Discussion on Deep Learning in Finance: Deep Portfolios

Catherine S. Forbes* and Worapree Maneesoonthorn†

November 30, 2016

Abstract

This paper is a discussion on the article by Heaton, Polson and Witte. The authors have demonstrated that this new discipline of financial analytics has great potential to improve prediction and classification in a financial data context. A challenge remains, however, to coherently synthesize the new framework together with traditional finance theories and contemporary financial econometric analyses. Our discussion here aims to illuminate some particular areas where a synthesis of the ideas from traditional finance together with the proposed financial analytics tools would provide potential users with an effective and practical framework. In particular, we feel that it is vital to articulate appropriate notions of risk, to outline where they arise and how they impact on financial decisions. Also important is to be able to have confidence that the resulting models both characterize the essential features of the data, as well as provide meaningful solutions to well-specified problems.

KEYWORDS: Risk-return tradeoff; Volatility clustering; Jumps; Transaction costs.

1 Introduction

We commend the authors for their contribution towards the application of deep learning techniques to the finance context. While the machine learning literature has focused on the development of such algorithms over many years, recent interest appears to be fueled by the emergence of big data. The authors have demonstrated that this new discipline of

*Corresponding author. Department of Econometrics and Business Statistics, Monash University. Email: Catherine.Forbes@monash.edu. Support is acknowledged from the Australian Resaerch Council Discovery Grant DP150101728.

†Melbourne Business School, University of Melbourne. Email: O.Maneesoonthorn@mbs.edu.

financial analytics has great potential to improve prediction and classification in a financial data context.

A challenge remains, however, to coherently synthesize the new framework together with traditional finance theories and contemporary financial econometric analyses. Our discussion here aims to illuminate some particular areas where a synthesis of the ideas from traditional finance together with the proposed financial analytics tools would provide potential users with an effective and practical framework. In particular, we feel that it is vital to articulate appropriate notions of risk, to outline where they arise and how they impact on financial decisions. Also important is to be able to have confidence that the resulting models both characterize the essential features of the data, as well as provide meaningful solutions to well-specified problems.

2 Measuring and Controlling Risk

The paper focuses on an algorithm that minimizes a given distance between hypothetical "deep" portfolio prices, constructed via combinations of nonlinear functions of asset prices. The only constraint applied enters through the regularization penalty $\lambda\phi(W, b)$. Yet the relative importance of this function, and its connection to traditional notions of risk, remain unclear.

In a modern portfolio theory (MPT) approach initiated by Markowitz (1952), an optimally efficient portfolio is the one that maximizes the expected return while maintaining a controlled and specified level of "risk", given by the variance of the portfolio return. While initially developed under the convenient assumption that individual asset returns are serially uncorrelated multivariate normal random variables, this approach has since been expanded to accommodate serial correlation and conditional heteroskedasticity, also known as time-varying co-variances. See Bauwens, Laurent and Rombouts (2006) for background, and Engle and Colacito (2006) and Boubaker and Sghaier (2013) for examples of portfolio optimization applications. An important feature of the traditional theory is that it is aligned with the intuitive and long-standing principle of diversification. It also asserts that if a portfolio lies on the *efficient frontier*, then the expected return cannot further be increased without accepting a higher level of risk.

Does the so-called deep portfolio theory (DPT) alluded to by the authors have a corresponding formalization for the risk-return trade-off? The authors suggest that the *deep frontier* is given by the collection of optimized portfolios according to the amount of regularization implied by the penalty function $\lambda\phi(W, b)$. Presumably then, this implies that $\lambda\phi(W, b)$ might be interpreted as a measure of risk, but it is not clear whether this is appropriate. Indeed, what would be the most suitable penalty function(s) for the purpose of measuring *portfolio risk*, particularly in this non-linear setting? The ultimate choice of this penalty function should, ideally, align with the investor's appetite for risk.

3 Characteristic Features of Financial Time Series

Individual time series of financial returns are well known to display distinctive features, including those characterized by strong positive autocorrelation in the second moment, known as volatility clustering, with occasional discontinuity caused by extreme events, or jumps. These long memory inducing features tend to occur either simultaneously, or with some small delay, in the returns of different assets, providing notions of systemic risk, volatility spillovers and mutually exciting jumps. See, for example, Brownlees and Engle (2016), Diebold and Yilmaz (2009) and Maneesoonthorn, Forbes and Martin (2016). While modeling these features parametrically can be cumbersome, with inferential methods for the inevitable nonlinear models challenging, the resulting output is, at least, interpretable. Such models provide conceptualization of the nature and sources of uncertainties that produce the characteristic features of the observed data.

While deep learning algorithms may implicitly capture characteristic data features, for example via the market map (Step I Auto-encoding) and the portfolio map (Step II Calibration), how can we be sure the extent to which the time series properties are adequately captured by the deep learning algorithm? Are other evaluative measures available that would adequately determine whether or not such attributes of the data have in fact been accommodated?

4 Probabilistic Interpretation

Thinking more about the residual deviations that result from the model fit, Steps III and IV require a balance be struck between the permitted errors from the auto-encoding and calibration stages. Is there a guideline in how these two errors should be weighted? What is the differential impact (or loss) relevant to the portfolio from auto-encoding errors as compared with those made from calibration errors? For example, would it generally be better to tightly control auto-encoding errors and be less strict with the calibration errors? Presumably different loss functions could be used for each component, thereby providing the possibility to interpret the corresponding aggregate losses.

We note that, conditional upon the optimized predictive model, as $\mathcal{L}((Y_i), \hat{Y}(X_i))$ is simply taken to be the sum of squared errors metric it is consistent with the assumption that calibration errors $(Y_i - \hat{Y}(X_i))$ have independent, mean zero, constant variance Gaussian distributions. Is there evidence to support this in the empirical setting? If not, could this error metric be modified to account for heteroskedasticity in the errors, or dependence between the errors and certain combinations of predictor variables?

5 Implementation of "Deep Portfolio Theory"

Our fundamental question here is related to the benefits that a "deep" portfolio might deliver. In the detailed example setting, since an exchange traded fund (ETF) exists for the IBB index, what is the benefit of constructing the deep portfolio rather than to simply invest in the corresponding ETF directly? Note that there exists an index-tracking ETF for most, if not all, major financial market indices. We are interested to hear the authors' view on this point.

Perhaps the deep portfolios could be used to develop new index funds, exchange traded or not. The authors suggest that the resulting univariate activation functions "can frequently be interpreted as compositions of financial put and call options on linear combinations of the input assets." However, the actual investment portfolio would need to be constructed using the output of the model, and so this interpretation (as compositions of puts and calls) does not appear to be guaranteed to hold for any *generic* univariate activation function, although it may hold true for some. Further, given that the algorithm uses nonlinear activation functions, how would the fund manager actually implement the "deep" portfolio strategy? Nonlinear payoffs are potentially difficult to mimic, and we anticipate it would often involve multiple positions, possibly including those in derivative markets. It remains unclear whether the benefits of constructing this complex approximating portfolio would outweigh the inevitably substantial transaction costs.

Moreover, the nonlinear nature of the deep portfolio carries with it some implicit risk factors (e.g. momentum, tail risks over and above a mean-variance trade-off) which may arise either within the auto-encoding or calibration stages, or both. See Harvey, Liechty, Liechty, and Müller (2010) for an example of how to manage higher order moments in a non-Gaussian optimal portfolio selection setting. The automated nature of the proposed deep learning algorithm does not lend itself to interpretation of these risk factors, which are important to both investors and fund managers as they come to terms with managing investment risks.

6 Conclusions

Not withstanding the above, the machine learning methodology advocated is, essentially, a "model-free" and "data driven" approach. While this seems to imply a substantial loss of interpretability, as noted by the authors, gains are available in terms of finding and validating nonlinear structure present in large data sets. We do feel, however, that a complete *deep portfolio theory* would give consideration to the underlying principles of diversification, the measurement of risk, a risk-return trade-off, ensuring that the features of the data and model-fitted errors are adequately characterized, the construction of actual nonlinear portfolios, and so on, and these will no doubt follow in due course.

We give our sincere thanks again to the authors for providing a very thought provoking paper.

References

[1] Bauwens, L., Laurent, S. and Rombouts, J. V. K. (2006) Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics*, 21 pp. 79-109.

[2] Boubaker, H. and Sghaier, N. (2013) Portfolio Optimization in the Presence of Dependent Financial Returns with Long Memory: A Copula Based Approach. *Journal of Banking & Finance*, 37, pp. 361-377.

[3] Brownlees, C. and Engle, R. F. (2016) SRISK: A Conditional Capital Shortfall Measure of Systemic Risk. Forthcoming in *Review of Financial Studies.*

[4] Diebold, F. X. and Yilmaz, K. (2009) Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets. *The Economic Journal*, 119, pp. 158-171.

[5] Engle, R. and Colacito, R. (2006) Testing and Valuing Dynamic Correlations for Asset Allocation. *Journal of Business & Economic Statistics*, 24 (2), pp. 238-253.

[6] Harvey, C. R Liechty, J.C. Liechty, M. W., and Müller, P. (2010) Portfolio Selection with Higher Moments. *Quantitative Finance*, 10, pp. 465-485.

[7] Maneesoonthorn, W., Forbes, C. S. and Martin, G. M. (2016) Inference of Self-exciting Jumps using High Frequency Measures. Forthcoming in *Journal of Applied Econometrics.*

[8] Markowitz, H. (1952) Portfolio Selection. *The Journal of Finance*, 7 (1), pp. 77-91.