Week 7

Keep researching on LSTM, including how input flows into/out from the LSTM unit, the architecture of LSTM (forget, input, output gates), how these additional gates helps LSTM to prevent gradient exploding and vanishing that RNN often suffers, understand backpropagation through time and so on.

I first started on watching some famous university’s youtube videos (MIT, Stanford) to understand the high level concept of LSTM (without diving into the math behind of it). Once I get the overall idea of LSTM, I started to dive deeper into the maths part of it through Medium and TowardsDataScience, including the derivation of the feed-forward and backpropagation in LSTM. I found that <https://medium.com/@aidangomez/let-s-do-this-f9b699de31d9> is the most help article among others in explaining backpropagation in LSTM as it provides numerical examples and pictures throughout the explanation, thus easier to understand and visualize.