Stock Portfolio Performance by Weighted Stock Selection

Mihir Bhatia, Anmol Madaan

4^{th} May 2018

Abstract

The aim of this project is to understand how a weighted stock selection model can affect portfolio performance parameters such as returns, risk and win rate that a portfolio manager could expect to achieve when developing weighted stock selection model. The project also aims to answer the question: What factors have the most influence on these output parameters? This predictive model could be used for any dataset that has been developed for another stock index along with it's simulated performance parameters to help managers make better decisions during the stock selection process for their portfolios as per the company's goals in terms of portfolio performance parameters and they desire to achieve.

1 Introduction

To understand the data and inferences in this project, one must know what a portfolio is. A portfolio is a grouping of financial assets such as stocks, bonds, commodities, currencies and cash equivalents, as well as their fund counterparts, including mutual, exchange-traded and closed funds. We also cover weighted stock selection, portfolio performance evaluation and portfolio optimization.

Weighted stock selection
Stock selection can be an arduous task for any quant firm or investment bank. Selecting from a large number of stocks just based on intuition alone or based on certain criterion such as expected returns, risk undertaken, P/E ratios cannot solve the problem of selecting from the large number of stocks available in the market. The criterion on which stock selection is going to take place are chosen and each of these criterion are given weights. Based on these criteria a scoring model is developed and the highest scoring stock is chosen. The weights need to be selected in a way that the portfolio manager's goals in terms of risk, return and win rates are achieved. The past performance of a stock is the only way that quantitative predictions of future stock performance can be carried out using machine learning algorithms. This is where weighted stock selection can help.

Portfolio Performance evaluation
Investment performance is the return on an investment portfolio. The investment portfolio can contain a single asset or multiple assets. The investment performance is measured over a specific period of time and in a specific currency. Investors often distinguish different types of return. Portfolio evaluating refers to the evaluation of the performance of the investment portfolio. It is essentially the process of comparing the return earned on a portfolio with the return earned on one or more other portfolio or on a benchmark portfolio.

Portfolio Optimization
Portfolio optimization is the process of choosing the proportions of various assets to be held in a portfolio, in such a way as to make the portfolio better than any other according to some criterion. The criterion will combine, directly or indirectly, considerations of the expected value of the portfolio's rate of return as well as of the return dispersion and possibly other measures of financial risk.

2 Background

2.1 Literature Review

Existing Literature:

- **Liu, Y.C., Yeh, I.C. Using mixture design and neural networks to build stock selection decision support systems. Neural Computing and Applications [6]**
This study aimed to more efficiently construct weighted scoring stock selection models to overcome the following disadvantages.
Current models cannot identify the relations between weights of stock-picking concepts and performances of portfolios.Current models cannot systematically discover the optimal combination for weights of concepts to optimize the performances.Current Models are unable to meet various investors preferences.

- **Leo Brieman, J.H. Friedman.(1995).Predicting Multivaraite Responses in Multiple Linear Regression.[2]**
The paper discusses about the advantages of utilizing correlations between response variables to improve prediction accuracy as compared to individual linear regression of each variable on the common set of predictor variables. It also discusses about a new method for the same i.e. Curd & Whey Method.

- **Cheng, W.L., Yeh, I.C.First and second order sensitivity analysis of MLP.[11]**
Multi-layered perceptrons (MLP) can build accurate classification and function mapping models.However, they have also been labeled a 'black box' because they provide little explanatory insight into the contributions of the input variables in the prediction model. This paper tries to derive the first- and second-order effect index and importance index through the differential statistical method.

- **Pande, A., Li, L., Rajeswaran, J. et al. Mach Learn (2017). Multivariate Boosted Trees[10]**
Multivariate trees are used to fit a novel flexible semi-nonparametric marginal model for longitudinal data. In this model, features are assumed to be nonparametric, while feature-time interactions are modeled semi-nonparametrically utilizing P-splines with estimated smoothing parameter. To to avoid overfitting, a relatively simple in sample cross-validation method is used, which can be used to estimate the optimal boosting iteration and which has the surprising added benefit of stabilizing certain parameter estimates. Our new multivariate tree boosting method is shown to be highly flexible, robust to covariance misspecification and unbalanced designs, and resistant to overfitting in high dimensions.

- **Hugh A. Chipman, Edward I. George, Robert E. McCulloch. Bayesian Adaptive Regression Trees (2008)[3]**
Bayesian Adaptive Regression Trees (BART) is a sum of trees model where each tree is constrained by a regularization prior to be a weak learner, and fitting and inference are accomplished via an iterative Bayesian backfitting MCMC algorithm that generates samples from a posterior. BART is a nonparametric Bayesian regression approach which uses dimensionally adaptive random basis elements. BART is very effective at removing variables that don't contribute to the model and hence is useful in variable selection. Moreover, the model was tried on 42 datasets where BART outperforms Linear Regression, MARS, Neural Nets and Boosting.

The research conducted in the Liu, Yeh's paper [11] has used neural networks to as a basis for analysis. Sensitivity analysis of the neural network is also done in order to interpret the results of the neural network that was developed. We use various simpler and interpretable models to find how different factors effect the return, risk and reward. Moreover, since there are 4 different datasets over various time periods, our aim is to obtain a trend in the effect various factors have over the 4 periods.

3 Data

Source: UCI Machine Learning Repository
Dataset Name: Stock Portfolio Performance Dataset
Institutions:

- Department of Information Management, Chung Hua University, Taiwan.
- Department of Civil Engineering, Tamkang University, Taiwan.

Description of the data: The stocks have been selected from the S&P500 stock index which is based on the market capitalizations of 500 large companies having common stock listed on NYSE and NASDAQ. TO develop the dataset Yi-Cheng Liu. I-Cheng Yeh makes use of the "Simplex Centroid Design of Experiments with Mixtures" in which we obtain $2^q - 1 = 63$ combinations of weights (since q=6). Using a stock market database, "backtesting" through which the performance of portfolios are obtained.

We will use this dataset to obtain a predictive model that helps us to determine these portfolio performance parameters by using weights as inputs. Moreover, the dataset is composed of 5 subsets of data, 4 which take into account different periods (5 years each) and 1 dataset that takes into account the overall performance (20 years) of a portfolio. We will train and test predictive models with all 5 datasets separately to give us an indication of whether the models are better at short-term or long-term dataset predictions.

3.1 Input Variable

There are 6 response variables that have been used as indicators of performance of a stock portfolio.

- **Large B/P** (x_1)
  The book-to-market ratio is defined as the value of book value to the the market price of the share.It comprises the Value factor of the portfolio. Higher the B/P ratio, Higher is the possibilty that the stock is undervalued.
- **Large ROE** (x_2)
  Return on equity is the amount of net income returned as a percentage of shareholders equity. Return on equity measures a corporation's profitability by revealing how much profit a company generates with the money shareholders have invested.It is a valuation metric for stocks. It is calculated by dividing the company's revenue by the market cap in the most recent year.
- **Large S/P** (x_3)
  Ratio of sales to the market capitalization of the company. It also comprises the Value factor of the portfolio. Higher the S/P ratio, Greater is the possibilty that the stock is undervalued.
- **Large Return Rate in last quarter**(x_4)
  A rate of return is the gain or loss on an investment over a specified time period, expressed as a percentage of the investment's cost. Gains on investments are defined as income received plus any capital gains realized on the sale of the investment.It Comprises the Momenturm factor of the portfolio. It appears as a Reversion in Short Term,Momentum in Mid-term and Reversion in Long Term.Momentum indicates that if return rate is high, it willcontinue to go high.
- **Large Market Value** (x_5)
  Market value is the price an asset would fetch in the marketplace. Market value is also commonly used to refer to the market capitalization of a publicly-traded company, and is obtained by multiplying the number of its outstanding shares by the current share price.It comprises the Scaling factor of the portfolio. There is a -ve relation between market capitalization of firm and Return Rate of a firm's stock.
- **Small Systematic Risk** (x_6)
  The reward-to-risk ratio is the expected return per "unit" of systematic risk, or, in other words, the ratio of the risk premium and the amount of systematic risk.It measures fluctuation in stock relative to market. It comprises the Risk factor of the portfolio.

SNo	Predictor	Meaning
1	Large B/P (x_1)	Ratio of book value to Market Price of the share.
2	Large ROE (x_2)	Net income returned as a percentage of share holder equity.
3	Large S/P (x_3)	Ratio of sales to the market capitalization of the company.
4	Large Return Rate in last quarter (x_4)	Net amount of discounted cash flow received on an investment.
5	Large Market Value (x_5)	Market capitalization of investment made through the portfolio.
6	Small systematic Risk (x_6)	It takes the risk-reward ratio of the investments.

Table 1: Summary of predictors

3.2 Response Variables

There are 6 response variables that have been used as indicators of performance of a stock portfolio.

- **Annual Return** (y_1)
  To define annual return, we need first need to know the meaning of simple return

$$SimpleReturn = \frac{CurrentPrice - PurchasePrice}{PurchasePrice} \tag{1}$$

  Annual return can be a preferable metric to use over simple return when you want to evaluate how successful an investment has been, or to compare the returns of two investments you've held over different time frames on equal footing. You can annualize simple return using:

$$AnnualReturn = (SimpleReturn + 1)^{\frac{1}{YearsHeld}} - 1 \tag{2}$$

- **Excess Return** (y_2)
  Excess returns are the return earned by a stock (or portfolio of stocks) and the risk free rate, which is usually estimated using the most recent short-term government treasury bill. For example, if a stock earns 15% in a year when the U.S. treasury bill earned 3%, the excess returns on the stock were $15\% - 3\% = 12\%$.

- **Systematic Risk** (y_3)
  Systematic risk, also known as **"un-diversifiable risk"**, is the uncertainty inherent to the entire market or entire market segment. It's also referred to as volatility. Systematic risk consists of the day-to-day fluctuations in a stock's price. Volatility is a measure of risk because it refers to the behavior of an investment instead of the reason for this behavior. Because market movement is the reason why people can make money from stocks, volatility is essential for returns, and the more unstable the investment the more chance there is that it will experience a dramatic change in either direction and thus higher the risk that is undertaken by investing in that stock.

- **Total Risk**(y_4)
  Total risk is the sum of the systematic and unsystematic risk. Unsystematic risk, also known as **diversifiable risk** is the type of uncertainty that comes with the company or industry you invest in.

$$TotalRisk = SystematicRisk + UnsystematicRisk \tag{3}$$

- **Absolute Win Rate** (y_5)
  Win Rate refers to the fraction of trades in which an investor makes a profit out of the total trades. This is an important number use along with the Risk/Reward ratio of an investment. A trade with low risk/reward ratio and high win rate is a desirable trade. An absolute win rate doesn't take into account the overall market performance.

$$Win_{abs} = \frac{n_1}{N} \tag{4}$$

  where,
  n_1 = No. of holding periods when the portfolio return rates are above 0
  N = Total no. of portfolio holding periods

- **Relative Win Rate** (y_6)
The relative win rate takes into indication the market performance. If the market performance is poor, then the absolute win rate will automatically decline. Thus relative win rates takes into account how the portfolio performance in comparison to the market.

$$Win_{rel} = \frac{n_2}{N} \tag{5}$$

where
n_2 = No. of periods in which the portfolio return rate beats the market return rate
N = Total no. of portfolio holding periods

The output variable data obtained from the simulations were normalized for the purpose of their application in neural networks. The normalization of data is done because stock markets have trends and the standard output data is influenced by the general trend of stock markets. We tried using both the standard and normalized data in our analysis because this data is representative of company portfolio performance independent of trends in the market. We represent the normalized output variables as: $y_{11}, y_{22}.....y_{66}$

SNo	**Response**	**Meaning**
1	Annual Return (y_1)	Expresses the stock's increase in value over designated period of time
2	Excess Return (y_2)	Investment returns from a security that exceed the riskless rate on a security perceived to be risk free
3	Systematic Risk (y_3)	Uncertainty inherent to the entire entire market segment
4	Total Risk (y_4)	Sum of Systematic and Unsystematic Risk
5	Absolute Win Rate (y_5)	Win Rate refers to the fraction of trades in which an investor makes a profit out of the total trades
6	Relative Win Rate (y_6)	The relative win rate takes into indication the market performance.

3.3 Exploratory Analysis

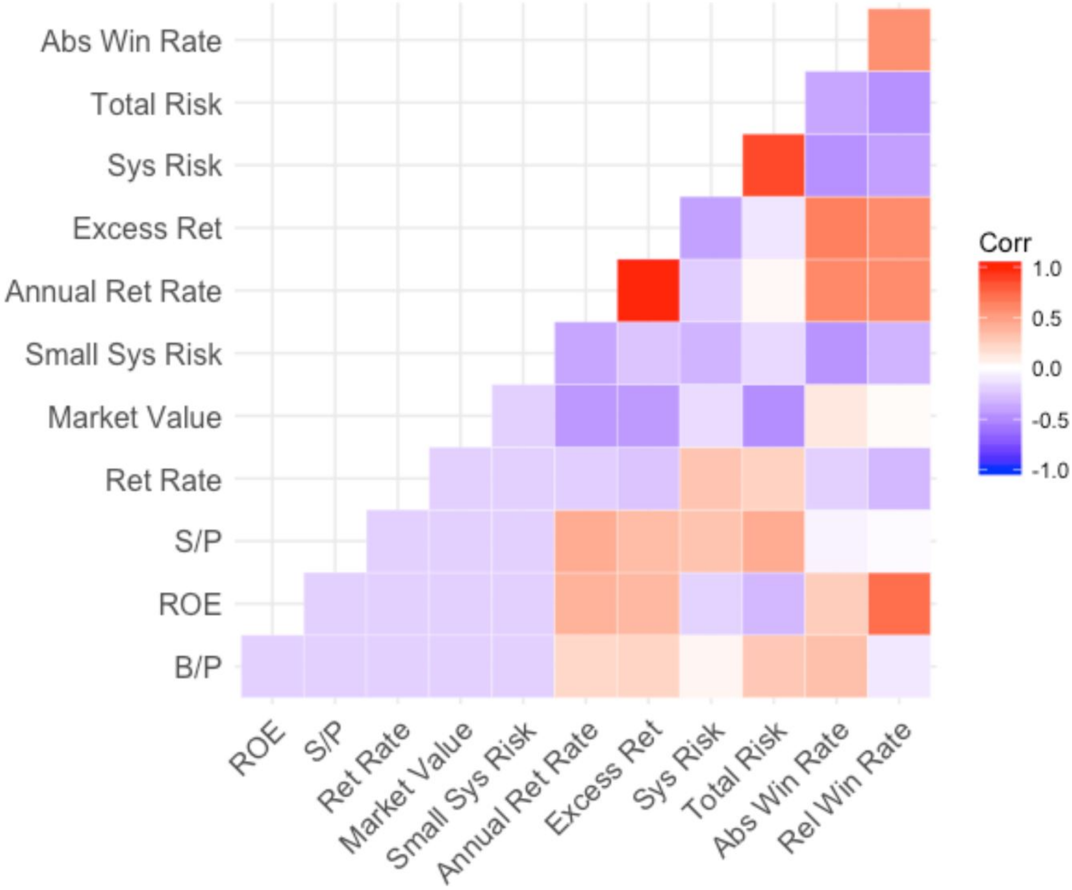

Figure 1: Correlation plot between the predictors and normalized response variables for the All Period dataset

From figure 1, we observe that the Annual Returns and Excess Returns are highly correlated. This makes sense because Annual Return is just the sum of what the market trend returns on an investment and the Excess return. Similarly, the Systematic risk and Total Risk are show high positive correlation. We observe that there are strong positive correlations between the Win Rates and Return Rates observed in the response variables. Moreover, we also observe strong negative correlations between Risks and Win Rates. This would mean that as managers aim for higher Win Rates with the investments made, they incur higher levels of Systematic risk and Total risk. There are no correlations observed amongst the predictors because the values imparted to these predictors are just weights associated with the factors that have been considered for the stock selection process. These weights have been uniformly assigned and hence we don't see any correlations among these predictors. As for response variables, there are both high positive and negative correlations amongst these responses. Hence models like linear regression, MARS and GAM that are used for modeling the data would have a tendency to give poor fits. But tree based models such as Random Forests, BART and MVTBoost might have superior performances.

Between the predictors and response variables, a strong positive correlation is observed between the predictor Relative Win Rate and response Return on Equity. This makes sense, because with improvements in a company's performance we can expect higher win rates and this performance improvement is indicated by a higher return on equity (which is a clear indicator of a company's current profits).

4 Methods

Since predictive models such as Support Vector Machines and Neural Networks would work well only with normalized data, it made sense for us to conduct our analysis and compare models based on their performance with normalized data ($y_{11}...y_{66}$)

Root Mean Squared Errors of Applied Models for normalized data (x 10^{-2})						
Model	Annual Ret.	Excess Ret.	Sys. Risk	Total Risk	Abs. Win Rate	Rel. Win Rate
Linear Model	7.89	9.45	9.86	8.85	8.8	9.63
Random Forests	4.10	4.25	4.73	5.28	4.95	5.80
GAM	9.68	9.40	21.53	22.00	13.96	15.26
MARS	4.96	5.56	6.10	5.75	7.11	8.20
BART	1.93	1.90	2.44	3.12	5.24	6.72
SVM	2.55	2.60	2.98	4.19	4.97	5.96
MVT Boost	17.99	2.22	1.36	14.31	4.57	2.18
Neutral Networks	3.23	3.10	3.01	2.89	5.39	7.16

Table 2: In-sample RMSE values for applied models for the all the periods

Root Mean Squared Errors of Applied Models for normalized data (x 10^{-2})						
Model	Annual Ret.	Excess Ret.	Sys. Risk	Total Risk	Abs. Win Rate	Rel. Win Rate
Linear Model	7.52	8.74	8.96	8.16	8.08	8.93
Random Forests	5.77	5.91	6.45	6.92	7.04	8.69
GAM	9.05	7.97	18.09	19.10	10.66	12.61
MARS	6.38	6.74	6.12	6.00	7.38	8.49
BART	3.43	3.36	4.23	4.83	6.65	7.59
SVM	4.23	4.01	3.97	4.95	6.20	8.54
MVT Boost	28.12	3.88	3.99	4.10	6.93	7.77
Neutral Networks	3.43	3.98	3.81	3.34	6.15	7.48

Table 3: Out-of-sample RMSE values for applied models for the all the periods

- **Generalized Linear Models**
  Generalized Linear Models (GLM) [9] are an extension of the general linear regression. In GLM, the response variable Y can be distributed according to some exponential family of distributions (e.g., Gaussian, Binomial, Poisson, Gamma or inverse-Gaussian), and the systematic component that relates the input variables to the response. ($Y = \sum_{i=1}^{p} \beta_i x_i$) can be linearized by using an appropriate link function. The model can be mathematically represented as :

$$Y_i \sim indep f(y_i) \tag{6}$$

$$f(y_i) = exp[((y_i\theta_i - b(\theta_i))/(a(\phi) - c(y_i, \phi))] \tag{7}$$

  where θ_i is the canonical parameter and ϕ is the dispersion parameter of the exponential family of distributions [7] .

- **Random Forests**
  Random Forest is an ensemble, data-miner which uses 'deep' (unpruned) decision trees as base learners [1]. It is a modification of applying bagging (bootstrap aggregating) to multiple classification and regression trees (CART), and averaging the predictions of the approximately uncorrelated trees to yield the final estimate.
  Random Forest model was unable to show any clear patterns in the data through variable importance plots and did not show any significant improvement in performance in comparison to generalized linear models.

- **Generalized Additive Models (GAMs)[5]**
Generalized Additive Models in which linear predictors depend linearly on unknown smooth functions of some predictor variables. The function for approximating the expectation of some an observed quantity, it could be written as:

$$g(E(Y)) = \beta_0 + f_1(x_1) + f_2(x_2) + ... + f_m(x_m) \tag{8}$$

- **Multivariate Adaptive Regression Splines (MARS)**
Multivariate Adaptive Regression Splines (MARS) [4] is a semi-parametric, flexible and adaptive procedure for regression and is suitable for high-dimensional (i.e., many input variables) datasets. The MARS model can be mathematically represented as:

$$f(X) = \beta_0 + \sum_{m=1}^{p} \beta_m h_m(X) \tag{9}$$

where $h_m(X)$ is a reflected pair of linear splines in C (or a product of two or more functions). β_0 is the intercept, and $beta_m$ are the model parameters that can be found by minimizing the sum of squared errors. MARS is built in a forward manner starting from a basis function of one (the intercept). It then uses a greedy algorithm for step-wise addition of reflected pairs of splines (in C) that lead to the largest reduction in the training error. The model is then pruned back using generalized cross validation (GCV) to avoid over-fitting.

- **Bayesian Adaptive Regression Trees (BART) [3]**
Bayesian adaptive regression trees are a sum-of-trees based model in which each tree is constrained by a regularization prior to be a weak learner, and fitting and inference are accomplished via an iterative Bayesian backfitting MCMC algorithm that generates samples from a posterior.

- **Multivariate Boosted Trees (MVTBoost)**
[10] describes multivariate boosted trees to fit longitudinal data. This algorithm gave fairly good predictions in comparison to other models. But it worked better with normalized data. Although, boosted trees don't usually require normalized data, we decided to develop predictions using normalized data as it would be easier to compare the performance of multivariate boosted trees to methods such as SVM and Neural Networks which do require normalized data.

- **Support Vector Machines (SVM)**
A Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples. In two dimensional space this hyperplane is a line dividing a plane in two parts where in each class lay in either side. The equation of the hyperplane can be given as:

$$w^T + b = 0 \tag{10}$$

The margin is width is $2/\|w\|$ and the learning problem is equivalent to unconstrained optimization problem over $\mathbf{w}$.

$$min\|w\|^2 + C\sum_{i}^{N} max(0, 1 - y_i f(x_i)) \tag{11}$$

Support vector machines are highly effective in high dimensional spaces but under perform when target classes (for classification problems) are overlapping i.e. kernel functions ($\phi(x)$) need to be used.

- **Neural Nets[8]**
Neural networks are a class of learning algorithms which involve multiple stages of prediction.

In its basic version, neural networks make predictions through one or more transformations (via hidden layers). Consider a vector of predictor variables $X = (X_1, X_2, ..., X_n)$. For a single-layer neural network, define the vector of variables belonging to the hidden layer as $Z = (Z_1, Z_2, ..., Z_k)$. Then for all $m \in 1, ..., K$

$$Z_m = \sigma(\alpha_0 + \alpha_m^T X), m = 1, 2..., M \tag{12}$$

$$T_k = \beta_0 + \beta_m^T Z \tag{13}$$

$$f(X_P) = g(T) \tag{14}$$

Where $\sigma()$ is a transformation function called the basis function, α_m are the parameters of the transformation function, β_m is the vector of weights which are typically estimated via optimization techniques such as gradient decent or the Newton's method with the goal of reducing the estimated sum of squared errors. $E_D = \sum_{i=1}^{n}(t_i - a_i)^2$

5 Model Inference

From Table 2 and Table 3, the RMSE values obtained for the outputs, we can conclude that BART and Neural Networks are the best performing models for the 20 year period. Analyzing model performances for the 5 year Period datasets, we conclude that BART's performance exceeds that of Neural Networks. Bayesian Adaptive Regression trees have been found to give superior performance in several datasets in comparison to other tree based models due to the considerations of prior distributions and MCMC algorithm that is used. Moreover, the model is robust to removal of predictors that might not be very important and gives accurate predictions even when predictors that might not be to important are removed.

The figure 2 consists of 6 bar plots that describe the importance of the predictors for each response variable. It is evident from all the plots that B/P ratio is by far the the most important predictor for all portfolio performance parameters. The reason for this could be that B/P ratio is an important predictor for a company's future performance i.e. it accurately reflects how the company might perform in the next quarter. A higher B/P ratio indicates that a stock is undervalued. This means that it's value is more likely to increase. Hence, a stock portfolio which prefers stocks with higher B/P values will see high performance. This is evident from figure 3 (i) where we see trend of increase in Annual Returns with increase in B/P ratio. An undervalued stock with high B/P ratio is also less risky to invest in and this is evident from figure 3 (iii). Also, from figure 3 (vi),we see that there is an increase in Absolute win rate with increase in B/P ratio. This is evident because an undervalued stock is likely to see increases in the coming quarters and hence there's higher win rates are observed.

From Figure 2, the second most important predictor for all response variables was Return on Equity (ROE). Although higher profits from a stock doesn't necessarily equate to better performance in the future, it does mean that the companies in the portfolio have a higher tendency to grow.

From figure 3 (ii), we see that Annual Return Rate from a stock portfolio decrease with increase the Market Value of a company. This is evident from the fact that investments in smaller companies with higher risks tend to bring in higher Annual Return rates. This is how most venture capital funds in the tech industry invest in start-ups which tend to give the high Annual Return Rates in comparison to well established large companies which maybe less riskier to invest in but don't attain high return rates. From figure 3 (v), we also observe that there is clear decreasing trend in Total Risk with involved with increase in market value of a company. This is evident with our initial hypothesis because companies with large market-caps would be less risky investments.

From figure 3 (iv) we observe that systematic risk is not affected by the S/P ratio (sales to price ratio). This doesn't align with our initial hypothesis, because higher sales had a higher likelihood of being under valued.

Variable importance plot

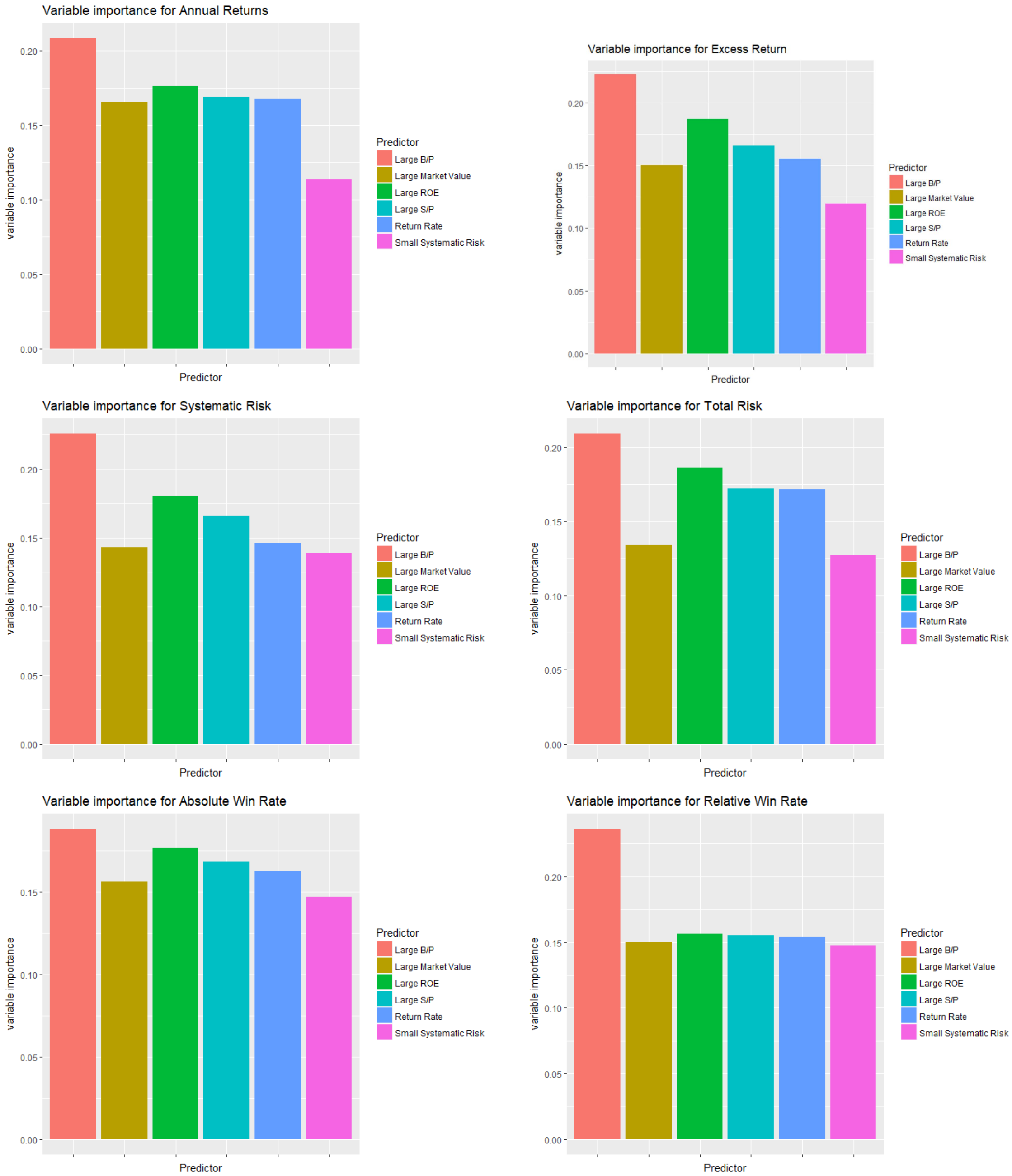

Figure 2: Variable Importance for all output variables with BART

Partial dependence plots

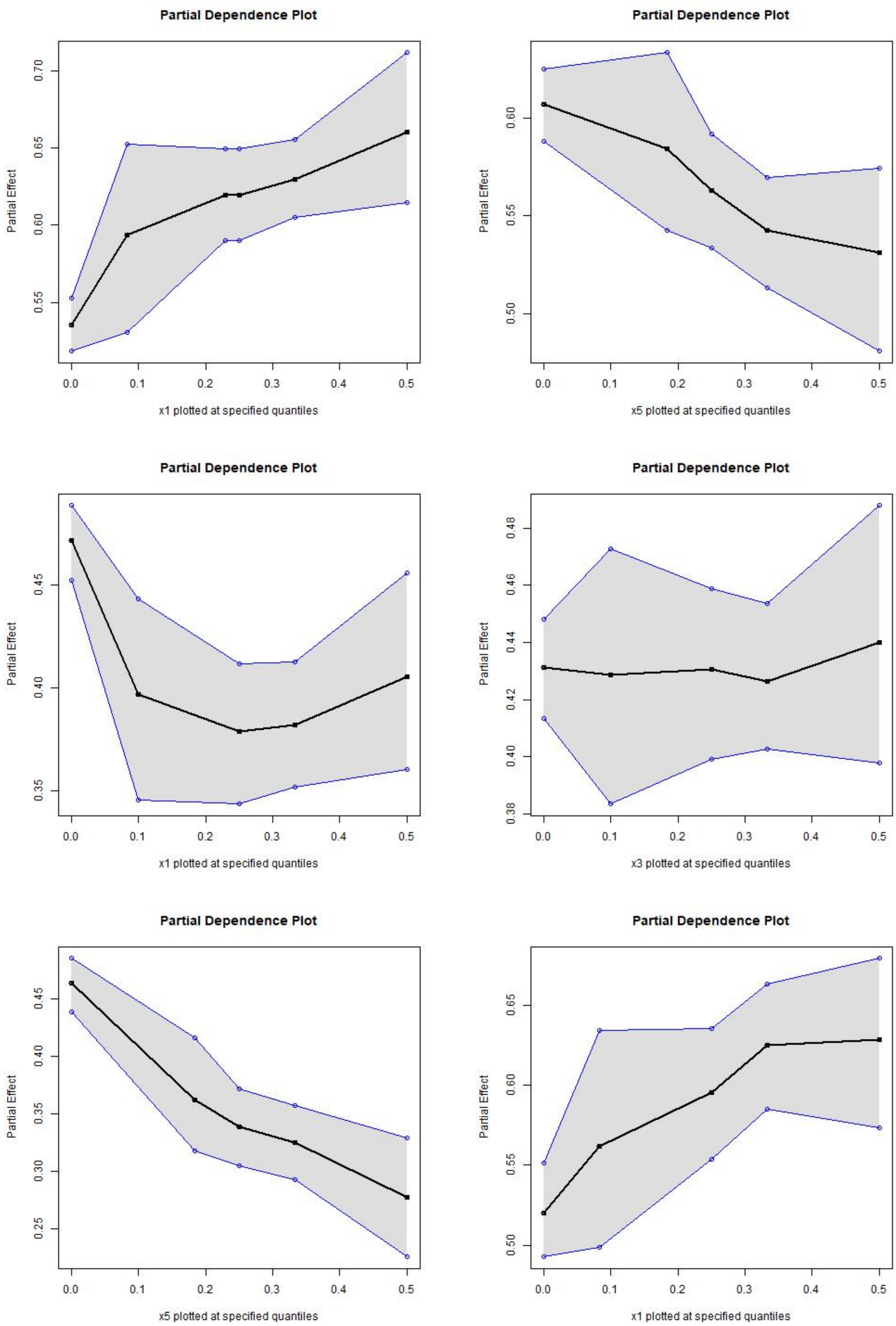

Figure 3: Important partial dependence plots (i)Annual returns v/s B/P ratio, (ii) Annual returns v/s Market Value (iii) Systematic Risk v/s B/P ratio (iv) Systematic Risk v/s S/P ratio (v) Total risk v/s Market value (vi) Absolute win rate v/s B/P ratio

Fitted v/s Actual values

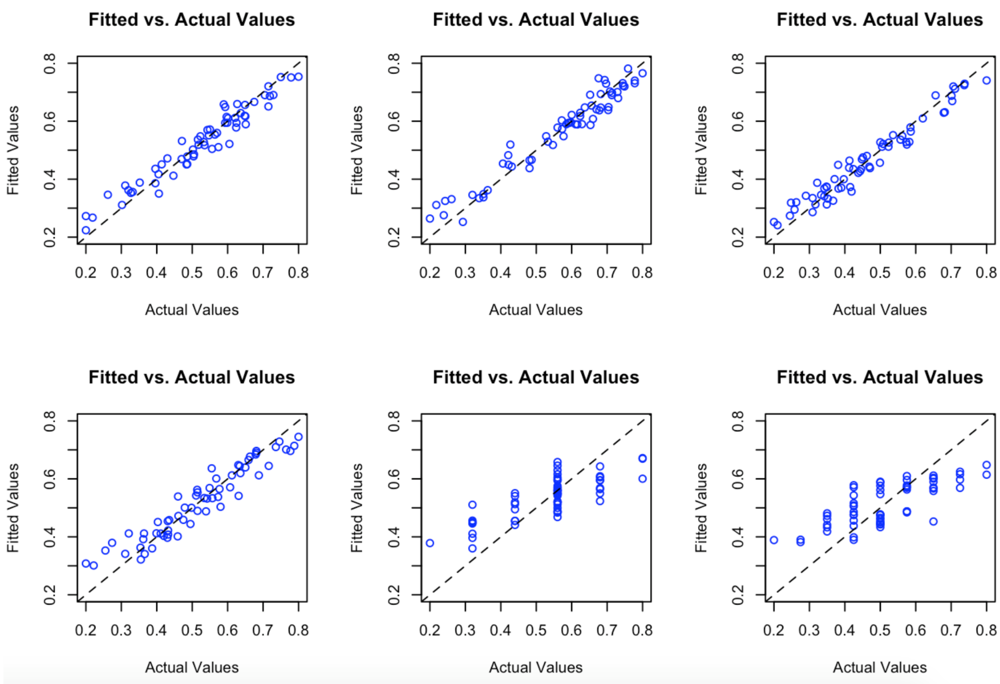

Figure 4: Actual v/s Fitted value plots for all i) Annual Returns ii) Excess Returns iii) Systematic Risk iv) Total Risk v) Absolute Win Rate vi) Relative Win Rate in the All Period Dataset

In figure 4, the actual v/s fitted values of all the response variables against the actual values have been plotted. We observe from this figure that BART gives a fairly good fit for Annual Return, Excess Return, Systematic Risk and Total Risk. Although, the fits for Absolute and Relative win rates are not very good. We find that neural networks give better fit than BART for these response variables. We still decide to go for BART because of the interpretibility of the model.

6 Conclusion

B/P ratio is the most important predictor for every portfolio performance parameter (risk, return and win rate) as is evident from the variable importance (Figure 2) and partial dependence plots (figure 3).For a porfolio manager aiming for higher annual returns, it is best to invest in companies with smaller market-caps. These investments will be riskier but the data shows that portfolios that favor such companies tend to perform well.

The BART model gives the best predictions amongst all the models we tried and test, but it doesn't predict Absolute Win Rate and Relative Win Rate too well even though it outperforms all other models except Neural Nets in predicting Win Rates. Most of the outputs obtained from the partial dependence plots of the BART model give inferences that align with our initial hypothesis of the trends in portfolio performance parameters with the respective input variables.

With neural networks, we think that hyper-parametric tuning good improve predictive model performance. Also, sensitivity analysis can be performed with neural networks to make them more interpretable.

References

[1] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[2] Leo Breiman and Jerome H Friedman. Predicting multivariate responses in multiple linear regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 59(1):3–54, 1997.

[3] Hugh A Chipman, Edward I George, Robert E McCulloch, et al. Bart: Bayesian additive regression trees. *The Annals of Applied Statistics*, 4(1):266–298, 2010.

[4] Jerome H Friedman. Multivariate adaptive regression splines. *The annals of statistics*, pages 1–67, 1991.

[5] Trevor Hastie and Robert Tibshirani. Generalized additive models: some applications. *Journal of the American Statistical Association*, 82(398):371–386, 1987.

[6] Yi-Cheng Liu and I-Cheng Yeh. Using mixture design and neural networks to build stock selection decision support systems. *Neural Computing and Applications*, 28(3):521–535, 2017.

[7] Peter McCullagh. Generalized linear models. *European Journal of Operational Research*, 16(3):285–292, 1984.

[8] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[9] John Ashworth Nelder and R Jacob Baker. *Generalized linear models*. Wiley Online Library, 1972.

[10] Amol Pande, Liang Li, Jeevanantham Rajeswaran, John Ehrlinger, Udaya B Kogalur, Eugene H Blackstone, and Hemant Ishwaran. Boosted multivariate trees for longitudinal data. *Machine learning*, 106(2):277–305, 2017.

[11] I-Cheng Yeh and Wei-Lun Cheng. First and second order sensitivity analysis of mlp. *Neurocomputing*, 73(10-12):2225–2233, 2010.