FE 800 Final Report

Network Analysis and Clustering for Better Portfolios

Group member:
Qinglin Yang
Zhaoyang Zhang

Academic Advisor:
Dr. Khaldoun Khashanah
Industry Advisor:
Cristian Homescu

May 16, 2021

Abstract

This project is aimed to construct a good portfolio by using clustering analysis and network analysis. The main methodologies we used in this project are centrality, minimum spanning trees, hierarchical risk parity and neutral network. We used the MST method to construct a portfolio. Next, we used centrality to measure the degree of stock concentration, and then to assign different weights to the portfolio. On the basis of MST, we used unsupervised learning machine learning method HRP to construct a new investment portfolio. By comparing the portfolios, we constructed with the benchmark portfolio, we finally get the better portfolio.

Key words: Neutral network, Equal weighted portfolio, ARIMA, Minimum Spanning Tree (MST), Hierarchical Risk Parity(HRP), Centrality, Machine Learning

1 Introduction

Markowitz (1952)[1] laid the foundation of modern portfolio theory. In this static framework, investors optimally allocate their wealth across a set of assets considering only the first and second moment of the returns' distribution. Despite the profound changes derived from this publication, the out-of-sample performance of Markowitz's prescriptions is not as promising as expected. The poor performance of Markowitz's rule stems from the large estimation errors on the vector of expected returns (Merton, 1980)[2] leading to the well-documented error-maximizing property discussed by Michaud and Michaud (2008)[3]. The magnitude of this problem is evident when we acknowledge the modest improvements achieved by those models specifically designed to tackle the estimation risk (DeMiguel et al., 2009)[4].

Recently, researchers from different fields have characterized financial markets as networks in which securities correspond to the nodes and the links relate to the correlation of returns (Mantegna, 1999)[5]. In spite of the novel and interesting insights obtained from these network-related papers, most of their results are fundamentally descriptive and lack concrete applications in the portfolio selection process.

Our data, methods and results part is divided into 2 parts. In the first part, it showed 2 analyses. At the beginning we tried to use a neural network to help optimize portfolio and forecast stock prices. We use covariance and stock returns to construct an efficient frontier, which can help us optimize our portfolio. Finally, we found it may not be the best analysis that we want. Therefore, we switched our mind to the second analysis and to show the stock change in the network. We get the idea that we could try to use network analysis to build a long-only portfolio [6]. In the data and methodology Part 2, we used S&P Index as benchmark and built the portfolio using Eigenvector and Betweenness centrality. Next, we used MST and centrality methods to distribute weights. Comparing with different methods algorithm, we think the betweenness centrality portfolio' performance is not so good, but all portfolios we constructed are better than S&P 100 index. In conclusion, we recommend that for risk seekers, investors should buy Eigenvector Centrality Portfolio. For risk averters, investors should buy an HRP portfolio.

2 Literature Review

2.1. Network-based Asset Allocation Strategies[7]

In this paper, they construct financial networks, in which nodes are represented by assets and where edges are based on long-run correlations. They use minimum spanning tree and centrality measures (betweenness, eigenvalue centrality). To improve risk-return characteristics of well-known return maximization and risk minimization benchmark portfolios, they propose simple adjustments to portfolio selection strategies that utilize centralization measures from financial networks.

To describe the topology of the financial networks, they use several measures of the relative importance of assets. They focus mainly on centrality measures, namely, betweenness and eigenvalue centrality. Each centrality represents a different measure, thus resulting in a numerical nodal attribute describing the importance of a node relative to others. The simplest centrality measure, which we do not use to construct our portfolios, is the degree centrality, which assigns each vertex to the number of incident edges; thus, the higher the degree of centrality, the more interconnection a vertex has with remaining vertices in a network.

2.2. Hierarchical Structure in Financial Markets[5]

A spanning tree is an acyclic connected subgraph containing all vertices (a graph with no circles) with a path connecting any two vertices. Requirement of a minimal such spanning tree refers to the values of edge weights. To satisfy the conditions of edge weight non-negativity, which allows for their interpretation as distances, Mantegna (1999) proposed a nonlinear decreasing transformation of correlations to be used for weights ($d_{i,j} = 2(1-\rho_{i,j})$). As the transform is decreasing, higher correlations translate into smaller distances. MST is thus a spanning tree with a minimum sum of weights of retained edges. This supports the notion of keeping only the most important edges in a graph intact. In a network of M vertices, an MST retains precisely M – 1 edges. The MST is extracted from the complete graph using Kruskal's (1956) algorithm.

2.3. The Hierarchical Equal Risk Contribution Portfolio[8]

The author built upon the fundamental notion of hierarchy, the “Hierar- chical Risk Parity” (HRP) and the “Hierarchical Clustering based As- set Allocation” (HCAA), the Hierarchical Equal Risk Contribution Portfolio (HERC) aims at diversifying capital allocation and risk al- location. HERC merges and enhances the machine learning approach of HCAA and the Top-Down recursive bisection of HRP.
HRP starts by reorganizing the covariance matrix to place similar investments together. Then, it employs an inverse-variance weighting allocation based on the number of assets with no further use of the clustering.
Briefly, the principle is to retain the correlations that really matter and once the assets are hierarchically clustered, a capital allocation is estimated.

3 Data and Methodology (Part 1)

3.1. Data for Equal Weighted

We got the data that are daily stock prices of 4 financial firms from 2005 to 2020 from Yahoo Finance. The financial firms are Morgan Stanley known as MS, JP Morgan known as JPM, Citigroup known as C, and Goldman Sachs known as GS.
We made some analysis of the data. Volatility, daily returns and exposure of risks are the most important factors to show how the market influences the stock prices visually. We can tell from Figure 1, Figure 2 and Figure 3 that there are some obvious fluctuations around 2009 and the beginning of 2020. It’s because the market was suffering from the economic crisis from 2008 and the pandemic in 2020. Therefore, we tried to form a portfolio to forecast the future stock prices as precisely as possible.

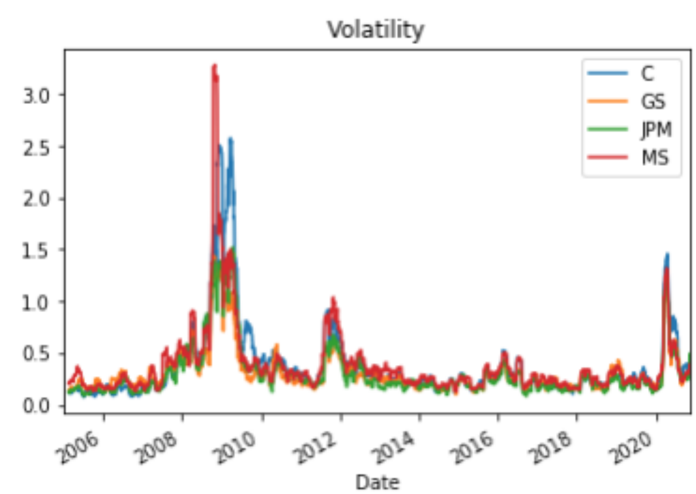

Figure 1 - Volatility of 4 stocks. There is a great peak around 2009 and a small peak in early 2020.

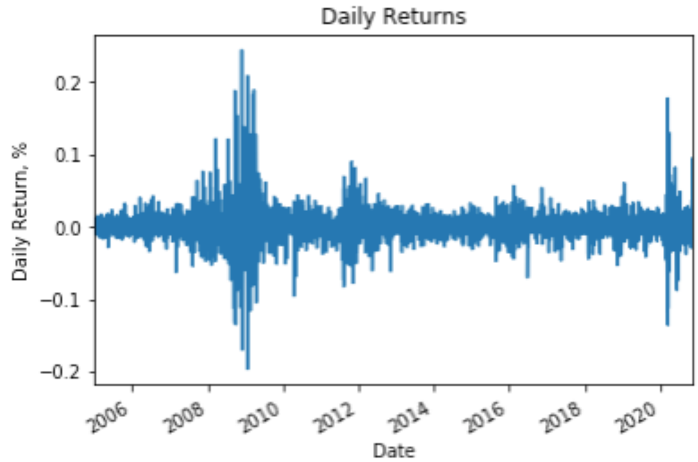

Figure 2 - Daily Returns of Stocks. There are 2 great fluctuations during 2009 and earlier of 2020.

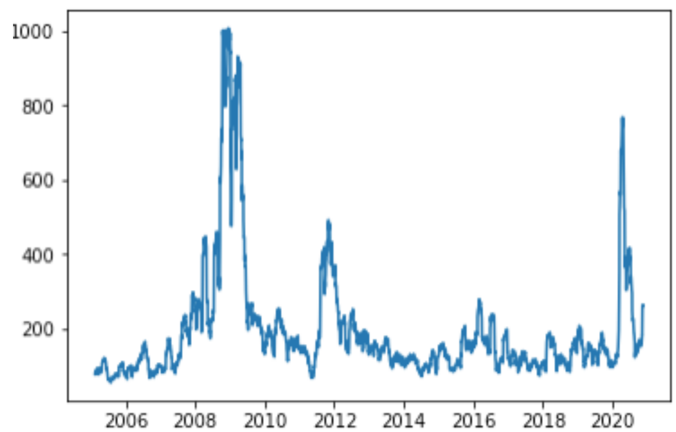

Figure 3 - Exposure of Risk. There are 2 great peaks around 2009 and 2020.

3.2. Methods and Results for Equal Weighted

We used neural network and ARIMA to predict stock prices and the neural network read asset prices as input and forecast the stock prices. Mean Squared error is to estimate the performance of the network method but the number is larger than 200 which means it’s not accurate and precise and far away from reality. Compared with neural network, the ARIMA model is much closer to the actual world trend. The mean squared error is as low as 5.07. we got the prediction illustrated as Figure 4 below. We can tell from the figure that the ARIMA follows the path of the actual line in each 4 companies but the neural network line is lower than the actual line except GS. The situation means the ARIMA model is better than neural network model when we predict the stock prices.

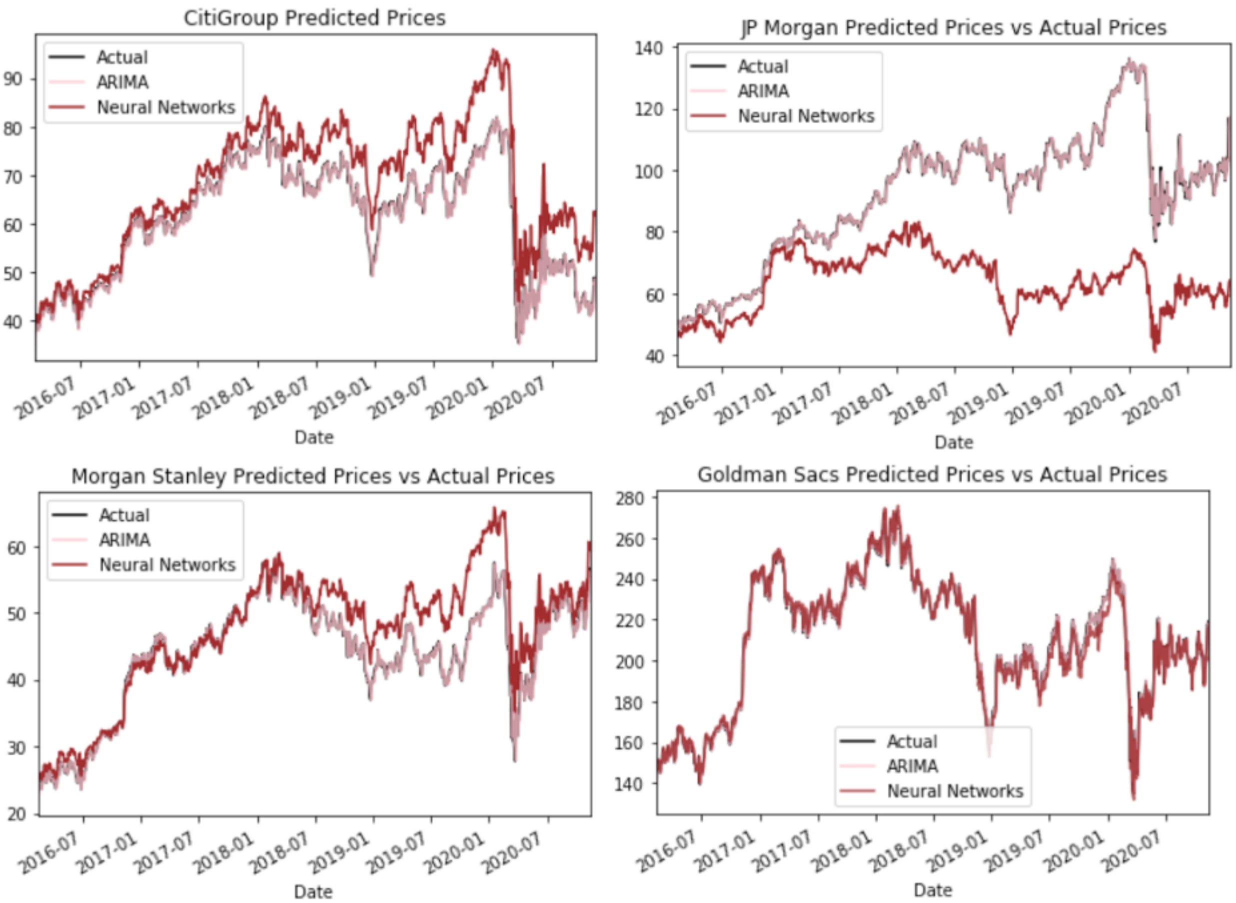

Figure 4: The Predictions Used Neural Network vs Actual vs ARIMA.

We get the information from Table 2 that the covariances of three situations have less difference. The two methods are close to the actual.

Table 1: Efficient Covariance Matrix of 3 Situation: Actual, Neural Network and ARIMA

Efficient Covariance Matrix Actual

	C	GS	JPM	MS
C	0.279985	0.133172	0.154679	0.178500
GS	0.133172	0.137316	0.106461	0.151068
JPM	0.154679	0.106461	0.152777	0.133869
MS	0.178500	0.151068	0.133869	0.265975

Efficient Covariance Matrix ARIMA

	C	GS	JPM	MS
C	0.156121	0.097755	0.093571	0.114804
GS	0.097755	0.089722	0.067648	0.088949
JPM	0.093571	0.067648	0.075437	0.078632
MS	0.114804	0.088949	0.078632	0.119353

Efficient Covariance Matrix NN

	C	GS	JPM	MS
C	0.123882	0.089083	0.087993	0.101259
GS	0.089083	0.102175	0.098562	0.085552
JPM	0.087993	0.098562	0.099983	0.083509
MS	0.101259	0.085552	0.083509	0.098664

The two models assign the same weight which are 0.6 in GS and 0.4 in JPM respectively. The efficient frontier of ARIMA and actual are very close but the neural network is lower than other two. ARIMA relative accuracy than neural networks predict the stock prices.

Table 1: Weight Size of Each Company in Three Different Situation.

Asset	Weights Actual	Weights ARIMA	Weights NN
C	0.0	0.0	0.0
GS	0.6	0.6	0.0
JPM	0.4	0.4	0.0
MS	0.0	0.0	0.0

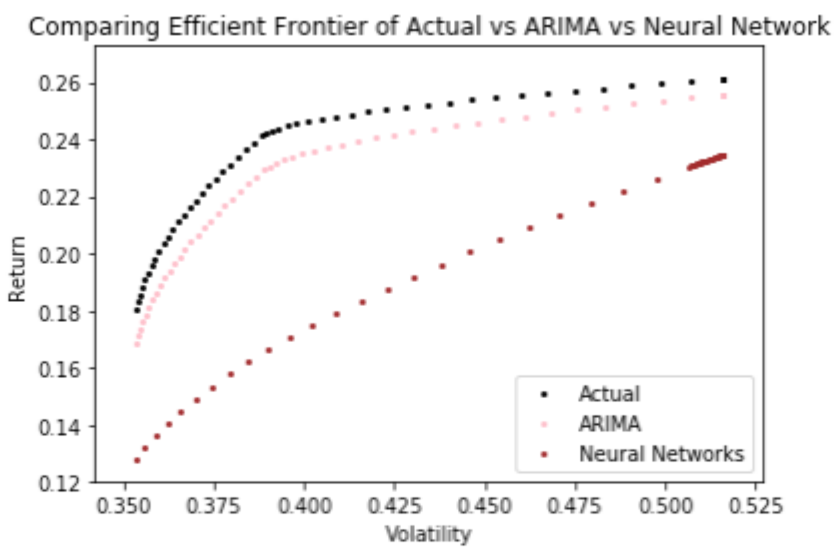

Figure 5: Efficient Frontier of Actual, ARIMA and Neural Network.

In summary, the first analysis gave us a direct visualization that the JMP and GS are the two most stable stocks compared with the other two.

Now we change our mind to another network analysis.

3.3. Data for Network Analysis

We got the historical daily stock data from Yahoo Finance from January 2, 2015 to the end of 2017. The Table 1 shows the screenshot of the data frame that includes more than 4,600 stocks.

Table 3: Head of the Daily Stock Data Frame.

Symbol	Name	Exchange	Category	Sector
^GSPC	S&P 500	SNP	Index	NaN
^DJI	Dow Jones Industrial Average	DJI	Index	NaN
AAPL	Apple Inc.	NMS	Electronic Equipment	Consumer Goods
BAC	Bank of America Corporation	NYQ	Money Center Banks	Financial
AMZN	Amazon.com, Inc.	NMS	Catalog & Mail Order Houses	Services

However, we only needed the adjusted closing prices, so we adjusted the data and the Table 4 below showed the partial adjusted closing prices stocks since 2015.

Table 4: Head of Adjusted Closing Prices of Stocks.

Date	AAPL	MSFT	AMZN	^GSPC	^DJI
2015-01-02	103.074181	43.134731	308.519989	2058.199951	17832.99023
2015-01-05	100.170410	42.738068	302.190002	2020.579956	17501.65039
2015-01-06	100.179840	42.110783	295.290009	2002.609985	17371.64063
2015-01-07	101.584595	42.645817	298.420013	2025.900024	17584.51953
2015-01-08	105.487686	43.900375	300.459991	2062.139893	17907.86914

3.4. Method and Result for Network Analysis

First, we constructed the network. We set each single node as a stock in the network. We used 50% as the threshold because the edges between nodes exist when the correlation of log returns exceeds the threshold. If less edges appear, the computation decreases in the network.

We set the training sets that include trading data with year 2016 and year 2017 and set test sets that include the year 2015. Next, we calculated the log returns. We made the correlation matrix with training data which will be leveraged in the construction of network.
We illustrated the network shown by Figure 8 below. On the left of the figure, there is a disjoint clustering in basic materials companies sector. Those companies provide general services with gold and other precious metals. Those companies with high correlated returns but they have less connection with broader. It means returns of those firms have no correlation with other equities in the network.

On the top left of the network, there is the healthcare related clustering combined by some biotech firms that have weaker connection with center mass, although the connection exists via few bridge equities. We can also see on the right side of the network, the basic materials clustering means non-utility energy companies covering the value chain from extraction to marketing and sales.

Otherwise, we moved our focus to the top right. Some utilities like water and electric form financial utility clustering and consumer goods firms compose a clustering in the right also. On the other hand, there is a large combination and merge of equities of different industries which contains primary firms in the financial sector in the center.

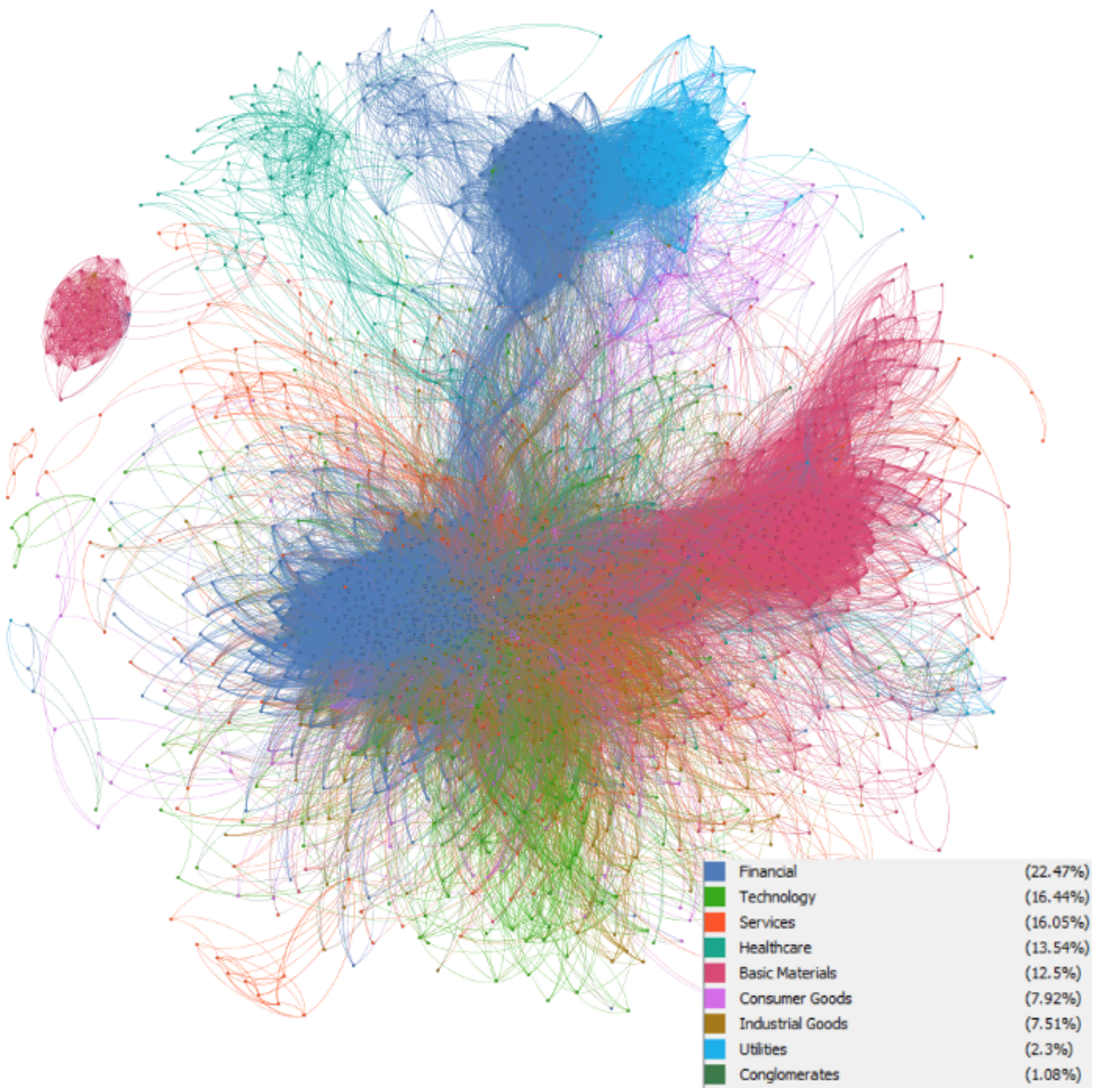

Figure 8: Network by Sector, which is the color-coding each node according to its sector.

There are three measures that we could use now and they are degree centrality, betweenness centrality and closeness centrality. The last two are complicated. Those three measures are correlated with each other. The easiest way to help us generate nodes is to use degree centrality because there are many different centrality methods that may not help us to recognize the crucial nodes. The centrality measure can build a connection between the number of edges and nodes that are important to us. The length of the shortest path between a node and other nodes is quantitative by closeness measure in the network. Meanwhile, the number of times playing a role as bridges is quantitative by betweenness in the network.

Our network method used the centrality method because we think optimize the weights assigned to each centrality measure in order to maximize the return of the resulting portfolio. Since centrality method identified a portfolio that marginally outperformed the S&P 500 over a three-year period using only one year of training data, though no additional detail is provided in the paper.

We used the zero centrality method to confirm the minimal correlation and then we will know what the historical performance is, which historical performance can be applied on a regular basis and a reference of future performance.

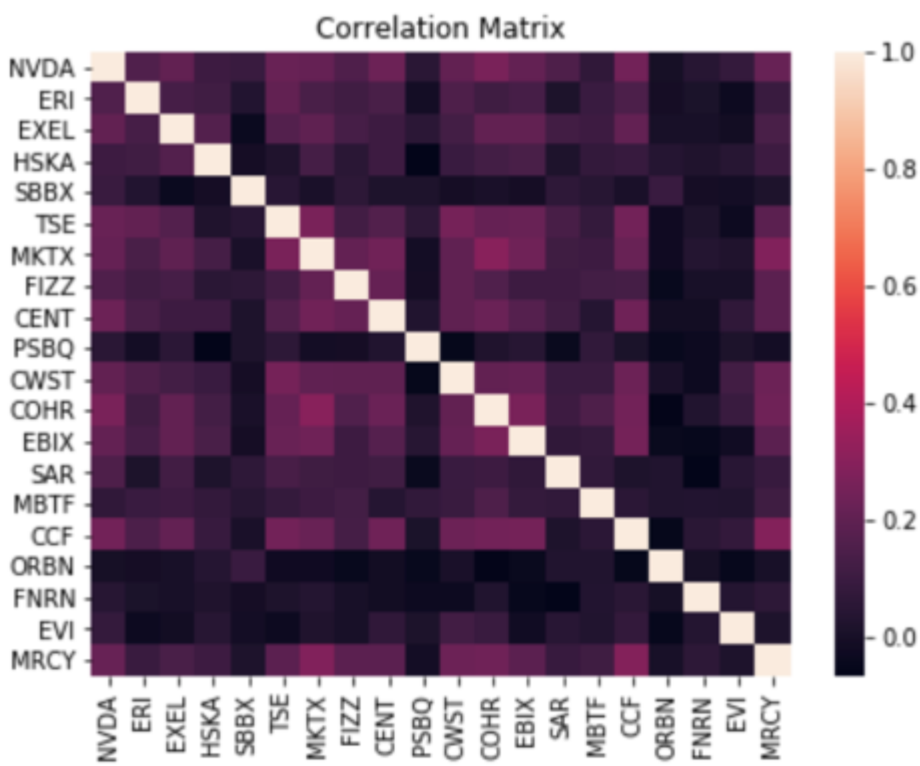

Figure 7: Correlation Matrix. A heatmap of the correlation and covariance matrices for selected stocks.

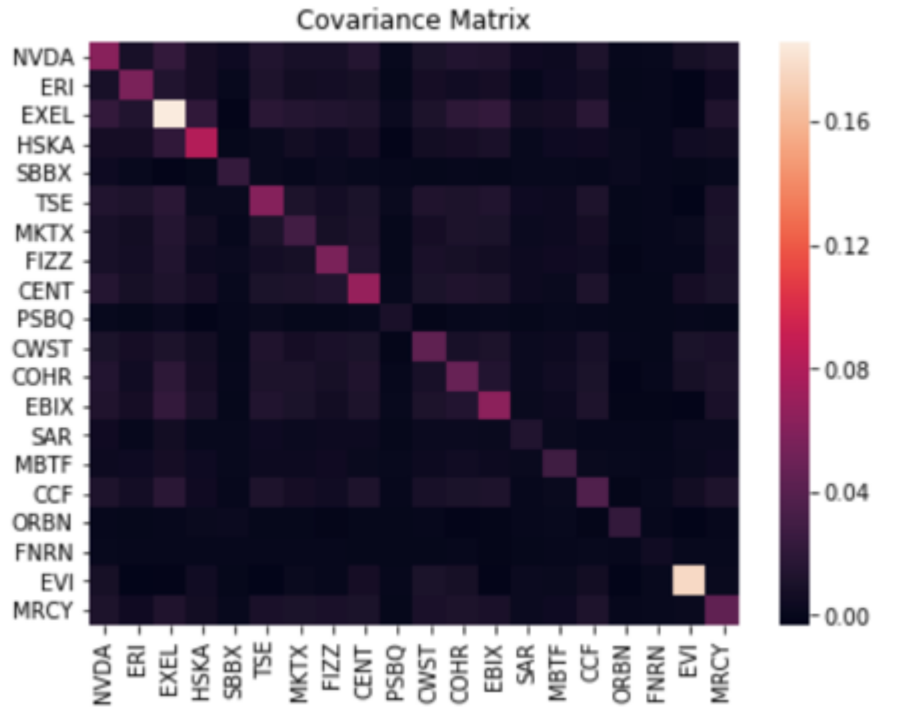

Figure 8: Covariance Matrix. A heatmap of the correlation and covariance matrices for selected stocks.

The efficient frontier showed as Figure 9. The sharpe ratio of the optimal portfolio is higher than 4.8 when there is no other research into the methodology.

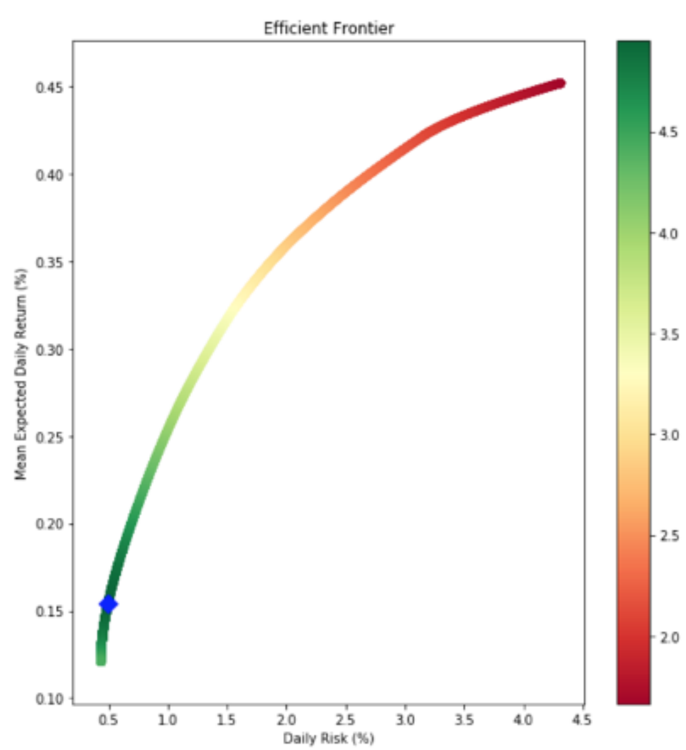

Figure 9: Efficient Frontier. It's impressive because the Sharpe ratio is much higher.

The allocation plot (Figure 10) here illustrated that the portfolio that we constructed is well diversified across 20 stocks.

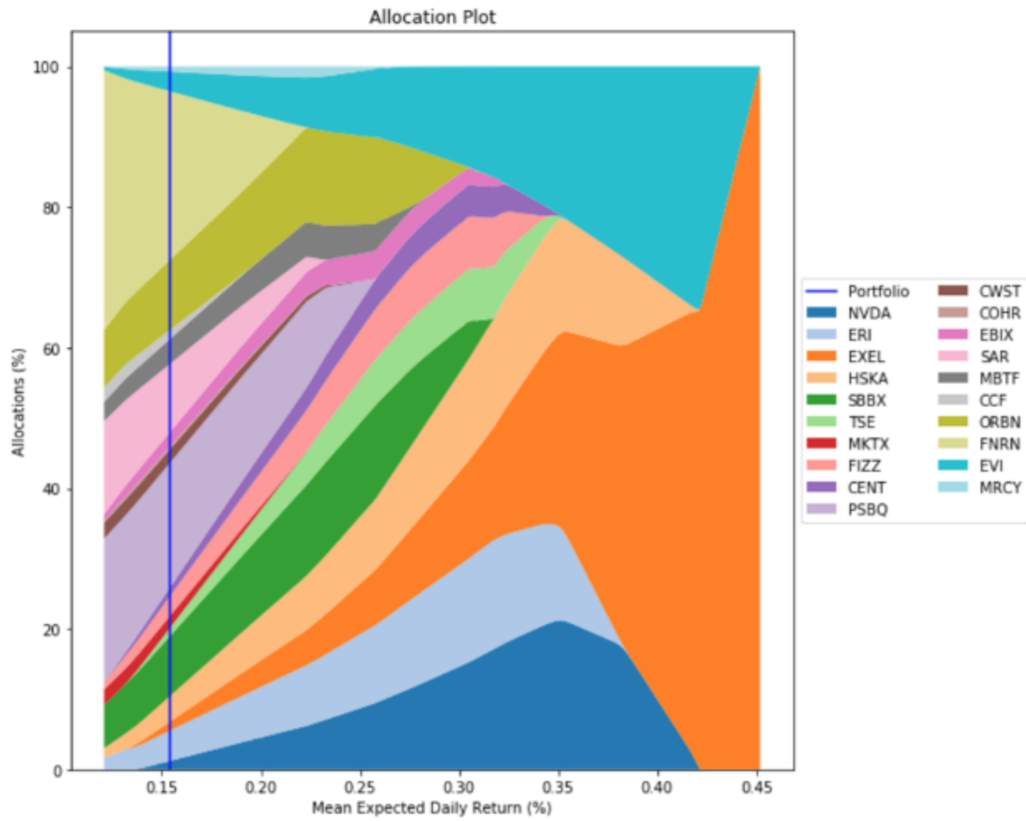

Figure 10: Allocation Plot

The time underperforms both indices two months out of the 12 assessed, while we were testing the zero centrality portfolio the market indices. The results are shown in Figure 11.

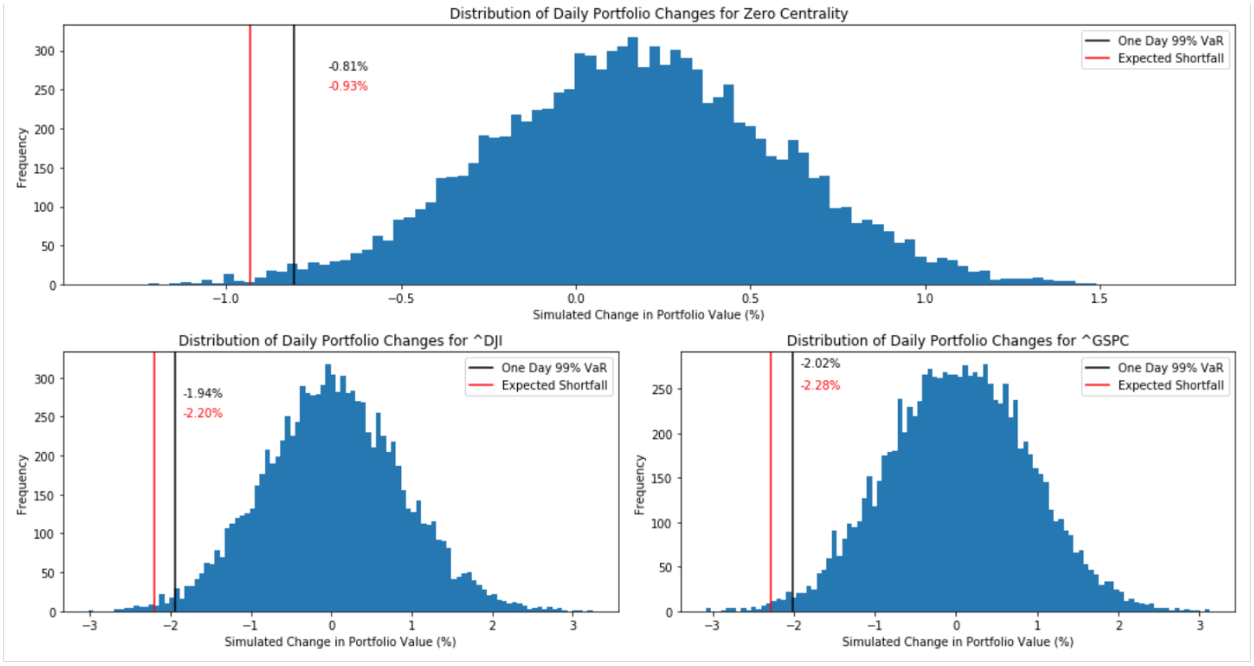

Figure 11: Distribution of Daily Portfolio Changes for Zero Centrality Measure.

The alpha of the portfolio in zero centrality is 15% (Table 5), which is relatively lower than normal we thought over the test period.

Table 5: Zero Centrality Alpha, Beta vs Benchmark

```
Alpha: 15.103
Beta: 0.365
Treynor Ratio: 61.932
```

In summary, after those two attempts, we found the centrality measure gave us direction and it's useful to analyze in the Part 2. We will provide a sufficient explanation of the portfolio and methods. There are some drawbacks that the data in this part choose only stocks in the US exchanges and may not give some good results, but that was a good try for us.

4 Data and Methodology (Part 2)

4.1. Data & Benchmark Portfolio

In this part, we used 100 stock data from S&P 100 Index from 2010-01-01 to 2020-01-01. Like S&P 500, S&P 100 is enough on behalf of the different industries, which could support us to do network analysis and clustering.

We calculated the stock return from the change of stock price. By return, we calculated the correlation and covariance between stocks. Using those data, we built three benchmark portfolios, Sharpe Ratio Portfolio, Global Minimum Variance Portfolio, and Equal Weight Portfolio. We used the S&P Index as a benchmark as well.

4.2. Mutual Information Distance

In order to obtain a distance statistic with certain desirable properties, we will use a Mutual Information Distance instead of Correlation or Covariance:

$$d(X,Y) = 1 - \sqrt{1 - \exp\left(-2I(X,Y)\right)}$$

The distance measure is always in between 0 and 1. If d(X,Y) = 1 implies X contains no information on Y and vice versa. In this case the distance measure takes the maximum value signifying that X and Y are far apart. If d(X, Y) = 0, there exists a perfect relationship between X and Y , or in other words, X and Y determine each other. Therefore, if there is a very close relationship between X and Y the distance measure is expected to be close to 0.

4.3. Minimum Spanning Tree (MST)

The mutual information distance (edge weights between nodes) between each asset (nodes of the network) constitute a fully connected network. For a universe of cardinality n ∈ N, this network has n(n − 1)/2. In our case, n = 100, we have 4950 pairs to think about.

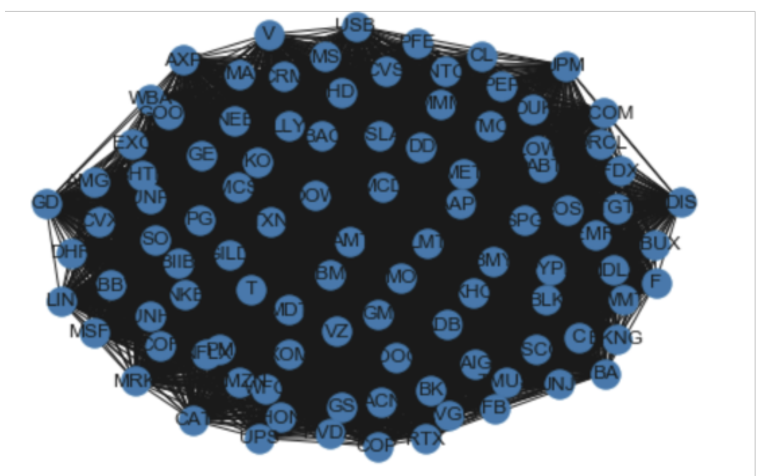

Figure 12 - Relation between Stocks before Using MST

Used MST, we get our network:

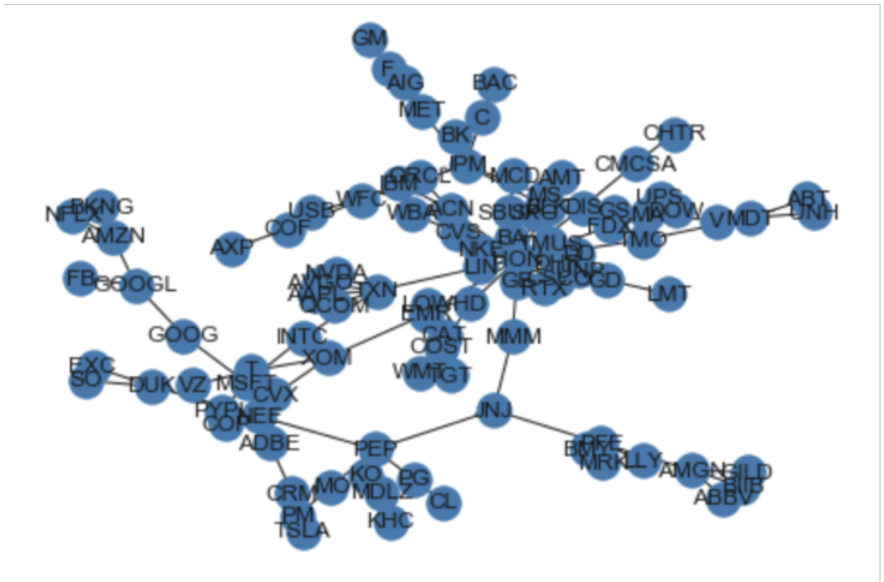

Figure 13 - MST of S&P 100 on 2021-05-10

From the above figure, we could obviously see the relationship between stocks, from 4950 pairs reduces to about 100 pairs.

4.4. Centrality Measures

To describe the topology of the financial networks, we use centrality as measures of the relative importance of assets. Each centrality represents a numerical nodal attribute describing the importance of a node relative to others.

In this paper, we use eigenvector centrality and betweenness centrality as our Measures:

```
features = pd.concat([cov_mtx, eigenvector_centrality_df, betweenness_centrality], axis=1)
features.head()
```

NVDA	PYPL	FB	...	CHTR	WFC	CL	GILD	AIG	KHC	ABBV	COP	eigenvector_centrality	betweenness_centrality
0.452763	0.404449	0.275029	...	0.291050	0.455932	0.337453	0.322238	0.372905	0.158619	0.290747	0.374904	0.008911	0.189033
0.385900	0.308912	0.243799	...	0.263149	0.367324	0.249690	0.240595	0.310077	0.168810	0.182157	0.318933	0.036275	0.000000
0.370495	0.320531	0.328987	...	0.250011	0.351391	0.249534	0.294449	0.325633	0.126120	0.212553	0.275760	0.000095	0.040198
0.408513	0.346554	0.346302	...	0.268284	0.407606	0.290421	0.321001	0.344386	0.161522	0.262278	0.317332	0.000425	0.078953
0.251276	0.195633	0.185731	...	0.171425	0.214311	0.152907	0.204037	0.163717	0.089266	0.158351	0.209536	0.000082	0.000000

Figure 14 - Stocks Centrality

4.5. Hierarchical Risk Parity (HRP)

Hierarchical Risk Parity (HRP) aims at diversifying risk allocation. It operates in three stages:

1. Minimum Spanning Tree (MST): This procedure allows to extract an MST and a hierarchical tree from a correlation coefficient matrix by means of a well-defined algorithm known as Single Linkage clustering algorithm.
2. Quasi-diagonalization: reorganization of the covariance matrix to place similar investments together.
3. Recursive bisection: the matrix diagonalization allows to dis- tribute weights optimally following an inverse-variance allocation between uncorrelated assets.

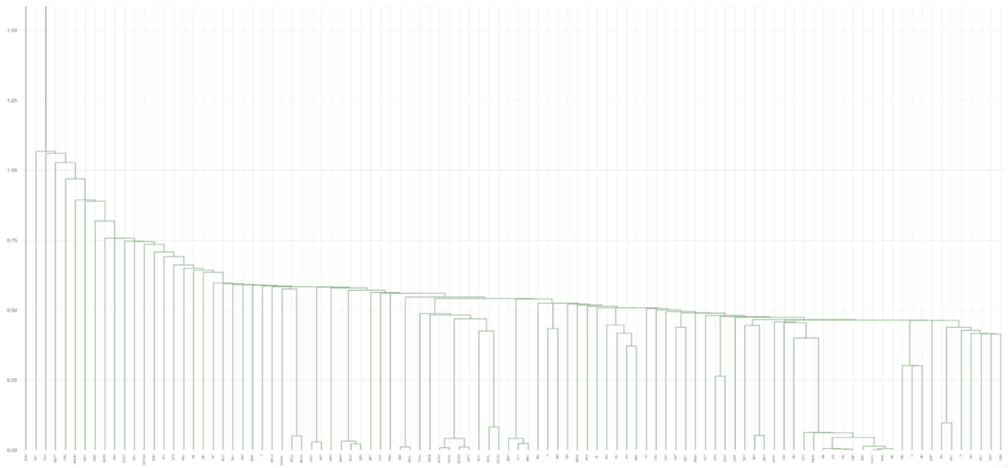

Figure 15 - Hierarchical Risk Parity (HRP) created by S&P 100 stocks

In this paper, we chose 0.6 as our threshold, and then, we get 20 clusters under the HPR.

```
clusters = fcluster(link, t=0.6, criterion='distance')
ticker_clusters = pd.DataFrame([mst_dist.iloc[sch.leaves_list(link)].columns, clusters], index=['Ticker', 'clusters
```

```
print("num of clusters:", ticker_clusters['clusters'].max())
```

```
num of clusters: 20
```

Figure 16 - Cluster Number

4.6. Portfolio Construction

We built 4 portfolios: Eigenvector centrality portfolio, Inverse Eigenvector centrality portfolio, Betweenness centrality portfolio, and HRP portfolio.

Eigenvector centrality portfolio and Betweenness centrality portfolio:
In these two portfolios, we allocate the weight by the value of centrality. If the centrality is bigger, then we put more weights on these stocks.

Inverse Eigenvector centrality portfolio:

In this portfolio, we inverse allocate the weight by the value of centrality. If the centrality is smaller, then we put more weights on these stocks.

HRP portfolio:
In this portfolio, we set the threshold at 0.6, and then get 20 clusters. We allocate 1/20 weights on each cluster. In each cluster, we use just two stocks, which means that if there is only one stock in this cluster, give them 1/20 weight, if three stocks, pick up two stocks and then give them 1/40 weight for each.

4.7. Performance Comparison

Compare with benchmark portfolios, we get this result:

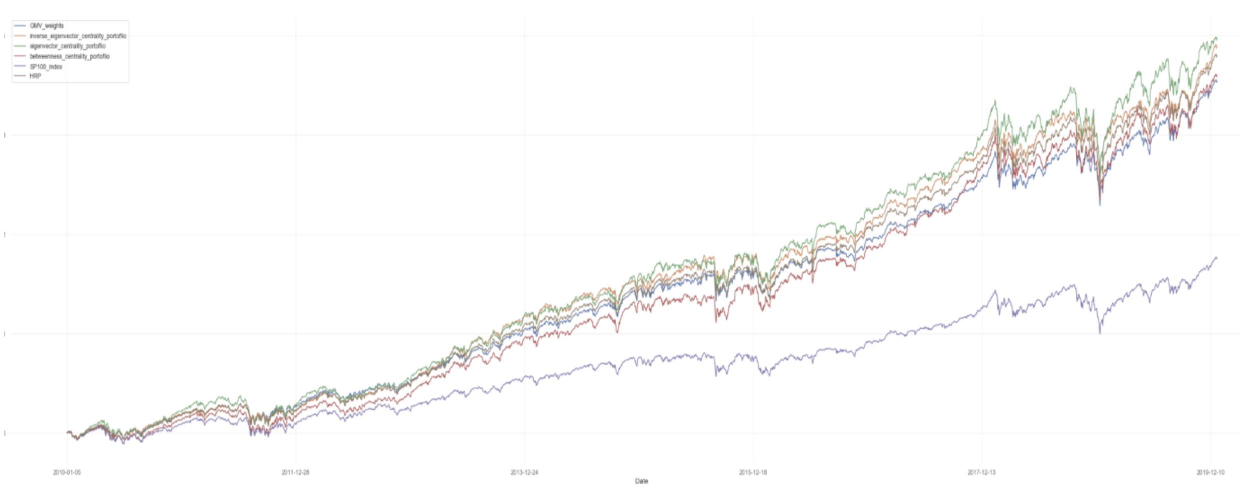

Figure 17 - Portfolios Cumulative Return from 2010 to 2020

From the Figure 17, we could observe that from top to bottom, it is Eigenvector Centrality Portfolio, Inverse Eigenvector Centrality Portfolio, HRP Portfolio, Betweenness Centrality Portfolio, GMV Portfolio respectively. All portfolios are better than S&P 100 Index, which is a very interesting phenomenon, and we will figure it out in the future. The top three are Eigenvector Centrality Portfolio, Inverse Eigenvector Centrality Portfolio, HRP Portfolio, which give us a very good performance.

4.7.1. Eigenvector Centrality Portfolio vs SPY

Then we compare the Inverse Eigenvector Centrality Portfolio and HRP Portfolio with SPY index separately.

Figure 18 - Eigenvector Centrality Portfolio Cumulative Return from 2010 to 2020

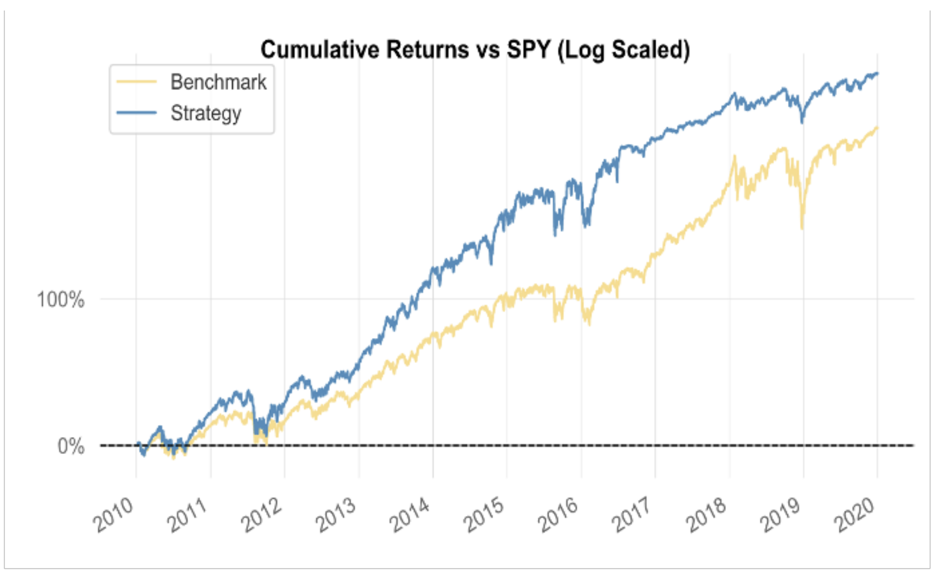

Figure 19 - Eigenvector Centrality Portfolio Cumulative Return (Log Scaled) from 2010 to 2020

Figure 20 - Eigenvector Centrality Portfolio end of ear Return from 2010 to 2020

Table 6 - Eigenvector Centrality Portfolio Key Performance Metrics

Key Performance Metrics

Metric	Strategy	Benchmark
Risk-Free Rate	0.0%	0.0%
Time in Market	100.0%	100.0%
Cumulative Return	397.48%	246.88%
CAGR%	17.42%	13.26%
Sharpe	1.09	0.92
Sortino	1.55	1.29
Sortino/√2	1.09	0.92

Max Drawdown	-22.66%	-19.35%
Longest DD Days	212	277
Volatility (ann.)	16.0%	14.69%
R^2	0.94	0.94
Calmar	0.77	0.69
Skew	-0.37	-0.43
Kurtosis	4.09	4.39
Expected Daily %	0.06%	0.05%
Expected Monthly %	1.35%	1.04%
Expected Yearly %	17.4%	13.24%
Kelly Criterion	8.65%	7.26%
Risk of Ruin	0.0%	0.0%
Daily Value-at-Risk	-1.59%	-1.47%
Expected Shortfall (cVaR)	-1.59%	-1.47%

Table 7 - Eigenvector Centrality Portfolio Worst 10 Drawdowns

Worst 10 Drawdowns

Started	Recovered	Drawdown	Days
2011-07-08	2012-02-01	-22.66%	208
2018-09-24	2019-04-04	-20.10%	192
2010-04-30	2010-11-02	-16.55%	186
2015-06-19	2015-10-22	-11.74%	125
2015-11-23	2016-03-17	-11.58%	115
2012-03-27	2012-09-07	-11.46%	164
2018-01-29	2018-08-29	-10.65%	212
2010-01-12	2010-03-05	-8.63%	52
2014-09-19	2014-10-30	-7.97%	41
2019-05-01	2019-06-20	-7.71%	50

Table 8 - Worst 10 Drawdowns

From above figures and Tables, we could clearly observe that Eigenvector Centrality Portfolio cumulative return is much higher than SPY (397.48% and 246.88%), and the volatility is nearly at the same level (16.0% and 14.69%). We could use Sharpe ratio to measure return against volatility. The Sharpe ratio of Eigenvector Centrality Portfolio is 1.09, and the SPY is just 0.92. In this case, although a little bit higher in volatility, most rational people will buy Eigenvector Centrality Portfolio not SPY Index.

4.7.2. HRP vs SPY

Compare to SPY, HRP is definitely better than SPY in many aspects:

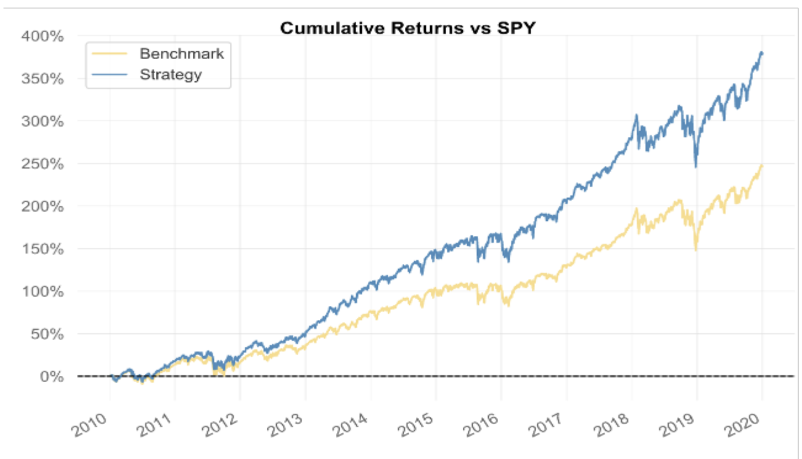

Figure 21 - HRP Portfolio Cumulative Return from 2010 to 2020

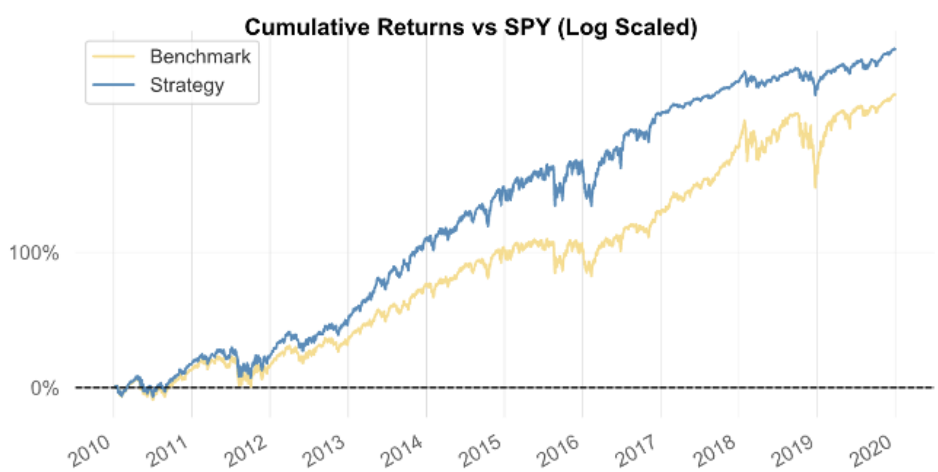

Figure 22 - HRP Portfolio Cumulative Return from 2010 to 2020

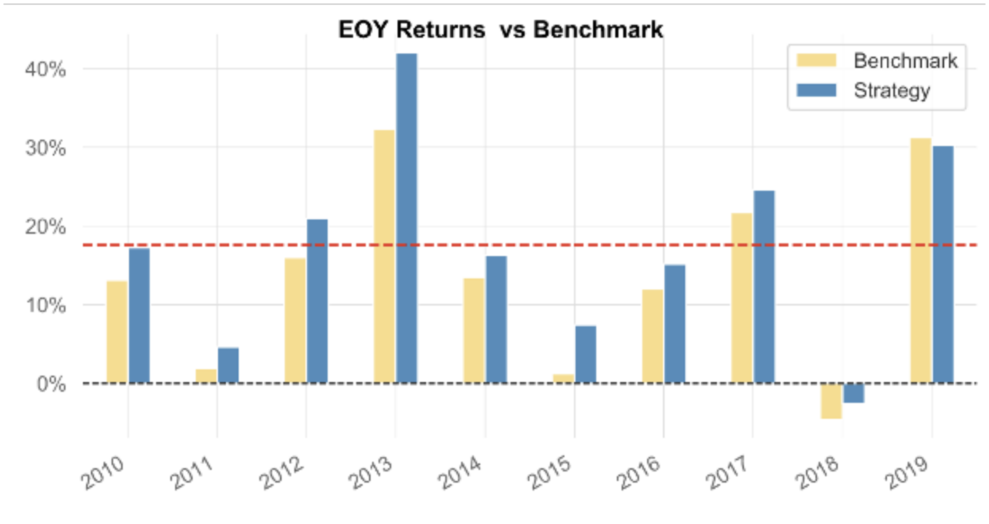

Figure 18 - HRP Portfolio End of year Return from 2010 to 2020

Table 9 - HRP Portfolio Key Performance Metrics

Key Performance Metrics

Metric	Strategy	Benchmark
Risk-Free Rate	0.0%	0.0%
Time in Market	100.0%	100.0%
Cumulative Return	379.57%	246.88%
CAGR%	16.99%	13.26%
Sharpe	1.18	0.92
Sortino	1.67	1.29
Sortino/√2	1.18	0.92
Max Drawdown	-17.31%	-19.35%
Longest DD Days	210	277
Volatility (ann.)	14.19%	14.69%
R^2	0.98	0.98
Calmar	0.98	0.69
Skew	-0.43	-0.43
Kurtosis	4.13	4.39

From above figures and tables, the cumulative return of HRP is 379.57%, slightly less than Eigenvector Centrality Portfolio, 397.48%. However, the volatility of HRP is the least. We could use Sharpe ratio to measure return against volatility. The Sharpe ratio of HRP is 1.18, and the SPY is just 0.92. In this case, although slightly less in return, all rational people will buy Eigenvector Centrality Portfolio not SPY Index.

5 Conclusion

We used a neural network to optimize our portfolio but not get a good result. Then we tried the Minimum Spanning Tree (MST) method. We use eigenvector centrality and betweenness centrality as our measures to construct three portfolios, Eigenvector centrality portfolio, Inverse Eigenvector centrality portfolio, Betweenness centrality portfolio. Based on MST, we used an unsupervised machine learning method Hierarchical Risk Parity (HRP) to clustering and then constructed another HRP portfolio. Compared with benchmark portfolios, GMV portfolio, Sharpe ratio portfolio, and S&P 100 Index, we concluded that Eigenvector centrality portfolio is the highest return portfolio, and HRP is highest Sharpe ratio portfolio. The betweenness centrality portfolio' performance is not so good, but all portfolios we constructed are better than S&P 100 index. We recommend that for risk seekers, investors should buy Eigenvector Centrality Portfolio. For risk averters, investors should buy an HRP portfolio.

In the future, we will try different equity, like bonds and real estate to test our conclusion. We also will use time series to back-test our results.

References

[1] Rubinstein, Mark. "Markowitz's" portfolio selection": A fifty-year retrospective." The Journal of finance 57, no. 3 (2002): 1041-1045.

[2] Merfon, R. "On estimating the expected return on the market." Journal of Financial Economics 8 (1980): 323-361.

[3] Michaud, Richard O., and Robert O. Michaud. *Efficient asset management: a practical guide to stock portfolio optimization and asset allocation*. Oxford University Press, 2008.

[4] DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?." *The review of Financial studies* 22, no. 5 (2009): 1915-1953.

[5] Mantegna, Rosario N. "Hierarchical structure in financial markets." *The European Physical Journal B-Condensed Matter and Complex Systems* 11, no. 1 (1999): 193-197.

[6] Sun, Wenyue, Chuan Tian, and Guang Yang. "Network analysis of the stock market." (2015).

[7] Výrost, Tomas, Štefan Lyócsa, and Eduard Baumöhl. "Network-based asset allocation strategies." The North American Journal of Economics and Finance 47 (2019): 516-536.

[8] Raffinot, Thomas. "The hierarchical equal risk contribution portfolio." Available at SSRN 3237540 (2018).