**Title**: RNN and LSTM

**Study Performed By:** Dipesh Gyawali

**Date:** 29- Jan 2020

## Objectives:

- To learn the radical concepts of RNN and LSTM
- To study the implementation methodology of LSTM in series of data
- To learn how RNN can be used in Computer Vision fields

## Findings:

Recurrent neural networks, or long short term memory(LSTM) are the most powerful and well known subset of artificial neural network designed to recognize patterns in sequences of data, such as numerical times series data emanating from sensors, stock markets and government agencies (but also including text, genomes, handwriting and the spoken word). What differentiates RNNs and LSTMs from other neural networks is that they take time and sequence into account, they have a temporal dimension.

Recurrent networks take as their input not just the current input example they see, but also what they have perceived previously in time. The decision a recurrent net reached at time step `t-1` affects the decision it will reach one moment later at time step `t`. So recurrent networks have two sources of input, the present and the recent past, which combine to determine how they respond to new data, much as we do in life. Recurrent networks are distinguished from feed forward networks by that feedback loop connected to their past decisions, ingesting their own outputs moment after moment as input. It is often said that recurrent networks have memory. Adding memory to neural networks has a purpose. There is information in the sequence itself, and recurrent nets use it to perform tasks that feed forward networks can't.That sequential information is preserved in the recurrent network's hidden state, which manages to span many time steps as it cascades forward to affect the processing of each new example. It is finding correlations between events separated by many moments, and these correlations are called "long-term dependencies", because an event downstream in time depends upon, and is a function of, one or more events that came before. One way to think about RNNs is this: they are a way to share weights over time. Just as human memory circulates invisibly within a body, affecting our behavior without revealing its full shape, information circulates in the hidden states of recurrent nets. The English language is full of words that describe the feedback loops of memory. When we say a person is haunted by their deeds, for example, we are simply talking about the consequences that past outputs wreak on present time. Given a series of letters, a recurrent network *will* use the first character to help determine its perception of the second character,

such that an initial `q` might lead it to infer that the next letter will be `u`, while an initial `t` might lead it to infer that the next letter will be `h`.

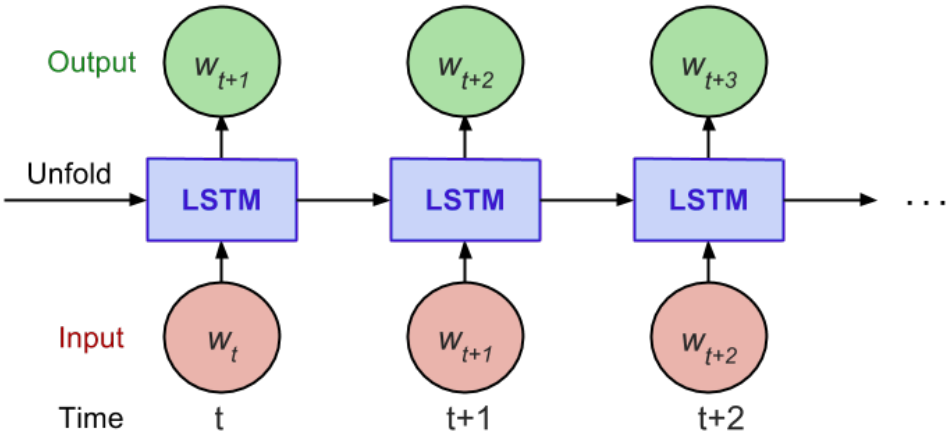
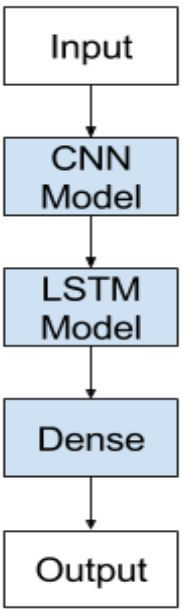



Fig: Recurrent Neural Network

## Usage of RNN(LSTM) in computer vision

Input with spatial structure, like images, cannot be modeled easily with the standard Vanilla LSTM. The CNN Long Short-Term Memory Network or CNN LSTM is an LSTM architecture specifically designed for sequence prediction problems with spatial inputs, like images or videos. The CNN LSTM architecture involves using Convolutional Neural Network (CNN) layers for feature extraction on input data combined with LSTMs to support sequence prediction. CNN LSTMs were developed for visual time series prediction problems and the application of generating textual descriptions from sequences of images (e.g. videos).



**Reviewed By:** Badal Shrestha

**Reviewed Date:** 29-Jan 2020

## Reviewer's Comment:

Good Description about RNN. Detailed description of LSTM network like workflow of network and its use case in real world scenario can be studied more. Further research is needed to be done on LSTM network.

*Note : Do research on Image captioning technique*