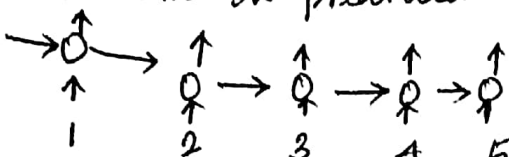


Natural Language Processing - Bird's eye view

- ① RNN. \rightarrow say a sequence of words 1, 2, 3, 4, 5
outcome is predicted for every word
- 
- context is passed across words.
- ② when the context is passed left to right & right to left
 \Rightarrow Bi-directional.
- ③ when a context from initial part of sequence is saved for later \Rightarrow Gated Recurrent Unit.
- ④ more sophisticated Gated Recurrent Unit \Rightarrow multidim.
 \Rightarrow LSTM.
- ⑤ Better method to represent text \Rightarrow Embeddings
Embeddings can be learned \Rightarrow representing words in multidimensions.
- ⑥ Transfer learning + embeddings + Bi-directional \Rightarrow BERT
Bi-directional encodes Representations from Transformers.
It's pretrained on Wikipedia & Books corpus.
requires fine tuning.
- ⑦ Sequence to sequence \Rightarrow Input \rightarrow encoder (LSTM) \rightarrow thought vectors \downarrow decoder (LSTM) \rightarrow output.
Sentence start token }
Sentence end token }
- ⑧ Attention model
preserving the hidden state of initial sequences also.
not just cell state (in case of LSTM).