Checkpoint 3 – Project Progress and Project Proposal
02/28/2017

**Group Members**:  Anmol Khanna       (6514-0549)

Hussain Bootwala (3925-9618)                    **Group Number**: 15

The following videos on YouTube were viewed to better our understanding of this topic:

- https://www.youtube.com/watch?v=iX5V1WpxxkY : This video is from Stanford cs231n course on deep learning. It goes about explaining what recurrent neural networks are and talks about its applications in machine translation systems and also talks about LSTM (long short term memory) networks.

- https://www.youtube.com/playlist?list=PL6Xpj9I5qXYEcOhn7TqghAJ6NAPrNmUBH&app=desktop : This is a video series of 92 short videos covering neural networks starting from what a perceptron is to the deep neural networks and then talking about natural language processing and explaining the use of encoders and decoders in the neural networks. I have stated with this videos series and hope to complete it by this week.

Apart from the videos, I am following a couple of blogs on neural networks :

- http://www.wildml.com/2015/11/understanding-convolutional-neural-networks-for-nlp/#more-348 : This is a blog that explains all of neural networks in detail and talks about the cases where which network needs to be applied. Also in this, it talks about the implementation challenges concerned with these multilayer networks.

- http://colah.github.io/posts/2015-08-Understanding-LSTMs/ : This is the second blog I am following. In this particular link, I learned about the LSTM(Long short term)networks. These LSTM networks are are a special type of RNN networks that are capable of handling long term dependencies that the RNN's lacked and hence these LSTM networks have gained much attention.

The following papers were referenced to further our knowledge on this topic:

- Neural Machine Translation By Jointly Learning To Align And Translate - https://arxiv.org/pdf/1409.0473.pdf: This paper talks about the how machine translation works and the idea behind it. It also talks about the encoders and decoders that are used in the traditional machine translation systems. It points out that though, the existing architecture is efficient but its inefficient when it comes to dealing with long length sentences and the vanishing gradient problem and it then discusses the approach for rectifying the same.

- Recurrent neural machine translation (Biao Zhang, Deyi Xiong and Jingsong su) – Like the previous paper, it also discusses the traditional encoders and decoders and how the RNN networks are built for machine translation and explains there working. It points out that the problem of handling long length sentences can be tackled using a dynamic approach based on the concept of traditional probabilities.

I have also started working on the coding part. That is just started to code the perceptron that Dr Banerjee explained to us in the class. I am using python and Numpy package in python for the same.