David William Edwards II

8/19/2018

AI Stock Market Prediction: Radial Basis Function vs LSTM Network

Abstract

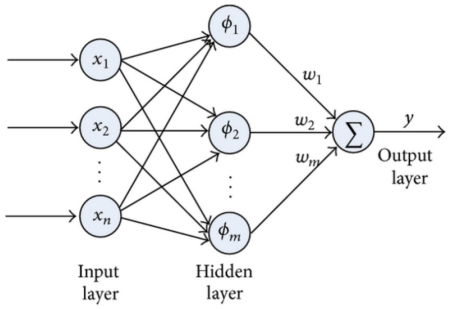
As advancements in deep learning methods continue, it is important to remember that adding complex methods does not guarantee accurate results. Simple machine learning methods can give better accuracy and are, generally, more time efficient. Experimenting with two of the most popular methods of stock market predicting will show the idea that complex methods do not guarantee highly accurate prediction. Testing will be accomplished by a radial basis function network as the simple method and a long short-term memory neural network as the complex method. Both methods will be used to predict the next day's stock price for AAPL and BTC-USD in July 2018.

Introduction

In the past years, complex neural networks have been getting more attention than simpler support vector methods. Albert Einstein's famous quote "*Everything should be made as simple as possible, but not simpler*" inspired the idea to show that sometimes newer complex methods do not always equal better results. Demonstrating this idea will be a popular simple machine learning method radial basis function (RBF), and the complex neural network method will be a long short-term memory (LSTM). There are many different methods to choose in both categories; these were chosen because of their popularity in predicting stock market prices.

RBF

The RBF consists of three different layers: the input, hidden, and output. Each layer is fully connected to the preceding layer. The input is responsible for passing the vectors of the input data to the hidden layer. The hidden layer units are activated based on the associated radial basis function (Gaussians). The output gives a linear combination of the hidden units. The network outputs are determined by how the hidden layers activation function and weights between the hidden and output layers react to the input vectors.(1)

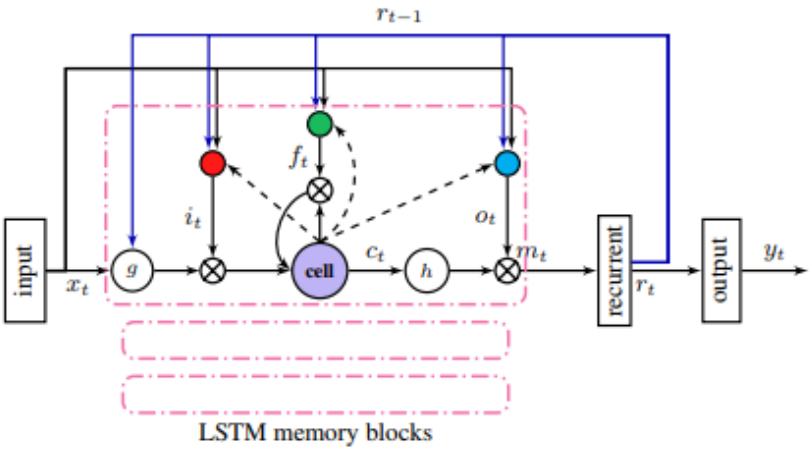


(3)

The RBF has the advantage over simpler methods (linear regression) by producing the linear combination of gaussians allowing an approximation of any function.

LSTM

The LSTM contains units called memory blocks. Each memory block contains memory cells that store the state of the network and additional units called gates. There are three different types of gates. The input gate controls the activations into the cell. The output controls the activation to the rest of the network. The forget gate can rest the memory of the cell.



(2)

LSTMs are most often fully connected to an output layer to make a prediction. The LSTM shows great promise in situations that involve time series data. This is one reason why they are used for making predictions in the stock market.

Experiment

The data for the experiment was taken from Yahoo Finance's historical data. The RBF does not require much historical data. Only a period of 30 days was used. The LSTM requires a longer length of data, so a 365-day period was used. The code for each model can be found at https://github.com/splgeo/Stock-Predictions. Two popular stocks were chosen to be used AAPL(Apple Inc.) and BTC-USD(Bitcoin price in USD). These two are popular for researchers to use as examples for models. The predictions were for the next day's closing stock price. The predictions of both models were compared to actual closing prices. Using Microsoft Excel, the error for each day was calculated, and an average was taken for both models. Time for running each model was logged using in program timers, started after the data is downloaded and stopped after the model outputs its prediction. The RBF will be running only using the CPU and the LSTM will use both the CPU and GPU. (4)

Results

Referencing *Table 1*, we can see the that the average error for AAPL for RBF is -1.06% and LSTM is 7.62%. The average error for BTC-USD for RBF is -1.57% and for LTSM is -9.05%. The average error between the two different stocks runs are -1.31% for RBF and -8.33% for LSTM. The maximum error for RBF is -13.14% (BTC-USD) and -24.27% (BTC-USD) for LSTM. The LSTM takes 107 seconds to run, 8 seconds for the first epoch and 1 second for the remaining 99. The RBF takes .08 seconds to run.

Table 1

AAPL	RBF Prediction	LSTM Prediction	Actual Close	RBF Error	LSTM Error
7/2/2018	187.28	182.15	187.18	0.05%	-2.69%
7/3/2018	190.10	162.61	183.92	3.36%	-11.59%
7/5/2018	185.95	169.01	185.40	0.30%	-8.84%
7/6/2018	190.43	176.58	187.97	1.31%	-6.06%
7/9/2018	188.75	171.48	190.58	-0.96%	-10.02%
7/10/2018	186.83	172.12	190.35	-1.85%	-9.58%
7/11/2018	185.94	177.04	187.88	-1.03%	-5.77%
7/12/2018	185.94	178.91	191.03	-2.66%	-6.34%
7/13/2018	192.23	176.10	191.33	0.47%	-7.96%
7/16/2018	184.03	183.59	190.91	-3.60%	-3.83%
7/17/2018	189.94	169.58	191.45	-0.79%	-11.42%
7/18/2018	190.46	168.94	190.40	0.03%	-11.27%
7/19/2018	188.12	174.53	191.88	-1.96%	-9.04%
7/20/2018	188.21	171.14	191.44	-1.69%	-10.60%
7/23/2018	189.10	174.14	191.61	-1.31%	-9.12%
7/24/2018	186.79	172.70	193.00	-3.22%	-10.52%
7/25/2018	187.27	179.67	194.82	-3.88%	-7.78%
7/26/2018	187.25	174.97	194.21	-3.58%	-9.91%
7/27/2018	189.51	178.20	190.98	-0.77%	-6.69%
7/30/2018	186.84	190.91	189.91	-1.62%	0.53%
7/31/2018	192.63	187.53	190.29	1.23%	-1.45%
			Avg Error	-1.06%	-7.62%

BTC-USD	RBF Prediction	LSTM Prediction	Actual Close	RBF Error	LSTM Error
7/2/2018	6339.14	6093.98	6509.58	-2.62%	-6.38%
7/3/2018	6615.55	5970.83	6590.06	0.39%	-9.40%
7/5/2018	6589.96	5659.62	6602.02	-0.18%	-14.27%
7/6/2018	6534.91	5829.88	6758.08	-3.30%	-13.73%
7/9/2018	6707.47	6511.46	6306.85	6.35%	3.24%
7/10/2018	6668.93	6340.56	6394.36	4.29%	-0.84%
7/11/2018	6387.52	6196.40	6253.60	2.14%	-0.91%
7/12/2018	6394.25	6371.53	6229.83	2.64%	2.27%
7/13/2018	6253.70	6208.94	6268.75	-0.24%	-0.95%
7/16/2018	6364.15	6340.08	7326.70	-13.14%	-13.47%
7/17/2018	6578.44	6116.89	7383.39	-10.90%	-17.15%
7/18/2018	7004.98	6394.07	7477.50	-6.32%	-14.49%
7/19/2018	7383.29	5819.32	7333.93	0.67%	-20.65%
7/20/2018	7477.39	5997.68	7405.40	0.97%	-19.01%
7/23/2018	7375.34	6357.92	8395.82	-12.15%	-24.27%
7/24/2018	7384.96	6823.07	8170.23	-9.61%	-16.49%
7/25/2018	7821.89	6912.45	7937.25	-1.45%	-12.91%
7/26/2018	8170.12	7420.08	8182.89	-0.16%	-9.32%
7/27/2018	7937.34	6861.31	8230.87	-3.57%	-16.64%
7/30/2018	8216.67	7969.68	7735.30	6.22%	3.03%
7/31/2018	8141.61	8548.70	7610.90	6.97%	12.32%
			Avg Error	-1.57%	-9.05%

Conclusion

It is clear to see that the simple RBF method results in an average accuracy that is 7.02% better than the LSTM method and is 1337.5% faster at the same time not requiring an expensive graphics processor. The RBF also completes its calculation in one hundredth the time of the LSTM's first step time. These results enforce the idea of only adding complexity when it is necessary. These results show that engineers deploying solutions should keep in mind that there may be a simpler solution to their problem.

Works Cited

1. https://www.csie.ntu.edu.tw/~yien/papers/tnn0485.pdf

2. https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/43905.pdf

3. https://www.researchgate.net/figure/Architecture-of-RBF-network_fig6_273610577

4. CPU Intel i7-4790k at 4.6GHz, GPU Nvidia GTX 1080ti, RAM 32GB DDR3 1600MHZ