

Predicting Excess Equity Returns Using Company Fundamentals

Jason Victor

May 11, 2009

1 Introduction

Most quantitative approaches to stock market analysis are referred to in the finance industry as “technical” approaches—that is, they rely solely on continuously observable market data in order to make bets. Thus, much of the reasoning that has gone into such approaches has been in the spirit of the *arbitrageur*, detecting subtle mispricing anomalies that can be exploited in the extremely short-term, often with little risk involved.

I am approaching the problem from a different standpoint. As opposed to relying on continuously observable market variables—stock prices, credit spreads, bond yields and so forth—I am focusing on a discrete data set: the financial statements of individual companies. Using this information, I will attempt to predict the return of that company’s common stock over the S&P 500 index over the next year. In industry parlance, this value is known as *alpha*.

I am using kernel support vector machines as my model of choice.

2 Features

Three documents make up a company’s annual financial statements: the cash-flow statement, the income statement and the balance sheet. Each provides a different look at the financial fundamentals underlying the firm’s operations. For example, the balance sheet provides detailed information on the assets and liabilities of the company, while the income statement enumerates the components that factor into the company’s profit (or loss).

Because these are raw values (in dollars), an SVM trained on the financials of any one company would not generalize to another company of a different size. Thus, in order to compare one company to another effectively, it is necessary to observe only *ratios* of these figures. These are known as *financial ratios*, and are a common tool in analyzing prospective investments.

I use several types of financial ratios in my feature vector.

- **Liquidity measurement ratios** describe a company’s access to cash and ability to cover its short-term obligations

- **Profitability indicator ratios** explain how well the company utilized its resources to generate shareholder value
- **Debt ratios** describe a company's debt load and mix of debt and equity financing
- **Investment valuation ratios** shed light on the price of a company's stock in the open market relative to the performance and breakup value of the firm

Furthermore, I include in the feature vector the excess return (alpha) of the given company's stock over the previous year.

3 Results

To test my method, I used the past three years of information for 50 prominent companies across several sectors. All test companies are industry leaders in market capitalization. I used the KSVM implementation in the Kernlab package for R, with a Gaussian RBF kernel and automatic parameter tuning.

All in all, 104 separate tests were performed, yielding a mean squared error of 0.06. However, what is perhaps more interesting from the point of view of portfolio optimization is the *accuracy* of the model. In our case, we define accuracy as the frequency with which the model correctly identified the direction of the output (i.e. whether the company under- or over-performed the S&P 500 index). The accuracy rate for the 104 tests is 72.1%.

The full output of the program (with the results summarized at the bottom) are available at <http://www.cs.dartmouth.edu/~jvictor/milestone-latex/output.txt>.

4 Conclusions

Based on the accuracy rate, I am delighted at the performance of this algorithm. Near-perfect accuracy is never possible in the world of equities, because unforeseeable events—everything from natural disasters to lawsuits—can greatly impact a company's stock price. So, an accuracy rate of 72.1% is fairly desirable.

However, there is plenty of work still to be done. The following are a few things I have planned:

- Test using a large data set (200 companies)
- Test sector-specific data sets (e.g. only technology companies, only banks, only pharma, etc.)
- Test using a data set of micro-cap stocks
- Train the SVM as a binary classifier (of whether a company beats the market)

- Expand the features to other financial ratios and technical variables (e.g. stock price volatility)
- Evaluate the efficiency of k-NN for the purpose of comparison