Making the Best Use of Review Summary for Sentiment Analysis

IE643 Course Project

DeepMind2.0

Swapnoneel Kayal 200100154
Yash Choudhary 200100173

September 27, 2022

Outline

What is Sentiment Analysis

Sentiment Analysis

It is a Natural Language Processing (NLP) technique that is used to determine whether given data is positive, negative or neutral. It is a sequence modelling problem.

Can we use a MLP like the one used in class?

No. Some issues with using a simple neural network are as follows:

- Variable size of input/output neurons
- Too much computation
- No parameters sharing

Applications

Where is sentiment analysis used?

They can be applied to any industry, from finance and retail to hospitality and technology. We list some of them below :

- Social Media Monitoring
- Brand Monitoring
- Voice of customer (VoC)
- Customer Service
- Market Research

Recurrent Neural Networks

LSTM : Long Short Term Memory

An RNN remembers each and every information through time. It is useful in time series prediction only because of the feature to remember previous inputs as well.

In traditional neural networks, all the inputs, and outputs are independent of each other, but in cases like when it is required to predict the next word of a sentence, the previous words are required and hence there is a need to **remember the previous words** (as the next word will depend on your previous input). Hence, it solves this problem of traditional neural network with **a hidden layer**.

Motivating the need for Bi-LSTM

Let us have a close look at the following two statements:

Statement 1

Swapnoneel loves **apple** because it keeps him healthy

Statement 2

Swapnoneel loves **apple** as it produces the best electronics

In statement 1, apple $\rightarrow$ fruit
In statement 2, apple $\rightarrow$ company
The words which come prior to apple are same, so until one looks at the words after apple, it is not possible to tell if apple is a fruit or a company

Limitation of Recurrent Neural Networks

A simple RNN is unidirectional

The word **apple** will be influenced only by the previous words that are "Swapnoneel loves". However, we just realized the importance of future words.

Bi-LSTM: Bi-directional Long Short Term Memory

Bidirectional recurrent neural networks $\rightarrow$ two independent RNNs put together. This structure allows the networks to have both backward and forward information about the sequence at every time step.

Working of Bi-LSTMs

In a bi-directional LSTM, the inputs run in two ways, one from **past to future** and one from **future to past** and what differs this approach from unidirectional is an LSTM that runs backward which allows you to preserve information from the future and hence, by using **two hidden states combined** you are actually able to preserve information from both past and future at any point in time.

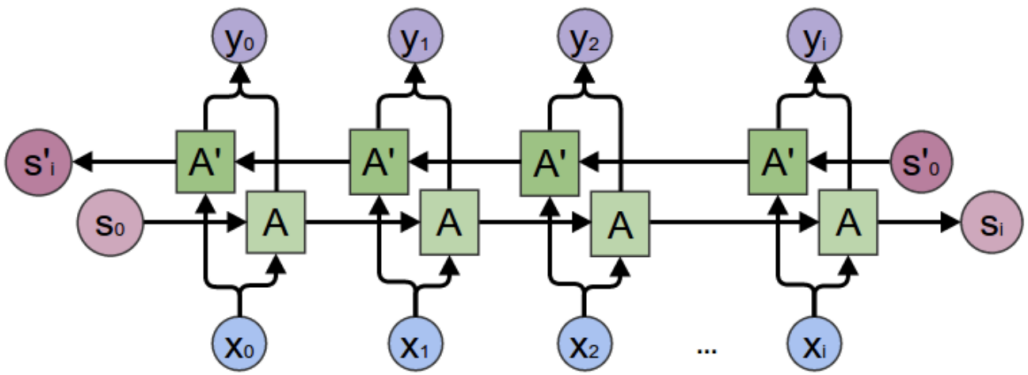

Figure: Schematic Representation of Bi-LSTM

Forward Propagation

- We apply forward propagation 2 times, one for the forward cells and one for the backward cells
- Both activations (forward, backward) would be considered to calculate the output $\hat{y}$ at time t

$$\hat{y}^{<t>} = g(W_y [\overrightarrow{a}^{<t>}, \overleftarrow{a}^{<t>}] + b_y)$$

Figure: Forward Propagation in Bi-LSTM

Forward Propagation

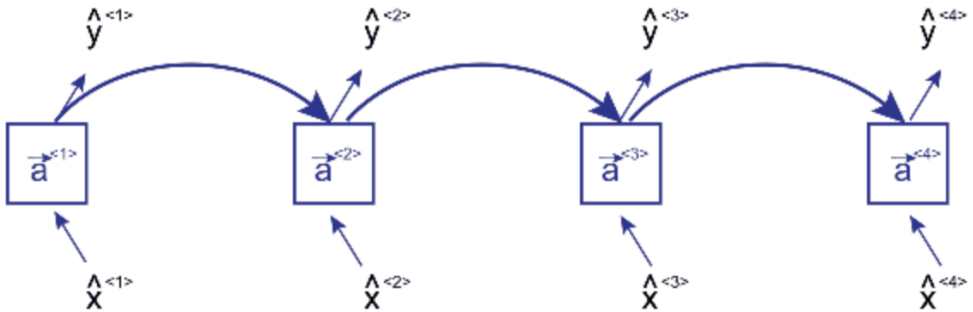

Figure: Forward RNN

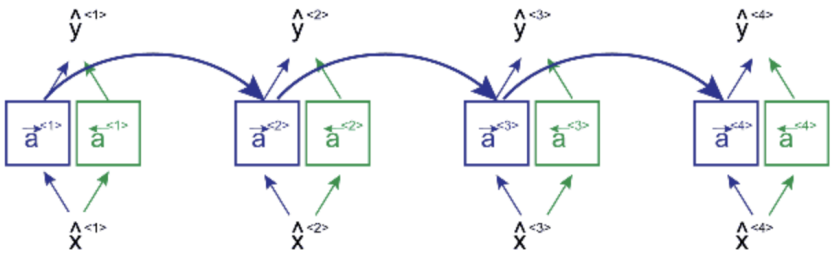

Figure: Adding another cells for the other direction

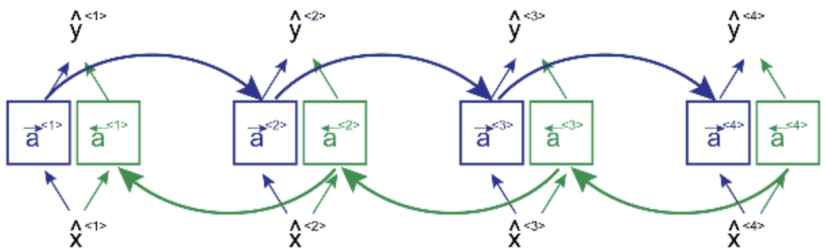

Figure: Connecting the backward cells

Backpropagation

Backpropagation Through Time (BPTT)

It is the application of the backpropagation training algorithm to recurrent neural network applied to sequence data like a time series.

The algorithm as follows:

1. Present a sequence of timesteps of input and output pairs to the network.
2. Unroll the network then calculate and accumulate errors across each timestep.
3. Roll-up the network and update weights.
4. Repeat.

BPTT

BPTT is computationally expensive!

BPTT can be computationally expensive as the number of timesteps increases.

If input sequences are comprised of thousands of timesteps, then this will be the number of derivatives that is required for a single weight update.

This can cause weights to vanish or explode (go to zero or overflow) leading to slow learning.

Conclusions

- A Bidirectional LSTM, or Bi-LSTM, is a sequence processing model that consists of two LSTMs: one taking the input in the forward direction, and the other in the backward direction
- Bi-LSTMs effectively increase the amount of information available to the network hence, improving the context given to the algorithm
- This is the model we plan on using for making the best use of Review Summary for Sentiment Analysis

References

- **Bi-LSTM :**
  www.medium.com/@raghavaggarwal0089/bi-lstm-bc3d68da8bd0
- **Differences Between Bidirectional and Unidirectional LSTM :**
  www.baeldung.com/cs/bidirectional-vs-unidirectional-lstm
- **Andrew NG's lecture on Bidirectional RNN :**
  www.youtube.com/watch?v=bTXGpATdKRY&t=330s
- **Backpropagation Through Time :**
  www.d2l.ai/chapter_recurrent-neural-networks/bptt.html