Journal Of Investment Management, Vol. 18, No. 2, (2020), pp. 5–8

www.joim.com

INTRODUCTION TO THE SPECIAL ISSUE ON MACHINE LEARNING

Charles Elkan[a]

Is machine learning useful in asset management? The answer is of course yes, but when and how? This special issue contains four papers that suggest answers to these questions. The papers show the way forward, and at the same time, they demonstrate some of the range of machine learning methods, so they will be of interest to practitioners and researchers in all areas of investment management.

Before discussing the papers individually, let us address, in the context of finance, the question of what machine learning actually is. By now there is a lot of consensus that there is continuity between the statistical and econometric methods that are conventional in finance and methods from outside finance that are commonly called machine learning. However, on the whole, ML methods are more complex than conventional methods. In particular, they tend to be nonlinear and to be designed for datasets that are large, heterogeneous, or both.

Methods that are more traditional are often based on the assumption that data has been generated by a particularly simple process, for example in a linear way, or by sampling from a standard probability distribution such as a Gaussian or a Poisson. In contrast, ML methods typically assume that the patterns to be found in data are nonlinear, and that a dataset is not necessarily a sample from any simple distribution.

Traditional methods tend to be appropriate for situations where there is not enough data to infer complex patterns reliably. In this case, valid inference is possible only if the phenomenon under analysis is relatively simple. However, in many application areas nowadays, we have very large datasets, and it is both useful and possible to fit a highly flexible model to the data. Often, these very large datasets contain heterogeneous types of data that previously have not been collected systematically, such as images and text. Traditional statistical and econometric methods do not apply to these types of data, while ML methods have been designed for them.

A related distinction between older quantitative methods and ML methods is that the latter focus more on making predictions and on achieving predictive accuracy, in situations where the true

[a]Goldman Sachs

data-generating process is unknown. In contrast, more traditional statistical methods focus on estimating the values of the parameters of a chosen model that is assumed to be correct.

Much of the difference between ML and other quantitative methods used in finance is in emphasis as much as it is in substance. For example, finance practitioners routinely refer to calibrating a model. This means setting the values of the free parameters of a model to make it consistent with observed prices or other information. In ML, a similar process is called training the model. The difference is that the number of free parameters in ML models is often orders of magnitude larger, while there are many more theoretical assumptions built into traditional models.

As a research area, ML has been an active subfield of computer science since the 1980s at least. If ML has been an active research area since the 1980s, why has it become so popular in the last five or ten years in business and in finance? One reason obviously is the greater availability of large datasets. Another reason is the success of what is called deep learning in many application domains, such as for speech recognition. Deep learning advances have led to new consumer product categories such as smart speakers, and the advances lend themselves to impressive demonstrations, such as winning a match against the human world champion in the game of Go. However, the problems where ML has been remarkably successful are very different from those encountered in finance, so it is still an open question to what extent deep learning, or ML in general, can be superior to previous methods in finance.

We can identify some stark reasons why ML in finance is challenging compared to ML for technological applications such as recognizing objects in images. Perhaps the most central reason is that financial domains tend to be non-stationary, meaning that the underlying patterns and relationships change over time. For example, the recent coronavirus pandemic has radically changed patterns of trading and volatility in the equity markets. One important source of non-stationarity is the competitive nature of finance. If a pattern exists that can be used to make a profit, trading to achieve that profit will tend to make the pattern disappear. Another important source of non-stationarity is the continual evolution of market structures. In equity trading for example, new regulations such as decimalization and MIFID, the growth and decline of different trading venues such as dark pools, changes in execution algorithms, and more, all make "the market" never the same.

A corollary of non-stationarity is often that relevant training examples are scarce. Suppose that we want to learn to recognize stop signs in images of streets. We can collect an almost unlimited number of independent examples, and we can trust that the characteristics that define stop signs are almost constant over time. However, suppose that we want to know which stock prices will increase when the price of oil increases. In different market regimes, the answer may be entirely different. It may be that only data from the recent past is useful, and if there has been a major event recently, even that data may not be useful for training a model, whether it is an ML model or a traditional model such as linear regression.

The paucity of relevant data affects humans and traditional methods as much as it does ML methods. However, simple methods with relatively few free parameters can be fitted to smaller datasets, and humans can bring to bear a wide range of general knowledge in addition to learning directly from a limited dataset. For example, humans can recognize the similarity between the quant quake of August 2007 and the unwinding of hedge funds in March 2020 at the peak of the coronavirus crisis, even though these were two rare episodes, and they involved mostly different funds given the time span between the episodes.

As mentioned above, the four papers in this special issue show the diversity of ways in which ML is relevant to investment management. The paper *Using Machine Learning to Predict Realized Variance* by Peter Carr, Liuren Wu, and Zhibai Zhang investigates ML methods to predict the future variance of the S&P 500 index. The current value of the VIX is essentially the market consensus view of this variance over the next 30 days, and no better predictor is known currently. The VIX is defined based on mathematical theory as a certain function of the observed prices of options on the index. The goal of this paper is to use ML to obtain an alternative empirical function of these prices that may be a better predictor than the theoretical function. The results are positive, but only a slight improvement is achieved over the VIX, despite a sophisticated application of modern ML methods. We can draw the conclusion that, consistent with much previous research, financial markets themselves are effective prediction mechanisms.

The paper *Dynamic Goals-Based Wealth Management Using Reinforcement Learning* by Sanjiv Das and Subir Varma considers the problem of investment allocation over a lifetime. The aim is to switch between asset classes in a way that is adaptive to market movements, with the goal of maximizing the probability that final wealth will be above a target at a certain horizon. This is a more realistic goal for many people than the mathematically simpler goal of maximizing expected wealth, or of finding the mean-variance efficient frontier. The paper first reviews dynamic programming for solving problems like this one in an exact, exhaustive way, then reviews ML methods called reinforcement learning that are applicable when market movements are known only via simulations. An implementation of a reinforcement learning method shows that it achieves the same solution as dynamic programming, while needing less in the way of explicit knowledge. The trade-off is that reinforcement learning requires a huge number of simulation runs, and the result is only as accurate as the simulation is.

The paper *On the Stability of Machine Learning Models: Measuring Model and Outcome Variance in Machine Learning* by Vasant Dhar and Haoyuan Yu looks for an empirical answer to the question of how much trust one should place in a model acquired from data. The answer, roughly, is that a model is not trustworthy if it is unstable, i.e., small changes in the training data can lead to large changes in the predictions made by the trained model. For binary classifiers, the authors then investigate how stability depends on the prevalence of the rare class and on the predictive power of the available features. The result of numerical experiments is that if the two classes are more unbalanced, then higher predictive power is needed for the trained model to be stable and hence reliable.

The aim of the paper *Can Machines "Learn" Finance?* by Ronen Israel, Bryan Kelly, and Tobias Moskowitz is to discuss ML in finance in general. The authors explain persuasively that many tasks in finance involve only a small number of historical observations; moreover, these observations are usually correlated, so they are not independently informative; and, any signal is weak compared to the noise. Together, these limitations make it impossible for any method to guarantee high accuracy. The authors then highlight the opportunity to use data more efficiently by building in assumptions that are given by economic or financial theory. In essence this is asking ML to model only the aspects of a domain that are not known a priori based on theoretical considerations. We see here a case of Fichtean dialectic, i.e., thesis, antithesis, and synthesis. Most traditional statistical methods are based on positing that a given simple model is true, but often in reality no such model is true. In contrast, many ML methods search over a very large space of complex models, but given the data limitations inherent in many financial applications, such large-scale search tends not to generalize well. The synthesis proposed in this paper is to use a model with parts that are posited to be true based on theoretical considerations, and parts that are learned from data using an ML method. The synthesis also shifts attention from earning alpha, which requires predicting the future better than other market participants do, to disentangling the factors that drive returns, which is useful even if other market participants do the same.