TEN APPLICATIONS OF FINANCIAL MACHINE LEARNING

Marcos López de Prado

This version: September 30, 2019

Marcos López de Prado is CIO of True Positive Technologies LP, in New York, NY, and Professor of Practice at Cornell University's School of Engineering, in Ithaca, NY. Website: www.truepositive.com. E-mail: mldp@truepositive.com

TEN APPLICATIONS OF FINANCIAL MACHINE LEARNING

ABSTRACT

This article reviews ten notable financial applications where Machine Learning (ML) has moved beyond hype and proven its usefulness. This success does not mean that the use of ML in finance does not face important challenges. The main conclusion from this article is that there is a strong case for applying ML to current financial problems, and that financial ML has a promising future ahead.

1. INTRODUCTION

Some of the most successful investment firms apply quantitative techniques, however those rankings rarely include firms that rely primarily on econometric methods (López de Prado [2016b]). The complexity of financial systems overly exceeds the modeling capability of these traditional quantitative methods. In addition, some of the most interesting datasets, such as satellite images, voice recordings, or news articles, are beyond the grasp of Econometrics (López de Prado [2019d]).

Around the turn of the 21st century, a few hedge funds began to experiment with machine learning (ML) approaches. While ML algorithms have the flexibility to identify complex patterns in financial datasets, they can be easily overfit, leading to false discoveries. Predicting financial outcomes is orders-of-magnitude more difficult than, for example, recognizing faces or driving cars. Sophisticated statistical techniques applied successfully to physical systems will fail when making predictions on financial time series.

The financial industry is presently being reshaped by technological advances founded on ML. Existing financial firms and disruptive new entrants are racing to explore, develop, and implement computer-assisted financial decision making. While there has been substantial hype around the application of ML to financial problems, there have also been remarkable successes. In this article, we review ten important investment problems for which ML players have delivered superior solutions.

2. FINANCIAL APPLICATIONS OF MACHINE LEARNING

ML has already demonstrated success and established a strong presence in many important financial applications. In this section, we review ten such critical use cases.

2.1. ASSET PRICING

Asset pricing is fundamentally a forecasting problem. Gu et al. [2018] explain that ML offers an improved description of asset price behavior relative to traditional methods. Those authors identify the best performing methods (trees and neural nets), and trace their predictive gains to their ability to model nonlinear predictor interactions that are missed by other methods. These authors conclude that the new risk premia extracted through ML can simplify the investigation into economic mechanisms of asset pricing, and justifies ML's growing role in factor investing. As one example, Exhibit 1 illustrates the use of classification trees to identify a taxonomy of stocks. Once each asset's peers are established, one can conduct relative value analysis in a more flexible way than afforded by traditional cross-sectional regressions. Carr et al. [2019] have found that ML algorithms predict realized volatility with significantly greater accuracy than standard econometric method.

[EXHIBIT 1 HERE]

A second application of ML in the context of asset pricing is the generation of synthetic datasets for backtesting investment strategies. For instance, once a variational autoencoder has learned the fundamental structure of the data, it can generate new observations that resemble the statistical properties of the original sample. This allows researchers to backtest investment strategies on historical series equivalent to thousands of years, and prevent overfitting to a particular sample realization. For a detailed analysis, see López de Prado [2019c].

2.2. RISK MANAGEMENT AND PORTFOLIO CONTRUCTION

Analytical hedging methods run into problems in presence of market frictions, such as transaction costs, market impact, liquidity constraints, risk limits, etc. Buehler et al. [2019] demonstrate the effectiveness of reinforcement learning approaches that are Greek-free and model free. These "deep hedging" methods consider many more variables and data points when making hedging decisions, and can generate more accurate hedges at greater speeds. They are purely empirical, and rely on very few theoretical assumptions.

In addition, financial series incorporate hierarchical relationships (like the tree structures embedded in region/sector/industry taxonomies) which are not recognized by traditional methods. To illustrate this point, Exhibit 2 displays a correlation matrix computed on 1,000 investment grade corporate bonds. An unsupervised learning algorithm recognizes highly correlated securities that should be clustered together. Risk and portfolio construction methods that fail to recognize the complex dependence structure in this data will often produce unstable and inefficient solutions.

[EXHIBIT 2 HERE]

López de Prado [2016a] introduced a new hierarchical clustering algorithm for the construction of investment portfolios. Monte Carlo experiments showed that the out-of-sample performance of this ML procedure beats the out-of-sample performance of Markowitz's critical-line algorithm and the inverse-variance allocation algorithm. An analysis of historical performance between these algorithms reached the same conclusion (Kolanovic et al. [2017]). López de Prado [2019b] reports Monte Carlo experiments in which the nested-clustering algorithm reduces the root-mean squared error of convex optimization solutions by more than 50%.

2.3. OUTLIER DETECTION

Outliers are observations produced by a stochastic process that differs from the process that generated the rest of observations within a sample. The appearance of outliers may or may not be associated with a structural break or regime switch. Regression methods are particularly sensitive to outliers. For example, Exhibit 3 shows how, in the context of cross-sectional studies, a small number of outliers can lead to the misclassification of a large number of securities. The few outliers tilt the regression line, leading to many securities being falsely labeled as cheap (false positives, in red), or falsely labeled as expensive (false negatives, in green). The random sample consensus (RANSAC) is one of several efficient algorithms that can isolate the outliers within a sample, hence preventing them from unduly biasing model estimates.[1] See Ruzgiene and Förstner [2005] for a general discussion of this method.

[EXHIBIT 3 HERE]

2.4. BET SIZING

Suppose that you have a primary model for making a buy-or-sell decision. That model does not need to be an ML algorithm —it can be fundamental, technical, or based on analyst expectations. Regardless of how the buy/sell decision is reached, one still needs to decide on the size of the bet (including the possibility of no bet at all). One approach is to derive some measure of confidence from the primary algorithm, however it is not always the case that side and size are optimally determined by the same model.

This is a situation that practitioners face regularly. A meta-labeling classification algorithm can learn bet sizing. First, we label the outcomes of the primary model as 1 (gain) or 0 (loss). Second, we train a classifier (such as a random forest, or a support vector machine) to learn how to predict those labels. The secondary (meta-labeling) model does not learn the side, it only learns the size. That is, it learns how to best use the primary model.

To understand how meta-labeling can turn a weak predictor into a strong predictor, recall that the Sharpe ratio associated with a binary outcome can be derived as

$$\theta[p, n, \pi_-, \pi_+] = \frac{(\pi_+ - \pi_-)p + \pi_-}{(\pi_+ - \pi_-)\sqrt{p(1-p)}}\sqrt{n}$$

where $\{\pi_-, \pi_+\}$ determine the payoff from negative and positive outcomes, p is the probability of a positive outcome, and n is the number of outcomes per year (see Section 15.3 of López de Prado [2018] for a derivation). When $\pi_+ \gg -\pi_-$, it may be possible to increase $\theta[.]$ by increasing p and the expense of n. Exhibit 4 shows how the secondary model can learn to separate positive outcomes from negative outcomes. As a result, the secondary algorithm reduces the incidence of negative events, even if that means accepting a few false negatives. The combined effect is a higher Sharpe ratio. The optimal threshold can be determined analytically, as a function of the meta-labeler's class skewness, and is not the result of a numerical search.

[EXHIBIT 4 HERE]

2.5. SENTIMENT ANALYSIS

Textual sentiment analysis encompasses the set of analytical tools used to identify whether a text implies a positive or a negative connotation. The analyzed texts typically fall within these categories: (a) Sentiment expressed by corporations, like public filings (10-Ks, 10-Qs), press releases or conference calls; (b) sentiment expressed by the media, like news articles or analyst

[1] For a video illustration, see https://youtu.be/vCsGzS9k0HI

reports; (c) sentiment expressed by the general public, like internet blogs or social media. Examples include Barker and Wurgler [2006], Tetlock [2007], and Mao et al. [2015].

Exhibit 5 shows the output from an NLP algorithm that has identified news articles containing information relevant to Tesla (TSLA US Equity). The blue bars tally the number of articles per day: the average is 458 articles/day, with a maximum of approx. 5000. The green bars count the daily number of articles expressing a positive sentiment, whereas the red bars report the daily number of articles expressing a negative sentiment. We can appreciate that strong sentiment imbalance is often followed by an immediate price reaction. Still, there appears to be some residual price momentum following the initial breaking news, which extends for weeks and months. A ML algorithm (e.g., a meta-labeling classifier) could be trained to identify when prices are most sensitive to sentiment imbalance.

[EXHIBIT 5 HERE]

2.6. FEATURE IMPORTANCE

One limitation of traditional regression methods is that they require researchers to solve two problems at once: (a) provide all variables involved in a particular phenomenon (variable selection); and (b) establish the correct structural equation that binds those variables, including all interaction effects (model specification). Under most circumstances, it is unreasonable to expect a researcher to have an answer to both questions prior to conducting an empirical analysis. A theory may hint an answer to the first question, but it will rarely nail the answer to the second. López de Prado [2019b] shows that a researcher will likely reject a true theory even if she knows the exact answer to (a), but her answer to (b) is only partially correct (e.g., she is right about the linear specification, but she misses one interaction effect). In other words, there is an econometric entanglement of (a) and (b): getting the variables right is conditional on getting the specification right and *vice versa*.

ML can be of great assistance to econometricians, by disentangling both questions. Feature importance methods, such as mean decrease impurity and mean decrease accuracy, provide and answer to (a) independent of (b). By providing a firm answer to (a), ML confirms the existence of a phenomenon, and lists the ingredients of the theory. It is then the job of the researcher to come up with a theory consistent with these empirical findings. Exhibit 6 provides an example of how mean decreased accuracy finds the variables that are important (prefix "I_") or redundant (prefix "R_"), while discarding noise variables (prefix "N_"). The algorithm returns the right answer in the presence of strong noise and multicollinearity, which are common attributes in financial applications.

[EXHIBIT 6 HERE]

2.7. CREDIT RATINGS AND ANALYST RECOMMENDATIONS

Human-generated credit ratings are the result of a complex logic that cannot be represented with a simple set of formulas. ML algorithms have been successful at replicating a large percentage of recommendations produced by credit rating agencies. As a result, years ago credit rating agencies started to use ML algorithms to support the work of their analysts. Exhibit 7 reproduces an example by Bacham and Zhao [2017]. The left figure shows a scatter plot of simulated bonds as a function of two features (X, Y), where defaults are colored in red. The middle plot shows that traditional methods fail at modelling this complex, non-linear relationship. The right plot shows that a random forest algorithm performs well. The same rationale can be applied to analyst recommendations in general.

[EXHIBIT 7 HERE]

2.8. EXECUTION

ML algorithms play an important role at modelling liquidity in over-the-counter markets, because they recognize the network topology inherent to trading systems (Easley et al. [2013]). For example, many investment grade bonds are not traded for days and even weeks. Kernel-based methods can identify "similar" trades based on their common features. The set of similar trades enables researchers to derive theoretical prices. Exhibit 8 shows the trade efficiency of buys (green) and sales (red) for similar bonds, according to kernel methods. A buy has efficiency 0 when it prints at the implied offer, and it has efficiency 100 when it prints at the implied bid. A sale has efficiency 0 when it prints at the implied bid, and it has efficiency 100 when it prints at the implied offer. Both have efficiency 50 at the mid. The distance between the red and the green line is the bid-ask spread that is being captured by liquidity providers.

[EXHIBIT 8 HERE]

2.9. BIG DATA ANALYSIS

The quantity and granularity of economic data has improved dramatically over the past few years. Administrative and private sector micro-level datasets offer an unparalleled insight into the inner workings of the economy. ML algorithms can help recognize the complex relationships involved in these datasets.

One successful example is the use of web-scraping algorithms by MIT's Billion Prices Project (BPP). This project collects daily price fluctuations associated with tens of millions of products sold by thousands of online retailers in almost 100 countries, which allows BPP to produce daily statistics of inflation and purchasing power parities. Specifically, BPP applies unsupervised ML to classify a large number of items into unique products (Berlotto [2016]). In countries where inflation statistics do not exist, are inaccurate, or are manipulated, BPP allows researchers to obtain real-time statistics with important implications for debt and foreign exchange markets. Exhibit 9 shows the official vs. BPP annual inflation estimates for the United States (left) and Argentina (right). While U.S. official statistics appear to be an accurate representation of online prices, Argentinian official statistics significantly and persistently underestimate prices paid online.[2]

[EXHIBIT 9 HERE]

2.10. CONTROLLING FOR EFFECTS AND INTERACTIONS

It is common in econometric studies to analyze the significance of a variable X in explaining a variable Y while controlling for the effect of a set of variables Z. Mullainathan and Spiess [2017] argue that, since researchers are not interested in "understanding" the effect of Z, they should not estimate their coefficients parametrically. Instead, they propose to replace Z with ML predictions of Y based on features Z. This semi-parametric approach will give researchers a better assessment of X's significance. The implication is that ML algorithms can complement econometric analyses by modelling complex interaction effects, involving hierarchical, non-linear and non-continuous relationships.

3. MACHINE LEARNING CHALLENGES

In the previous section, we have seen that ML is an established technology with important and matured applications in finance. In combination with Big Data and Supercomputing, ML has a growing role to play in economics and financial research. At the same time, we must be cognizant of the challenges that practitioners face when they attempt to expand the list of applications beyond the ten use cases listed earlier.

3.1. INTERPRETABILITY OF RESULTS

The use of ML encompasses a trade-off: the algorithm will reveal the features involved in a phenomenon, however the analytical expression that binds those features together may remain hidden from us. ML can provide the ingredients of a theory, but the researcher still needs to figure out the theory that utilizes those ingredients. The absence of an algebraic equation does not mean that ML solutions are black-boxes. On the contrary, Molnar [2019] and López de Prado [2019b] show how feature importance analyses such as MDI and MDA provide a greater level of transparency than the traditional method. For instance, *p*-values are easily gamed, leading to the profusion of false positives (and false negatives) in financial economics.

The good news is that researchers will keep their jobs for the foreseeable future. Even though ML is of invaluable assistance in building theories across all scientific fields, the task of turning discovered features into a testable equation is non-trivial. Part of the reason is that, in some instances, such equation does not exist. Take gene expression: There is little hope for finding a formula that will allow us to predict a person's physical attributes based on their individual genes. The same gene plays different roles in combination with other genes. ML algorithms are well equipped to model this complexity, even though the findings of the algorithms may not be

[2] For additional examples of Big Data and ML in economics, see Einav and Levin [2014] and Varian [2014]. Cavallo and Rigobon [2016] forcefully advocate for the use of alternative datasets by economic researchers and policy makers.

translatable into an equation. That does not make ML discoveries less truthful, however these discoveries may be harder to interpret in classical terms.

3.2. RISK OF OVERFITTING

A model is said to be overfit when it performs well in-sample but performs poorly out-of-sample. The discrepancy between in-sample and out-of-sample performance is known as generalization error. All statistical methods are susceptible to overfitting, traditional techniques and ML methods. There are two kinds of overfitting: train-set overfitting and test-set overfitting. Fortunately, ML researchers have developed powerful tools to control and reduce this risk.

3.2.1. TRAIN-SET OVERFITTING

Train-set overfitting results from choosing a specification that is so flexible that it explains not only the signal, but also the noise. The problem with confounding signal with noise is that noise is, by definition, unpredictable. An overfit model will produce biased predictions with an unwarranted confidence, which in turn will lead to poor performance out-of-sample (or even in a pseudo-out-of-sample, like in a backtest).

ML researchers are keenly aware of this problem, which they address in three complementary ways. The first approach to correct for train-set overfitting is evaluating the generalization error through resampling techniques (such as cross-validation) and Monte Carlo methods. The second approach to reduce train-set overfitting is regularization methods, which prevent model complexity unless it can be justified in terms of greater explanatory power. Model parsimony can be enforced by limiting the number of parameters (e.g., LASSO) or restricting the model's structure (e.g., early stopping). The third approach to address train-set overfitting is ensemble techniques, which reduce the variance of the error by combining the forecasts of a collection of estimators. For example, we can control the risk of overfitting a random forest on a training set in at least three ways: (a) cross-validating the forecasts; (b) limiting the depth of each tree; and (c) adding more trees.

3.2.2. TEST-SET OVERFITTING

Test-set overfitting occurs when a researcher conducts multiple tests and reports only the best outcome. This is a problem, because if the researcher conducts enough tests, eventually a false discovery will be made.

There are essentially two ways of addressing this concern. The first approach is to control for the number of experiments conducted, and derive from that the probability that the discovery is false (sometimes known as the family-wise error rate, or FWER), see López de Prado [2019a]. The second approach is to generate new, synthetic datasets that replicate the properties of the observed dataset, and test on them, see López de Prado [2018]. This is useful, because it is much easier to overfit one dataset than a multiplicity of them. These synthetic datasets can be generated via resampling or via Monte Carlo. Resampling means reusing the observed data to generate new datasets. Monte Carlo involves estimating the properties of the process that generated the observed dataset, and then produce an entirely new dataset that matches those properties.

3.3. REDUCED DATASETS

Some of the most powerful ML techniques require large amounts of data. Although the breadth and depth of economic datasets has improved dramatically over the past decade, economic datasets still suffer from a number of limitations. Some of these limitations result from the need of safety, privacy and confidentiality, or the lack of economic incentives to collect and process the information, etc. Other limitations arise from the very nature of certain economic problems, for example that macroeconomics inherently requires time series data and new observations are slow to bring new information (due to persistence in economic cycles). Even if those issues were resolved, the problem remains that economics is a dynamic field, where laws and institutions evolve and change with society.

Out of the ten financial applications we described earlier in this paper, the problem of reduced datasets impacts primarily asset pricing. For the other applications, we can distinguish between three alternative cases: (a) applications that enjoy an abundance of data. Examples include sentiment analysis, credit ratings, execution and Big data applications that often have large datasets comprising millions of examples; (b) applications that allow researchers to conduct randomized controlled experiments, where causality mechanisms can be established. Examples include execution and sentiment extraction. We may reword a news article and compare an ML's prediction with a human's conclusion, controlling for various changes. Likewise, we may experiment with the market's reaction to alternative implementations of an execution algorithm; (c) applications that require no data at all. For instance, risk analysis, portfolio construction, outlier detection, feature importance and bet sizing methods can be developed to satisfy appealing mathematical properties.

4. CONCLUSIONS

Financial firms have successfully applied ML techniques to a wide range of problems. Not only ML methods have proven their ability to perform better than traditional approaches, but also have been able to solve tasks that were beyond the reach of classical tools.

This does not mean that financial ML does not face challenges. In particular, when used improperly, ML tools may appear to produce opaque answers, with a high risk of overfitting to small datasets. Researchers have learned to overcome these difficulties in the wide range of applications listed in this article.

The availability of rapidly growing datasets, computing power and algorithms offer the opportunity to model the complexity of financial systems. These trends can only be expected to continue. To remain relevant, it is important for financial researchers to embrace this change of paradigm, and become familiar with this new set of tools.

REFERENCES

Bacham, D. and J. Zhao (2017): "Machine Learning: Challenges, Lessons, and Opportunities in Credit Risk Modeling." Moody's Analytics Risk Perspectives, Vol. 9, July. Available at https://goo.gl/Bmz1E9

Bailey, D., J. Borwein, M. López de Prado, and J. Zhu (2014): "Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-Of-Sample Performance." *Notices of the American Mathematical Society,* Vol. 61, No. 5, pp. 458-471.

Baker, M. and J. Wurgler (2006): "Investor sentiment and the cross-section of stock returns." *Journal of Finance*, Vol. 61, No. 4, pp. 1645-1680.

Bertolotto, M. (2016): "Real Exchange Rates using Online Data." Universidad de San Andres. Working Paper.

Buehler, H., L. Gonony, J. Teichmanny, B. Wood, B. Mohan, and J. Kochems (2019): "Deep Hedging: Hedging Derivatives Under Generic Market Frictions Using Reinforcement Learning." Working paper. Available at https://arxiv.org/pdf/1802.03042.pdf

Carr, P., L. Wu, and Z. Zhang (2019): "Using Machine Learning to Predict Realized Variance." Working paper. Available at https://arxiv.org/pdf/1909.10035.pdf

Cavallo, A. and R. Rigobon (2016): "The Billion Prices Project: Using Online Prices for Measurement and Research." NBER Working Paper No. 22111, March.

Easley, D., M. López de Prado, and M. O'Hara (2013): *High Frequency Trading: New Realities for Traders, Markets and Regulators*. Risk Books, First edition.

Einav, L. and J. Levin (2014): "Economics in the age of big data." *Science*, Vol. 346, No. 6210, November.

Gu, S., B. Kelly and D. Xiu (2018): "Empirical Asset Pricing via Machine Learning." *Chicago Booth Research Paper No. 18-04*. Available at https://ssrn.com/abstract=3159577

Harvey, C. and Y. Liu (2015): "Backtesting." *Journal of Portfolio Management*, Vol. 42, No. 1, pp. 13-28.

Kolanovic, M., A. Lau, T. Lee, and R. Krishnamachari (2017): "Cross asset portfolios of tradable risk premia indices. Hierarchical risk parity: Enhancing returns at target volatility." White paper, Global Quantitative & Derivatives Strategy. J.P. Morgan, April 26.

López de Prado, M. (2016a): "Building Diversified Portfolios that Outperform Out-Of-Sample." *Journal of Portfolio Management*, Vol. 42, No. 4, pp. 59-69.

López de Prado, M. (2016b): "Mathematics and Economics: A reality check." *Journal of Portfolio Management*, Vol. 43, No. 1, pp. 5-8.

López de Prado, M. (2018): *Advances in Financial Machine Learning*. Wiley, First Edition.

López de Prado, M. (2019a): "A Practical Solution to the Multiple-Testing Crisis in Financial Research." *Journal of Financial Data Science*, Vol. 1, No. 1, pp. 99-110.

López de Prado, M. (2019b): *Machine Learning for Financial Researchers*. Cambridge University Press, First edition.

López de Prado, M. (2019c): "Tactical Investment Algorithms." Working paper.

López de Prado, M. (2019d): "Beyond Econometrics: A Roadmap Towards Financial Machine Learning." Working paper. Available at https://ssrn.com/abstract=3365282

López de Prado, M. and M. Lewis (2018): "Detection of False Investment Strategies Using Unsupervised Learning Methods." Working paper. Available at https://ssrn.com/abstract=3167017

Mao, H., S. Counts and J. Bollen (2015): "Quantifying the effects of online bullishness on international financial markets." Statistics paper series, European Central Bank.

Molnar, C. (2019): Interpretable Machine Learning: A Guide for making Black-Box models explainable. Available at https://christophm.github.io/interpretable-ml-book/

Mullainathan, S. and J. Spiess (2017): "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives*, 31(2), pp. 87–106.

Ruzgiene, B. and W. Förstner (2005): "Ransac for outlier detection." *Geodezija ir Kartografija*, Vol. 31, No. 3, pp. 83-87.

Tetlock, P. (2007): "Giving content to investor sentiment: The role of media in the stock market." *Journal of Finance*, Vol. 62, No. 3, pp. 1139-1168.

Varian, H. R. (2014): "Big data: New tricks for econometrics." *The Journal of Economic Perspectives*, Vol. 28, p. 25.

EXHIBITS

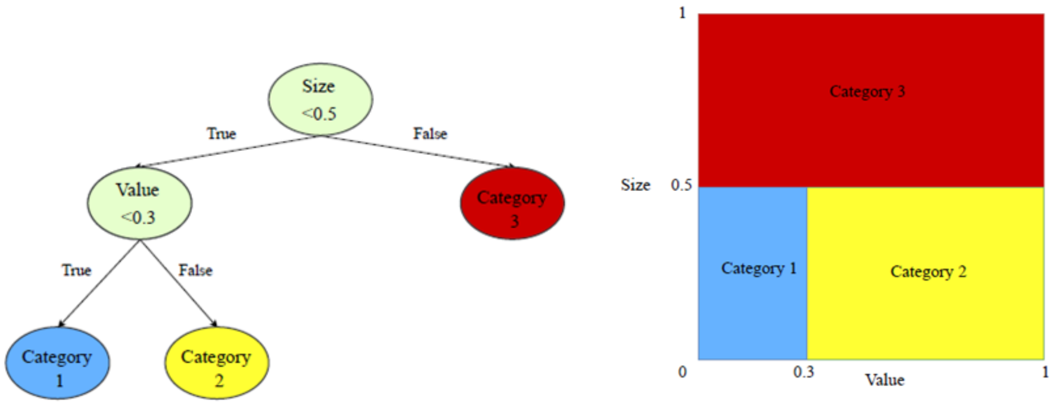

Exhibit 1 – Classification trees offer a flexible structure for the relative valuation of peer instruments. Source: Gu et al. [2018]

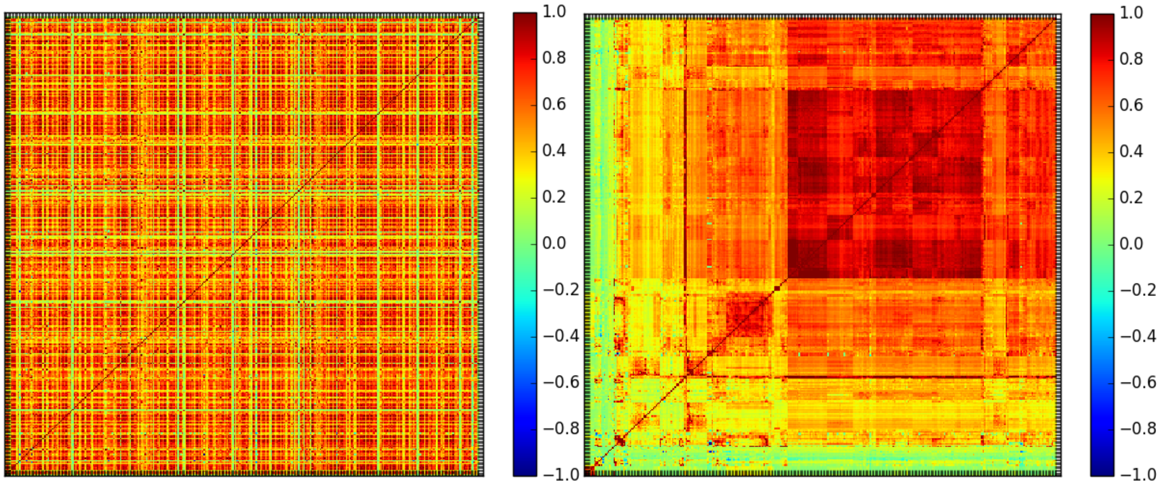

Exhibit 2 – A correlation matrix of investment grade corporate bonds, before (top plot) and after (bottom plot) clustering. Source: López de Prado [2018]

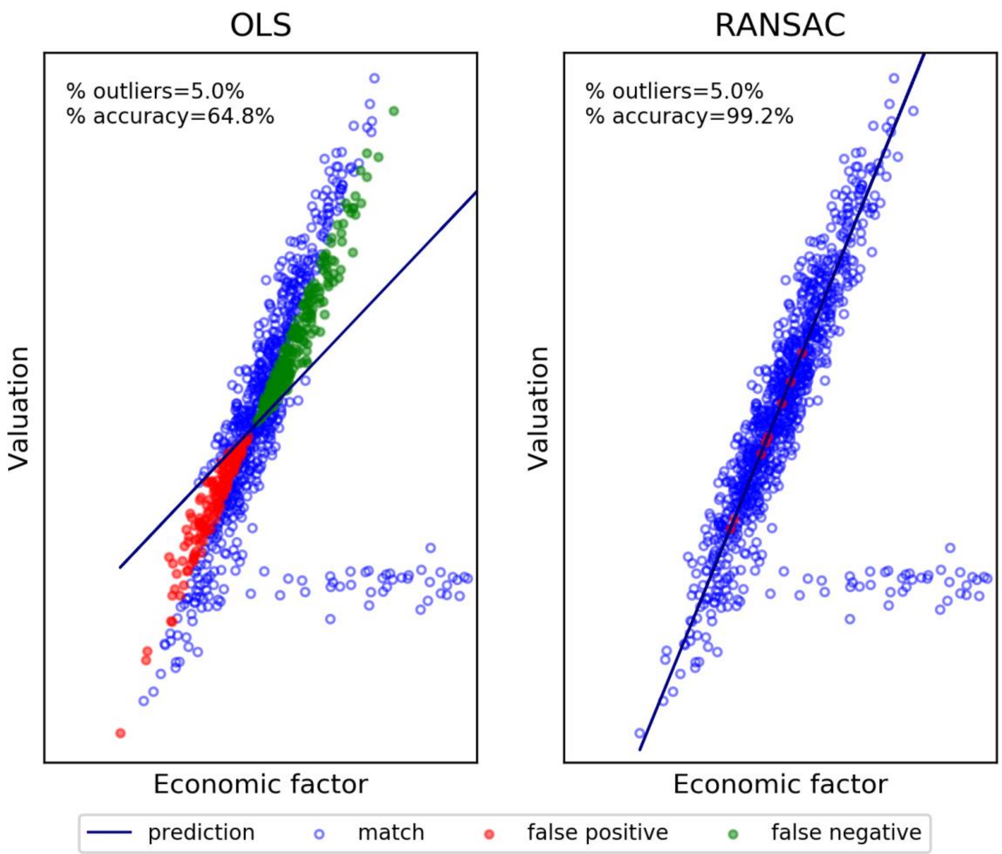

Exhibit 3 – In this cross-sectional regression, outliers accounting for 5% of the sample cause OLS to misclassify 35.2% of the observations. In contrast, if we apply OLS excluding the outliers detected by RANSAC, the number of misclassified observations is only 0.8% (mostly borderline cases)

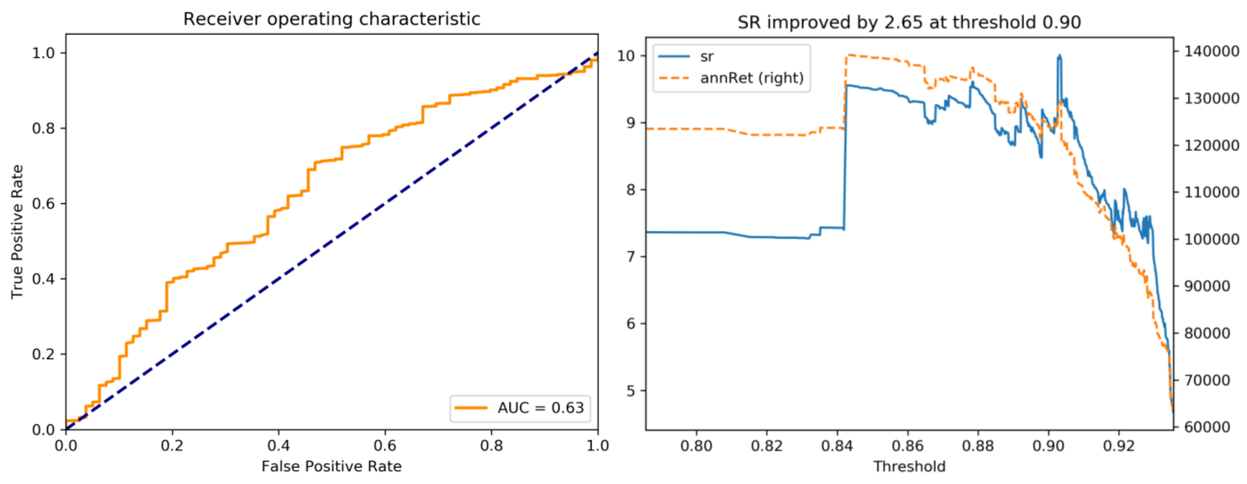

Exhibit 4 – In meta-labeling, the primary model determines $\{\pi_{-}, \pi_{+}\}$, and the secondary model regulates $\{p, n\}$. In the above example, a strategy's Sharpe ratio increased by 2.65 thanks to meta-labeling's ability to avoid the largest losses

Exhibit 5 – Number of news, and extracted sentiment, associated with TSLA US Equity (Tesla). Source: Bloomberg Terminal, AQR.

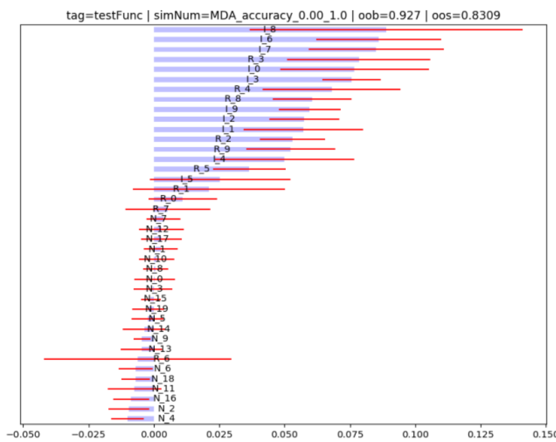

Exhibit 6 – A mean decrease accuracy algorithm applied to a simulated dataset with strong noise and multicollinearity. Source: López de Prado [2018].

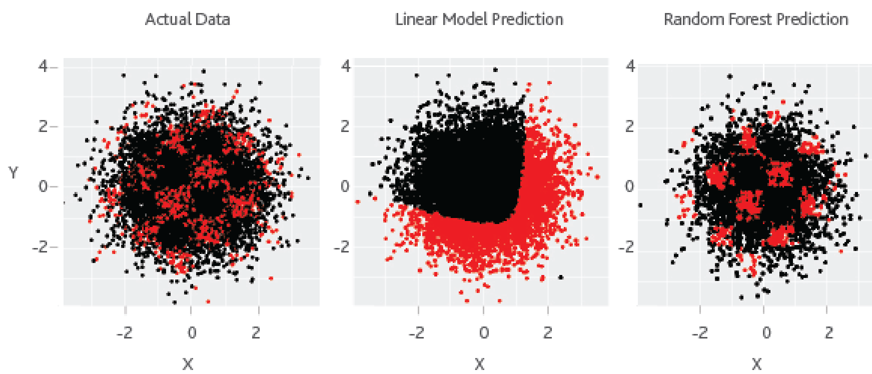

Exhibit 7 – A linear binary model (center) fails to reproduce the complex structure of the simulated data (left). A random forest (right) accurately recognizes that structure. Source: Bacham and Zhao [2017]

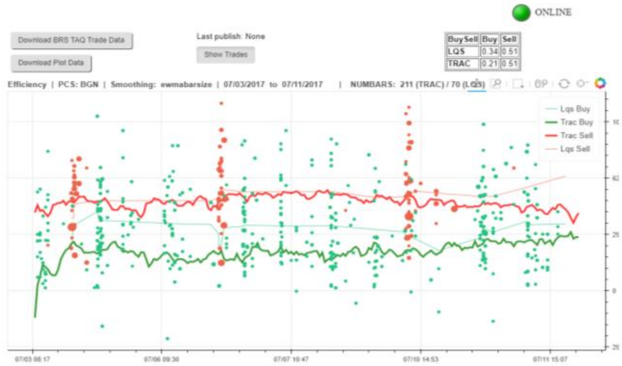

Exhibit 8 – ML-based analysis of the liquidity provision in the investment grade corporate bond market

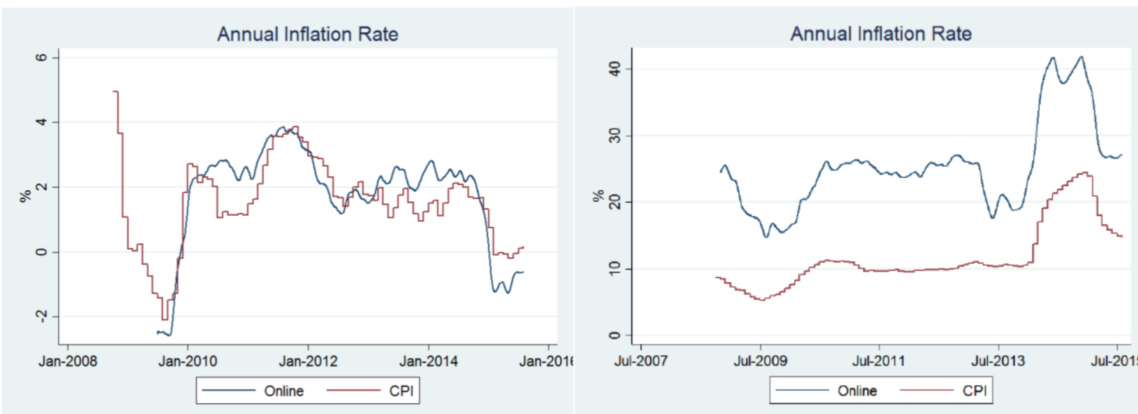

Exhibit 9 – Official vs. BPP annual inflation statistics for the U.S. (left) and Argentina (right). Source: Cavallo and Rigobon [2016].