Temporal Fusion Transformer Models for Predicting Stock Behavior

By Andrew Lys*

Your abstract here.

In portfolio design, a very common, and profitable strategy is called quantitative value investing. This strategy consists of designing some objective measure, ranking stocks based on this measure, and investing into the top ranked companies. The measure is derived from expert knowledge about how markets function, data, or some combination of the two. Portfolio design, in general, has been a field slow to adopt machine learning techniques, generally forgoing the data driven aspect of constructing these measures in exchange for more traditional financial measures. Examples of these measures include Trailing Twelve Month Returns (TTM Returns) and Earnings Before Interest and Taxes/Enterprise Value (EBIT/EV). These measures are often called value factors, as in the Fama-French value factor model. A clear drawback of this method is that there are countless different value factors an analyst could choose. For instance, what if it's the case that ranking stocks based on Trailing Eighteen Month Returns, instead of TTMs, is the secret to beating the market? Of course this example is silly, but in this manner, in their use as feature engineers, Neural Networks serve as an obvious next step to improve this strategy.

In traditional computer vision learning techniques, like boosted decision trees for facial recognition (Yoav Freund and Robert E. Schapire 1995), the main hurdle was always the development of better methods to extract relevant features. With the advent of convolutional neural networks, able to extract relevant features automatically, computer vision was effectively revolutionized (Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton 2017). Why can't Neural Networks do the same for factor engineering? Humans are bias-prone. When we pick factors, we are unintentionally choosing factors that are not statistically clean. Analysts have subconscious reasons for selecting factors, no matter how hard they can try to avoid these biases. For instance, if an analyst truly believes in a company, they might unintentionally choose factors that bias them towards that company, but will ultimately lead to lower returns overall. For reasons like this, Neural Networks have the capacity to revolutionize forecasting in the same way as they did image recognition.

Financial fundamental data is in the form of time series data, and thus admits a very natural sequence structure that can be leveraged by Recurrent Neural Networks (RNNs). (John Alberg and Zachary C. Lipton 2018) investigated the

* Lys: University of Chicago, andrewlys@uchicago.edu

use of Long Short Term Memory (LSTM) RNNS in forecasting fundamental data. These are RNN architectures where there are two streams of data passing through the recurrent neural network, the long term and short term streams. In this paper, we will investigate the use of Temporal Fusion Transformers (TFTs), as proposed by (Bryan Lim, Sercan O. Arik, Nicolas Loeff and Tomas Pfister 2020). We use the implementation provided by PyTorch Forecasting.

TFTs are a novel architecture developed by Google in collaboration with the University of Oxford. They lev

I. Data and Data Pre-Processing

A. Raw Data

B. Pre-Processing

C. Transformation and Deseasonalization

II. Model Framework

A. Temporal Fusion Transformer

B. Quantile to Ordinal Conversion

III. Experimental Results

IV. Conclusions

REFERENCES

Alberg, John, and Zachary C. Lipton. 2018. "Improving Factor-Based Quantitative Investing by Forecasting Company Fundamentals."

Freund, Yoav, and Robert E. Schapire. 1995. "A desicion-theoretic generalization of on-line learning and an application to boosting." 23–37. Berlin, Heidelberg:Springer Berlin Heidelberg.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2017. "ImageNet classification with deep convolutional neural networks." *Commun. ACM*, 60(6): 84–90.

Lim, Bryan, Sercan O. Arik, Nicolas Loeff, and Tomas Pfister. 2020. "Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting."

Mathematical Appendix