Empirical Asset Pricing via Deep Learning Algorithms

Kaan Yolsever

Vancouver School of Economics

26-03-2019

Outline

Motivation

- Banks, hedge funds and investment banks have been using machine learning algorithms for a long time.
- Academic research in the area is nascent.
- Long Short Term Memory Network(LSTM) is of special importance to stock prediction because they handle time series data well.

LSTM

LSTM's are explicitly designed to avoid the long-term dependency problem. Hochreiter and Schmidhuber (1997)

Related Literature

- Shihao et al. (2018) show that deep learning algorithms can reach positive R^2 in predicting the stock prices, outperforming traditional linear models. Their paper suggests that complex methods outperform more naive methods.
- Fama and French (2015) shows that portfolios formed by 5 factors, namely operating profitability, size, investment factors (aggressive or conservative), book-to-market equity ratio, and market risk outperform their 3-factor portfolios.
- Carhart (1997) shows that momentum has significant predictive power when analyzing the returns from mutual funds.

- 122 different fundamental and technical variables and almost 2 million observations spanning 1978 to 2017 is fetched using Green et al. (2012)'s SAS code. The code combines COMPUSTAT, CRSP, and IBES. The data set includes 94 characteristics (61 of which are updated annually, 13 updated quarterly, and 20 updated monthly).
- The second data resource is the 8 macroeconomic variables in Goval and Welch (2004)
- I also incorporate the investor sentiment data set by Baker and Wurgler (2006). The data set has 14 variables.

  This adds up to $94+8+14 = 118$ variables per company per month.

Data ctd.

No.	Acronym	Firm characteristic	Paper's author(s)	Year, Journal	Data Source	Frequency
1	absacc	Absolute accruals	Bandyopadhyay, Huang & Wirjanto	2010, WP	Compustat	Annual
2	acc	Working capital accruals	Sloan	1996, TAR	Compustat	Annual
3	aeavol	Abnormal earnings announcement volume	Lerman, Livnat & Mendenhall	2007, WP	Compustat+CRSP	Quarterly
4	age	# years since first Compustat coverage	Jiang, Lee & Zhang	2005, RAS	Compustat	Annual
5	agr	Asset growth	Cooper, Gulen & Schill	2008, JF	Compustat	Annual
6	baspread	Bid-ask spread	Amihud & Mendelson	1989, JF	CRSP	Monthly
7	beta	Beta	Fama & MacBeth	1973, JPE	CRSP	Monthly
8	betasq	Beta squared	Fama & MacBeth	1973, JPE	CRSP	Monthly
9	bm	Book-to-market	Rosenberg, Reid & Lanstein	1985, JPM	Compustat+CRSP	Annual
10	bm_ia	Industry-adjusted book to market	Asness, Porter & Stevens	2000, WP	Compustat+CRSP	Annual
11	cash	Cash holdings	Palazzo	2012, JFE	Compustat	Quarterly
12	cashdebt	Cash flow to debt	Ou & Penman	1989, JAE	Compustat	Annual
13	cashpr	Cash productivity	Chandrashekar & Rao	2009, WP	Compustat	Annual
14	cfp	Cash flow to price ratio	Desai, Rajgopal & Venkatachalam	2004, TAR	Compustat	Annual
15	cfp_ia	Industry-adjusted cash flow to price ratio	Asness, Porter & Stevens	2000, WP	Compustat	Annual
16	chatoia	Industry-adjusted change in asset turnover	Soliman	2008, TAR	Compustat	Annual
17	chcsho	Change in shares outstanding	Pontiff & Woodgate	2008, JF	Compustat	Annual
18	chempia	Industry-adjusted change in employees	Asness, Porter & Stevens	1994, WP	Compustat	Annual
19	chinv	Change in inventory	Thomas & Zhang	2002, RAS	Compustat	Annual
20	chmom	Change in 6-month momentum	Gettleman & Marks	2006, WP	CRSP	Monthly
21	chpmia	Industry-adjusted change in profit margin	Soliman	2008, TAR	Compustat	Annual
22	chtx	Change in tax expense	Thomas & Zhang	2011, JAR	Compustat	Quarterly
23	cinvest	Corporate investment	Titman, Wei & Xie	2004, JFQA	Compustat	Quarterly
24	convind	Convertible debt indicator	Valta	2016, JFQA	Compustat	Annual
25	currat	Current ratio	Ou & Penman	1989, JAE	Compustat	Annual
26	depr	Depreciation / PP&E	Holthausen & Larcker	1992, JAE	Compustat	Annual
27	divi	Dividend initiation	Michaely, Thaler & Womack	1995, JF	Compustat	Annual
28	divo	Dividend omission	Michaely, Thaler & Womack	1995, JF	Compustat	Annual
29	dolvol	Dollar trading volume	Chordia, Subrahmanyam & Anshuman	2001, JFE	CRSP	Monthly
30	dy	Dividend to price	Litzenberger & Ramaswamy	1982, JF	Compustat	Annual
31	ear	Earnings announcement return	Kishore, Brandt, Santa-Clara & Venkatachalam	2008, WP	Compustat+CRSP	Quarterly

Figure 1: Sample Features, taken from Shihao et al. (2018)

Data

- The data points are scaled using a min-max scaler to be in the (-1,1) range.
- 50% of the data set is used for training, 35% for validation, and 15% for training. This correspond to: 25 years for training, 9 years for validation, and 6 years for testing.

Overarching Model

- Fully connected neural network with 3 hidden layers. In the first layer there are 32 neurons, second layer has 16, and the third layer has 8 neurons
- After each layer we apply a 'relu' function.

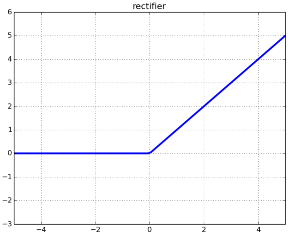

Figure 2: ReLU function

Overarching Model ctd.

As in Shihao et al. (2018), this paper describes an asset's excess returns as an additive prediction error model 1:

$$r_{i,t+1} = E_t(r_{i,t+1}) + \epsilon_{i,t+1} \tag{1}$$

where

$$E_t(r_{i,t+1}) = g(z_{i,t}) \tag{2}$$

and evaluates the prediction accuracy with the R^2 specified in that paper:

$$R^2_{oos} = 1 - \frac{\sum_{i,t\in\tau_3}(r_{i,t+1} - \hat{r}_{i,t+1})^2}{\sum_{i,t\in\tau_3} r^2_{i,t+1}} \tag{3}$$

Simple Neural Network

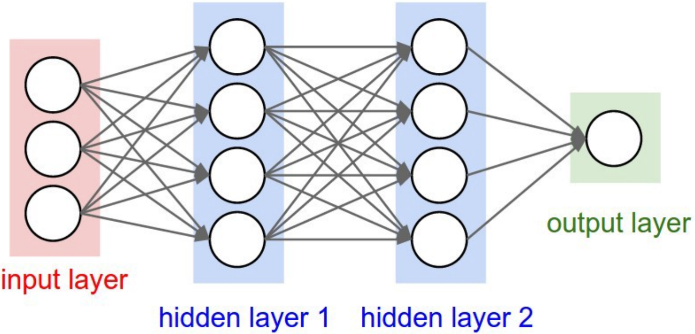

Figure 3: A neural network

Simple Neural Network ctd.

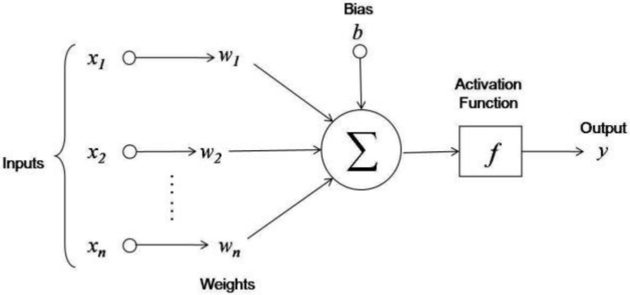

Figure 4: Basic NN Neuron Operation

Long Short Term Memory Network

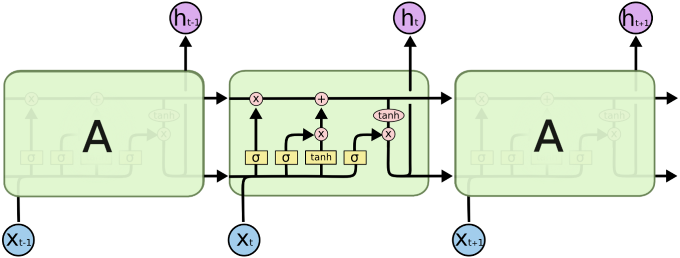

Figure 5: Basic LSTM Neuron Operation

Predictions

Figure 6: LSTM Predictions for 6 months

Predictions ctd.

Test RMSE: 0.084

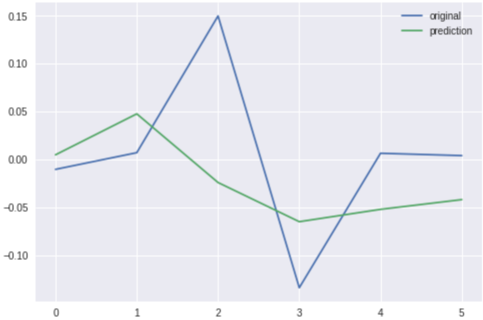

Figure 7: NN Predictions for 6 months

Variable Importance

Weight	Feature
0.0010 ± 0.0034	mom36m
0.0010 ± 0.0037	mom6m
0.0008 ± 0.0005	salerec
0.0005 ± 0.0007	pps
0.0004 ± 0.0009	mve_m
0.0004 ± 0.0007	mom12m
0.0002 ± 0.0004	quick
0.0002 ± 0.0003	roeq
0.0001 ± 0.0002	depr
0.0001 ± 0.0001	Returns
0.0001 ± 0.0001	prccq
0.0001 ± 0.0002	aeavol
0.0001 ± 0.0006	std_turn
0.0001 ± 0.0002	dy
0.0001 ± 0.0001	pchsale_pchxsga
0.0001 ± 0.0000	cinvest
0.0000 ± 0.0004	roavol
0.0000 ± 0.0001	pchcapx_ia
0.0000 ± 0.0001	invest
0.0000 ± 0.0001	cash
… 107 more …	

Figure 8: Variable Importance Ranking

First 5 are,respectively, 36 month-momentum, 6-month-momentum, sales-to-receivables, financial statement score, size. These are in parallel with financial literature: Carhart (1997), and Fama and French (2015)

Conclusion

- Stateless LSTM predictions align with those of simple NN. This means that the additional predictive power is not enough to justify the significantly higher computational cost.
- I observed that the interactions terms which add 800 features to the data set as in Shihao et al. (2018) do not contribute much to the results.
- The networks were able to learn which of the variables are most important for prediction.
- I had to introduce $L1$-regularization for a simple neural network which was not necessary for LSTM. This suggests that LSTM has a self-regularizing effect.

References

Carhart, Mark M. "On Persistence in Mutual Fund Performance". *The Journal of Finance* 52.1. ISSN: 00221082, 15406261 (1997): 57–82. Web. <http://www.jstor.org/stable/2329556>.

Hochreiter, Sepp and Jürgen Schmidhuber. "Long Short-term Memory". *Neural computation* 9. DOI: 10.1162/neco.1997.9.8.1735 (Dec. 1997): 1735–80. Print.

Goval, Amit and Ivo Welch. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction". Working Paper Series 10483. DOI: 10.3386/w10483 (May 2004). Web. <http://www.nber.org/papers/w10483>.

Baker, Malcolm and JEFFREY Wurgler. "Investor Sentiment and the Cross-Section of Stock Returns". *The Journal of Finance* 61.4. DOI: 10.1111/j.1540-6261.2006.00885.x (2006): 1645–1680. Print.

Green, Jeremiah, John R. M. Hand, and Frank Zhang. "The Supraview of Return Predictive Signals". *Review of Accounting Studies* 18. DOI: 10.2139/ssrn.2062464 (May 2012). Print.

Fama, Eugene F. and Kenneth R. French. "Dissecting Anomalies with a Five-Factor Model". *The Review of Financial Studies* 29.1. ISSN: 0893-9454. DOI: 10.1093/rfs/hhv043 (Aug. 2015): 69–103. Web. <https://dx.doi.org/10.1093/rfs/hhv043>.

Shihao, Gu, Kelly Bryan T., and Dacheng Xiu. "Empirical Asset Pricing via Machine Learning". *Chicago Booth Research Paper No. 18-04; 31st Australasian Finance and Banking Conference* (2018). Print.