Quantitative Tactical Asset Allocation Using Ensemble Machine Learning Methods

By Kemal Oflus

Summary:

There is a strong interest in beating the S&P 500 Index benchmark among active portfolio managers. We propose a new method to add a 2-layer augmentation to relative strength and momentum based active portfolio management methods; first layer is to add a filtering mechanism to add a momentum filter in the recommendation engine and second is to include a multi level- multi layer machine learning method to integrate an ensemble model to improve decision making process. The ensemble model consists of gradient boosted decision trees and neural network models. Our initial results show that it is possible to beat the SP500 benchmark index by 600 basis points (in the calculations industry standard trading costs are included) as it is demonstrated by comparing the overall performance of the proposed method.

Introduction:

The paper is outlined as below:

- Review, verify and present relative strength based momentum strategies work with historical data from Dr. Kenneth French's website.

- Apply the strategy developed using method above to current available ETFs to replicate and verify that it is possible to replicate the success of the method

- Introduce a filter to make better use of the momentum strategies

- Augment the new system with multi-level; multilayer neural network to fine tune the decisions to take a position with the system defined ETF.

All of the replication portfolios assume a $10,000 US portfolio value with $17 round trip trading cost per month.

Review of the relative strength/momentum method

The idea for this research came after reading "Relative Strength Strategies for Investing" [1] by Mebane Faber of Cambria Investment Management. In his paper, Mr. Faber, states that relative strength/momentum based methods have been used by many fund managers and demonstrates that these methods continue to work since late 1920s. For the completeness of analysis, we also replicated relative strength/momentum based strategy with 10 industry monthly return data retrieved from website maintained by Kenneth French [2]. This dataset contains monthly returns of 10 industries listed below since 1926.

Table 1: Industry Descriptions

Industry	Description
Consumer Non-Durables	Food, Tobacco, Textiles, Apparel, Leather, Toys
Consumer Durables	Cars, TV's, Furniture, Household Appliances
Manufacturing	Machinery, Trucks, Planes, Chemicals, Off Furn., Paper, Com Printing
Energy	Oil, Gas, and Coal Extraction and Products
Business Equipment	Computers, Software, and Electronic Equipment
Telecom	Telephone and Television Transmission
Shops	Wholesale, Retail, and Some Services (Laundries, Repair Shops)
Health	Healthcare, Medical Equipment, and Drugs
Utilities	
Other	Mines, Constr., Bld. Mt., Trans, Hotels, Bus Serv., Entertainment, Finance

Implementation/Backtesting:

Every month order the monthly return of each sector and select the top sector for the current month. Invest the portfolio on the last months winning sector and repeat this each month. Here is an example:

Let's say today is May 12; the latest information I have is the monthly returns up to end of April. We will order the monthly return from each industry for the month of April; return as calculated by dividing price at the end of April over the price at the end of March and subtracting 1.

$$monthlyReturn = 100 * \left[\left(\frac{price(t-1)}{price(t-2)}\right) - 1\right]$$

Most portfolio strategists –for backtesting purposes- would use the return of May (current month) to decide to take a position for the next month but we believe this is an unrealistic assumption since there is no way of knowing the exact return for the current month until the month ends. We propose a moth lag so that the decision to take a position at the end of current month is available during the current month; before the month closes.

To demonstrate that this system works we tested the basic momentum based relative strength strategy with Dr. French's data[3]. Also we truncated the data to only include dates back to 1950 so that we can compare it to S&P 500 returns.

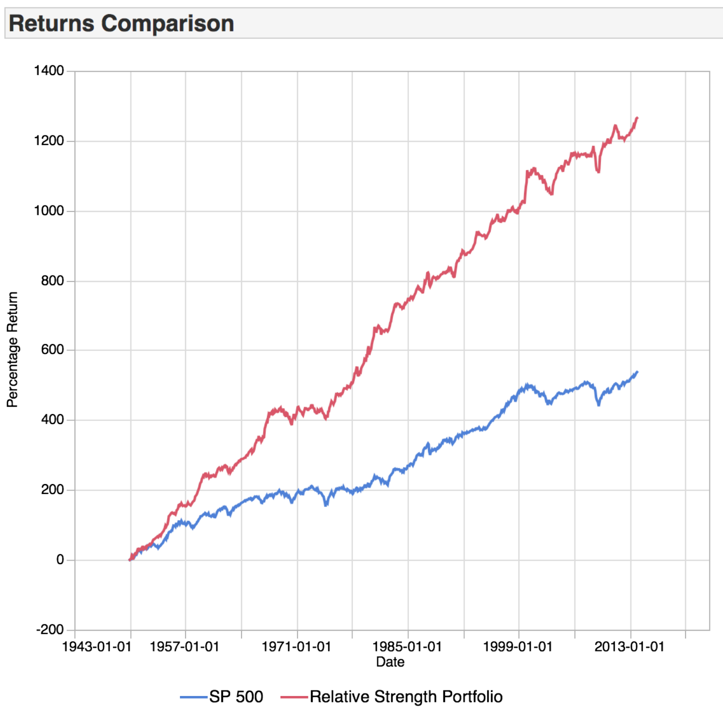

Figure 1: Comparison of Basic System Returns

The system has outperformed the buy and hold S&P500 index 69% of the time.

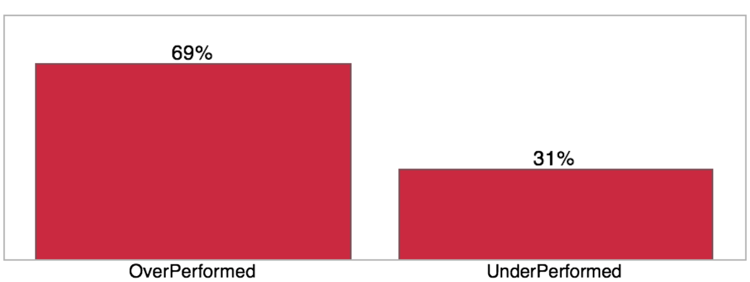

Figure 2: Over/Under Comparison

It can be seen from the table below that the system continues to perform extremely well not only in the 50s but also today.

Table 3: Comparison of Annual Returns

period	Average SP500 Return	Average System Return	outPerformance
1950s	13.32	24.76	11.44
1960s	5.04	16.79	11.75
1970s	2.85	17.20	14.35
1980s	13.28	24.17	10.90
1990s	15.23	24.33	9.09
2000s	-1.44	12.20	13.64
2010s	13.66	18.53	4.86
All	8.40	19.82	11.42

Replication:

We wanted to verify that the system can be replicated today by using Sector SPDRs and iShares ETFs that represent broad indices. The following ETFs were used and data is truncated to include a time period since 2003 because of the extent of these ETFs. Also, in all our analysis we assumed a portfolio value of $10,000 with a round trip transaction cost of $17 per transaction. The transaction costs are in line with most brokers charging $7.95 to buy/sell ETFs.

Table 4: Industry Descriptions

ETF	Industry/Index
XLY(4)	Consumer Discretionary
XLP(4)	Consumer Staples
XLE	Energy
XLF	Financials
XLV	Health Care
XLI	Industrials
XLB	Materials
XLK	Technology
XLU	Utilities
TLT(5)	20+ Year Treasury Bond
IYR	US Real Estate
EEM	MSCI Emerging Markets
EFA	MSCI EAFE

As it can be seen from the plot below the system continues to perform well compared to buy and hold S&P 500 with ETFs available today.

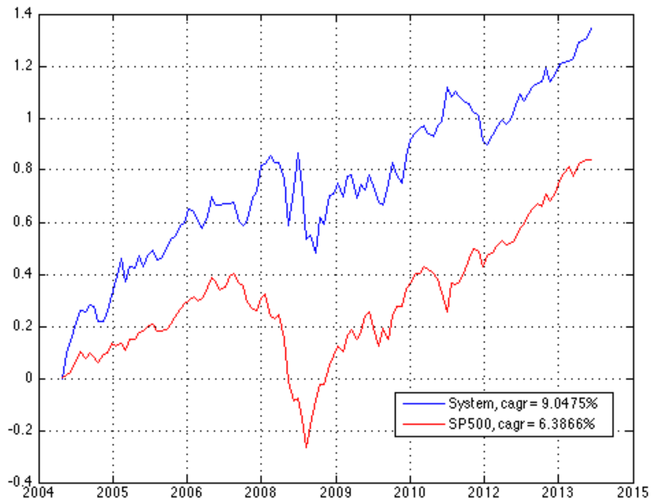

Figure 3: Comparison of Replicated System Returns with SP500

Introduce a Momentum Filter

To improve upon the results presented above, we introduced a momentum-filtering scheme in the form of a moving average filter to help us guide either to take the proposed position or not given the momentum factor of the recommended ETF. The position was only taken if the cumulative return of the given ETF is above its moving average; otherwise no position was taken.

As it can be seen from the plot below, the application of moving average filter improved the CAGR by 150 basis points by increasing from 9.04% to 10.58%.

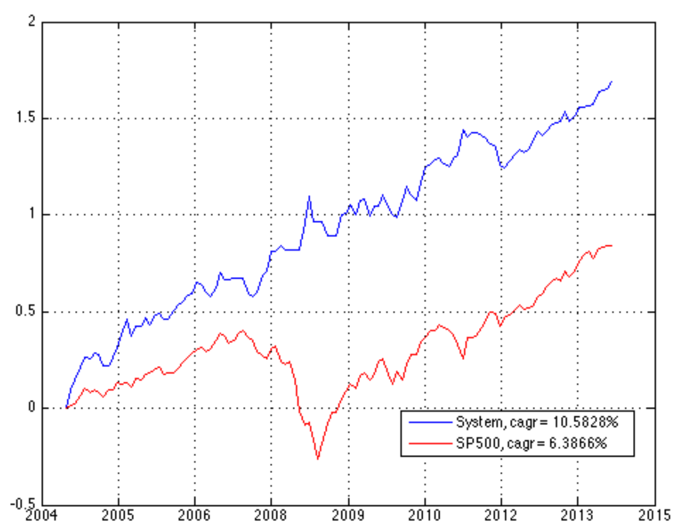

Figure 4: Comparison of Replicated System with Filter over SP500

Introduce an ensemble machine learning system

Armed with historical data and a simple system that we can implement, next we wanted to see if we can build a model to predict, to identify the underlying structure in our decision making scheme. What we are looking for is a mathematical model with the information available prior to taking position; referring back to the definition of the strategy - we have the monthly returns of each industry/sector ETF and we also have selected ETF given the momentum filter results along with the historical results of each of the trades that we took in the past. Without the predictive model we would take a position at that selected ETF. Our predictive model will act as the final review board to decide if the position should be taken.

Before moving into model description, a few critical points to keep in mind and pitfalls to avoid. The goodness, the appropriateness or the strength of most mathematical models are defined by their rSquare values. In the pursuit of getting a better rSquare value, it is possible to overfit your model; in a sense to memorize your historical data and expect everything to be the same in the future. In reality, this is a flawed assumption, similar to driving your car by looking at the rear view mirror.

We also didn't want to use any set time frame since the results wouldn't be as representative. For example; if we tested the model in 2005 it will be a different economic climate than it was in 2009 and the model wouldn't generalize. . To avoid this, we randomly divided our dataset to validate our model. We built our model and tested with the sections of the data that we held back and hidden from the model.

Our final model had 80% prediction accuracy for the randomly selected validation dataset with a compound CAGR of 12.14%.

The model above is a multilevel model where in the first and second levels we create surrogate models to represent intermediary variables that will be the input to the third level multilayer gradient boosted neural network model. These surrogate models include gradient boosted decision trees and multilayer neural networks.

Table 5: Overall Comparison of Annual Returns

Year	%Return - SP500	%Return - original	% Return - ensemble model
2003	12.45	0.00*	0.00*
2004	10.46	26.16	21.20
2005	5.04	16.37	29.48
2006	14.95	22.30	29.79
2007	5.45	-6.80	13.13
2008	-42.87	51.54	51.54
2009	25.93	-1.19	9.31
2010	15.80	8.84	35.91
2011	3.02	25.22	29.24
2012	15.44	-4.34	6.27
2013	28.66	18.35	17.77
2014	2.41	12.45	12.45

*First Year results are not included due to filtering

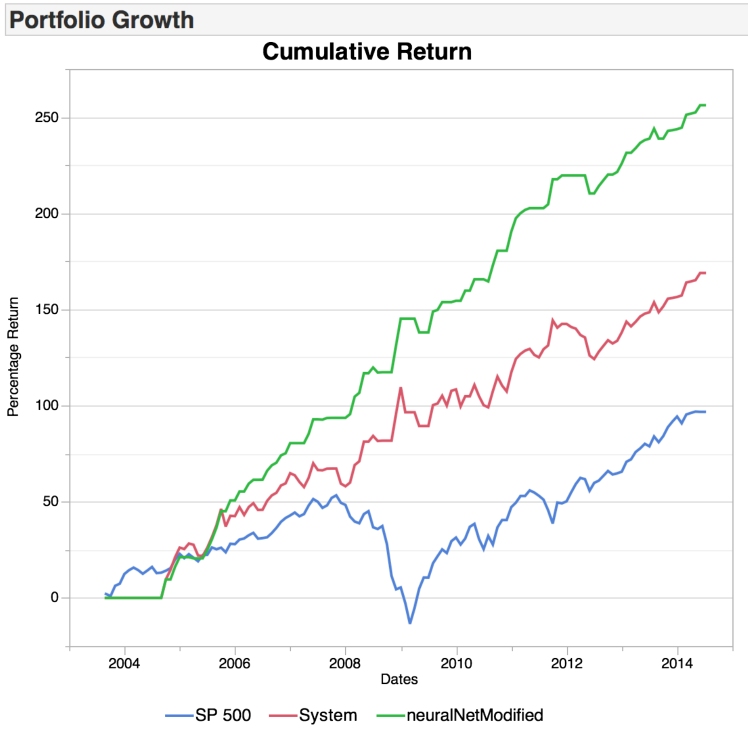

Figure 5: Neural Network Portfolio Growth

Summary

We demonstrated that relative strength methods can be improved with momentum filters and novel adaptations of surrogate models with ensemble methods. With the proposed improvements, we achieved close to 600 basis points improvements; almost doubling the SP500 CAGR rate since 2003 (see table below)

Table 6: CAGR Comparison

	S&P 500 (Buy and Hold)	Original (no filter)	Original (with Filter)	Updated Method with Ensemble Models
CAGR (%)	6.39	9.05	10.58	12.14

Further work will involve generalizing this method to other asset classes; and to weekly and daily returns.

References:

1 – Faber, Mebane T., Relative Strength Strategies for Investing (April 1, 2010). Available at SSRN: http://ssrn.com/abstract=1585517 or http://dx.doi.org/10.2139/ssrn.1585517

2 - Kenneth French website (for data); http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/10_Industry_Portfolios.zip

3 - Kenneth French web site (for data definitions); http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_10_ind_port.html

4 - Sector SPDR ETFs. Definitions can be found at http://www.sectorspdr.com/sectorspdr/sectors

5 - iShares ETFs. Definitions can be found at http://www.ishares.com/us/