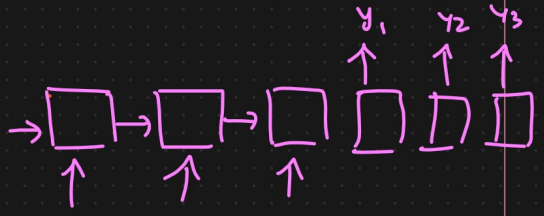
Encoder And Decoder

1. Simple RNN → Vanishing Gradient Problem
2. LSTM RNN →
3. GRU RNN → } Long Short Term Memory.
4. Bidirectional RNN ←

Type of RNN

1. Many to Many RNN



Encoder And Decoder

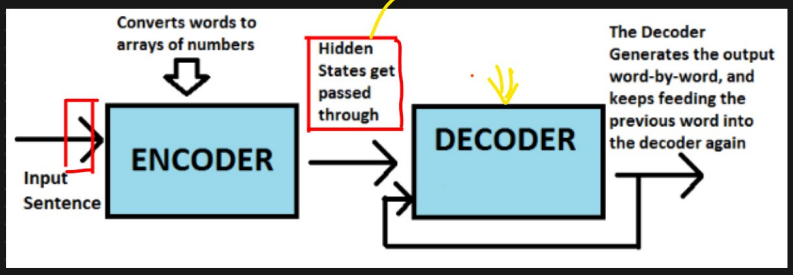
Sequences I/p O/p Sequence Of Words

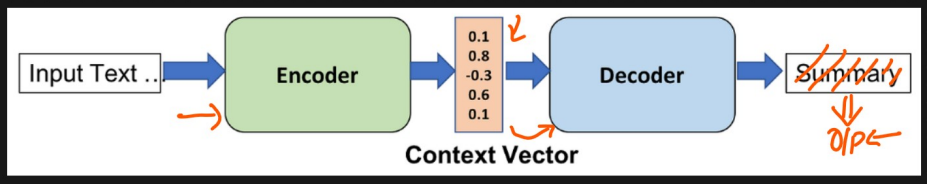
Eg: One Language To Other

English → French

Eg: Linked Chat → Hi, How are you?

Simple Working





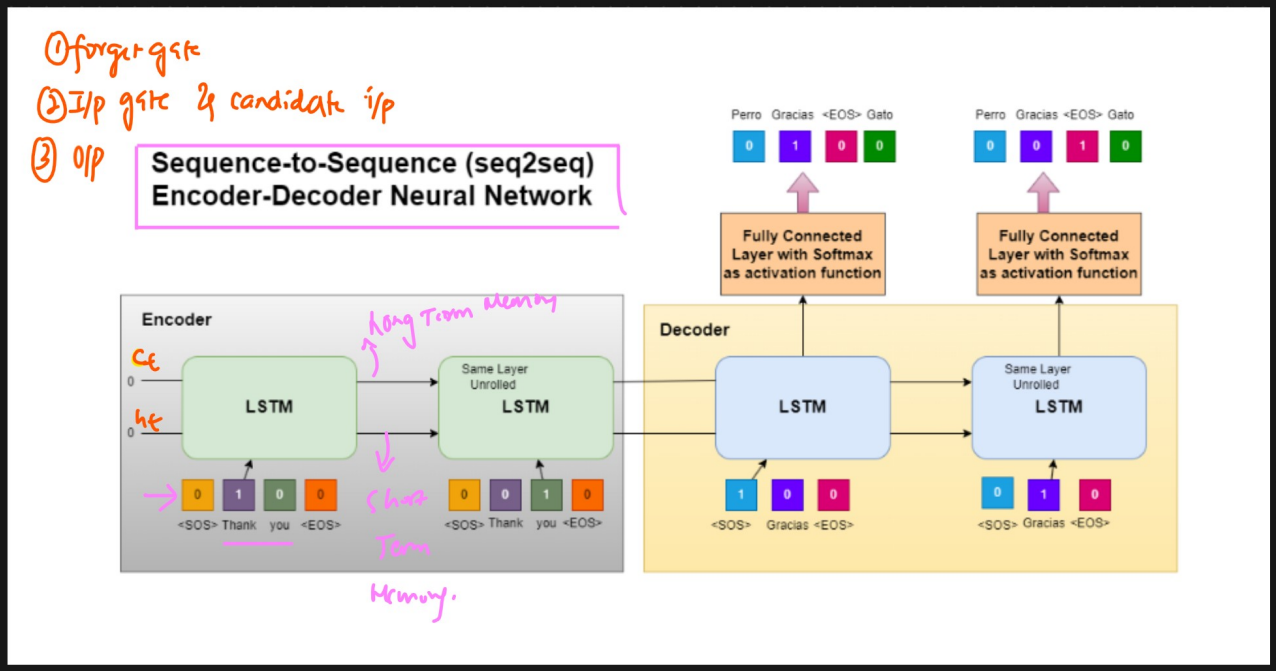
1. Encoder ⇒ I/p ⇒ Context Vector ⇐ Vectors
2. Decoder ⇒ ⇒ O/p

Usecase

1. Language Translation
2. Text Generation
3. Text Suggestion.

RNN → Vanishing Gradient Problem

① forget gate
② I/p gate & candidate i/p
③ O/p



Dataset

English: <SOS> Thank you <EOS>.

True O/p: French {Training}: Gracias

<SOS> GRACIAS → <EOS> . GRACIOUS <EOS>.

$\hat{y}$ [0.1, 0.6, 0.3] [0.1, 0.3, 0.6] $\hat{y}$ [- - -]

$[y\ \hat{y}]$ like

Encoder → Context Vectors

Long Term Memory

Short Term Memory

Softmax Softmax

Decoder

Ct, ht

ONE ⇒ ⇓ Context of this Sentence.

<SOS> Thank you <EOS>. <SOS> GRACIOUS

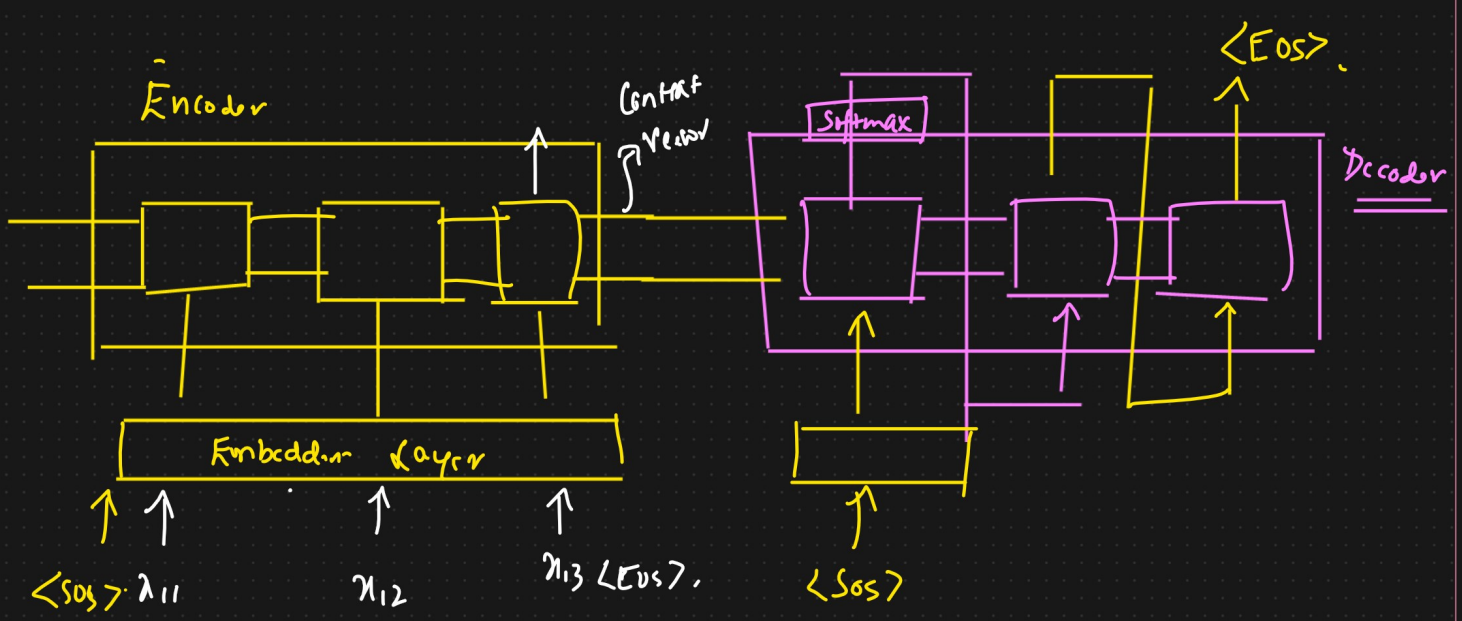
t=1 t=2 t=3 t=4 ← t=1 t=2 t=3

$\hat{y}$ [- - -] [- - -] [- - -]

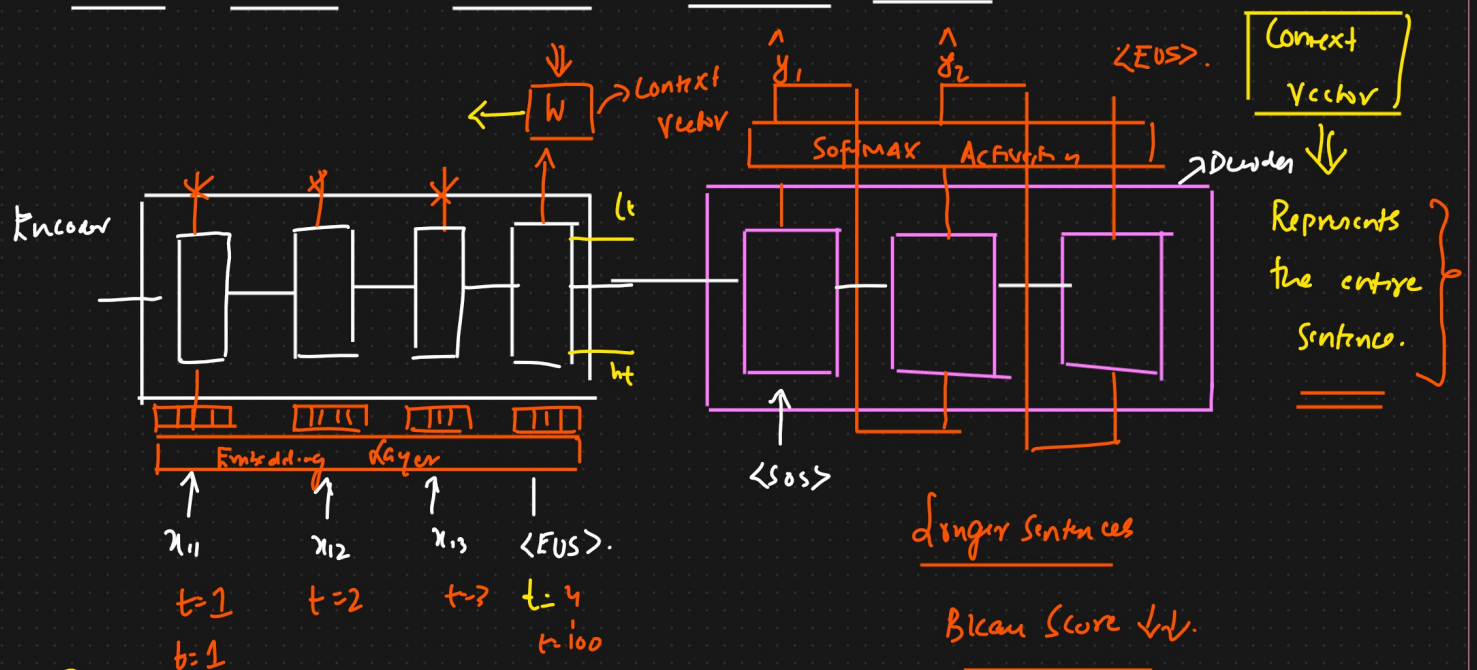
y [. . .] [. . .] [. . .]

Optimizer

Loss $(y-\hat{y})^2$



Problems With Encoder-Decoder Seq2seq Architecture

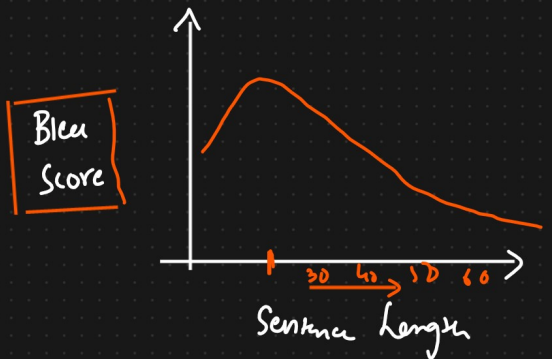


Context vector
Represents the entire sentence.

Longer Sentences

Bleau Score ↓↓.

Reasons : Sentences of varying length



⇒ Seq to Seq Data

(*) Attention Mechanism → Seq2Seq Network

Longer paragraph → {context vector}.

+

{context}