Aug 4, 12:24 AM

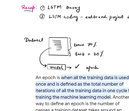
RAF Data Science Masters S.O.

Ty Michael, Ty Zyan, Prabhath, Ut Pulk & Hardeep, amazing Mentors.

Recap ① LSTM: Long
② LSTM using - sentiment-input representation

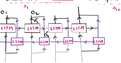
Tutorial

An epoch is when all the training data is used at once and is defined as the total number of iterations of all the training data in one cycle for training the machine learning model. Another way to define an epoch is the number of passes a training dataset takes around an algorithm. Wikipedia



Bi-directional LSTM

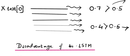
example: I was born in France. I can speak French fluently



$O_t = x_t, x_{t-1}, x_{t+1}, x_{t+2}$

Bi-directional → have dependencies on both past & future sequence

model. predict



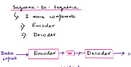
Disadvantage of the LSTM

① Complex
② Slower
③ More training time
④ Computationally expensive

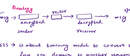
Bow → Bag of word - Problem
long term dependencies
LSTM → Vol

Sequence - to - sequence

↳ 2 main components
↳ Encoder
↳ Decoder



W2V
↳ Deep representation

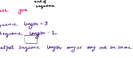


Seq2Seq → is about training models to convert sequences from one domain to another sequence. for another domain.

① English → Encoder → Decoder → Spanish
machine/ language translation

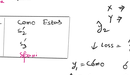
② Text summarization

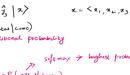
Seq2Seq



input sequence length = 3
output sequence length = 2

NB input & output sequence length maybe not the same.





$x \rightarrow$ Como
$y \rightarrow$ Estas
$\hat{y}_2$
$\downarrow$ Loss = $\hat{y}_1 - y_1$

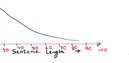
y_1 = Como
$\hat{y}_1$ = Como

Acc = $y_1 - \hat{y}_1$

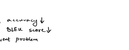
Decoding // Prediction

$\hat{y} = \langle \hat{y}_1, \hat{y}_2, \hat{y}_3 \mid x \rangle$ $\quad x = \langle x_1, x_2, x_3 \rangle$

Goal $y = P(\text{Estas} \mid \text{Como})$
↳ Conditional probability



Problems with S2S / Encoder-Decoder model



Say length → BLEU ↓ after 30 words
↳ accuracy of encoder & decoder decrease if input sequence length is high

Disadvantage

① Say length↑, accuracy↓
② , ↑, BLEU score↓
③ Vanishing Gradient problem

During Training

During Inference

Seq2Seq

① Input seq length → memory, BLEU score
② context, memory →