Effectiveness of LSTM Model for Financial Time Series Prediction: Evidence from New Zealand

Wajira Dassanayake

Let us view this video

Source: https://www.youtube.com/watch?v=OJn4vJN1ko
1 Introduction

- Stock market is one of the most fascinating, sophisticated and complex financial markets whose movements are influenced by multitude of interwoven macroeconomic, historical and spontaneous factors.
- Globalisation and financial market integration intensifies further complexities.
- Accurate prediction of stock price/index movement becomes important.
- A considerable research attention has been devoted to stock market index/price prediction.
2 Purpose

• To evaluate the forecasting effectiveness of LSTM (a deep learning model)
• To compare and contrast the predictive efficacies of the devised LSTM with ARIMA and HWES

3 Methodology

• Scenario: Assume we are watching Den of Thieves.
• To understand and to forecast the next scene/event we don’t recall everything from the beginning. We rely on the recent experiences happening in the movie and learn from them.
• This means our thoughts have persistence of memory.
• A conventional neural network is unable to learn from the previous events because the information does not pass from one step to the next and this a major shortcoming.
• RNN learns information from immediate previous step.
3 Methodology

• Scenario: Assume we are watching **Den of Thieves**.

• RNN learns information from immediate previous step.

A RNN can be thought of as multiple copies of the same network, each passing a message to a successor. Sometimes we do not need our network to learn only from immediate past information. This is the issue with RNN.

RNN limitations were investigated in depth by Hochreiter & Schmidhuber (1997) and introduced **LSTM**.

Source: [Understanding LSTM and its quick implementation in Keras for sentiment analysis](https://towardsdatascience.com/understanding-lstm-and-its-quick-implementation-in-keras-for-sentiment-analysis-af100685497)
3 Methodology

• Scenario: Assume we are watching Den of Thieves.

• RNN learns information from immediate previous step.

\[ h_t = \text{Output} \]
\[ x_t = \text{Input} \]
\[ A = \text{Neural network} \]

Sometimes we do not need our network to learn only from immediate past information. This is the issue with RNN.

LSTMs are used for speech recognition, handwriting recognition, language modelling, sentiment analysis, text prediction, financial time series prediction.


**Forget Gate**
- Controls what information to throw away from memory.
- Decides how much of the past you should remember.

**Update/Input Gate**
- Controls which information will be used to modify the memory.
- Decides how much of this unit is added to the current state.

**Output Gate**
- Conditionally decides what to output from the memory.
- Decides which part of the current cell makes it to the output.

LSTM had a three step process which is given at the figure that shows LSTM module has 3 gates named as Forget gate, Input gate, Output gate.
3 Methodology

• **Software used:**

  • **Python** provides well-structured and tested environment for LSTM.
  
  • Well known **deep libraries** “Keras”, “TensorFlow” and “Theano” are utilised.
  
  • Additional libraries (“**math**, “**numpy**”, “**matplotlib**”, “**pandas**”, “**sklearn**”) which are required for the implementation were installed.
  
  • **Spyder** the **scientific python development environment** available with **Anaconda** is used for development after installing the required packages.

Source: [https://upload.wikimedia.org/wikipedia/commons/7/7e/Spyder_logo.svg](https://upload.wikimedia.org/wikipedia/commons/7/7e/Spyder_logo.svg)

3 Methodology

• **Data used:**

  • **S&P/NZX50 Index** data are extracted from S&P Dow Jones Indices produced by S&P Global. The index is developed to capture the overall performance of the 50 largest stocks listed on the Main Board (NZSX) of New Zealand’s Exchange (NZX).
  
  • We use daily price series of **S&P/NZX50 Index** from **2009** to **2017** having a total number of **2173** observations.
  
  • To evaluate the performance of each model configured, the first **1521** observations (approximately **70%**) are used as the **training** sample and the rest (**652**) of the observations are used for **prediction/testing** purposes.
3 Methodology

Time plot of Data used:

We use daily price series of S&P/NZX50 Index from 2009 to 2017 having a total number of 2173 observations.

3 Methodology

• Epochs and batch sizes: To determine the best fit LSTM network, different numbers of epochs and batch sizes were used combined with RMSE. Adopting a manual experimentation process, the best fit model was found at 550 training epochs with a batch size of 152.

\[ RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (X_t - \hat{X}_t)^2} \]

• Learning curves: Learning curves show the relationship between training set size and the RMSE evaluation metric on the training and validation/test sets. Learning is on y-axis and time/experience on the x-axis.
3 Methodology

Section of the devised algorithm

4 Results

**Learning curve**

A good fit is identified by a training and validation loss that decreases to a point of stability with a minimal gap between the two final loss values.

- training loss decreases to a point of stability
- Test/validation loss decreases to a point of stability and has a small gap with the training loss

RMSE: 18.659
4 Results

LSTM Actual vs Prediction

RMSE: 18.659

550 training epochs with a batch size of 152

5 Conclusion

Taking a technical analysis perspective, the study used LSTM network, a deep learning method.

550 training epochs with a batch size of 152. Predictive power is strong with a RMSE of 18.659.
6 Limitation

- Selection of ideal epoc size and batch size for generalised optimum model should be arrived in an automated process.
- I plan to devise a strategy to generate an automated process to capture the optimum epoc and batch size to make my algorithm an iterative optimization algorithm.
- Possibly I will focus on “Transformers”, a relatively new neural network architecture, gaining popularity as more efficient algorithm in comparison LSTM.

End of the presentation

Thank you

Questions?