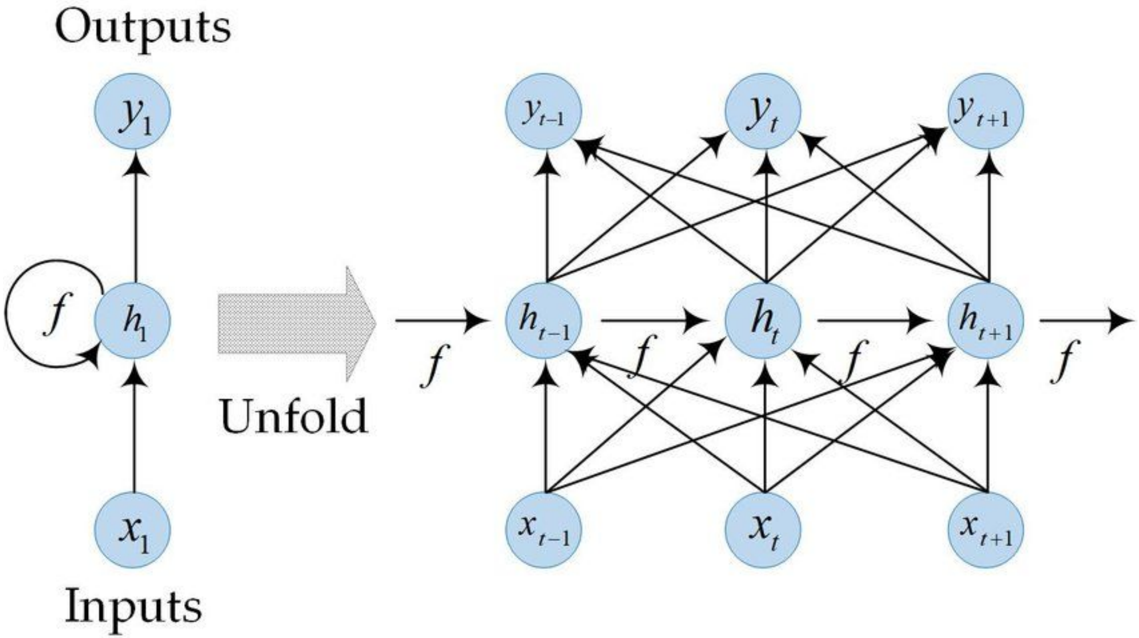
Natural Language Processing

Aanal Sonara
Mentor - Jigar Makwana

Recurrent Neural Network

- A normal neural network takes in a fixed size input network.

- RNN address the issue of variable size input vectors.

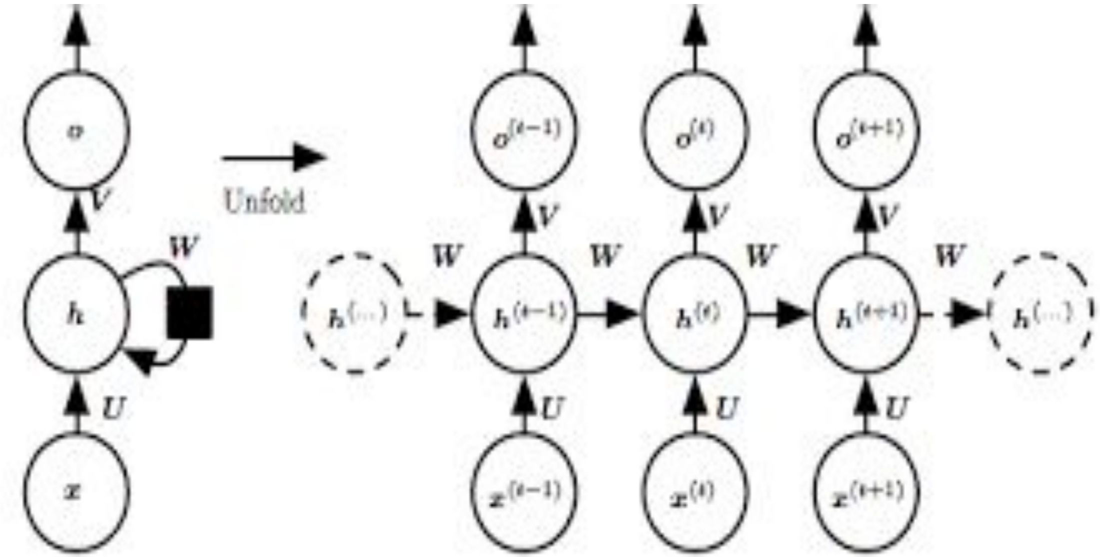


Recurrent Neural Network

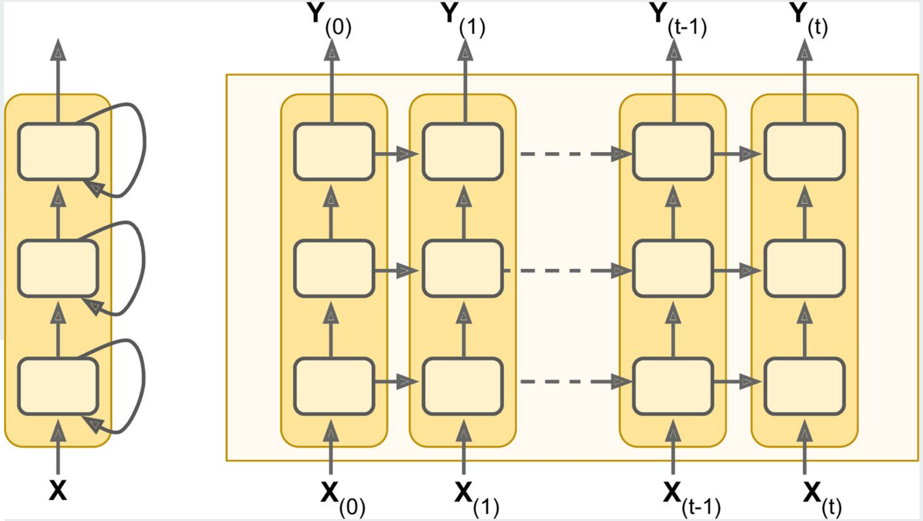
- Recurrent Neural Network remembers the past and it’s decisions are influenced by what it has learnt from the past. Recurrent Neural Network remembers the past and it’s decisions are influenced by what it has learnt from the past.
- A neural network becomes “recurrent” when you repeatedly apply the transformations to a series of given input and produce a series of output vectors.
- There is no pre-set limitation to the size of the vector.

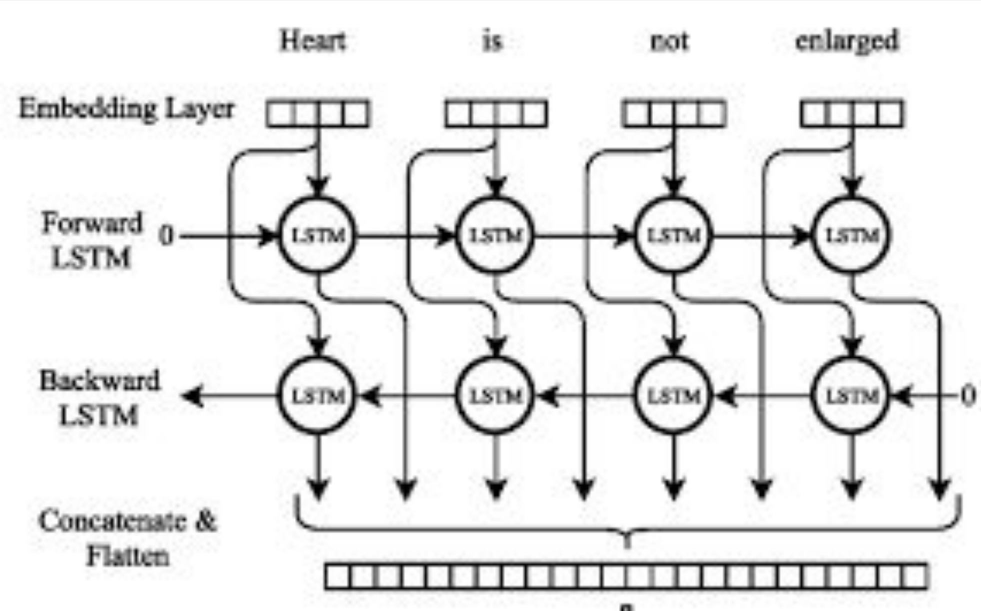
Parameter Sharing

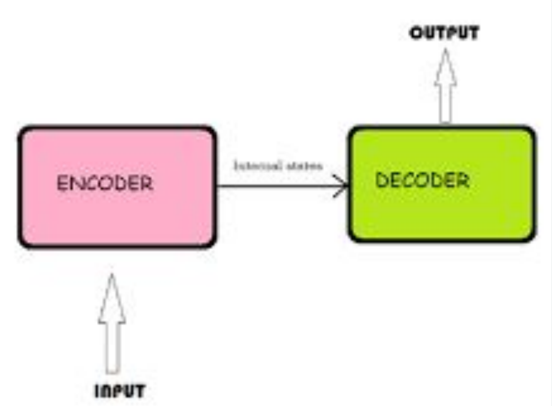
The “hidden state” that links one input to the next. The hidden state captures the relationship that neighbors might have with each other in a serial input and it keeps changing in every step, and thus effectively every input undergoes a different transition.



Types of RNNs

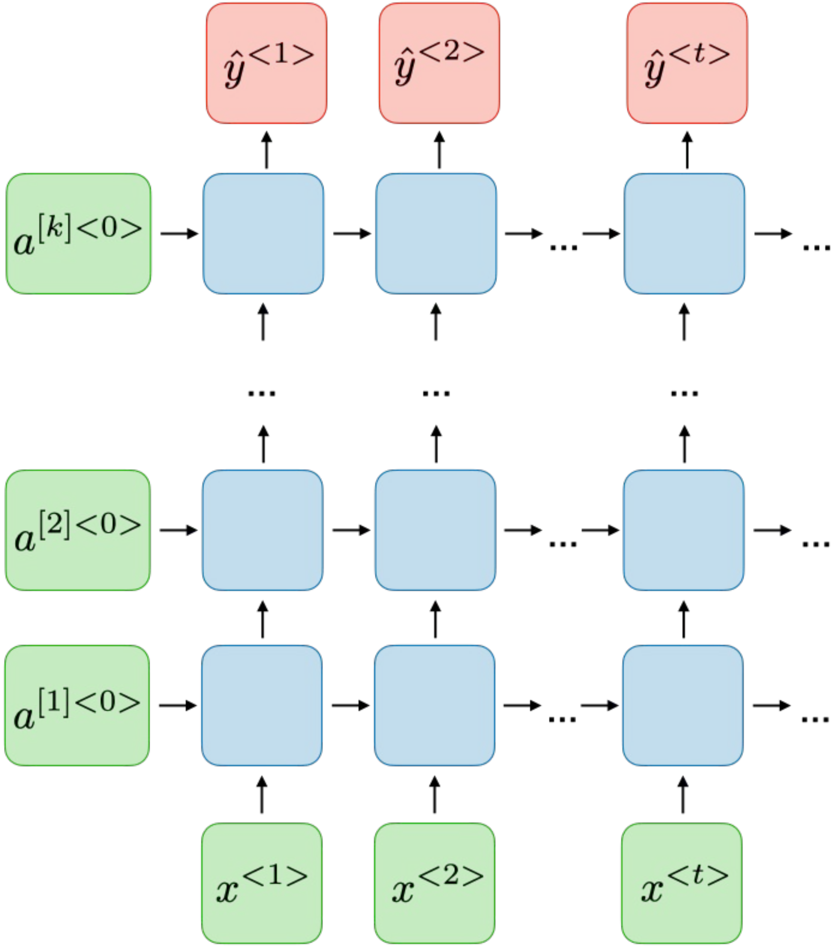






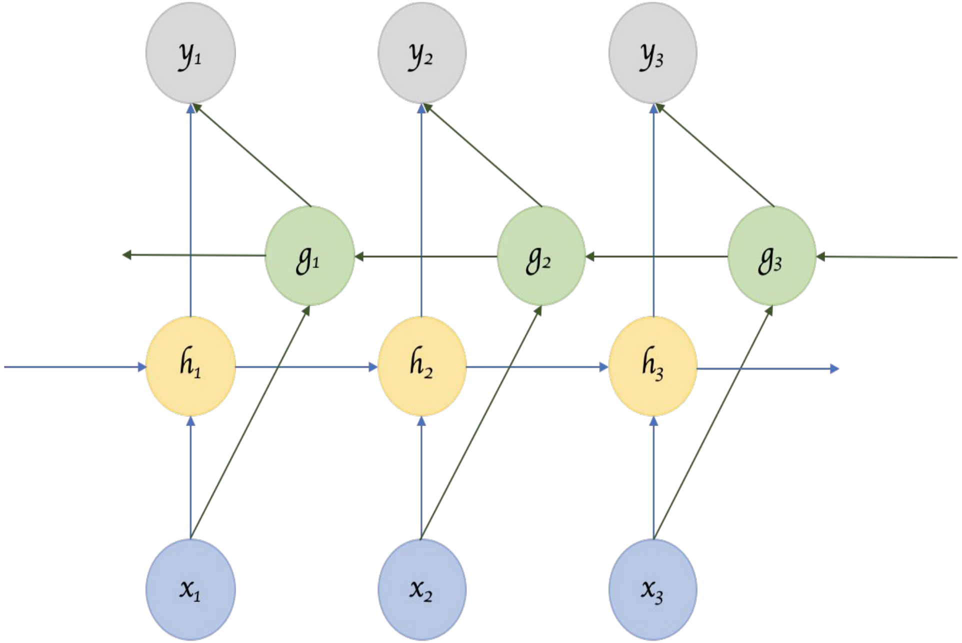
Deep RNN

1) Add hidden states, one on top of another, feeding the output of one to the next.
2) add additional nonlinear hidden layers between input to hidden state.
3) Increase depth in the hidden to hidden transition .
4) Increase depth in the hidden to output transition.



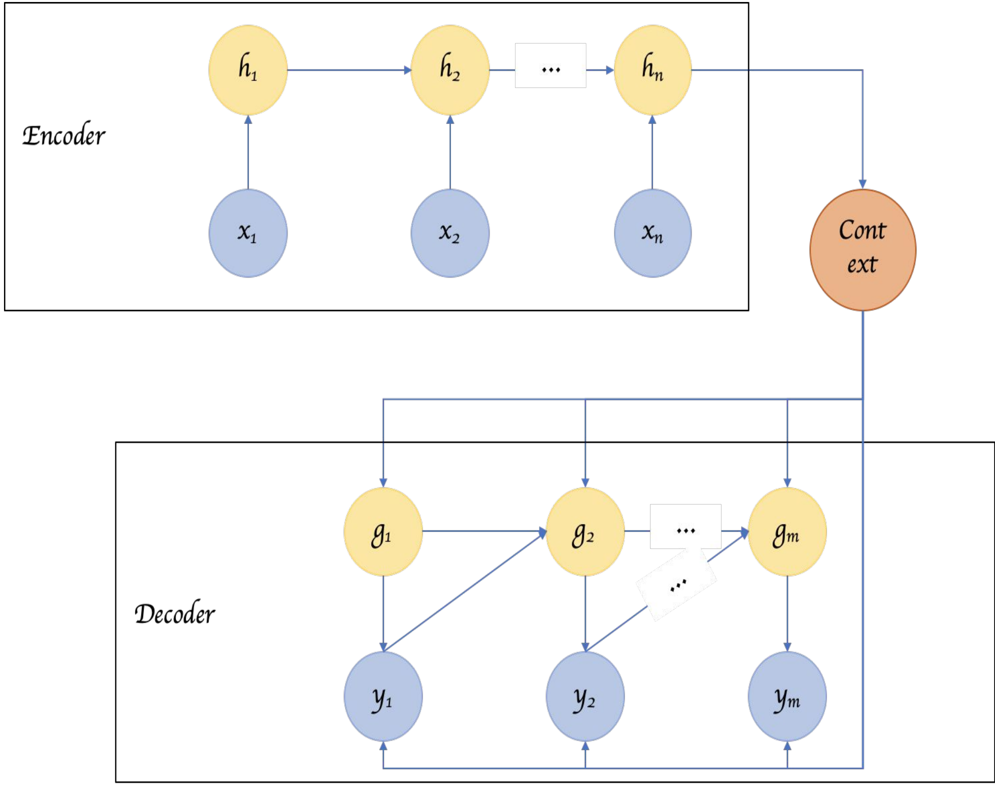
Bi-Directional RNN

1) It’s not just about learning from the past to predict the future, but we also need to look into the future to fix the past.
2) Useful in In speech recognition and handwriting recognition tasks
3) Useful to capture full context of data
4) For these applications we use bidirectional RNN which look into the ahead and backward to capture more context of data.



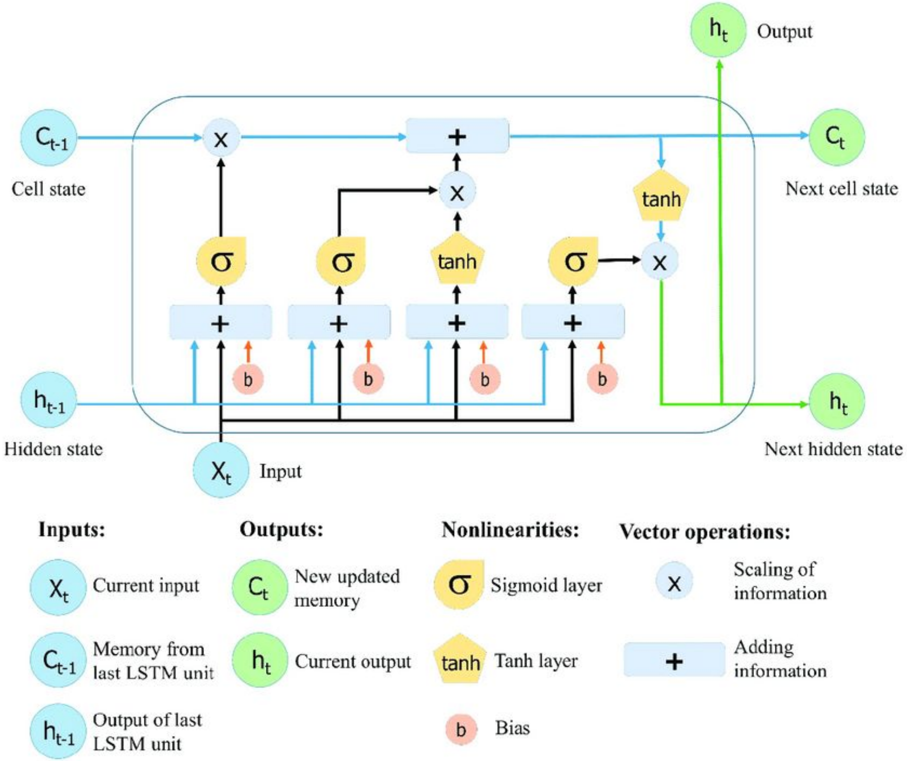
Encoder Decoder Sequence to Sequence RNN

- Used a lot in translation services.
- one an encoder that keeps updating its hidden state and produces a final single “Context” output.
- Context is then fed to the decoder, which translates this context to a sequence of outputs.
- Another key difference in this arrangement is that the length of the input sequence and the length of the output sequence need not necessarily be the same.



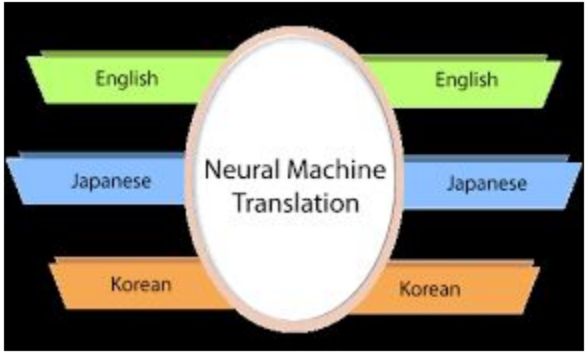
Long Short Term Memory (LSTM)

- LSTMs are designed to avoid the long-term dependency problem
- Sometimes, we only need to look at recent information to perform the present task.
- But there are also cases where we need more context.
- LSTMs are used to address this problem



Applications

- Generating Text
- Sentiment Analysis
- Speech recognition
- Text Summarization
- Call centre Analysis (automated customer service)
- Chatbots
- Machine Translation


English
English
Neural Machine Translation
Japanese
Japanese
Korean
Korean