RockCreek

Future of Investment Management

Quantitative Models & Machine Learning

Rock Creek Group

November 14, 2016

Quantitative and Machine Learning models are becoming extremely important to the asset management industry. Recent surveys have shown that investors have allocated more capital to strategies based on these models than any other hedge fund strategy in 2016. While Machine Learning models have made great strides they are in their relative infancy. Their use raises a number of ethical and regulatory issues much like those than exist with the use of AI based technology in areas such as self-driving cars. This short paper seeks to illustrate the use of a Machine Learning model to answer a simple question – under what market environments will hedge funds outperform a 60% equity/40% bond portfolio?

I. Introduction

Traditionally, capital allocation decisions have been based on asset classes, such as allocating 60% of the capital to equities and the remaining balance 40% to fixed income in a portfolio. Consequently, capital allocation decisions have been called asset allocation decisions. Over time, investment firms have expanded the asset classes in which they invest. For example, the original set of domestic equity and fixed income asset classes has been expanded to include high yield bonds, emerging market equities, commodities, currencies, private equity, real estate and hedge funds. One problem with asset class-based grouping of investments is that it based on security or instrument type. Similar instruments, i.e., stocks of US companies, in the same asset class can behave quite differently in different market environments. Consequently, the asset class approach has not been the most effective approach for making capital allocations for any given forecasted market environment

The alternative to asset class based allocations is the risk factor based allocation paradigm. In this approach, the investments are grouped based on their exposure to risk factors. Larger capital allocations are made to risk factors expected to generate high returns in a future market environment with lower allocations to risk factors expected to do poorly. Besides beta (also known as the market risk factor) which is well known, academicians have identified a wide range of risk factors such as market cap, value-growth, momentum etc., for partitioning investments. Risk factor exposures are widely used in both quantitative and Machine Learning models.

Quantitative models have been used for capital allocation decisions for decades. They have proven to be particularly helpful in allocating capital across investments. The availability of inexpensive computing power (cloud computing) and lots of data (big data) has enabled investment professions to build extremely complicated quantitative models that can: take into account tens of risk factors, optimize over hundreds of securities, adapt to both normal and non-normal distributions, incorporate both linear and non-linear models, and be dynamic over time.

The inputs to quantitative models are generally in pre-defined formats known as “structured” data, examples of which include time series of returns, and exposure to risk factors. Most real life data, however, such as newspaper articles, company announcement, and internet blogs, are unstructured data. This data is not directly “understood” by computers, and usually requires human intervention before it can serve as input for computers. The developments in language processing capabilities now help computers process unstructured data. Computers now “read” Federal Reserve announcements and make trade decisions. For example, Rock Creek has started using these technologies to update in house databases on managers and markets on a real time basis (Rock Creek’s Project Mercury). These technological developments enable computers to access large volumes of both structured and unstructured data for dynamic decision making.

Quantitative models define causal relationships between inputs and outputs, in which the input/output variables as well as the mathematical algorithms relating these variables are defined by human beings. A well-defined quantitative model may be precise in relating inputs and outputs in back tests, but quite wrong with predictions in out of sample tests. The experience of the model builders is sometimes used to change the input variables or their relationships to output variables to improve performance. However, the constant updating of models can lead to over-fitting, and to poorer predictive capabilities. Sophisticated qualitative models generally have many more input variables, and tend to be better predictors of future investment outcomes.

Unlike quantitative models, machine learning models are purely predictive models. In this case the relationship between inputs and outputs may or may not be causal. As before, the inputs and outputs are defined by human beings but the mathematical algorithms relating the variables are developed by the machine. The machine may or may not use all the input variables provided when it develops a model. Moreover, the machine suitably updates its mathematical algorithms, i.e., it learns’, as new input and output data become available. Machine learning models can more easily incorporate non-linear relationships between inputs and outputs than quantitative models - because of the inability of human model builders to identify these relationships.

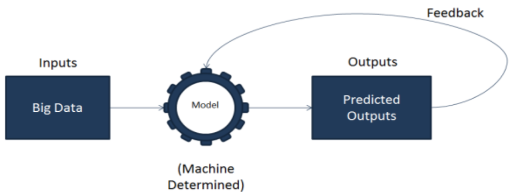


Figure 1: Machine Learning Model

We provide a simple example that illustrates the manner in which machine learning models can be used to answer a couple of simple questions such as:

- What are the important inputs for determining when hedge fund indices will do better than a 60%equity/40% bond portfolio?
- Do the values of these inputs accurately answer why hedge fund indices have underperformed the 60%/40% over the past year or two?
- Are hedge fund indices likely to do better than a 60% equity/40% bond portfolio over the next few months?

II. Example: Predicting Hedge Fund Alpha

The objective in this exercise is to construct a Machine Learning Model (MLM) to predict the difference in return between HFRI Equity Hedge index and a 60% MSCI ACWI Index/40% Cash portfolio. The MLM was provided with input data on the following 17 factors that have been identified in the literature as having predictive power for equity or hedge fund returns: The MLM used data from January 1994 to July 2016 (271 months) on these 17 factors.

Market Factors	Trend Factors	Higher Moments
US Equity	Equity Trend	Cross-section Correlation
Size Spread	Bond Trend	Realized Volatility
Value-Growth Spread	Short Term Interest Trend	Implied Volatility
Emerging Market Equity	FX Trend	
Government Bond	Commodity Trend	
Fixed Income		
Credit Spread		
Currency Exchange Rate		
Commodity		

Hedge fund alpha is defined as the difference between six month rolling returns for the HFRI Equity Hedge Index and the 60% MSCI All Country World Index / 40% Cash portfolio. The MLM is an out of sample predictive model and seeks to predict hedge fund alpha over the next six months using data on all or a subset of the 17 factors, over the previous six months. The MLM was built using the IBM Watson supercomputer. The benchmark to evaluate the MLMs performance is an OLS regression model using the same 17 factors. The prediction power of the MLM and benchmark model is defined as the percentage of times the model correctly classifies the sign of the hedge fund alpha over the subsequent six month period. The five factors with highest prediction power on an individual basis are:

Factor	Correlation	Prediction Power
Cross-section Correlation	-0.59	35.0%
Size Spread	-0.19	3.7%
Government Bond Trend	0.17	3.0%
Short Term Interest Rate Trend	0.15	2.1%
US Equity Trend	0.12	1.5%

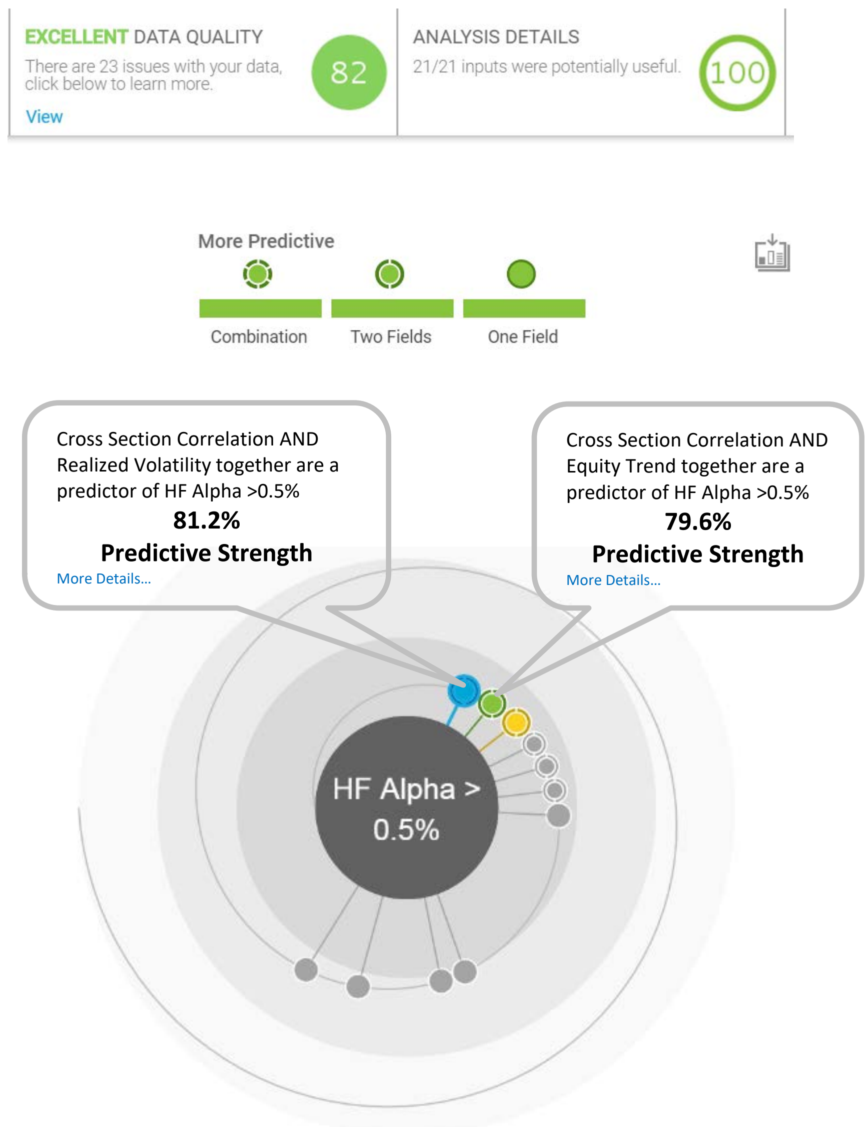
In an OLS regression the prediction power of all five factors was 38%, but none of the factors other than cross-section correlation[1] was significant. In the MLM, the most important prediction factor was also cross-section correlation with a prediction power of 38%. When MLM used two factors, cross-section correlation and realized volatility[2] its prediction power increased to 82%. Realized volatility was not one of the top five predictive factors in the benchmark model. In the MLM, the use of these two factors was machine determined and not prescribed in advance. The robustness of the MLMs predictions tended to improve as the model was trained with new data.

III. Conclusions

The MLM model, used two factors (from the 17 provided), namely cross sectional correlation and realized volatility to predict the sign of the hedge fund alpha with a probability of 82% over a subsequent six month period. The values of these two variables at the end of Q4, 2015 were jointly predicting that the hedge fund alpha would be negative the first half of 2016. This prediction did come to pass. The values of these two variables started changing during the first two quarters of 2016. At the end of Q3, 2016 the variables had values that suggest hedge funds should do better than a 60% equity/40% bond portfolio going forward.

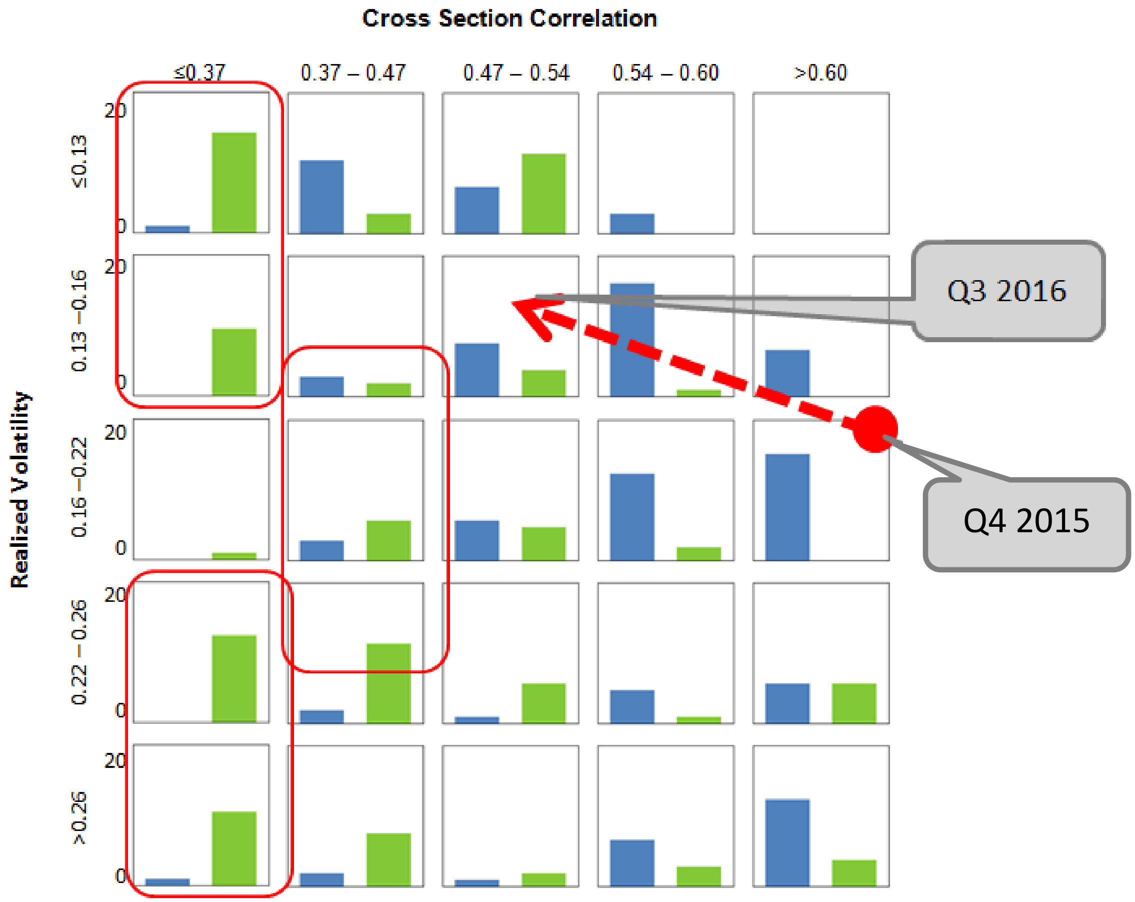
[1] *Cross-section correlation: The average of the correlations between daily returns of S&P 500 components and that of S&P 500 total return index over 6 months.*
[2] *Realized Volatility: The volatility of the daily S&P 500 total return index over a 6 month period.*

Figure 2: Top Predictors of Hedge Fund Alpha



Source: Rock Creek

Figure 3 : Market Environment for Hedge Fund Outperformance



Source: Rock Creek

Important Disclosures

Confidential – Not for public disclosure.

The information contained herein is for the recipient's exclusive use only, and may contain information that is proprietary, constitutes a trade secret, and/or is subject to copyright. This information is confidential and may not be reproduced, distributed, copied or used for any other purpose. It cannot be disclosed to any third party or used for any purpose unrelated to the investor's investment portfolio. By accepting such information, the recipient agrees that it will, and it will cause its directors, partners, officers, employees and representatives, to use such information only to evaluate its potential interest in the securities or services described herein and for no other purposes. By accepting such information, recipient agrees not to divulge any such information to any other party. It may not be reproduced, distributed, copied or used for any other purpose unless specifically authorized in advance by The Rock Creek Group, LP.

This material reflects the opinions of the authors and not necessarily that of Rock Creek.

This material is intended only to facilitate your discussions with Rock Creek as to the opportunities available to our clients, is not intended to be used as a general guide to investing, or as a source of any specific investment recommendations, and makes no implied or express recommendations concerning the manner in which any client's account should or would be handled, as appropriate investment strategies depend upon the client's investment objectives. The discussions herein concerning general economic conditions and political developments are not intended to be used as a general guide to investing, or as a source of any specific investment recommendations, and Rock Creek makes no implied or express recommendations or warranties concerning the manner in which any account should or would be handled, as appropriate investment strategies depend upon the investor's unique investment objectives. As such, the information contained herein has been prepared solely for general informational and is not an offer to buy or sell or a solicitation of any offer to buy or sell any security or to participate in any strategy. If any offer of securities is made, it shall be made pursuant to a definitive offering memorandum (the "Offering Memorandum") in accordance with the terms set forth in the Offering Memorandum, prepared by or on behalf of the relevant fund, which would contain material information not contained herein and which would supersede this information in its entirety.

Neither Rock Creek nor any of its affiliates or employees makes any representation or warranty, express or implied, as to the accuracy or completeness of the information contained herein and nothing contained herein shall be relied upon as a promise or representation as to past or future performance. Past performance is not indicative of future performance.

The analysis, discussions, and commentary presented herein include information that has been prepared by independent third parties and made publicly available. Rock Creek has not verified and is not liable or responsible for the completeness or accuracy of such information. Accordingly, neither Rock Creek nor any of its affiliates or employees makes any representation or warranty, express or implied, as to the accuracy or completeness of the information contained herein and nothing contained herein shall be relied upon as a promise or representation as to past or future performance. Past performance is not indicative of future performance.

A portfolio and how it is advised and managed is subject to a client's investment objectives and restrictions and material economic and market factors that may differ from the scenarios presented herein. As such, no guarantee that your portfolio will be subject to the application of same or similar inputs, factors, and analytics as presented herein. Furthermore, no guarantee that your portfolio will achieve same or similar returns, performance, allocations, and exposures if the factors described are applied to your portfolio as presented herein.

With respect to the indices referenced herein, the volatility of any referenced indices may be materially different from that of an investor's account's portfolio. In addition, the portfolio's holdings may differ significantly from the securities that comprise the indices. The indices have not been selected to represent appropriate benchmarks to compare the portfolio's performance, but rather are disclosed to allow for comparison of the portfolio's performance to that of well-known and widely recognized indices. Information contained herein regarding performance of any index or security is based on information obtained from the indicated sources as of the specified dates, but there is no guarantee as to the accuracy of such information. The underlying exposures, specifically securities, derivatives, or hedges in the selected benchmark index or indices, may vary substantially from that of the portfolio(s) presented.

Discussions and calculations regarding potential future events and their impact on the portfolio are based solely on historic information and estimates and/or opinions, are provided for illustrative purposes only, and are subject to further limitations as specified elsewhere in this material. No guarantee can be made of the occurrence of such events or the actual impact such events would have on the portfolio's future performance. In addition, the opinions, forecasts, assumptions, estimates, and commentary contained in this material are based on information provided to Rock Creek on both a formal and informal basis. Further, any such opinions, forecasts, assumptions, estimates, and commentary are made only as of the dates indicated, are subject to change at any time without prior notice and cannot be guaranteed as accurate.

Any opinions, forecasts, assumptions, estimates and commentary herein are made only as of the dates indicated, are subject to change at any time without prior notice, and cannot be guaranteed as accurate. Rock Creek has no obligation to provide any updates or changes to any such opinions, forecasts, assumptions, estimates, and commentary or to any data or information contained herein.