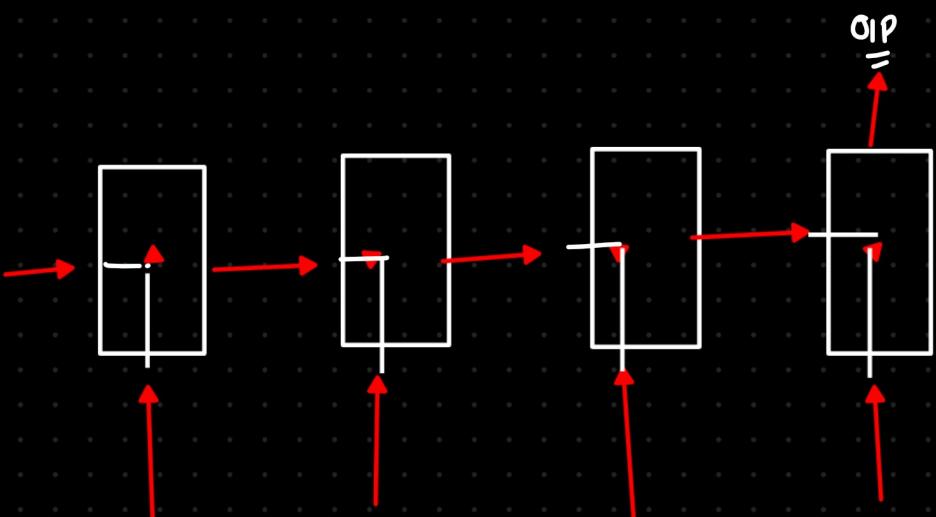
2



3

NLP

text
sequence

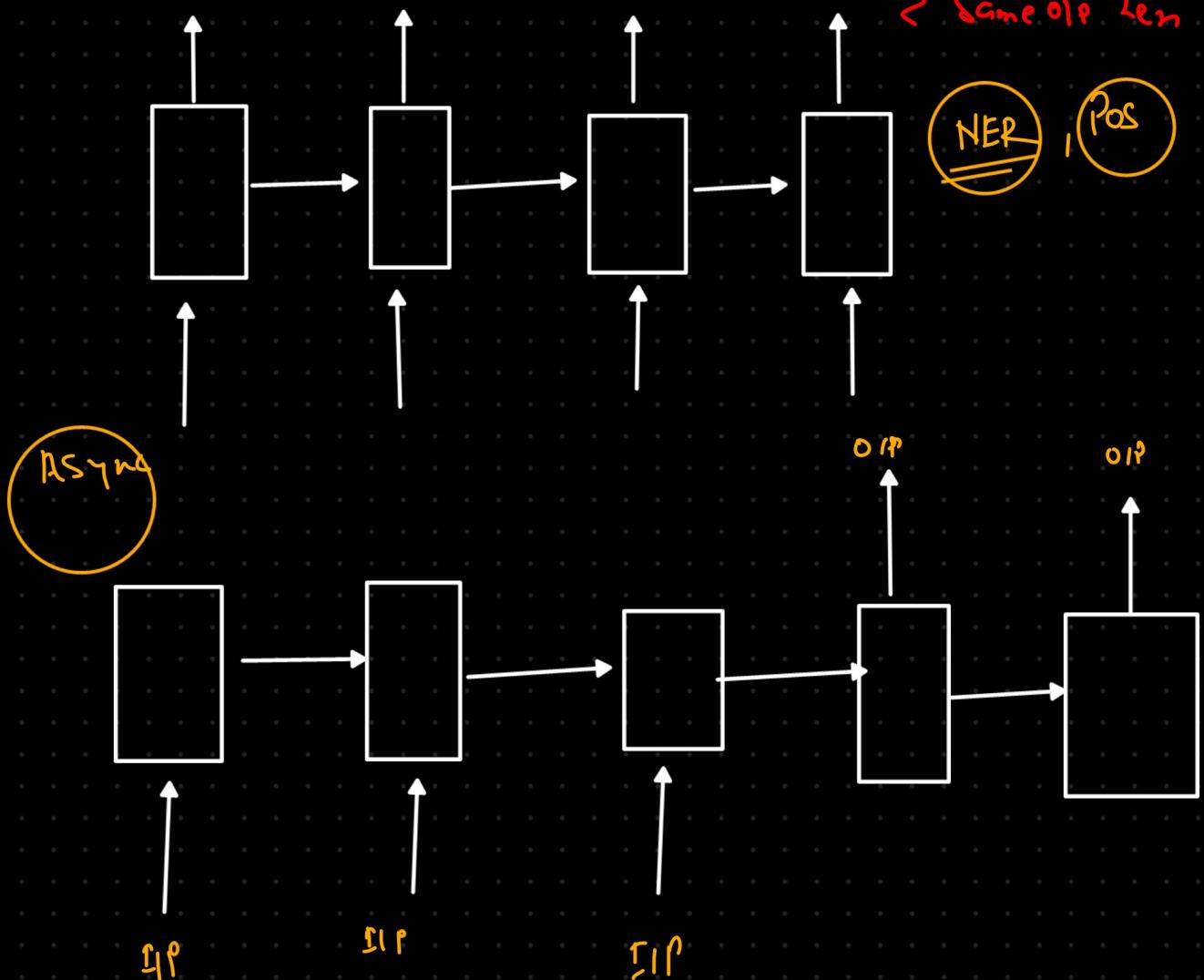


text classification

② many to many

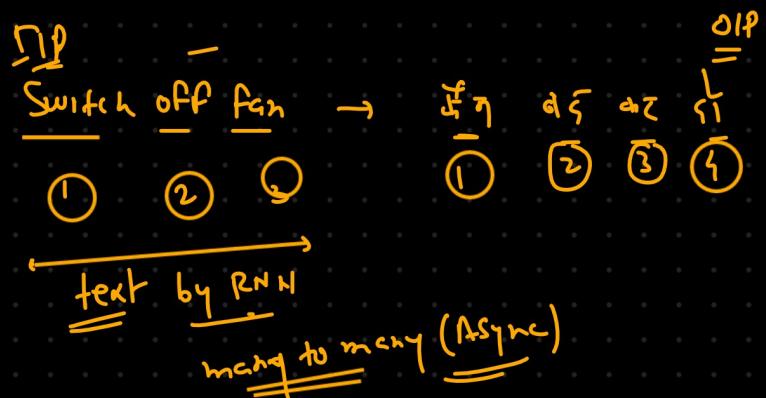
Synchronization (same length)

{ Same IIP Len }
Same OIP Len }

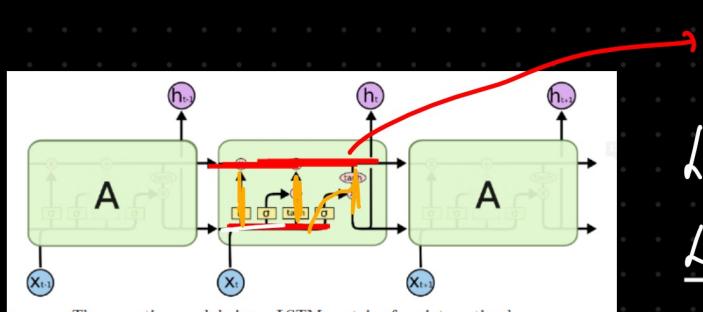
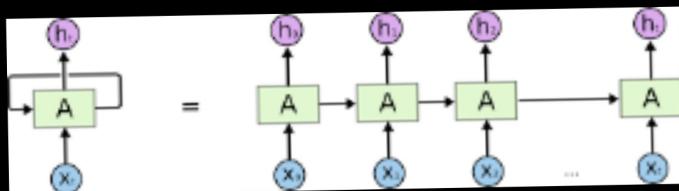


{ Diff IIP Length
Diff OIP Length }

Machine translation

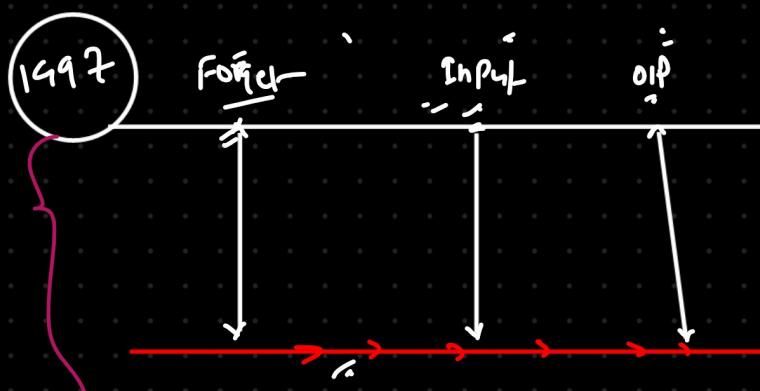
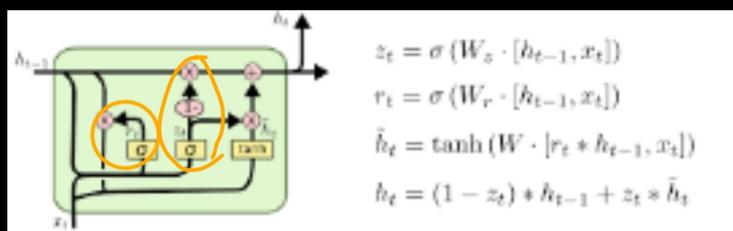
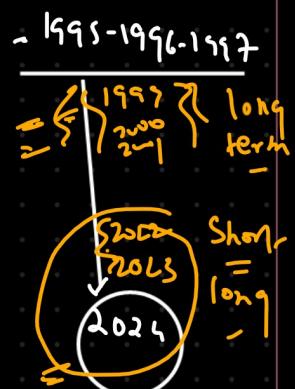


- ② LSTM
- ③ GRU

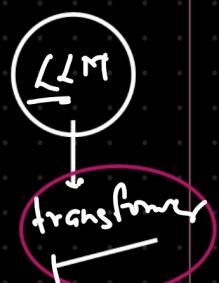


long short-term memory
long-short term

1995-2024



Long term
context
Short term



{ Long term text }
Short term text

Sequence

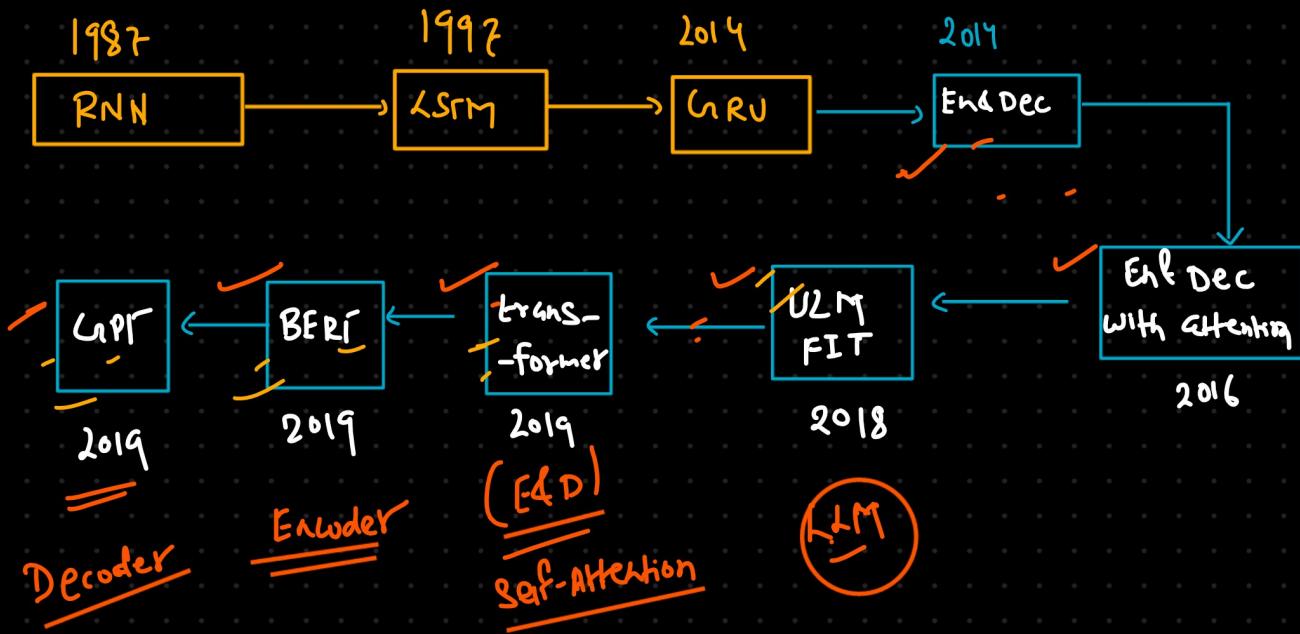
LSTM
1st, 2nd, 3rd, 4th

parallel
trans

GPT-3 Optimized LSTM Graph \Rightarrow 2019

- 1 Reset
- 2 Update

Data Science



e&d2014:<https://arxiv.org/abs/1409.3215>

transformer:https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fdbd053c1c4a845aa-Paper.pdf

ulmfit:<https://arxiv.org/abs/1801.06146>

bert:<https://arxiv.org/abs/1810.04805>

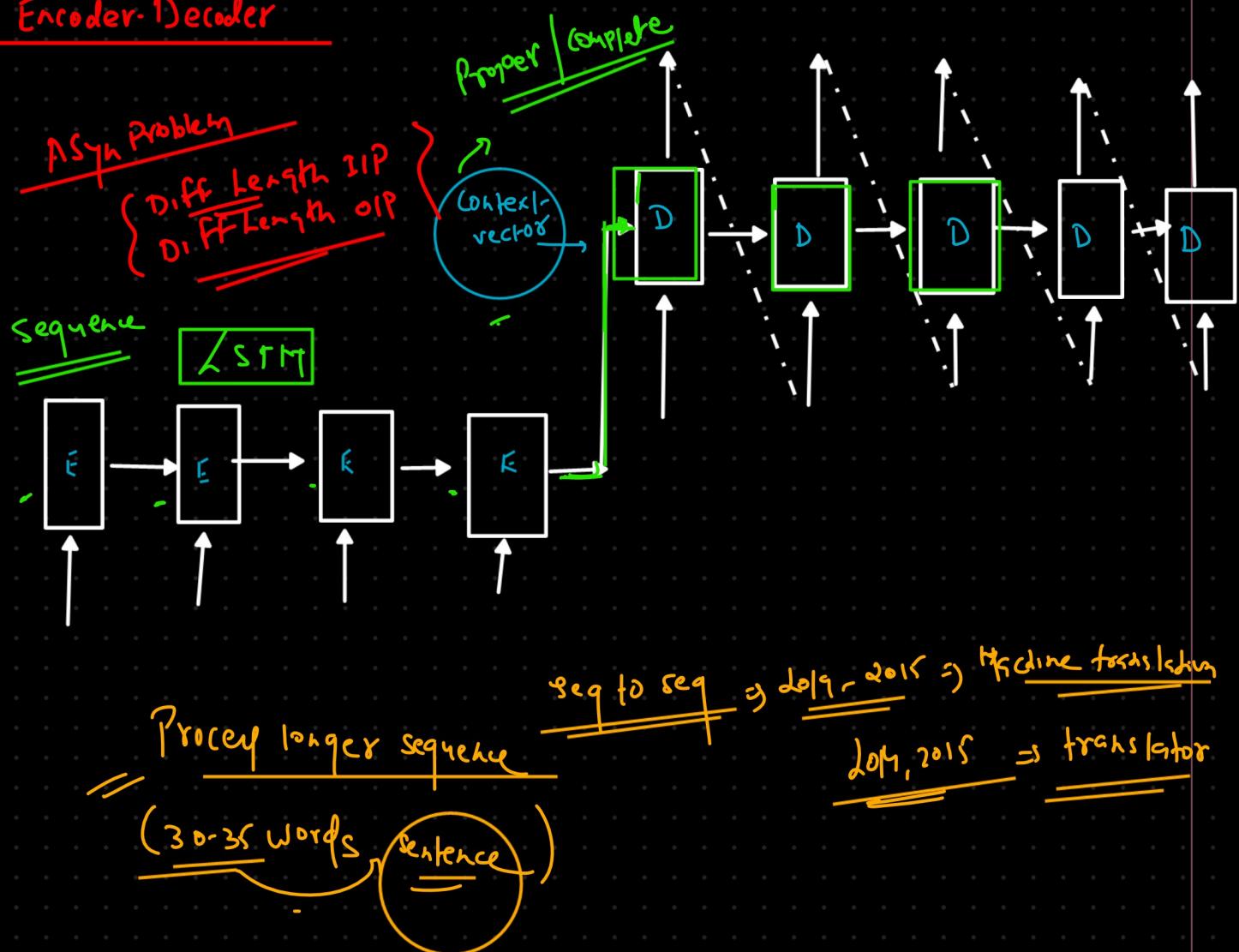
neural machine translation with attention

gpt1: improve language understanding by generative pre training

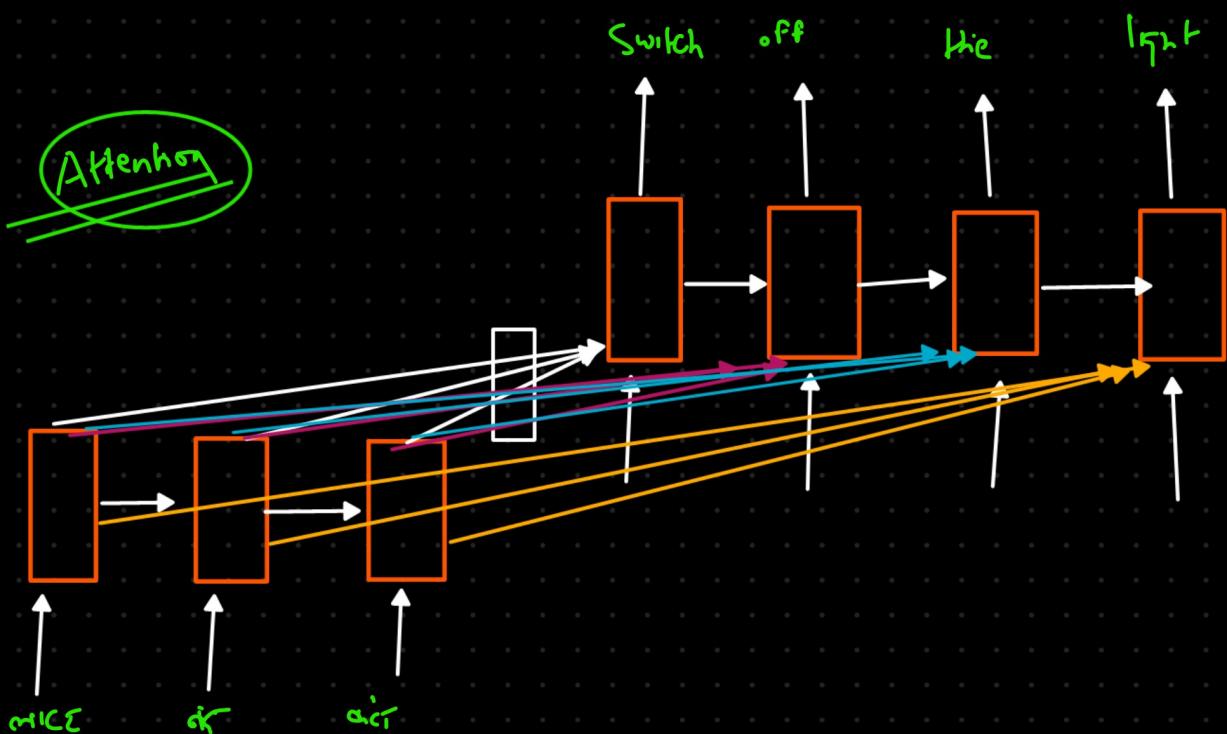
gpt2: language models are unsupervised multitask learner

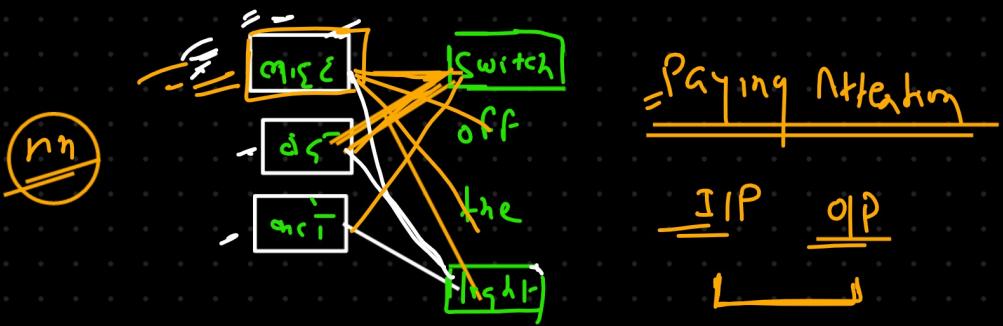
gpt3: language model are few shot leaner

Encoder-Decoder



Encoder-Decoder with attention





= Paying Attention

IIP OIP

LM - FIT \Rightarrow transfer learning in NLP
Fine-tune

