Proof Protocol for a Machine Learning Technique Making Longitudinal Predictions in Dynamic Contexts

ACM KDD presentation August 2015 Sidney AU

Kevin B. Pratt, Chief Scientist*
Kevin.Pratt@zzalpha.com
www.ZZAlpha.com

*This work is not associated in any way with Kevin Pratt’s other role as Senior Analytics Scientist at Teradata Corporation / Aster Big Data group.
Our challenge

• In 2010, we believed we had developed a machine learning technique that would daily pick stocks whose price would increase in the short-term.

• If true, this result violates a Nobel prize winning orthodoxy: The Efficient Market Hypothesis

• How do we **scientifically prove** our technique?
  • Generally, how does one prove frequently erring, longitudinal predictions in dynamic contexts?

PS: We do not consider the Marketing Dept. mantra “see if it sells” as sufficient proof.
Proof protocol – the big picture

“We think it very interesting to teach machines to understand behaviors that humans cannot describe.

But, how do we know they understood?”
Components of ZZAlpha Learning Architecture in the Cloud

- Daily price and volume
- Weekly fundamentals
- IPOs
- Mergers

Public data acquisition

Filter and Validate

Machine learning engine

Feature transforms and derivations

Archive and working store

Audit store

Timestamp Certification

Newsletter delivery system

Daily newsletters

Results Evaluator

ZZAlpha LTD. Consistent returns from machine learning
ZZAlpha® Machine Learning Engine Recommendations
for trading on Jun 30 2015

LONG Portfolio SP500

<table>
<thead>
<tr>
<th></th>
<th>Cap</th>
<th>Sector</th>
<th>Industry</th>
<th>SIC</th>
<th>Vol(10 d)</th>
<th>Last EA</th>
<th>Liquidity</th>
<th>Close</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMZN</td>
<td>$204014 m</td>
<td>Amazon.com</td>
<td>Services</td>
<td>5091</td>
<td>2396 k</td>
<td>2015-04-22</td>
<td>$1181 m</td>
<td>429.86</td>
</tr>
<tr>
<td>ARG</td>
<td>$8002 m</td>
<td>Aigas</td>
<td>Basic Materials</td>
<td>5159</td>
<td>912 k</td>
<td>2015-04-29</td>
<td>$89 m</td>
<td>105.48</td>
</tr>
<tr>
<td>CVS</td>
<td>$119422 m</td>
<td>CVS Health Corp</td>
<td>Services</td>
<td>5112</td>
<td>552 k</td>
<td>2015-04-30</td>
<td>$591 m</td>
<td>84.81</td>
</tr>
<tr>
<td>EMN</td>
<td>$12450 m</td>
<td>Eastman Chemical Company</td>
<td>Chemical Manufacturing</td>
<td>5169</td>
<td>1709 k</td>
<td>2015-04-29</td>
<td>$143 m</td>
<td>82.60</td>
</tr>
<tr>
<td>NFLX</td>
<td>$39503 m</td>
<td>Netflix</td>
<td>Services</td>
<td>5841</td>
<td>3646 k</td>
<td>2015-04-14</td>
<td>$2375 m</td>
<td>645.62</td>
</tr>
</tbody>
</table>

**RECENT PERFORMANCE**

<table>
<thead>
<tr>
<th></th>
<th>5 Yr (annualized)</th>
<th>One Yr</th>
<th>Yr to Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>LONG Portfolio SP500 (5)</td>
<td>21.4</td>
<td>12.0</td>
<td>6.8</td>
</tr>
<tr>
<td>LONG Segment Avg SP500</td>
<td>16.1</td>
<td>0.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Benchmark: S&amp;P 500 Index</td>
<td>14.0</td>
<td>0.1</td>
<td>2.4</td>
</tr>
</tbody>
</table>

Notes regarding performance calculations. Performance shown as of 2015-06-12. Performance reflects realized returns and thus approximate value of the portfolio. Returns assume purchase (or sale short) on day recommended at price equal to reported opening price. Returns assume sale (or cover) 5 trading days later at price equal to reported opening price on that that later date. Returns assume that for a given day, an equal dollar value is invested in each stock in the recommended portfolio for that day. Returns are not reduced for commissions, spread or slippage. Segment average assumes all stocks in the segment (see specifications of segments in ZZAlpha portfolio descriptions) were traded using the above assumptions. Missing data, erroneous data, delayed data and corporate events may have affected calculations. Returns for 'SHORT' assume an equal dollar value is invested in each stock in the recommended portfolio for that day, and sold short on the next trading day. Returns are not reduced for commissions but are reduced for interest on borrowed money. Returns are not reduced for the cost of shorts. Returns are calculated on a gross basis. The blended returns of all segments in the portfolio are used to calculate the overall portfolio returns. Returns for 'TOTAL' assume an equal dollar value is invested in each stock in the recommended portfolio for that day, and sold short on the next trading day. Returns are calculated on a gross basis. The blended returns of all segments in the portfolio are used to calculate the overall portfolio returns. Returns for 'TOTAL' assume an equal dollar value is invested in each stock in the recommended portfolio for that day, and sold short on the next trading day. Returns are calculated on a gross basis.
Proof protocol components

1. Contemporaneously verifiable discrete predictions
2. Deterministic computability of repetitive longitudinal application of predictions
3. Imposition of realistic costs and context constraints during evaluation
4. Exposure to diverse contexts
5. Statistically significant excess benefits relative to a priori benchmarks and Monte Carlo trials
6. Insignificant decay of excess benefits
7. Controlled risk and absence of pathologies
8. Extended duration real-time trial “in the wild”
Proof protocol

1. *Contemporaneously verifiable discrete predictions*

A prediction must be specific. The set of cohorts for which predictions will be made must be defined in advance. The deadline or trigger for each prediction must be established. In order to validate these predictions, all should be consolidated, unalterably archived and time-stamped.

**What we did:**

The daily consolidated recommendations for the 320 sets are unalterably time stamped by Digistamp (e-timestamp) prior to market open.
Proof protocol

2. Deterministic computability of repetitive longitudinal application of prediction

What we did:

A rolling, cumulative Results Evaluator (as described in paper) applied to all recommendations, using pre-defined purchase (according to recommendations) and sale (5 trading days later).
Proof protocol

3. Imposition of realistic costs and context constraints during evaluation

What we did:
The Results Evaluator (as described in paper) uses the open price (determined by the pre-market auction) as the tradable price. Commissions are paid on every transaction. A minimum block size is applied. Trades can only be made with cash in the account and proceeds available for withdrawal under exchange rules. Stocks must exceed $3 price and 80k volume for inclusion. As variations in cash occurred, the portfolio was rebalanced daily by adjusting new purchase quantities.
Proof protocol

4. Exposure to diverse contexts

What we did:
The Results Evaluator operates on 41 different portfolios – Sectors, Sub-sectors, S&P groups, Capitalization ranges, Liquidity ranges, Capitalization rank, and All Capitalization.

The daily recommendations span 3 years. Historical testing spanned 10 yrs. However, there was only 1 market correction (10% drop from prior peak) (Energy in 2014) among the contexts in the certified 3 years.
5. Statistically significant excess benefits relative to a priori benchmarks and Monte Carlo trials

What we did:

The Results Evaluator was constructed to operate identically on either a) the machine learning recommendations or b) random selections from the market segment or c) benchmarks (either an index or an ETF exchange traded fund tracking a segment that collects and reinvests dividends for greater returns than an index).
Proof protocol

6. Insignificant decay of excess benefits

An engine that does not learn well may become volatile, oscillate or “revert to the mean.”

What we did:

For 41 market segments, longitudinally across the 3 certified years and the 10 historical years:
Proof protocol

7. Controlled risk and absence of pathologies

Humans have broad, varied, sophisticated concepts of pathologies. Effective **visualizations** enable rapid application of those to machine results.

What we did:

For 41 market segments, longitudinally across the 3 certified years and the 10 historical years: 1000 + graphs and visualizations.
Recommendation portfolio results and benchmark
S1.BasicInd 1/1/2012-12/31/2014

Cumulative %

Comparison Distribution of Returns
S1.BasicInd 1/1/2012-12/31/2014

Relative Frequency

Recommendations Avg= 0.444
XLB Avg= 0.285

Recommendation portfolio results (daily min, max, avg) and benchmark
S1.BasicInd 1/1/2012-12/31/2014

% change each day
This small is corner counts!

Prediction response for top 5
8. Extended duration real-time trial “in the wild”

History can only partially prepare us (and our machines) for the unknown future.

What we did:

With a pot of $400k, trade daily stocks solely in the S&P500 list, with a 5 day hold, using only the machine learning recommendations in an ordinary retail online non-margin brokerage account. Typically complete trades in the first 15 minutes after market open. Track results using monthly brokerage statements. ($ results exceeded evaluator determination of recommendations which exceeded SPY ETF which exceeded SP500 Index)

Log every operational and trading issue.

Key Surprises: 1) An experienced trader can obtain results better than shown by the Results Evaluator.

2) We have not needed to turn any knobs in 5 years.
Thank you.

www.ZZAlpha.com
www.ConservativeStockPicks.com