Lock, Stock, and Barrel

Stock Portfolio Analysis

*Capstone Final Report*

ABSTRACT

The risk return tradeoff is known to improve with portfolio diversification. Diversification is beneficial because covariance between stocks is generally smaller than the variance of individual stocks, and it lowers risk. The expected return of a portfolio is a weighted average of the returns. These tenets of portfolio theory form the basis of the data science project to optimize portfolio structures while aiming to improve the risk return tradeoff. Kernel PCA analysis was used to optimize the variance of all 500 stocks of the S&P 500. The Euclidian distance was calculated between all 500 stocks on the first two principle components in order to maximize the variance with a selected stock. Based on the Kernel PCA analysis the selected stocks are optimized using the SciPy optimization function to select the optimal portfolio based on the return and volatility ratio. The optimization function selects the remaining preferred stocks and assigns the weight for the optimal portfolio. The model was validated using 12 fold cross validation with a resulting prediction accuracy of 0.98.