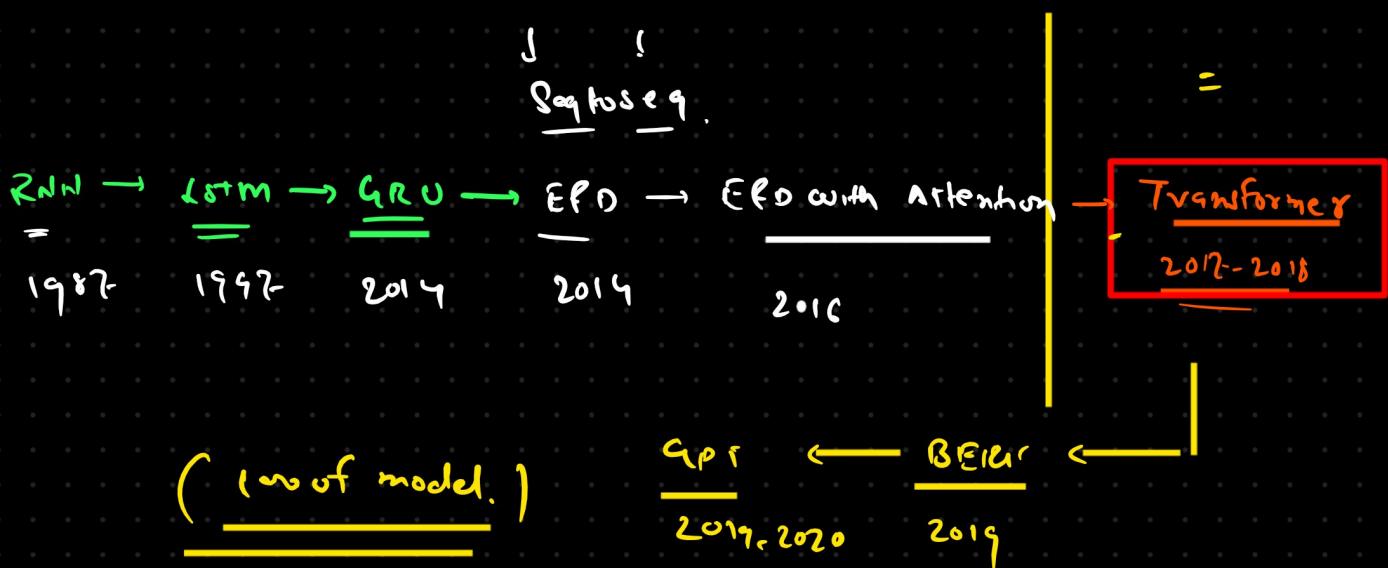
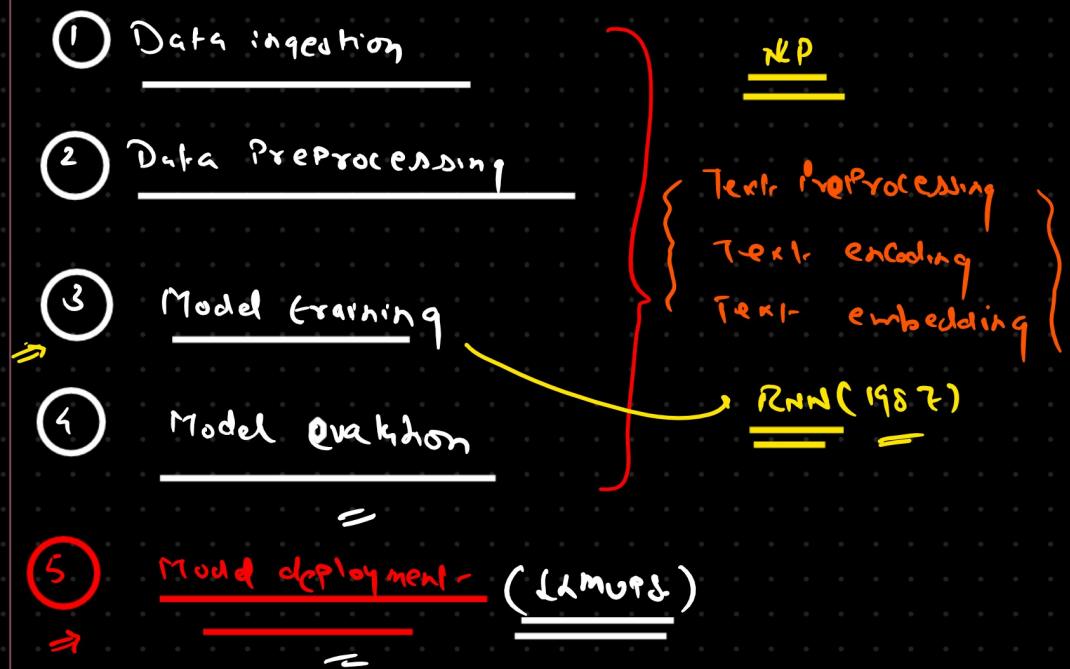
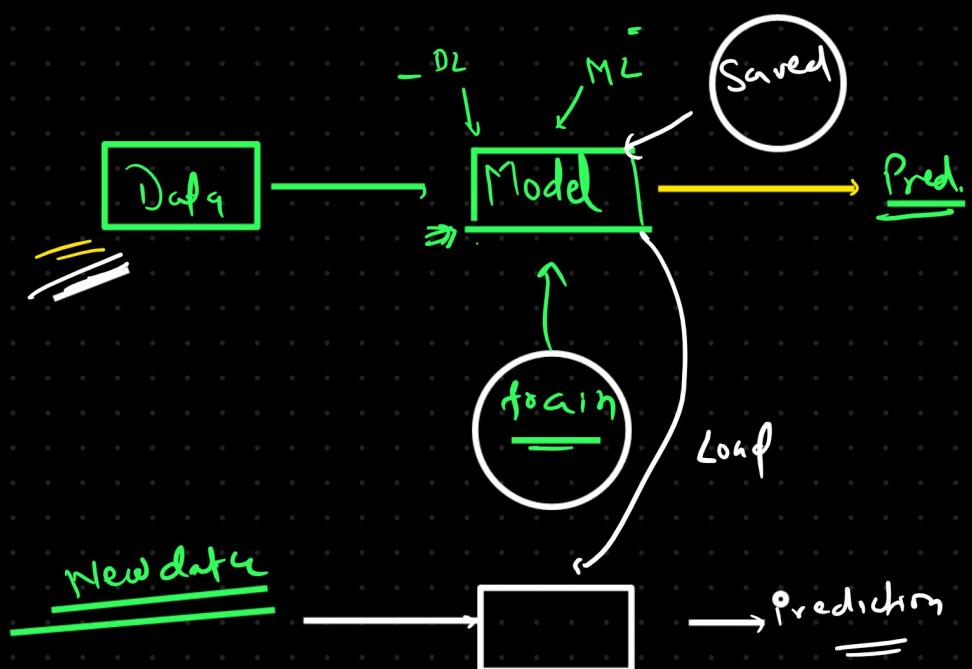
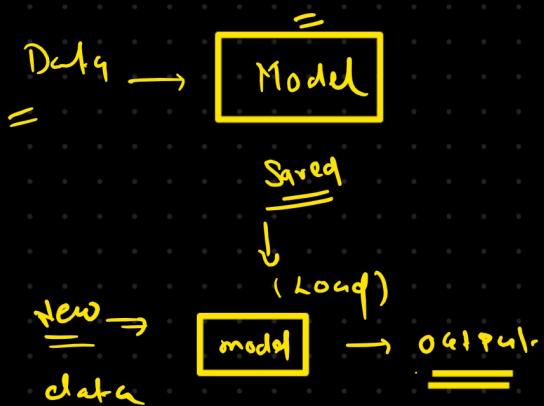


- ① RNN / LSTM / GRU
  - ② Encoder & Decoder
  - ③ Encoder & Decoder with attention
  - = ④ Transformer  $\Rightarrow$  3,4 }  $\Rightarrow$  Bert, GPT, ...  
- ⑤ Pytorch
  - ⑥ Inferencing with LLM
- RAG, Agent, LLMops



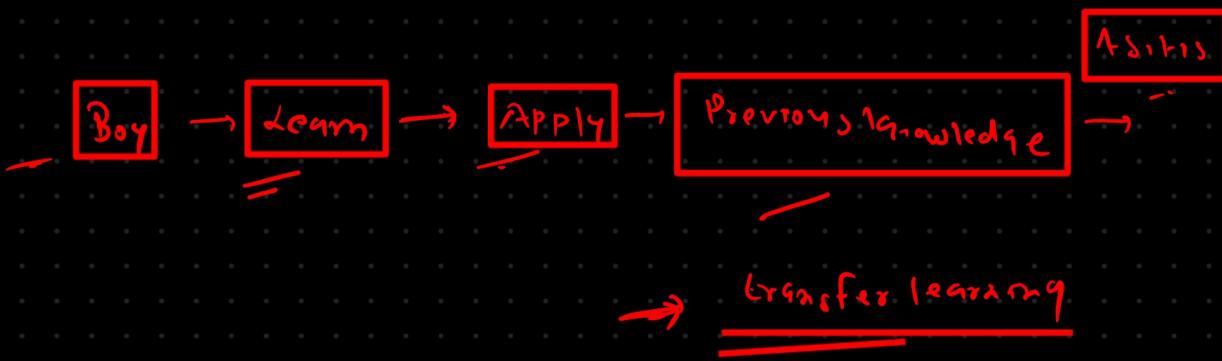
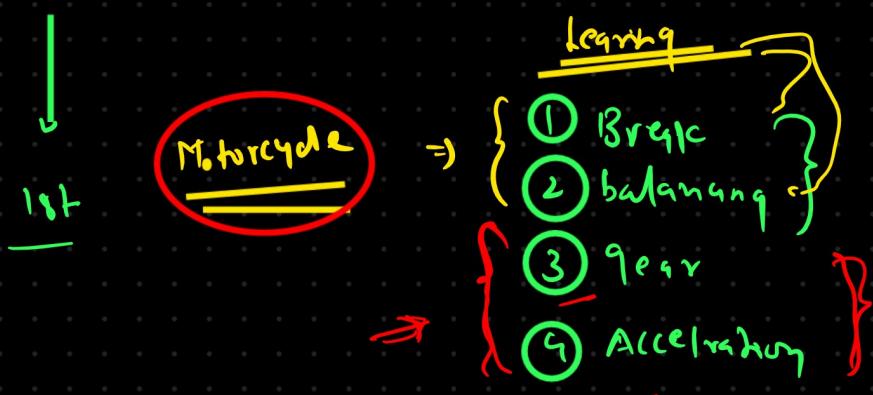
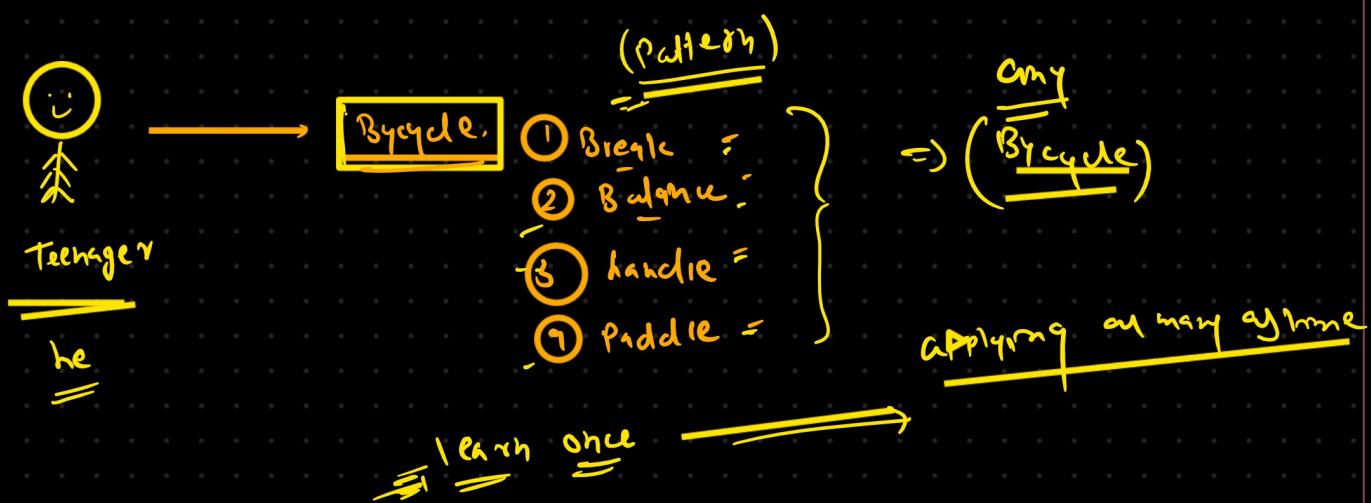
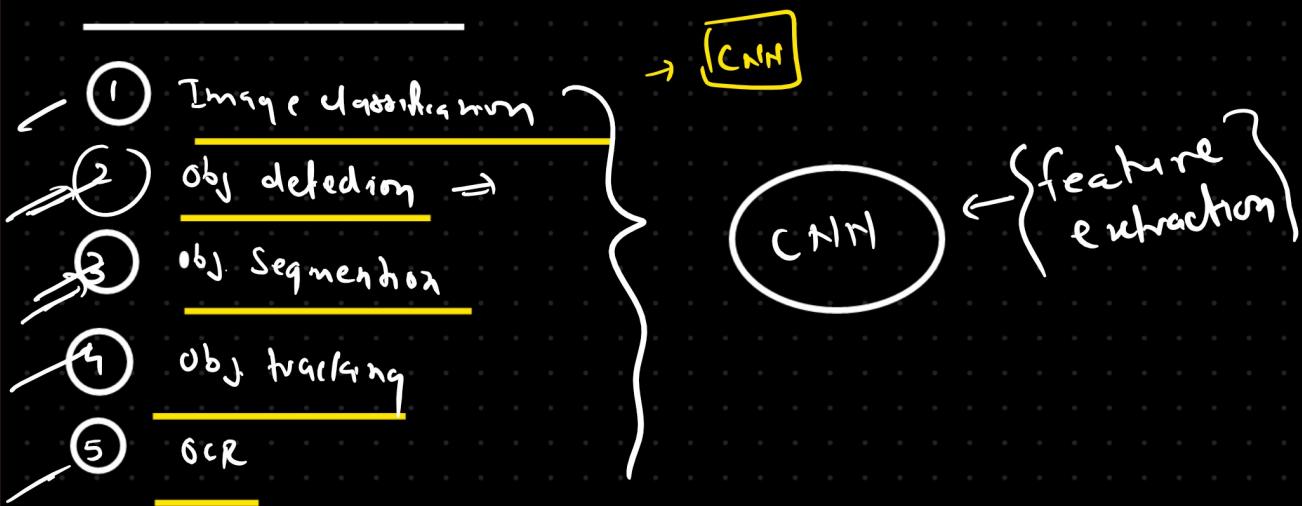
- 1 Why TL was not possible in NLP?
- 2 Gimp research Paper
- 3 GPT-2 Transformer
- 4 Self Attention

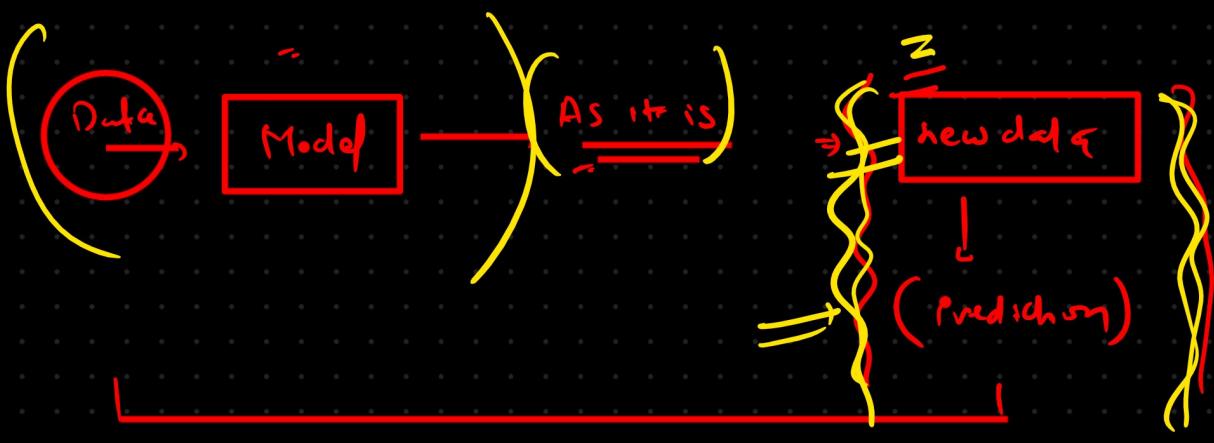
## Transfer Learning





## Computer vision

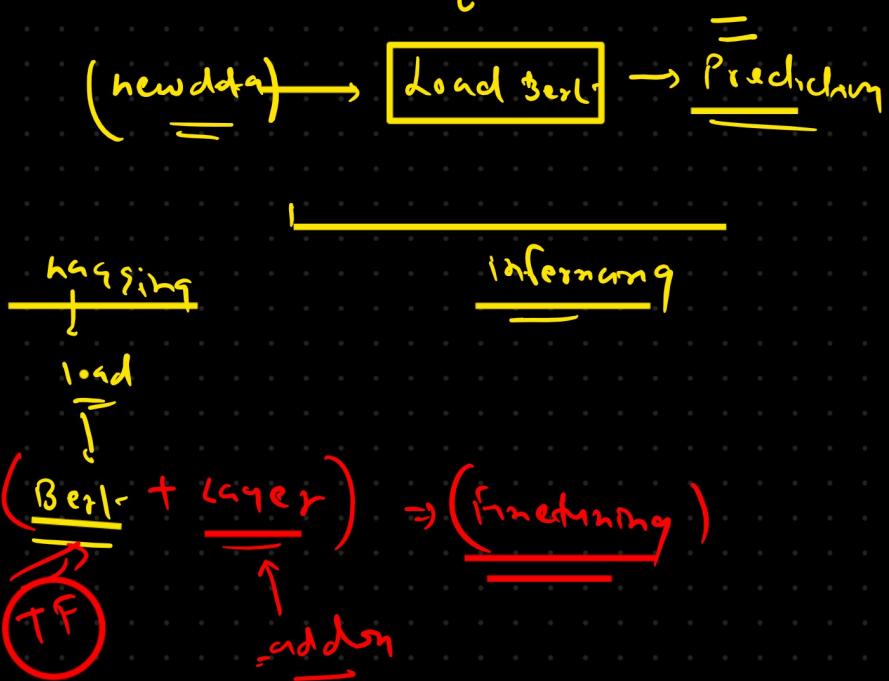




① transfer Learning  
(Inferencing)

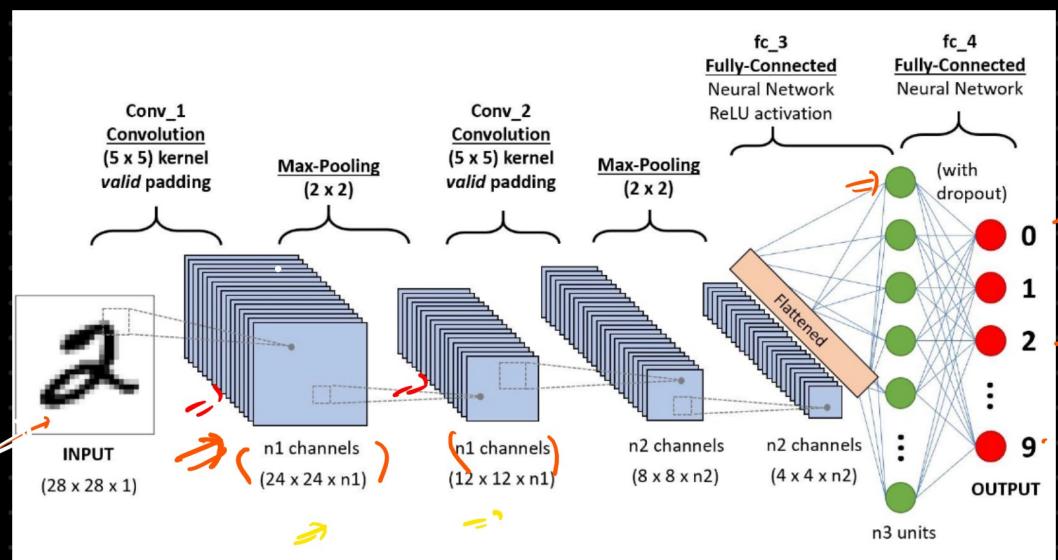
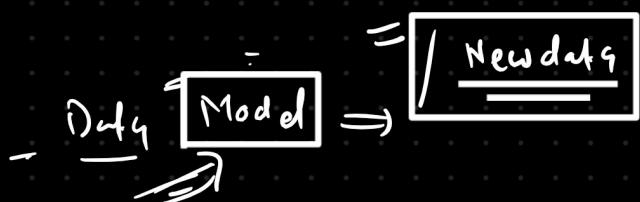
Data  
→ Model ← training

② Fine-tuning

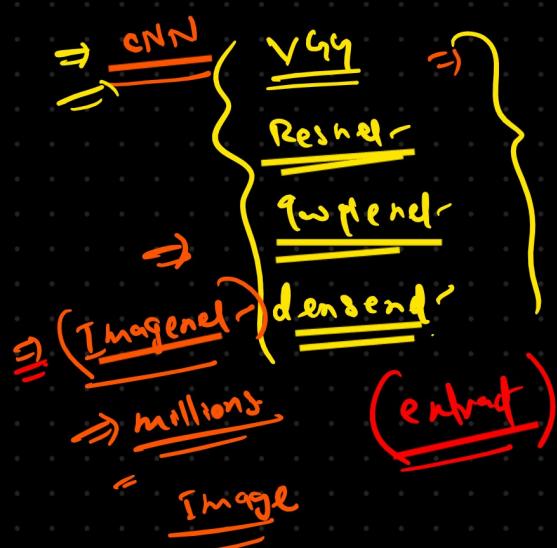


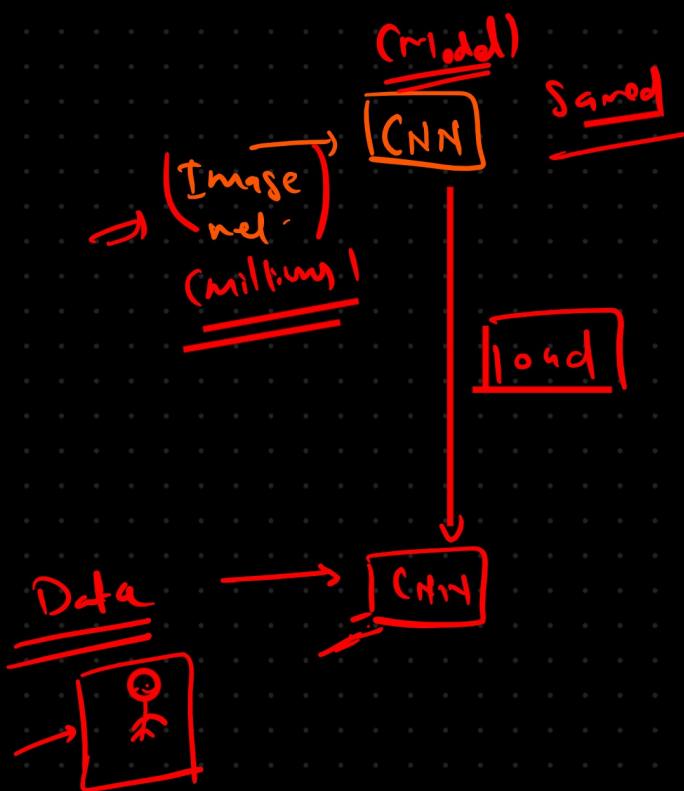
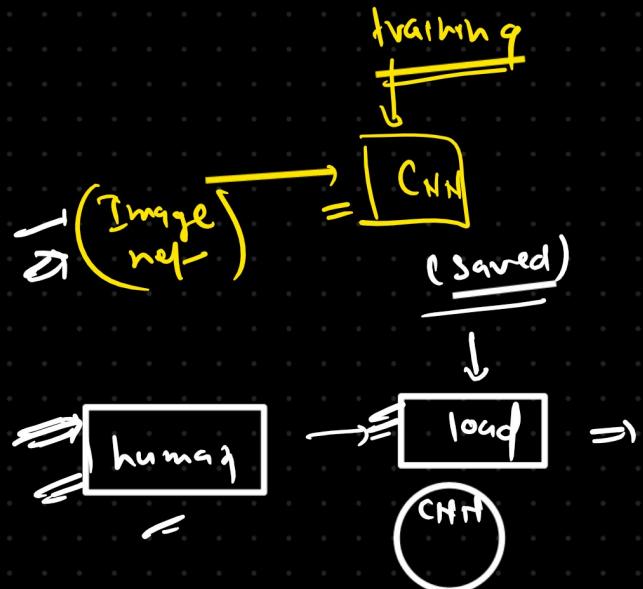
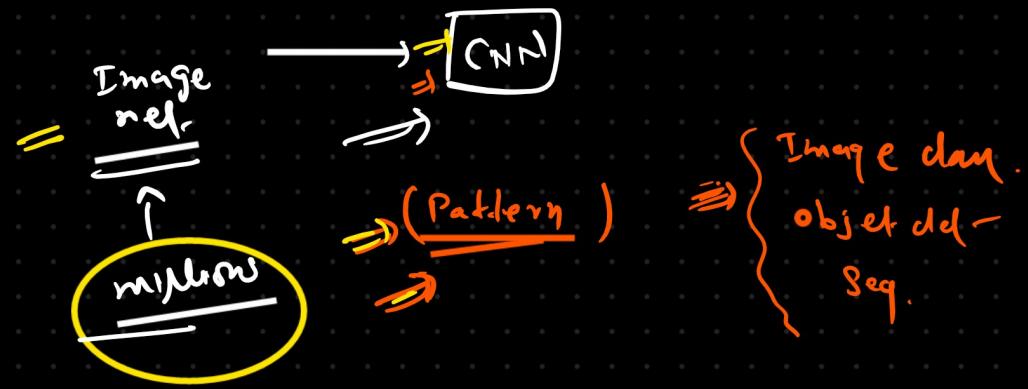
# Google word2vec

↓  
16

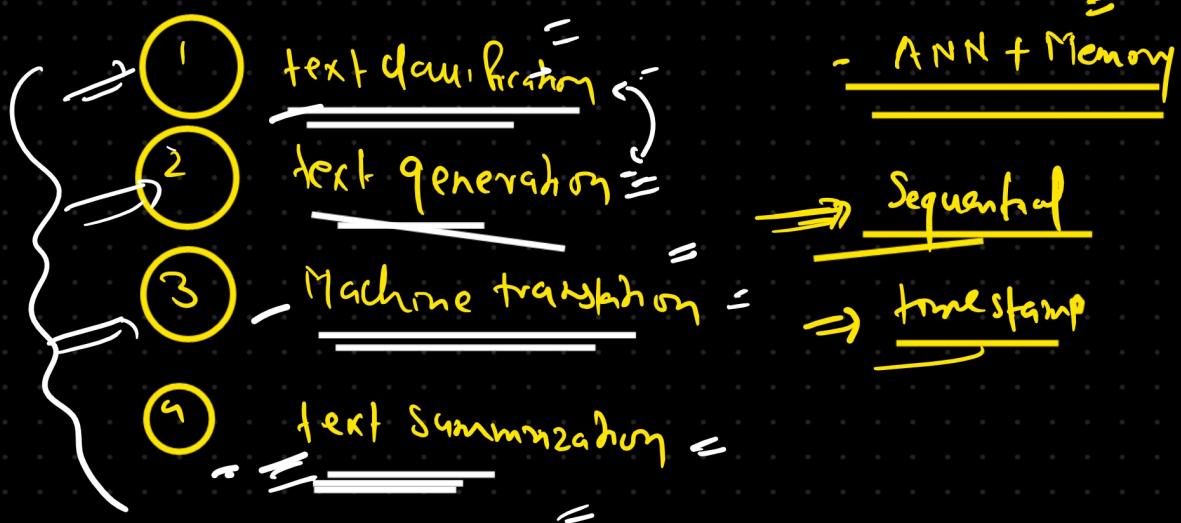


10





NLP = RNN, LSTM, GRU ← ~~lot's of~~

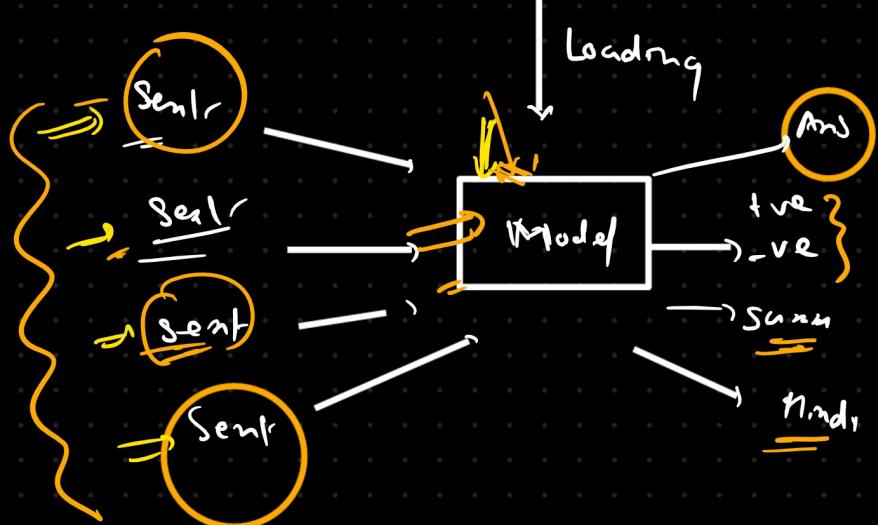
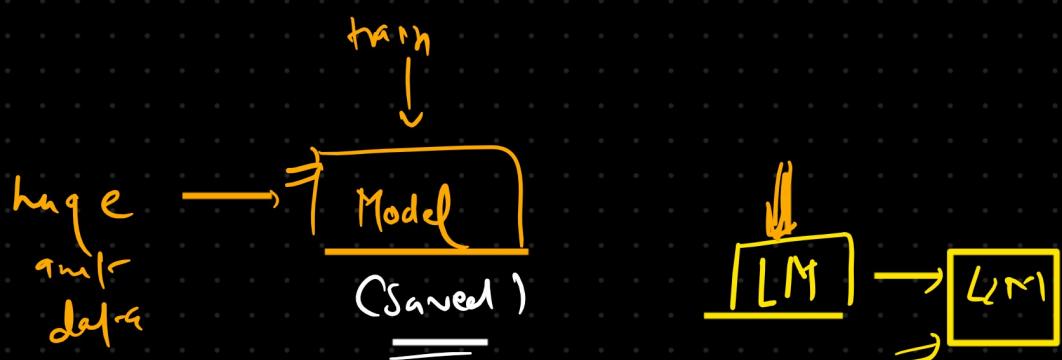
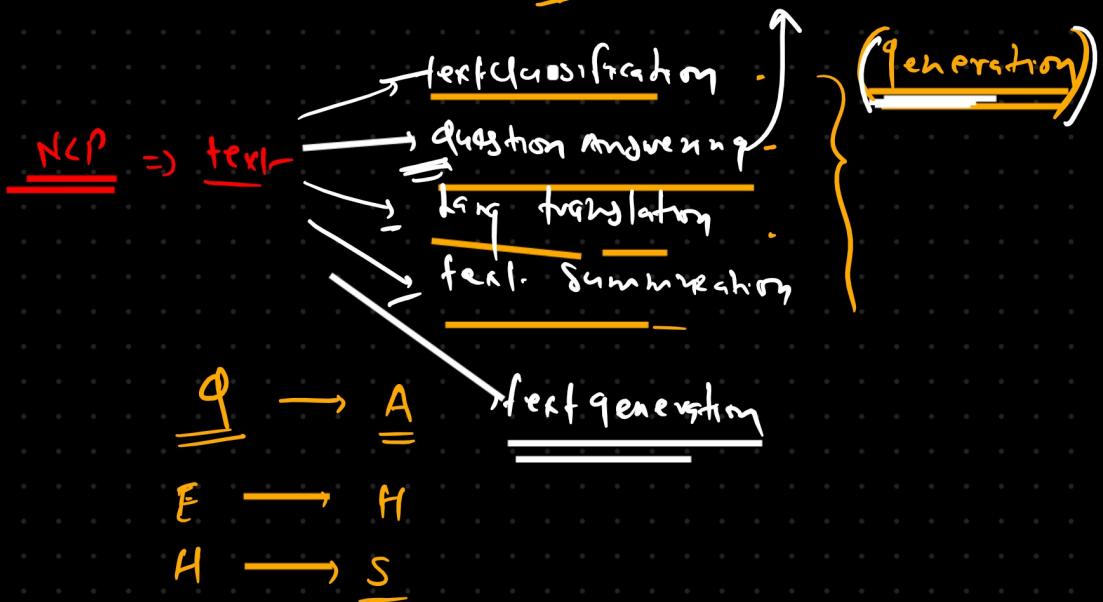


- ⇒ 1 large am. data X  
⇒ 2 task exclusiveness

## Transformer

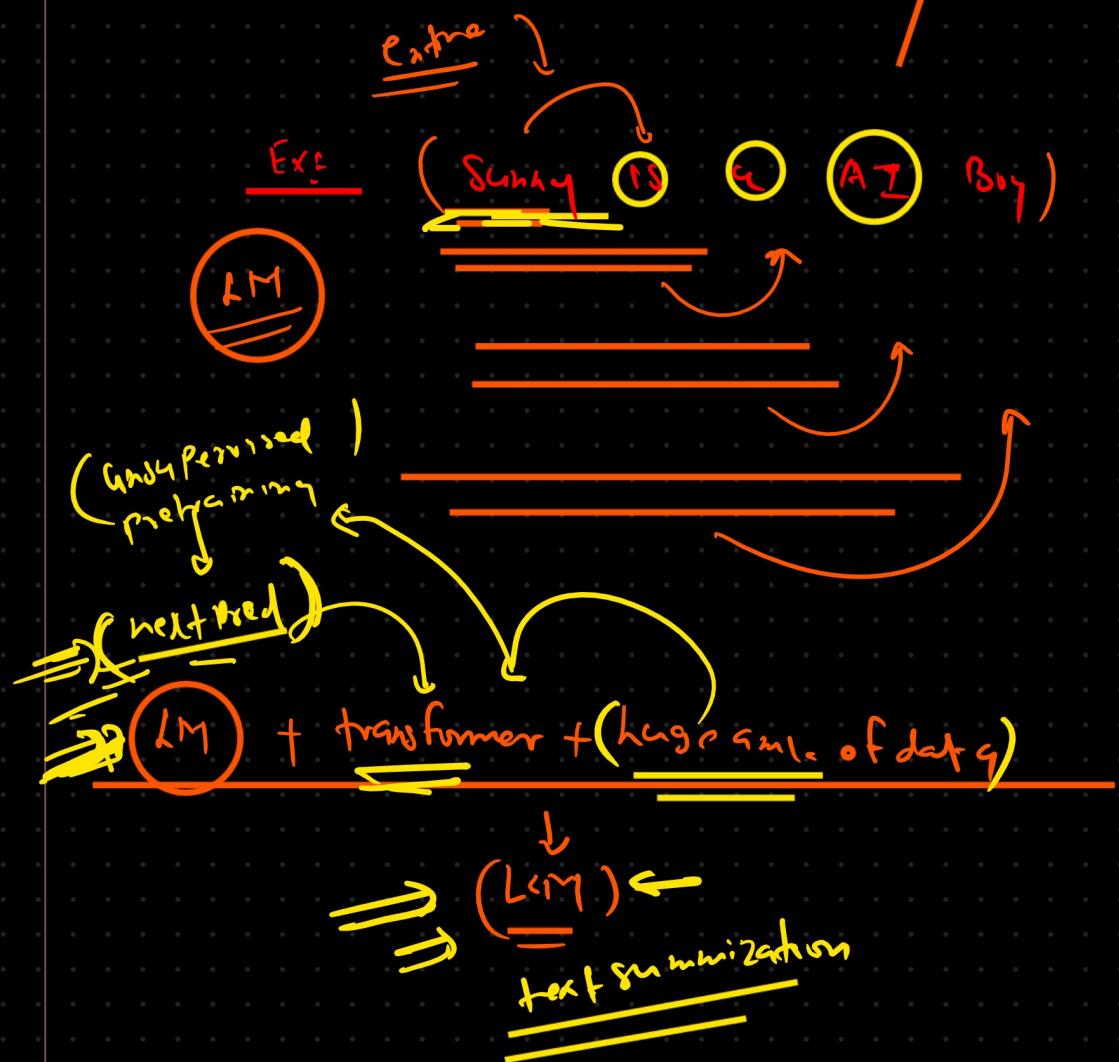
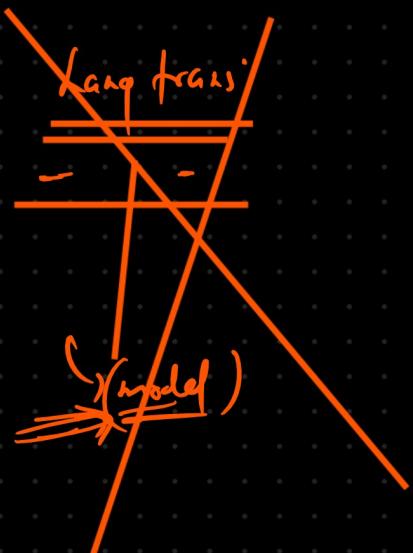
- 1 RNN, LSTM, GRU ✗
- 2 Parallel data ✓ ↘ ↙
- 3 huge contr. ✗
- 4 Training extensive ✓

= Conversational AI



Nextword Pred

Date



RNN, LSTM, GRU  $\times$

EFLD, EFLD with attention

transfer learning  
finetuning

$\hookrightarrow$  huge amf-

