Applying ML to the All Weather Portfolio

Steve Donahue

An All Weather Portfolio is a simple model for diversifying a retirement portfolio, which captures some of the gains of the stock market, while insulating the portfolio from severe losses. This is accomplished by investing in Stock Indices, Intermediate Term Bonds, Long Term Bonds, Gold, and Broad Basket Commodities in the respective ratios (30 / 40 / 15 / 7.5 / 7.5), and rebalancing the whole portfolio periodically. In a previous data exploration and simulation exercise, one of the stand out, top performing All Weather Portfolios was composed of the following respective funds: VIGRX, IEF, WHOSX, SGDDX, DBC, which had a cumulative annual growth rate of 10% over the past 11 years. It's worst performing year also outperformed similar portfolios, losing only 9.3% where other portfolios lost 15% or even 20%.

If the hypothesis of the All Weather Portfolio is correct, there are relationships between these different investment sectors which compensate for dips in the stock market and shield the portfolio from major losses. It is possible that these relationships are reactive to the lowering of index value, but this project will apply machine learning to determine if the relationships between these sectors might be used to predict upcoming losses in the stock market.

Methodology

The data to be trained and predicted will be the anticipated weekly change in VIGRX fund value. Previous work has been published predicting increases or decreases in a single company's stock, but not without much precision. In keeping with the philosophy of the All Weather Portfolio, this Machine Learning project will only focus on losses which might be seen as significant (2% or more in one week). Such losses only account for about 14% of the data, but will likely be easier to predict than smaller losses. When the models are trained, all weeks with losses equal to or worse than 2% and all weeks which do not experience those incremental changes will be flagged with 0 and 1 respectively. Additionally, Naive Resampling, SMOTE and ADASYN methods will be performed to develop training sets with observations appropriate for a 2% loss.

The features will include average changes in each of the aforementioned sectors, using funds from the sample space identified in the project which precedes this one: Data Science and Retirement Planning, Resampling the All Weather Portfolio. Additional features will be trailing averages for both the VIGRX value, and trailing averages from each of the identified sectors. Durations of 2, 3, 6, 9, 12, and 15 weeks were arbitrarily chosen for these trailing averages.

Finally, the algorithms trained on those features will be tested using the past 34 weeks of market information (from June 1st, 2018 to Feb 1st, 2019), in which 9 weeks experienced significant VIGRX losses of at least 2% value. While this might not be enough data to statistically assess the quality of the trained algorithms, it will provide some helpful context for investors who would like to use this tool when making decisions.

Algorithms and Results

The full results from the described method are appended at the end of this report. Below are the results which simultaneously had the best results in the train/test/split evaluation while also maintaining at least a .40 F1 score for identifying market downturns in the last 34 weeks. The following results are the F1 scores for predicting a significant loss of 2% or more / not predicting a significant loss of 2% or more.

	Train/Test/Split	Last 34 Weeks
LogReg, untuned, with SMOTE:	.59/.64	.42/.68
RandomForest, untuned, with SMOTE:	.84/.83	.42/.78
RandomForest, tuned, with SMOTE:	.65/.64	.48/.70
LogReg, untuned, with ADASYN:	.63/.66	.42/.68
RandomForest, tuned, with ADASYN:	.63/.66	.48/.70

Next Steps

This project shows that aggregate information across multiple sectors of the NYSE can be used to predict losses of 2% or more in at least one stock index (VIGRX). Within this project's context, error analysis should be performed for the above algorithms to determine how often a mischaracterized loss of 2% or more is actually a loss of some other amount. If money is shifted from an index fund because of such an error, it would still be useful for avoiding losses. It might also be easier to apply ML to other stock indices, and while they may not grow as fast as VIGRX, the ability to avoid losses might make them more appealing. Different trailing average lengths may be used as features. Finally, other data sources might be included to provide more features, including Federal economic data and statistics.

Summary of Findings

Classifier	Imbalanced-Learn Adjustment	F Stats on Train/Test/Split July 2007 - May 2018	F Stats on June 2018 - Feb 2019
SVC	none	.00/.93	.00/.85
LogReg Classifier	none	.00/.93	.00/.85
Gradient Boost	none	.17/.93	.00/.85
Decision Tree	none	.23/.91	.44/.80 untuned 0.33/0.86 tuned
Random Forest	none	.21/.92	.20/.86 untuned .00/.85 tuned
SVC	SMOTE	.46/.68 untuned .91/.92 tuned	.36/.70 untuned .00/.85 tuned
LogReg Classifier	SMOTE	.59/.64 untuned .64/.63 tuned	.42/.68 untuned .11/.68 tuned
Gradient Boost	SMOTE	.84/.84 untuned .82/.83 tuned	.29/.68 untuned .32/.73 tuned
Decision Tree	SMOTE	.81/.79 untuned .82/.81 tuned	.21/.69 untuned .13/.75 tuned
Random Forest	SMOTE	.84/.83 untuned .65/.64 tuned	.42/.78 untuned .43/.71 tuned
SVC	ADASYN	.33/.67 untuned .93/.94 tuned	.42/.68 untuned .00/.85 tuned
LogReg Classifier	ADASYN	.63/.66 untuned .67/.67 tuned	.40/.65 untuned .18/.65 tuned
Gradient Boost	ADASYN	.86/.83 untuned .84/.83 tuned	.30/.71 untuned .20/.67 tuned
Decision Tree	ADASYN	.83/.79 untuned .75/.75 tuned	.10/.63 untuned .32/.73 tuned
Random Forest	ADASYN	.88/.86 untuned .63/.66 tuned	.35/.78 untuned .48/.70 tuned

Note that all five classifiers were also run on a sample set subject to Naive Resampling, and while the train-test-split results were consistently in the low 80’s to low 90’s, they unanimously failed the recent data exercise with F1 scores of .00/.85