LOMBARD ODIER
INVESTMENT MANAGERS

Machine learning for the prediction of stock returns

– For professional investor use only –

05/20

May 2020

Machine learning is able to discover more complex predictive relationships beyond those used in simplistic factor-based models.

p.05

At a glance

- Once associated with science fiction, machine learning has come a long way over the past decade and is now integrated in everyday life. At LOIM, we jumped on the machine learning train early on. It was not an easy journey and we share our experience and views in this note.
- Implementing quantitative models in asset portfolios is the bread and butter of systematic investments, and machine learning can help with detecting more complex predictive relationships beyond traditional factors.
- However, the application of machine learning is far from straightforward, and especially in finance. There are several challenges to overcome, from the forward-looking bias to discovering spurious relationships. We believe that only a combination of strong financial expertise with a good knowledge of the data science will produce a strategy that performs not only on the paper.
- Our research efforts did bear fruit. We are pleased to announce a new investible long-short equity strategy as a building block of ARP offering. The strategy seeks to achieve the dual objective of generating performance and promoting responsible companies through application of ML techniques to traditional fundamental and technical stock characteristics, as well as ESG data.

For professional investor use only. Please see important information at the end of this document.

This document has been prepared by:

Alexey Medvedev
Portfolio Manager, Systematic Equities

Laurent Joué
Senior Portfolio Manager, Commodities and Alternative Risk Premia Strategies

Marc Pellaud
Senior Portfolio Manager, Commodities and Alternative Risk Premia Strategies

Cesar Moura
Senior Analyst, Alternative Risk Premia

For more information please contact:

Clément Leturgie
Product Specialist, 1798 Alternatives
c.leturgie@lombardodier.com

Combining strong financial expertise with a good knowledge of data science.

Introduction

A mere decade ago, for most of us Machine learning ('ML') would be associated with pure science fiction. While ML methods have been known for quite a while, it was only recently that the explosive growth of computational power accessible at low costs facilitated the integration of ML in all aspects of our life. Nowadays, numerous algorithms are running behind the scenes filtering spam emails and enhancing our selfies. Armies of developers have coded and optimised ML algorithms while academics provide plenty of educational material, making ML accessible to a wide audience.

Media has further popularised ML and AI promising bright future where machines perform all routine work, and even challenge humans in some creative activities. It is not surprising that ML has also become a very hot topic in finance. Over recent years we witnessed an avalanche of research papers discussing the applications of ML to finance, including risk modelling, portfolio construction, contextual analysis, and, of course the most intriguing one, the prediction of asset returns. At the same time, we are aware of only a handful of successful implementations of advanced ML methods in actual portfolios.

At LOIM, we jumped on this train early on with the development of a proprietary ML model (which we will refer to as 'LOIM-ML') for stock scoring that we currently incorporate in our systematic strategies. Truth be told, this was not an easy journey. During this time, we experienced multiple frustrations and setbacks while discovering all the various pitfalls of implementation of ML on financial data. In this note we would like to share our experience and views.

Systematic investment without human biases

Implementing quantitative models in asset portfolios is the bread and butter of systematic investment. An example of such an approach in equities is factor-based investing, which consists of building diversified portfolios with intentional exposures to certain systematic factors deemed to be rewarding.

In general, the systematic approach is a disciplined application of a certain quantitative process. On paper, such an approach avoids as much as possible human interventions thus ensuring the consistency of live performance with back-tested results. That said, even if the investment process is truly systematic, it has been designed by humans. Efforts spent on the research model are often not very transparent and stay behind the scene. It is not an uncommon scenario when excessive data mining leads to models that work well on historical data, yet disappoint in production.

Machine learning can take the systematic approach to a whole new level by minimising human participation at the stage of strategy development. Instead of manually searching for the best model, the researcher implements an ML algorithm that does this job in a more productive way. Of course, machines will never succeed on their own. Even if all the tedious work is left to machines, the role of humans becomes critical in the design of ML algorithms and data management. We will come back to this later in the paper.

Traditional fundamental analysis relies on simple manipulations, with financial ratios to judge whether stocks are fairly valued or not. In the systematic space, factor-based investing follows a similar route by grouping financial ratios into buckets called factors. Intuitively, it is difficult to believe that any simple predictive relationships persist in the market and not arbitraged away. While we see no reason to question the long-term premia of factors, systematic investors should be prepared to persevere through long periods of factors' ups and downs before their patience is rewarded. The recent disappointing performance of factor-based strategies is a timely reminder of this.

Machine learning is here to help again. ML methods are designed to detect more complex predictive relationships, which are less likely to be easily arbitraged away, and therefore, are able to produce more stable returns. We expect that these new sources of alpha will provide a long-awaited diversification for systematic investment, which is almost fully dominated by factor-based strategies. The flip side of the increased model sophistication is the loss of its transparency. This is typically the main criticism of ML as, traditionally, model interpretability is an integral part of model validation. In our view, transparency is not critical for models "learnt" by machines since the automated process of model development minimizes human biases and judgement errors.[1] That being said, it is possible to gain an insight into how an ML model works through the analysis of its behavior.

Machines will not work without human

Advances in computer technology and the availability of open source tools makes the application of Machine Learning seemingly straightforward. This is far from being true in general, and especially in financial applications.

[1] Interpretable Machine Learning is a new field that tries to address the problem of transparency where the model error can have severe consequences like, for example, self-driving cars.

Financial theory teaches us that markets are efficient, therefore, asset returns are fundamentally difficult to predict. We should not be overoptimistic to expect that the application of ML techniques is going to change this dramatically. Instead, what we can hope to achieve is a prediction accuracy that is good enough to produce a decent performance when implemented in a diversified portfolio of stocks. Good news is that only a tiny forecasting power is sufficient. For example, a model predicting monthly stock returns with a hit ratio as low as 51% is enough to generate a Sharpe ratio of almost three[2] ! The low forecasting power of financial models is what makes the application of ML in finance especially challenging where every decision counts.

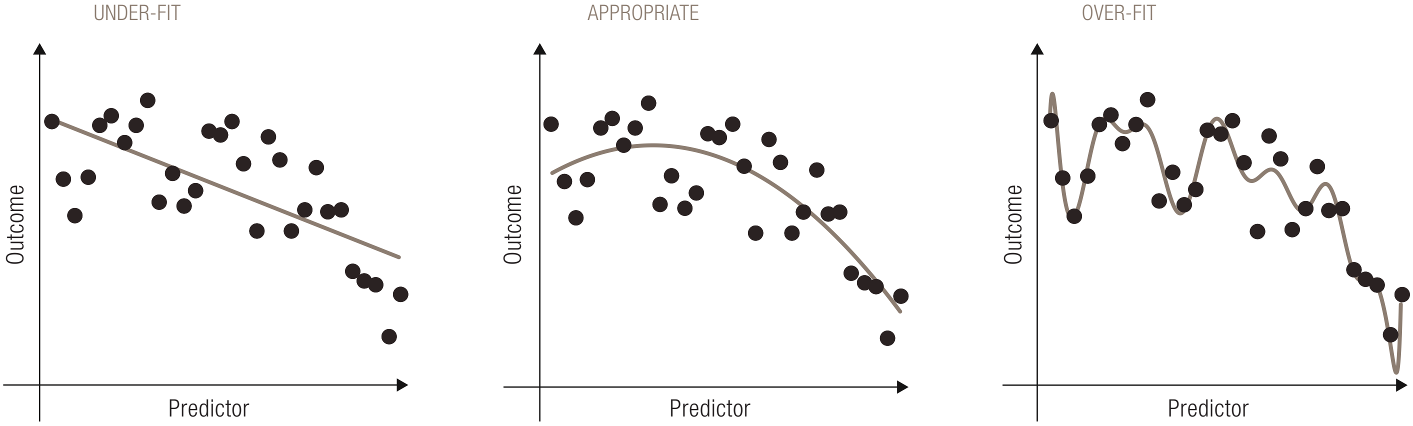
Probably, the most important decision is the choice of ingredients to be used in the model. A naïve approach to modelling stock returns would be to throw all the possible stocks' attributes into the machine and let it do its magic. Unfortunately, this is not going to work. First and foremost, we do not possess a sufficient number of observations to deal with a large set of "raw" predictors. In image recognition, for example, we are able to estimate or train a model on an arbitrary large set of observations since experiments can be repeated as many times as needed. In finance, we have to rely on historical observations only, which are naturally very limited. The dynamic nature of financial markets makes the reality even more complicated as we have to focus on recent observations that are most relevant for the current state of the markets.

In our view, the selection of model predictors should start by clearly formulating the hypothesis about the underlying model. In LOIM, we believe that stock returns are driven by fundamental and technical forces, which is consistent with a long-standing culture of stock picking. Inspired by a vast academic and practitioners' literature as well as our internal expertise, we selected a total of 70 predictors of both technical and fundamental nature. All data went through a well-thought standardisation procedure making predictors comparable across stocks and over time. Due to the dynamic nature of financial markets, magnitudes that seemed to be excessive in the past might be new normal today.

One of the biggest challenge of ML in finance is avoiding the forward-looking bias, which is a situation when the model learns to predict the past from the future and not the other way round. The forward looking bias occurs when the data used for model training was actually not known before the stock returns were realized. A typical example is when a company quarterly earnings are used to predict stock returns from the end of the same quarter onward. In real-life, quarterly figures are never available immediately. In fact, this data may be even revised several times far in the future, which amplifies the forward looking bias when one uses the last revision. Machine learning techniques are powerful enough to detect and take advantage of even slightest forward looking biases resulting in overly optimistic back-tested results that cannot be replicated in live. In LOIM, we avoid the forward-looking bias by taking into consideration the exact time when each quantity became publicly known or apply a conservative time lag when the exact date is unknown.

Machine learning algorithms are able to find arbitrary complex relationships in data, which is both a benefit and a curse. If the learning process is not properly controlled, the algorithm may easily "overfit" data, which means discovering spurious relationships that are unlikely to hold in the future. This is even more important in finance where predictive relationships may disappear due to ever changing markets. Exhibit 1 illustrates the concept of overfitting using an example where an ML attempts to find a relationship between a single predictor and an outcome. The "appropriate" model captures well the non-linear relationship in data without becoming over sophisticated.

EXHIBIT 1 FINDING THE RIGHT BALANCE



Source: LOIM calculations, May 2020. For illustrative purposes only.

[2] Source: LOIM calculations, based on the universe of stocks in the MSCI World index. For illustrative purposes only.

The right level of model complexity is achieved by fine-tuning the learning process, which is called *cross-validation*. The idea behind this process is to split observations into two parts: training and testing sets. The training set is used for model training, while testing set is left out for evaluating the model performance. This data separation mimics what is happening in reality where the model is estimated on the past data while being implemented in the future. Therefore, the training set should strictly precede the testing set in time. The optimal configuration of the learning process is such that it yields the best performing model on the testing set and not the training one. The cross-validation is a very computer intensive process as it requires training multiples models with different configurations of the learning process. At LOIM, we use cloud computing to perform cross-validation tasks.

How do they learn?

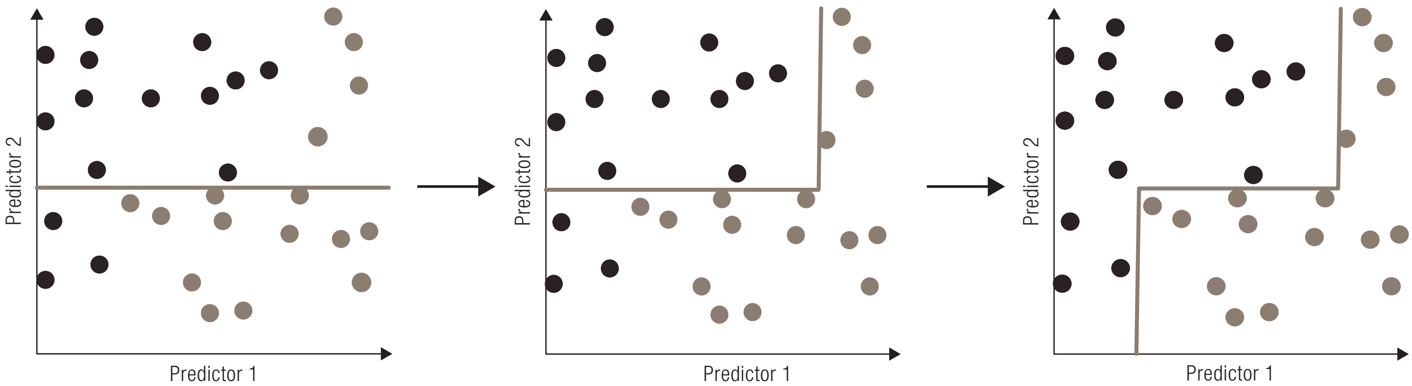
A good practice to introduce a complex subject such as Machine learning is to start with a high level overview of different approaches before going into details of particular algorithms. In this paper, we decided not to do this as it is likely to be confusing for someone who is not familiar with the subject. Instead, we directly introduce a particular methodology called *Gradient Boosting Machines* (GBM).

Despite its somewhat confusing name, Gradient Boosting Machines is based on an intuitively straightforward idea of building a complex model by aggregating multiple simple ones. In most applications, these simple models, also called *weak learners*, take the form of a decision tree – a model whose prediction can be visually represented by a tree diagram. In its simplest form, the decision tree has only two possible outcomes depending on whether a certain condition is satisfied or not. While such a tree is obviously too simple to make an accurate prediction, an aggregation of multiple trees like that can fit an arbitrary complex relationships in data.

Exhibit 2 illustrates how GBM works for a classification task with two predictors. The process starts by training the first weak learner, which classifies observations based on a single criteria involving values of Predictor 2. Next, the second weak learner is trained to explain the failure of the first one. The new classification model incorporates two criteria, one per each predictor. During each subsequent learning cycle, a new weak learner is trained to improve or “boost” the accuracy of the model built so far. The process continues until the desired maximum number of learning cycles is reached. In our example, after the third cycle, the model with three criteria perfectly classifies the data.[3]

What we have described above is, of course, an oversimplified version of GBM. In real life applications, GBM receives many enhancement that help fine-tuning the learning process, which can be a game changer especially in financial applications. Ultimately, the design of the ML algorithm and the cross-validation procedure are those human decisions that determine the success of machines. On our side, we did a comprehensive analysis of a number of different Machine Learning models and selected GBM for the LOIM-ML predictor model. This is also in line with GBM being one of the top choices among practitioners.[4]

EXHIBIT 2 GBM: FROM WEAK LEARNERS TO A STRONG MODEL



Source: LOIM calculations, May 2020. For illustrative purposes only.

3 Any further learning will improve the classification in terms of the likelihood of the prediction as the outcome of the model is the probability of each class.

4 A forthcoming technical paper will provide a more comprehensive analysis of different alternatives.

LOIM-ML model: new alpha or old factors?

Our LOIM-ML model makes predictions based fundamental and technical stock characteristics, which overlap with definitions of traditional equity factors. One may wonder whether the model discovers new sources of alpha or it simply rediscovers the "old" factors. To answer this question Exhibit 3 shows the full performance of a long-short portfolio built using the LOIM-ML model score,[5] and the part explained by time-varying portfolio exposures to five major factors: Value, Quality, Momentum, Low beta and Small size.[6]

Knowing that factors have performed well in the past, it would be surprising if machines did not learn their predictive abilities. Indeed, we observe that the exposure to factors explains about one third of the total return of the strategy since 2004. However, it is also clear that from 2016 onward, factors do not add any value while the "new alpha" continues to deliver a robust performance. Overall, LOIM-ML model meets our expectations that Machine learning is able to discover more complex predictive relationships beyond those used in simplistic factor-based models.

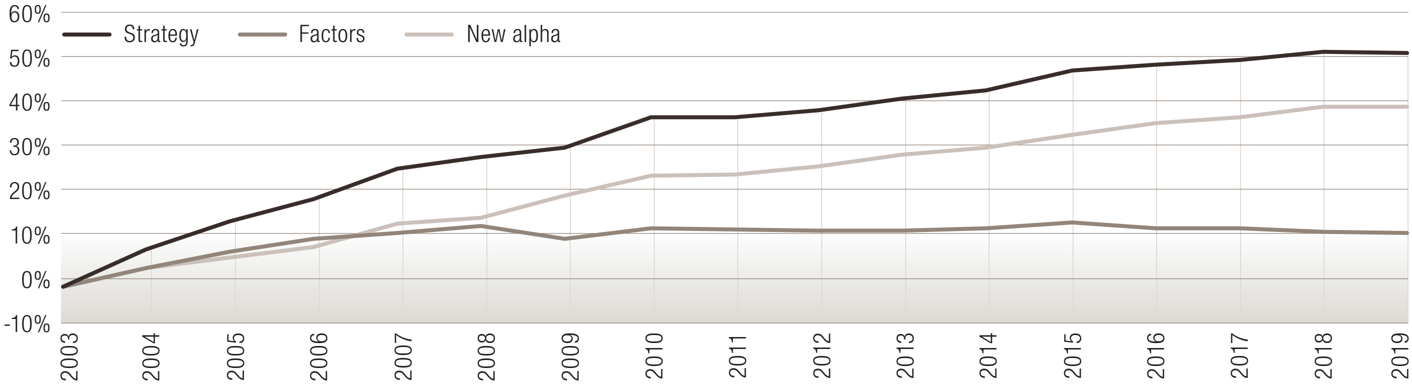
Machine learning and ESG

Recently, we have witnessed the emergence of multiple alternative data sources that offer new insights on the business of companies. Economic, Social and Governance (ESG) is probably the most distinguished data providing an assessment of various ESG issues in the form of quantitative scores.

The key distinction of ESG data from traditional fundamental and technical characteristics is that it is a qualitative in nature, and therefore tends to be sensitive to the scoring methodology. For example, it is not uncommon that different ESG vendors disagree substantially in their assessments of ESG issues. In our view, ESG data provides a rich alternative set of stock characteristics that can be very useful for return prediction. However, a short history of the data and its qualitative nature prohibits implementing complex predictive models.

At LOIM, we built a proprietary ESG-QUANT score using elements of Machine learning. Our ESG-QUANT score represents an index of original ESG scores with weights reflecting their capacity to predict stock returns. We require that each ESG score has a positive or no contribution to the aggregate score, which is a reflection our conviction that companies with fewer ESG-related issues are likely to outperform. The methodology is described in more details in our recent white paper "**ESG Alpha: Doing Well While Doing Good**".

EXHIBIT 3 LOIM-ML MODEL LOOKS BEYOND FACTORS



Source: LOIM-ML. For illustrative purposes only.

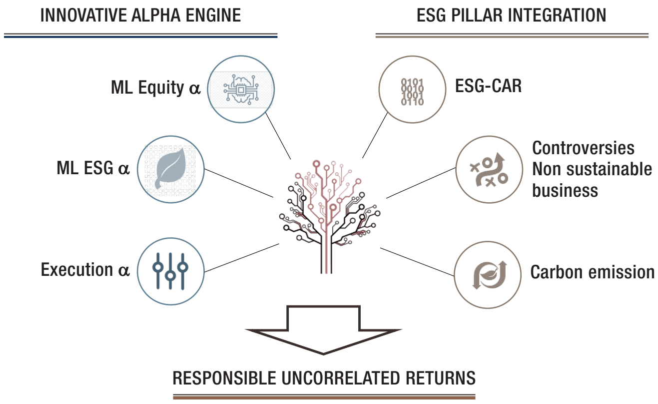
[5] This is an unconstrained long-short portfolio where stock weights are proportionate to their scores from LOIM-ML model.

[6] We used proprietary definitions.

What we offer

In our everyday life, we are witnessing an increasing adoption of data-intensive technologies and a rising interest in data science. Finance is not an exception. Machine learning opens new horizons for systematic investment, however its successful application requires much more effort than just hiring a bunch of data science graduates. In this paper, we highlighted the main promises and challenges of using ML techniques for prediction of asset returns. The main message we wanted to deliver here is that only a combination of strong financial expertise with a good knowledge of the data science will produce a strategy that performs not only on the paper.

EXHIBIT 4 LOIM-ML EQUITY ESG ALPHA STRATEGY



Source: LOIM. For illustrative purposes only.

Our research efforts in Machine learning did bear fruit. We are pleased to announce a recent launch of a new investable long-short equity strategy as a building block of the ARP offering. This strategy seeks to achieve a dual objective of “doing well” (generating performance) and “doing good” (promoting responsible companies) by exploring opportunities both on the long and on the short side. The application of ML techniques to ESG data (ESG-QUANT score) establishes the essential balance between the two objectives. ML approach to fundamental and technical data further enhances the strategy “doing well” side (LOIM-ML model).

The success of any systematic strategy depends just as much on the quality of the model as on the efficiency of portfolio construction and execution. A model that looks attractive on the paper will not always lead to a decent strategy as the alpha gets quickly “eaten” by multiple layers of deductions such as direct and indirect transactions costs, dividend taxes, shorting fees and operational costs. In LOIM we optimize the strategy execution using an advanced procedure that allows achieving given performance and impact objectives in a most diversified and cost efficient way.

Exhibit 4 presents an overview of the main building blocks of the strategy. A forthcoming paper will review the implementation process from the model development to portfolio construction and execution. In the meantime, we will be happy to provide any details and explanations on request.

For professional investor use only. Please see important information at the end of this document.

Important information

For institutional investors only.

This document has been issued by Lombard Odier Funds (Europe) S.A. a Luxembourg based public limited company (SA), having its registered office at 291, route d'Arlon, 1150 Luxembourg, authorised and regulated by the CSSF as a Management Company within the meaning of EU Directive 2009/65/EC, as amended; and within the meaning of the EU Directive 2011/61/EU on Alternative Investment Fund Managers (AIFMD). The purpose of the Management Company is the creation, promotion, administration, management and the marketing of Luxembourg and foreign UCITS, alternative investment funds ("AIFs") and other regulated funds, collective investment vehicles or other investment vehicles, as well as the offering of portfolio management and investment advisory services.

Lombard Odier Investment Managers ("LOIM") is a trade name.

This document is provided for information purposes only and does not constitute an offer or a recommendation to purchase or sell any security or service. It is not intended for distribution, publication, or use in any jurisdiction where such distribution, publication, or use would be unlawful. This material does not contain personalized recommendations or advice and is not intended to substitute any professional advice on investment in financial products. Before entering into any transaction, an investor should consider **carefully** the suitability of a transaction to his/her particular circumstances and, where necessary, obtain independent professional advice in respect of risks, as well as any legal, regulatory, credit, tax, and accounting consequences. This document is the property of LOIM and is addressed to its recipient exclusively for their personal use. It may not be reproduced (in whole or in part), transmitted, modified, or used for any other purpose without the prior written permission of LOIM. This material contains the opinions of LOIM, as at the date of issue.

Neither this document nor any copy thereof may be sent, taken into, or distributed in the United States of America, any of its territories or possessions or areas subject to its jurisdiction, or to or for the benefit of a United States Person. For this purpose, the term "United States Person" shall mean any citizen, national or resident of the United States of America, partnership organized or existing in any state, territory or possession of the United States of America, a corporation organized under the laws of the United States or of any state, territory or possession thereof, or any estate or trust that is subject to United States Federal income tax regardless of the source of its income.

Source of the figures: Unless otherwise stated, figures are prepared by LOIM.

Expected returns are estimates of hypothetical average returns of economic asset classes derived from statistical models. There can be no assurance that these returns can be achieved. Actual returns are likely to vary.

Models may be misspecified, badly implemented or may become inoperative when significant changes take place in the financial markets or in the organization. Such a model could unduly influence portfolio management and expose to losses.

Although certain information has been obtained from public sources believed to be reliable, without independent verification, we cannot guarantee its accuracy or the completeness of all information available from public sources.

Views and opinions expressed are for informational purposes only and do not constitute a recommendation by LOIM to buy, sell or hold any security. Views and opinions are current as of the date of this presentation and may be subject to change. They should not be construed as investment advice.

No part of this material may be (i) copied, photocopied or duplicated in any form, by any means, or (ii) distributed to any person that is not an employee, officer, director, or authorised agent of the recipient, without Lombard Odier Funds (Europe) S.A prior consent.

©2020 Lombard Odier IM. All rights reserved.

17 96

LOMBARD ODIER
INVESTMENT MANAGERS

www.loim.com