Day 16, 18th Jan 24

NLP Data Science Masters 2.0

→ By Dr Rajan Deshmukh, IIT Delhi + Mumbai University Alumni

Recap ① seq2seq → Theory
② coding implementation

Topic → ① Attention
Vvip ② Transformer
↳ Generative AI, SOTA
↳ ① Overview ② Dataset details ③ notes, any details

★ text generation by transformer :

① RNN →

sequence length short

eg → sentiment analysis → eg The food is delicious → +ve

Disadv → Vanishing gradient problem
Vvip
e.g. length↑

② LSTM → adv → long term memory dependency
e.g. length↑

disadv → e.g. length↑↑ , ...↓↓
Vvip

③ Bi-directional RNN/LSTM

past & future e.g. → dependence

④ seq2seq
↳ encoder & ↳ Decoder
Language translation

ds e.g. sequence

context X

⑤ Attention

⑥ Transformer → 2017 → Google Brain & UOT

Attention model

The animal didn't cross the road, it was very tired

context → it → refer ?

syntactic info X
context X
relevance X
} by transformer

I love work in France, I love drink French family

attention going first Attention map

$P(y|x_1, x_2, x_3, x_4)$

Transformer

application → chat GPT (backbone → transformer)
chat bot
google translate
Voice assistance
Siri, Alexa.

Embedding

Dog → X

→ $\begin{bmatrix} 0.25 \\ 0.13 \\ 0.11 \end{bmatrix}$ Numerical representation

Tokenize

I	LOVE	NLP

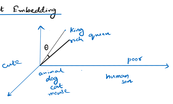
↓

23	965	215

text

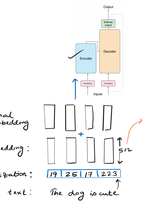
token IDs

Input Embedding

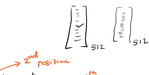


Vector space (or) embedding space

$\begin{bmatrix} 0.23 \\ 0.11 \\ 0.93 \end{bmatrix}$ King $\begin{bmatrix} 0.24 \\ 0.17 \\ 0.98 \end{bmatrix}$ queen $\begin{bmatrix} -0.17 \\ 1.23 \\ 1.23 \end{bmatrix}$ poor



tokenization → w_1 w_2 w_3 w_4 w_5
19 26 50



seq1 : The dog is cute
seq2 : This animal looks like a dog

position or Order of the words
↳ v. imp

Positional Embedding / positional Encoding

vector that gives context based on position of word in sentence

$PE_{(pos,2i)} = \sin(pos/10000^{2i/d})$
$PE_{(pos,2i+1)} = \cos(pos/10000^{2i/d})$

Dog → 23 → $\begin{bmatrix} 0.37 \\ 0.23 \\ 0.93 \end{bmatrix}$ + $\begin{bmatrix} 0.19 \\ 0.23 \\ 0.67 \end{bmatrix}$ = $\begin{bmatrix} 0.52 \\ 0.84 \\ 0.12 \end{bmatrix}$

text → 2

Embedding of "dog" Vector encoding of position of "dog" in that sentence Final Embedding of "dog" with context & position in that sentence

Encoder