



# Financial News in Predicting Investment Themes

Andrew Han

Department of Computer Science, Stanford University

Stanford | ENGINEERING  
Computer Science

## Motivation and Problem

### Questions

- Can neural models learn relationships between financial news and relative asset returns?
- What are more and less effective ways in encoding large corpuses of documents, such as news articles?
- Can neural models learn to ascertain the relevance of particular news articles?

### Datasets

Fama and French [1], as well as Applied Quantitative Research (AQR) [2], one of the world's largest quantitative hedge funds, have made data on daily factor returns public.

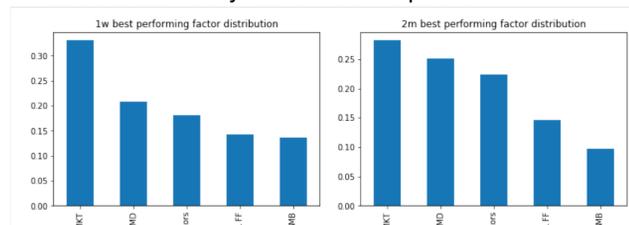


Figure 6: Best performing factors, sampled weekly and monthly

## Approach

### Architectures

- Baseline feed-forward neural network.
- Recurrent neural network with attention.
- Convolutional embedding with attention.

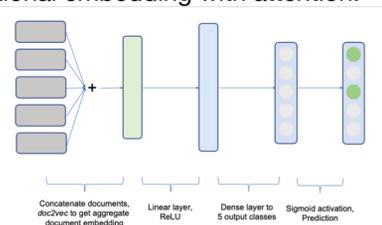
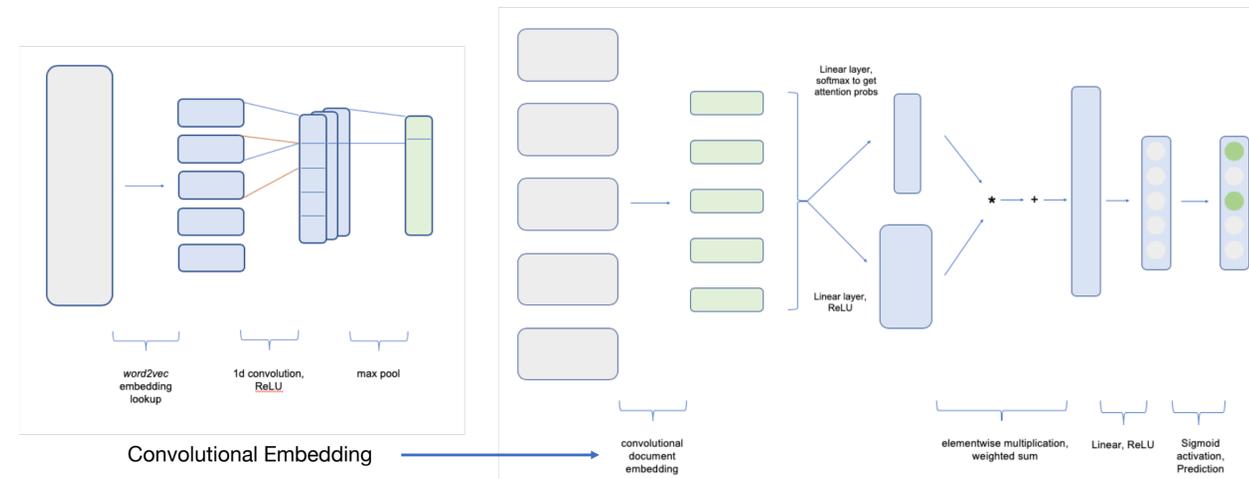


Figure 3: Baseline Feed-Forward Neural Network

### Evaluations

- Five factors: **market, size, value, momentum, volatility**
- Multilabel predictions
- Varied forward-looking prediction windows (1w, 6m)

## Process



## Results

### Summary

Model	Accuracy (6m)	Accuracy (1w)
Random guessing	69%	69%
Baseline NN, doc2vec embeddings	90%	<b>80%</b>
RNN with attention, doc2vec embeddings	80%	70%
Attention network, conv. embeddings	<b>95%</b>	73%

### Takeaways

The baseline NN and attention network w/convolutional embeddings seemed to outperform GRU with attention.

GRU, being a more complex model, was prone to overfitting.

Additionally, predictions were more accurate on a longer time scale, suggesting more noise in financial markets in short-term.

### Selected Charts

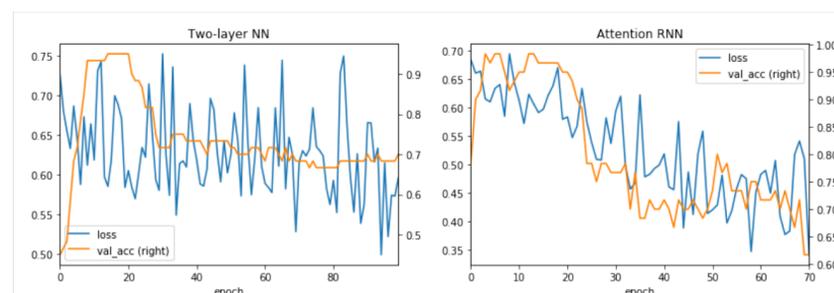


Figure 7: RNN and Baseline NN training loss and validation accuracy on monthly data

## Analysis

### Qualitative

#### Most and least attended documents, December 2007:

How Many Mutual Funds Routinely Rout the Market? Zero The bull market in stocks turned six last Monday, and despite some rocky stretches — like last week, when the market fell — it has generally been a very pleasant time for money managers, who have often posted good numbers.

Deep Freeze on Great Lakes Halts Cargo Shipments THUNDER BAY, Ontario —The trip to pick up a load of iron ore powder in Conneaut, Ohio, was supposed to take four days by way of the Great Lakes.

## Conclusions

- Using relatively little data, models using a convolutional embedder and doc2vec yielded document embeddings that, when fed to other neural architectures, were mildly predictive of factor returns.
- This suggests some relationship between financial news and short-term investment themes.
- Further, insight from the attention mechanism also suggests that, when supervised against empirical forward asset returns, neural network methods are able to differentiate between news article that are more or less relevant to financial markets.

## References

- <sup>1</sup> Fama, French. Current Research Returns. [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data\\_Library/ff\\_5\\_factors\\_2x3.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/ff_5_factors_2x3.html)
- <sup>2</sup> Applied Quantitative Research. Betting Against Beta: Equity Factors, Daily. <https://www.aqr.com/Insights/Datasets/Betting-Against-Beta-Equity-Factors-Daily>