Time Series Prediction Stock Market Prices

Priyank Shah

Pratikkumar Patel

Anirvan Maiti

Time Series Data - Examples

- Customer Behavior
- Seasonal Patterns
  - Ticket Sales
  - Polar Ice Caps Data
  - Weather Patterns
- Stock Prices

Problem Definition

- Factors Impacting Stock Market Prices
  - Company Performance (Quarterly Results)
  - Future of business
  - Momentum
  - News (Market, Company or Sector, Regulations)
  - ...
- Prediction very difficult because of these factors
- Our focus of analysis is to identify trends using Neural Nets
  - Recurrent Neural Nets
  - LSTM
  - WaveNet by DeepMind?
- Enables further extension to include other features like
  - Volume
  - Social Media Buzz
  - Breaking News

Data Set

AAPL

- Data from Nasdaq
- Time Period from Sept-2011 to Mar-2017
- Tick Prices

Pre-Processing to reduce number of price entries in a day

- Weighted Average using volume
  - 1 Min
  - 5 Min
  - 15 Min
- Stock-Split Adjustments (06/06/2017 : 1 to 7)

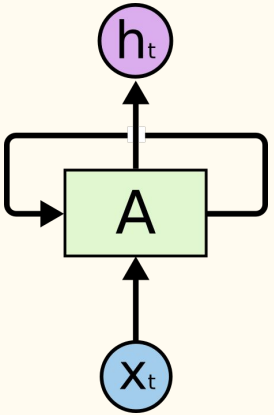
Models

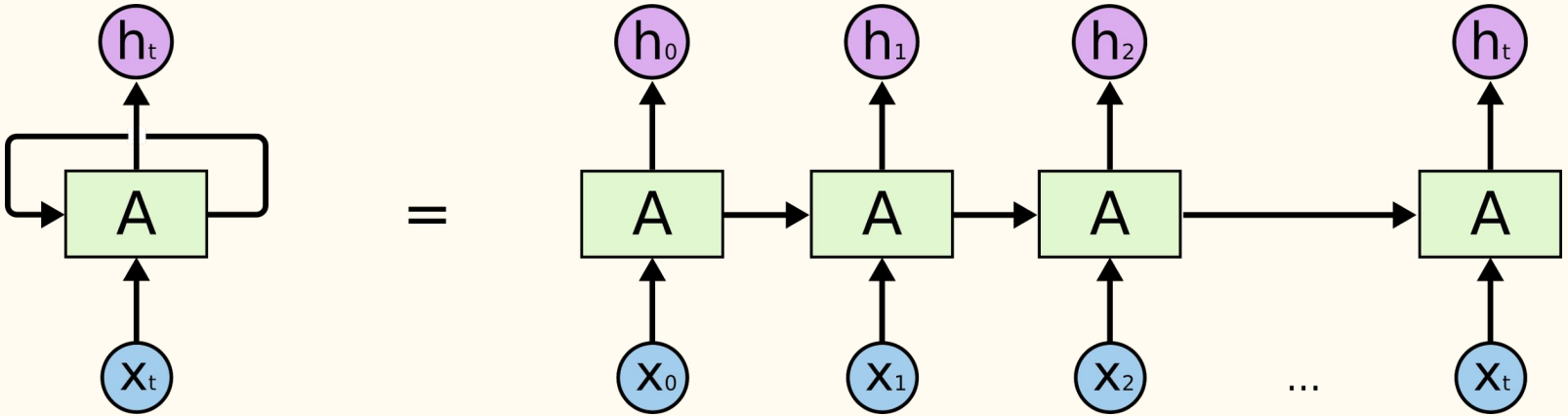
What we have tried?

- Long Short Term Memory Networks(LSTM)

RNN (Recurrent Neural Network)

- Before we move on to LSTM,
- What is RNN?
- Problem with RNN

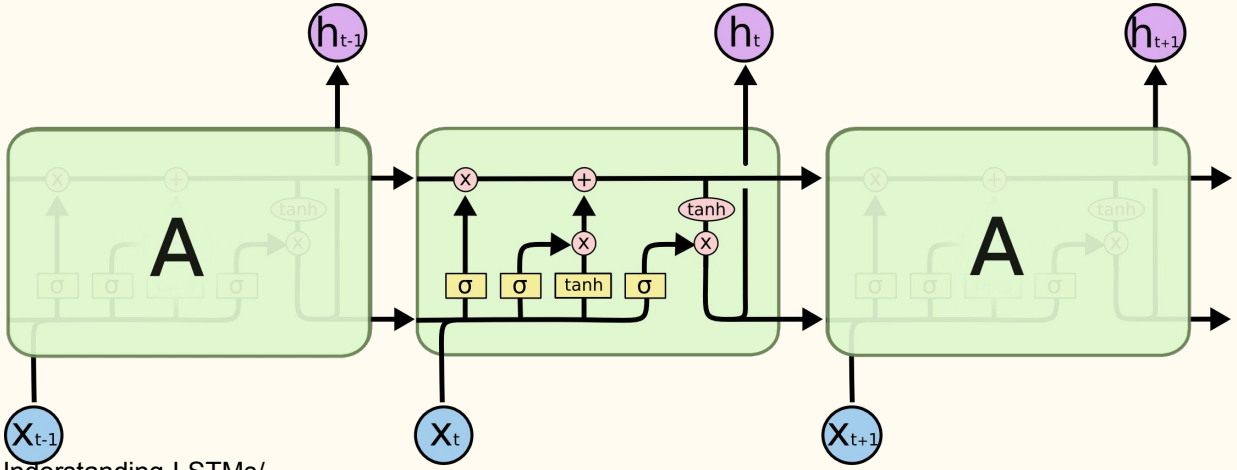




Source : http://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM

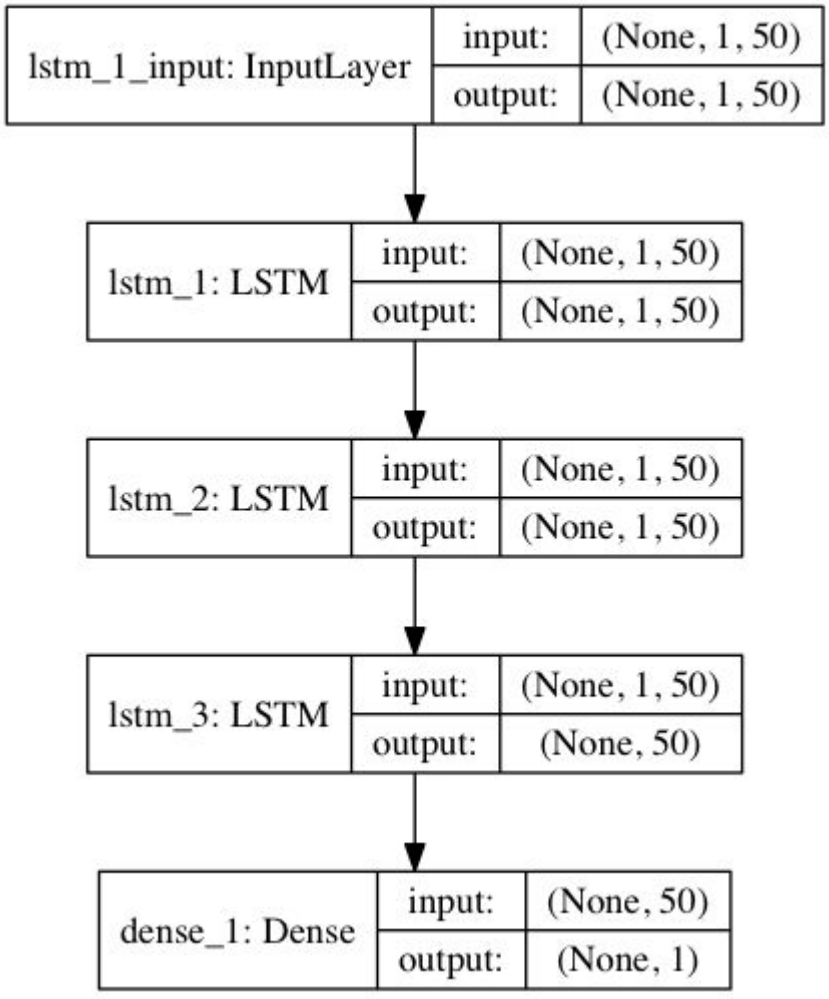
- Why LSTM?
- A cell has ability to remove or add information to the cell state, carefully regulated by structures called gates.
- Gates are composed out of a sigmoid neural net layer and a pointwise multiplication operation.



Source : http://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM - Model

- MinMaxScaler (range 0 - 1)
- Look Back Parameter
  - Helps generate data
  - Moving window of length look_back
  - Decides input shape
- LSTM Layers
- Dense Layer
- Inverse Scaling
- Loss - Mean Squared Error
- Epochs - 30 to 50



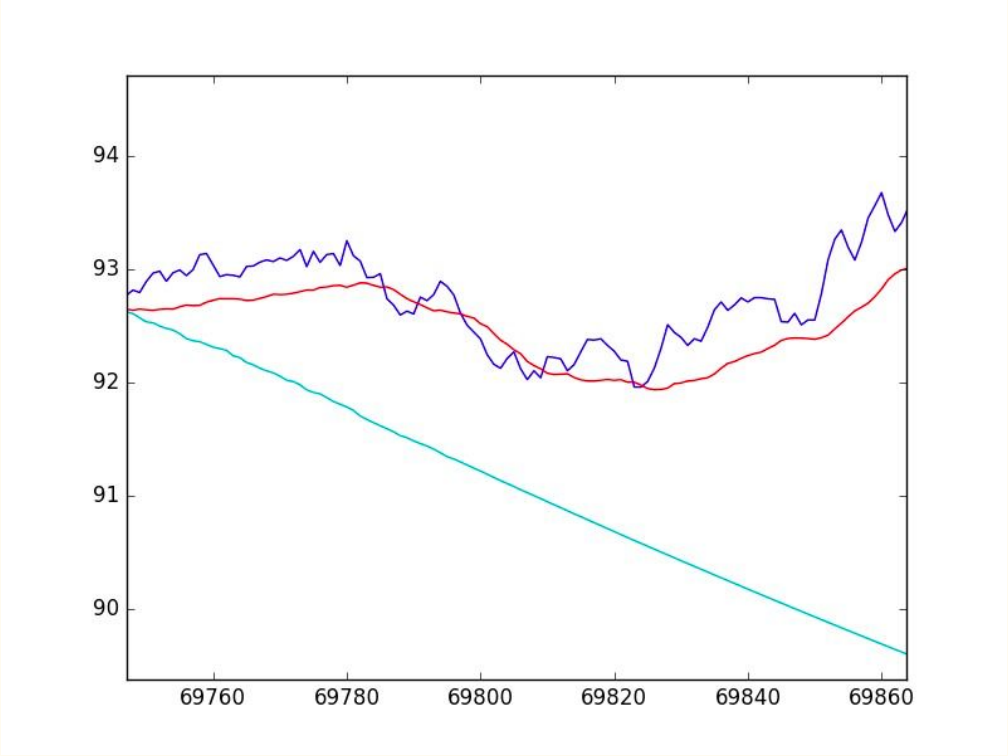
LSTM - Training & Testing Results

	Model 1 1 LSTM Layer	Model 2 2 LSTM Layers	Model 3 3 LSTM Layers
Look Back 50 Epochs 30	Train Score: 0.53 RMSE Test Score: 0.63 RMSE Test Score: 14.91 RMSE	Train Score: 0.46 RMSE Test Score: 0.45 RMSE Test Score: 28.98 RMSE	Train Score: 0.45 RMSE Test Score: 0.56 RMSE Test Score: 57.02 RMSE
Look Back 100 Epochs 30	Train Score: 0.48 RMSE Test Score: 0.45 RMSE Test Score: 28.39 RMSE	Train Score: 0.41 RMSE Test Score: 0.92 RMSE Test Score: 15.30 RMSE	Train Score: 0.54 RMSE Test Score: 0.67 RMSE Test Score: 25.38 RMSE
Look Back 50 Epochs 50	Train Score: 0.39 RMSE Test Score: 0.36 RMSE Test Score: 31.22 RMSE	Train Score: 0.40 RMSE Test Score: 0.43 RMSE Test Score: 29.86 RMSE	Train Score: 0.35 RMSE Test Score: 0.39 RMSE Test Score: 56.66 RMSE

The Moment of Truth...

LSTM - Prediction

- RED - Moving Window
- CYAN - Prediction over Prediction
- BLUE - Ground Truth



LSTM - Conclusions

Look Back Dependency

- Shorter the window
  - Better for short term prediction
- Longer window
  - Worse for short term
  - Better Long term prediction

Model Dependency

- No of LSTM Layers
  - More layers means better short-term performance
  - Long term performance is adversely impacted

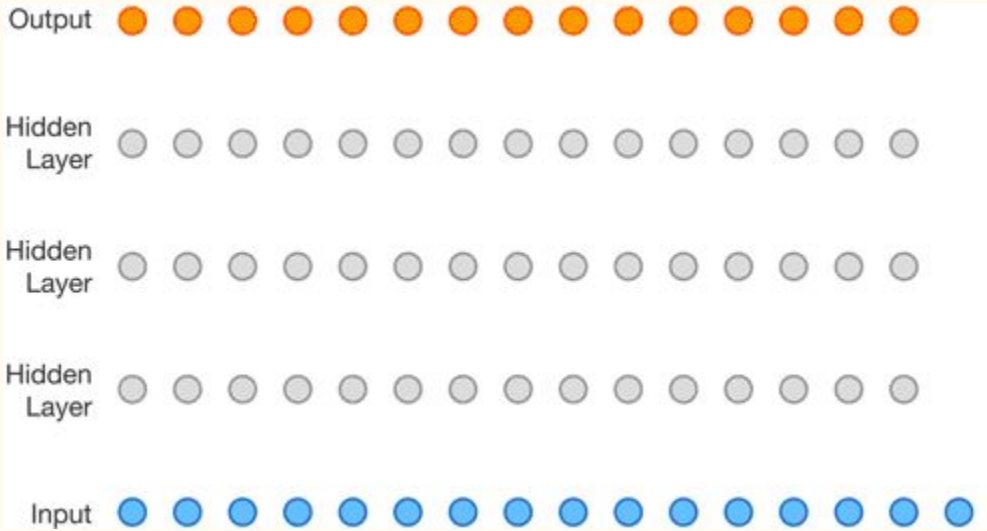
Models

What we have tried?

- WaveNet by DeepMind

WaveNet

- Why WaveNet?
- Built for Audio where Prediction for every samples is influenced by all previous ones
- PixelNet applied to 1D data



Source : https://deepmind.com/blog/wavenet-generative-model-raw-audio/

WaveNet - Model

```
"dilations": [1, 2, 4, 8, 16, 32, 64, 128, 256, 512,
              1, 2, 4, 8, 16, 32, 64, 128, 256, 512],
                                        #A list with the dilation factor for each layer.
"residual_channels": 64,                #How many filters to learn for the residual.
"dilation_channels": 32,                #How many filters to learn for the dilated convolution.
"quantization_channels": 1000,          #How many amplitude values to use
```

WaveNet - Training & Results

- Difficulties
  - Data Preparation
  - Non - convergence
  - Time Consuming
- Testing In Progress

Questions?