Machine Learning Methods Application In Variable Moving Average Trading Rules

Zhiyu Ma, Chen Hu, Hai Wei

1. Project Outline

1.1 General Scheme of Technical Trading Analysis with Machine Learning

- Step 1: Forecast prices of certain financial instruments by studying past prices;
- Step 2: Practiced certain technical trading rules based on the forecast results;
- Step 3: Complete related summary statistics concerning security trading.

1.2 Our Approach

- Construct forecasting based on statistical learning methods: boosting, Bayesian, and committee to reduce the data-snooping bias introduced by the arbitrary selection of the parameters in technical trading rules, and to avoid the element of subjectivity that the procedure involves, and then compute the returns of the trading strategies generated by the above methods;
- Set up a filter to reduce trading frequency, construct statistical learning methods: filtered boosting/Bayesian/committee, and then compute the returns of the generated strategies again.

2. Machine Learning Methods Applied in Studying Securities Prices

2.1 Boosting Approach

Boosting is a learning algorithm that combines the outputs of many weak classifiers to produce a powerful committee. A weak classifier is one which performs just slightly better than random guessing. Boosting refers to the general problem of producing an accurate classifier by combining rough and moderately inaccurate weak classifiers. The most popular one is proposed by Freund and Schapire (1997) called "AdaBoost.M1". A classifier $h_t(x)$ is a function that produces a prediction taking one of the two values $\{-1, +1\}$, where x is a set of predictor variables. The predictions are then combined through a weighted majority vote to produce the final prediction. More specifically, the algorithm takes as an input a training observation set $Z = \{(x_1, y_1), ..., (x_M, y_M)\}$, where $x_i \in X$, $y_i \in \{-1, +1\}$. As Freund and Schapire (1999) show, the AdaBoost.M1 algorithm is as follows:

First, initialize $w_1(i) = \frac{1}{M}$, $i = 1, 2, ..., M$, and then for $t = 1, ..., T$, follow the algorithm:

- Step1: Fit a classifier $h_t(x)$ to the training data using a weight $w_t(i)$;
- Step 2: Compute $\varepsilon_t = \sum_{i=1}^{M} w_i I(y_i \neq h_t(x_i)) \Big/ \sum_{i=1}^{M} w_i$;
- Step 3: Compute $\alpha_t = \frac{1}{2}\log\frac{1-\varepsilon_t}{\varepsilon_t}$;
- Step 4: Update $w_{t+1}(i) = w_t(i)\exp(-\alpha_t y_i h_t(x_i)) / Z_t$, $i = 1, 2, ..., M$.

Here, Z_t is a normalization factor. The final result is: $H(x) = \mathrm{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x_i)\right)$.

2.2 The Bayesian Model-Averaging Approach

Given a set of candidate models $h_t, t = 1, ..., T$, Bayesian model averaging consists of taking a weighted average of the individual predictions, with weights proportional to the posterior probability of each model.

The weights are the rates of success $s_t \Big/ \sum_{j=1}^{T} s_j$ of the model $h_t, t = 1,...,T$ on the training set

$Z = \{(x_1, y_1),...,(x_M, y_M)\}$, where $x_i \in X$, $y_i \in \{-1,+1\}$, and $s_t = \sum_{i=1}^{M} I(y_i \neq h_t(x_i))$.

2.3 Committee Approach

The committee method considers a simple arithmetic average of the predictions from each model, giving equal probability to each model.

3. Variable Moving Average (VMA) Rules

VMA rules are based on the comparison of a short-term moving average of prices to a long-term moving average. A VMA (S, L, B) rule emits buy (sell) signals when the S-day moving average of prices exceeds (is less than) the L-day moving average of prices by at least B%. The choice of a band around the moving average can help reduce the number of trades when the short-term and long-term moving averages get close to each other (See Brock et al. 1992). In this project, we choose S = 1, 2, 5 or 10; L equals from 50 to 200 by a step of 10 and B equals from 0 to 10% by a step of 1%. Then we have 704 VMA rules in total, and apply various methods to combine these rules.

4. Signal Filtering

In real transaction, the high frequency of trades is expensive and inefficient. Thus, we need to construct an additional filter in the signals provided by all the learning methods shown above. The success of the filters would justify the existence of some systematic trends in the prices which are not explained by the random walk model. In this project, the filters for boosting, Bayesian, and committee all have the same form, i.e., for a given a number $\varepsilon > 0$:

- $\varepsilon < \sum_{t=1}^{T} \alpha_t h_t(x_i)$. If we are out of the market, a buy signal is generated; if we are in the market, the trading rule suggests we should continue holding the market;
- $\sum_{t=1}^{T} \alpha_t h_t(x_i) \leq -\varepsilon$. If we are in the market, a sell signal is generated; if we are out of the market, we continue holding the risk-free security;
- If $-\varepsilon < \sum_{t=1}^{T} \alpha_t h_t(x_i) < \varepsilon$, no signal is generated and we maintain the previous position.

Here, in our project, ε is set as the 0.05 quintile of previous 100 signals.

5. Trading Action Set

We consider a simple trading action set: if our methods give a forecasting that the next day's index will increase, we invest all of our money holding the index; on the contrary, we invest all of other money in the U.S. 3 month Treasury bond market. The one-sided transaction costs are 0.1%, which can be achieved by large institutions nowadays.

6. Empirical Result

The main result we get is that in the bullish market, the (filtered) committee/Bayesian methods give good results and are similar to B&H strategy, and all of them are better than (filtered) boosting methods. However, in bearish market filtered boosting method performs best and the performance is much better than all other strategies. We will show our conclusion from the next three results.

Result 1: We first calculate the forecasting error rates of all trading strategies (see Exhibit 1).

Exhibit 1: Forecasting Error Rates

Methods	Forecasting Error Rates (01/041982-11/102008)	Forecasting Error Rates (06/01/1999-05/30/2002)
Committee	0.4737	0.5295
Filtered Committee	0.4742	0.5275
Bayesian	0.4730	0.5265
Filtered Bayesian	0.4742	0.5295
Boosting	0.4886	0.4975
Filtered Boosting	0.4867	0.4995

We find that in general the (filtered) Boosting does not give a better forecasting accuracy compared with other methods. However, in volatile market, they can still give a reasonable forecasting accuracy (>50%) while other forecasting methods performs very poor (<47.4%).

Result 2: To analyze the efficiency of our approaches, we use Sharpe ration (see Exhibit 2). From Exhibit 2, we can see that B&H Dow Jones index has the highest average Sharpe ratio from 1982-2008. However, our filtered boosting algorithm works much better than all other strategies including B&H in bearish market 1999-2002. When we consider the time interval from 1999-2007, filtered boosting algorithm gives 0.5132, much larger than those given by all other methods.

Exhibit 2: Sharpe Ratios

Year	DJ index	committee	F. commit	Bayesian	F. Bayesian	Boosting	F. Boost
1982	0.1092	0.7643	0.6974	0.8622	0.7643	-0.4610	-0.8668
1983	0.9939	0.9939	0.9939	0.9939	0.9939	0.9939	0.9939
1984	-1.0028	-1.3425	-1.2665	-1.1294	-1.2665	-1.5079	-1.4339
1985	1.5387	1.5387	1.5387	1.5387	1.5387	1.3319	1.1167
1986	0.9478	0.9478	0.9478	0.9478	0.9478	0.7891	0.9478
1987	-0.2256	-0.5998	-0.1832	-0.5998	-0.1832	0.2621	-0.5806
1988	0.1658	-0.4270	-0.5552	-0.4048	-0.5552	-0.8153	-0.7622
1989	1.2862	1.2862	1.2862	1.2862	1.2862	1.2862	1.2862
1990	-1.4545	-2.4330	-2.6198	-2.4330	-2.6198	-2.2384	-2.1497
1991	1.2023	0.8750	0.7609	0.8750	0.7609	-0.2023	-0.0512
1992	0.1037	0.1037	0.1037	0.1037	0.1037	-0.8530	-0.5082
1993	1.1201	1.1201	1.1201	1.1201	1.1201	0.6779	0.6967
1994	-1.0990	-2.1172	-1.6242	-2.1172	-1.6400	-2.0738	-1.7664
1995	2.5984	2.3833	2.2916	2.3833	2.2916	2.3810	2.4052
1996	0.9800	0.9800	0.9800	0.9800	0.9800	0.4998	0.4915
1997	1.2731	1.2731	1.2731	1.2731	1.2731	1.22731	1.2731
1998	0.2437	-0.2304	-0.4910	-0.2622	-0.0406	0.5991	0.5846
1999	0.3628	-0.0356	0.1101	-0.0356	0.1104	0.3182	0.5259
2000	-0.2028	-1.9847	-1.8214	-1.7392	-2.0208	0.8219	0.4567
2001	-0.7639	-1.7500	-1.8183	-1.6678	-1.8183	-0.2312	-0.1783
2002	-0.7629	-1.1909	-1.3722	-1.1654	-1.3722	-0.8442	1.0112
2003	1.2265	1.8520	1.8688	1.8520	1.8198	1.4443	1.4974
2004	0.9447	0.9447	0.9447	0.9447	0.9447	0.5845	0.6885
2005	0.4537	0.4537	0.4537	0.4537	0.4537	0.4537	0.4537
2006	0.6270	0.6270	0.6270	0.6270	0.6270	0.1515	0.2007
2007	-0.0367	-0.4097	-0.0367	-0.3354	-0.2086	-0.3116	-0.0367
Average	**0.4088**	0.1393	0.1619	0.1674	0.1650	0.1665	0.1644
Average(99-07)	0.2054	-0.1659	-0.1160	-0.1184	-0.1627	0.2652	**0.5132**

Result 3: In exhibit 3-5, we show the plots of how our various trading strategies work. Exhibit 3 shows that all of the committee and Bayesian methods performs similar in Bullish market compared with B&H

strategy. Among them filtered Bayesian performs best. However, none of them performs well in volatile/bearish market. Exhibit 4 shows that boosting methods gives worse performance than buy & hold strategy in the bullish market. However, when we observe the plots in Exhibit 5, we find that (filtered) boosting methods perform much better than buy & hold strategy (and buy & hold actually works much better than all other methods) in the volatile/bearish market.

Exhibit 3: (Filtered) Bayesian/Committee Performance

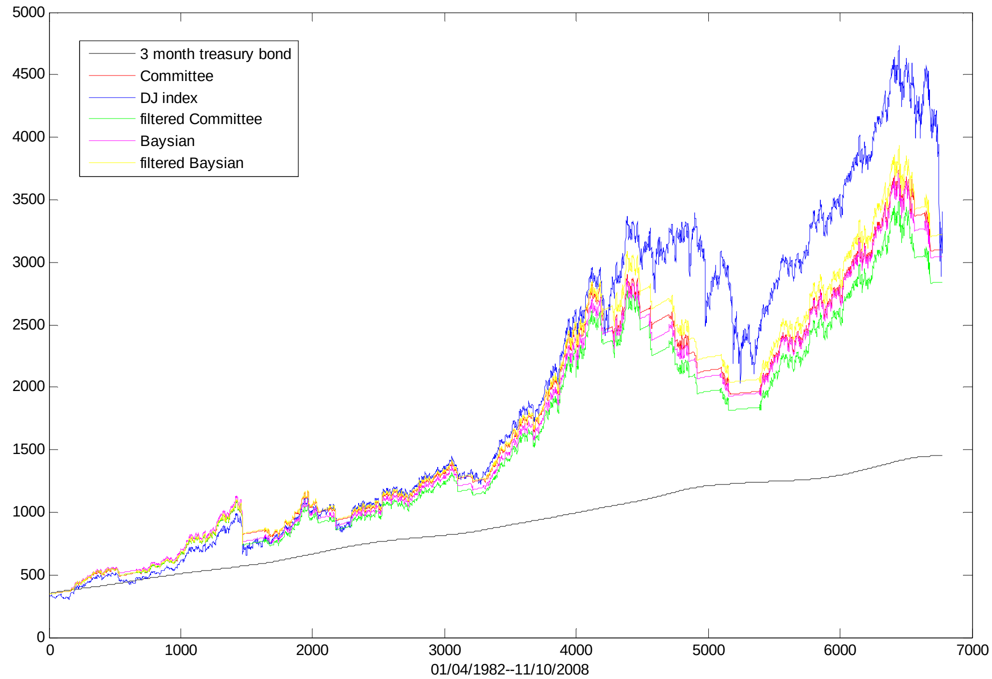


Exhibit 4: (Filtered) Boosting Performance

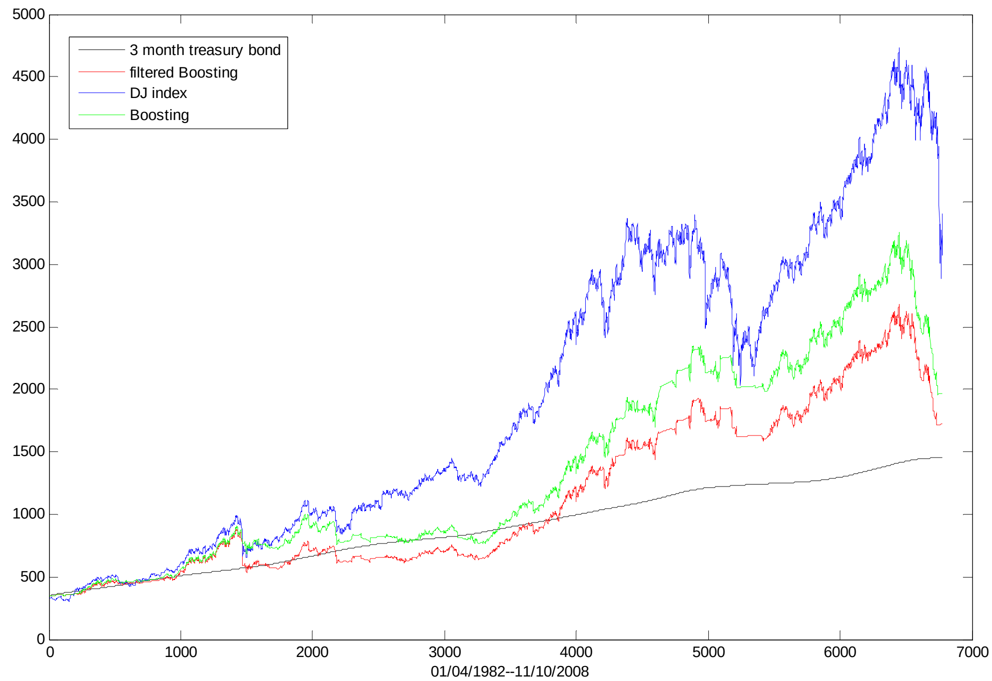
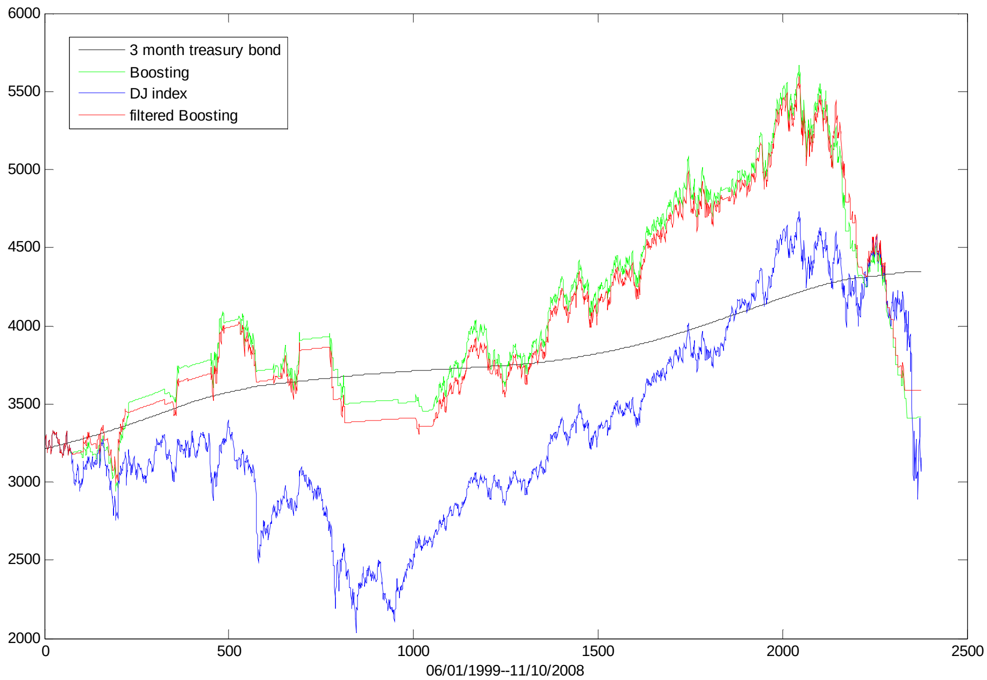


Exhibit 5: (Filtered) Boosting Performance in Bearish Markets



7. Conclusion

Beating the market is a very difficult task. In this project we use different methods to combine a number of variable moving average trading rules to try to forecast the Dow Jones composite index move, and use a simple trading action set: hold the index or hold the 3 month Treasury bond. Our empirical study shows that although in general no one method can beat the market from a long time interval (1982 - 2008), in bearish market the filtered boosting method can give a much better result compared with buy & hold index or any other strategy in terms of absolute return and Sharpe ratio. This result may illustrate other researchers to find a profitable trading strategy based on this boosting method in the future.

References

Andrada-Felix J, Fernandez-Rodriguez F. 2008. Improving moving average trading rules with boosting and statistical learning methods. Journal of Forecasting.

Bessembinder H., Chan K. 1998. Market efficiency and the returns to technical analysis. Financial Management.

Brock W, Lakonish J, LeBaron B. 1992. Simple technical rules and the stochastic properties of stock returns. Journal of Finance.

Freund Y, Schapire RE. 1997. A decision-theoretic generation of on-line learning and application to boosting. Journal of Computer and System Sciences.

Freund Y, Schapire RE. 1999. A short introduction to boosting. Journal of the Japanese Society for Artificial Intelligence.

Hastie T, Tibshirani R, Friedman J. 2001. The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer: New York.

Luenberger D. 1998 Investment Science Oxford University Press.