Using LSTM in Stock prediction and Quantitative Trading

Zhichao Zou zzou@Stanford.edu

Zihao Qu quzihao@Stanford.edu

Video: https://www.youtube.com/watch?v=K3w_zsz1qYI

I. Motivation

- Accurate prediction on finance market will yield attractive profit
- Traditional methods rely on subjective views on economy and company's future direction
- Quantitative methods start to leverage machine learning to avoid human subjectivity and emotion

Figure 1. Deep learning is expected to help avoid stock plunge and generate profit.

II. Data & Features

Historical daily stock price and quarterly reported corporate accounting statistics were collected from Kaggle Open Dataset and merged by date.

- Train set: 2004 to 2011
- Dev set: 2012
- Test set: 2013
- Features: adjusted close price, trade volume, 'Debt-to-Equity Ratio', 'Return on Equity', 'Price-to-Book' ratio, etc.

date	adj_close	volume	DE Ratio	Return on Equity	Price/Book	Profit Margin	Diluted EPS	Beta
2013-12-17	0.950181	0.019849	0.652485	0.01414	0.046009	0.626554	1.0	0.24031
2013-12-18	0.964769	0.036554	0.652485	0.01414	0.046009	0.626554	1.0	0.24031
2013-12-19	0.966209	0.023068	0.652485	0.01414	0.046009	0.626554	1.0	0.24031
2013-12-20	0.980317	0.062588	0.652485	0.01414	0.046009	0.626554	1.0	0.24031
2013-12-23	0.994504	0.024452	0.652485	0.01414	0.046009	0.626554	1.0	0.24031
2013-12-24	0.991310	0.000000	0.652485	0.01414	0.046009	0.626554	1.0	0.24031
2013-12-26	0.996816	0.014948	0.652485	0.01414	0.046009	0.626554	1.0	0.24031
2013-12-27	0.997737	0.020691	0.652485	0.01414	0.046009	0.626554	1.0	0.24031
2013-12-30	0.988978	0.012428	0.652485	0.01414	0.046009	0.626554	1.0	0.24031
2013-12-31	1.000000	0.015446	0.652485	0.01414	0.046009	0.626554	1.0	0.24031

Figure 2. An example of input data structure

III. Models

Both classic time series and deep learning models were experimented:

- Linear model (ARIMA)
- Single-LSTM
- Stacked-LSTM
- Attention-LSTM

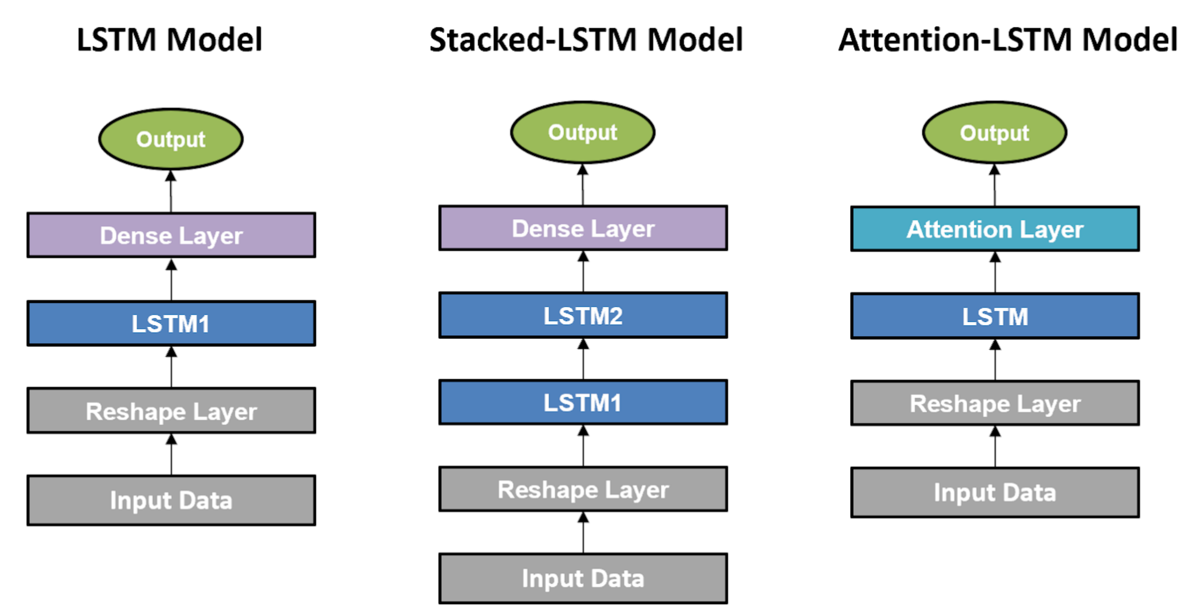

Figure 3. Illustration of three deep learning architects

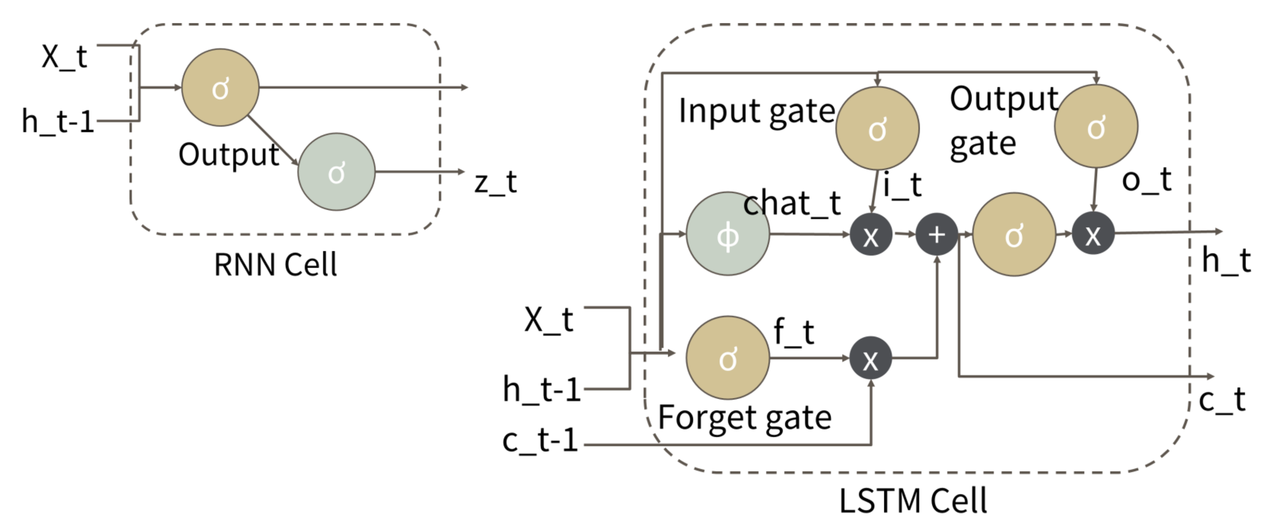

Figure 4. Illustration of typical RNN unit and LSTM unit

Besides prediction model, we also built two trading strategies to check the capability of gaining profit based on modelling results:

- Long-Only Strategy
- Long-Short Strategy

Two metrics were used to evaluate model performance:

- Mean-Squared-Error
- Total Annual Return on test set

IV. Results

Mean-Squared-Error:

- ARIMA: (p, d, q) = (1,0,1), MSE = 0.0546

ARIMA didn't perform well on predicting non-linearity and long term results.

- Attention-LSTM outperformed among three deep learning models as expected.
- Single layer LSTM performed better than Stacked-LSTM

Mean Squared Error ($*10^{-3}$)			
Stock Ticker	**LSTM**	**Stacked-LSTM**	**Attention-LSTM**
XOM	25	25	11
GE	18	2	3
WMT	128	163	76
MSFT	40	27	16
IBM	28	7	6
AAPL	54	61	44
GOOG	35	37	26
GS	12	16	18
PFE	32	64	49
JNJ	57	123	41

Figure 5. The summary of MSE on test set for three deep learning models.

Total Annual Return:

- Only Single-LSTM and Attention-LSTM were evaluated
- Both models significantly outperformed the benchmark

Trading Strategy Annual Return			
Annual Return	**LSTM**	**Attention-LSTM**	**Benchmark**
Long-Only Strategy	173%	266%	30%
Long-Short Strategy	99%	263%	30%

Figure 6. The summary of MSE on test set for three deep learning models.

References

[1] Framework used: Keras, tensor flow.

[2] Toxic comment classification challenge, https://www.kaggle.com/qqgeogor/keras-lstm-attention-glove840b-lb-0-043/log.

[3] Ryo Akita, Akira Yoshihara, Takashi Matsubara, and Kuniaki Uehara. Deep learning for stock prediction using numerical and textual information. In *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, pages 1–6. IEEE, 2016.

[4] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[5] JB Heaton, NG Polson, and Jan Hendrik Witte. Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry*, 33(1):3–12, 2017.

V. Discussion

1. Attention-LSTM is the best:

- A-LSTM brings the advantage of selecting the important and relevant information
- Significant better performance on stocks which experienced large fluctuation, demonstrating the significance of relying on the context
- The result of total annual return further emphasized the advantage of selecting most relevant info with the attention model

2. Single-LSTM beats Stacked-LSTM

- Opposite to our expectation as we assume complex architecture can learn non-linearity of stock trend better
- Stacked-LSTM doesn't generalize well on unseen data
- Stacked-LSTM is more suitable for predicting classification problems rather than continuous time series

VI. Future Work

- Denoised data with Attention-LSTM: one difficulty of predicting stock arises from its non-stationary behavior, denoised data may improve the accuracy
- Incorporate high frequency data rather than quarterly reported statistics: one good example is the recent stock plunge due to COVID19, feeding timely info may help improve model performance.