Machine Learning and Data-Intensive Systems

Babynin Andrey

Moscow, 2023

Graph Neural Networks for Stock Portfolio Optimization

Table of Content

Purpose of the study

I. Explore the abilities of GNN in solving tone of the fundamentals problem in Finance – asset allocation task

II. Present a novel approach with GNN and end-to-end portfolio optimization

III. Test whether the introduction of Prioritized Experience Replay (PER) from Reinforcement Learning (RL) can meaningfully impact the behavior of a model

Painting of Venice by Canaletto

Believe me, no. I thank my fortune for it—
My ventures are not in one bottom trusted,
Nor to one place, nor is my whole estate
Upon the fortune of this present year.
Therefore my merchandise makes me not sad.

Merchant of Venice, Act I, Scene 1
William Shakespeare

Literature overview

I. It (portfolio allocation) started with Harry Markowitz (“*Portfolio Selection*”, 1952)
   I. … or even before him (“*The origins of the mean-variance approach in finance*”, 2007)

II. … graph theory has long been viewed as a key to understanding the **interpedendence** of financial markets…
   I. Financial risk is concentrated at the center of the graph (“*Spread of risk across financial markets*”, 2013)
   II. Uncorrelated bets are associated with inverse graph centrality measures (“*Network diversification for a robust portfolio allocation*”, 2022)

III. … and it provided a point of view to look at the stock markets differently
   I. … by constructing graphs out of shareholder ties (“*Incorporating Corporation Relationship via Graph Convolutional Neural Networks for Stock Price Prediction*”, 2018)
   II. … or out of supply chains (“*Exploring Graph Neural Networks for Stock Market Predictions with Rolling Window Analysis*”, 2019)

IV. … while neural networks were used as a tool to predict single stock performance (“*Deep Learning for Stock Market Prediction Using Technical Indicators and Financial News Articles*”, 2018)

V. … the ultimate task is to predict and optimize assets portfolio (the idea behind an **end-to-end optimization**)
   I. Predict risk contributions and optimize (“*End-to-End Risk Budgeting Portfolio Optimization with Neural Networks*”, 2021)
   II. Predict returns and optimize (“*Distributionally Robust End-to-End Portfolio Construction*”, 2022)

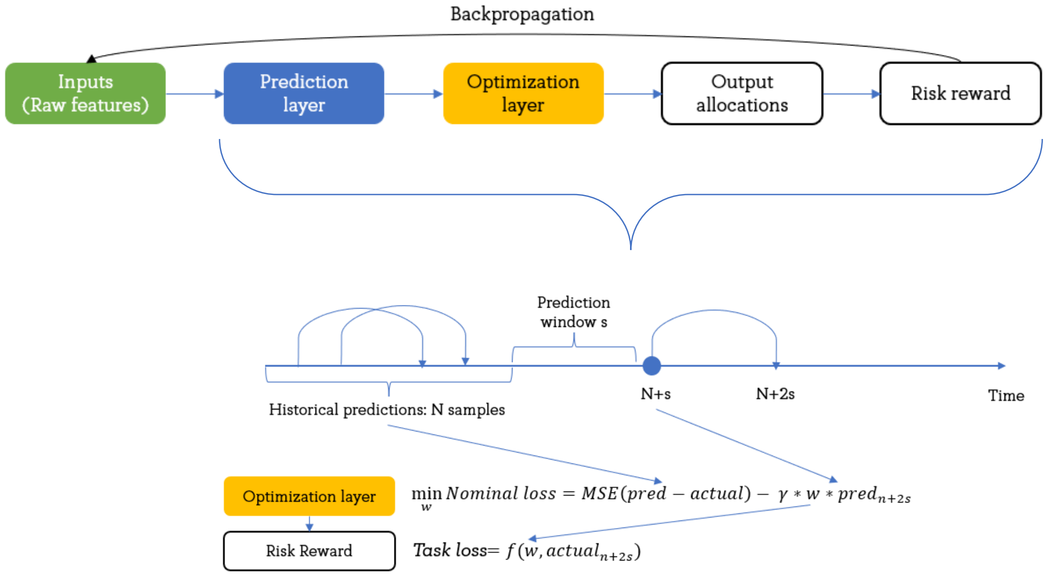
Methodology & algorithm explanation

I. End-to-end optimization is based on combining prediction and optimization layers into a single model.

II. In order to account for a model risk – there is a *nominal loss function*, in order to optimize the ultimate financial goal – a *target loss function*.

III. To make model more robust to potential outliers, I implemented a Prioritized Experience Replay (PER) buffer from Reinforcement Learning.

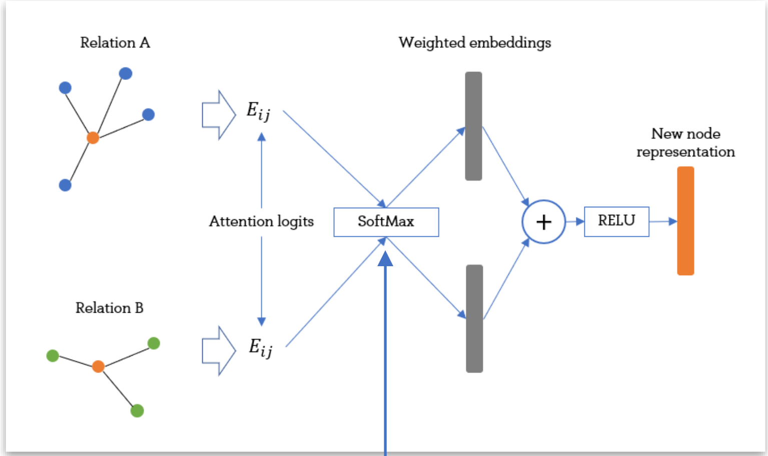
IV. Weights are obtained using convex optimization (CVXPYLayers library)



Model set-up

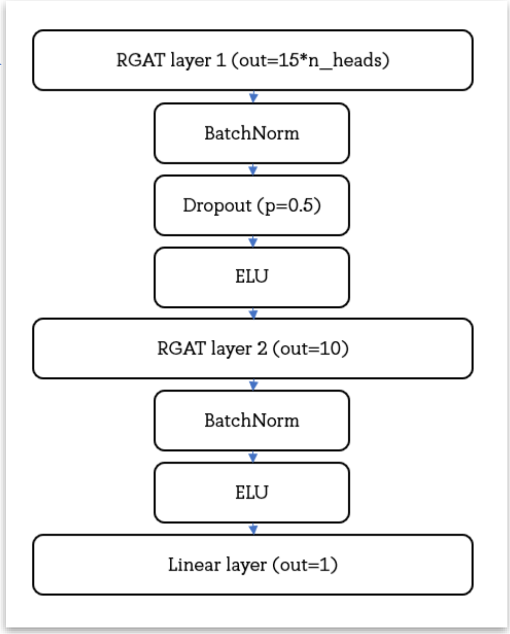
- Model consists of two logical part: prediction and optimization layers.
- Prediction layer consists of 3 layers: two RGAT layers and final Linear layer.
- There is 1 head per relation type

For each node, the model computes an attention score for each of its neighbors. In RGAT, the attention score is also dependent on multiple levels of relations, allowing it to incorporate relational reasoning.



Across-relation attention mechanism

$$e_{i,j}^{r} = \frac{\exp(e_{i,j}^{r})}{\sum_r \sum_k \exp(e_{i,k}^{r})}$$



Data collection

ETFs are great in representing markets because:

I. They provide embedded diversification by itself

II. Help to mitigate the survivorship bias and ceased tickers

Data represented by a collection of US ETFs representing different universes of assets. The general idea to select **the widest representation of financial markets:**

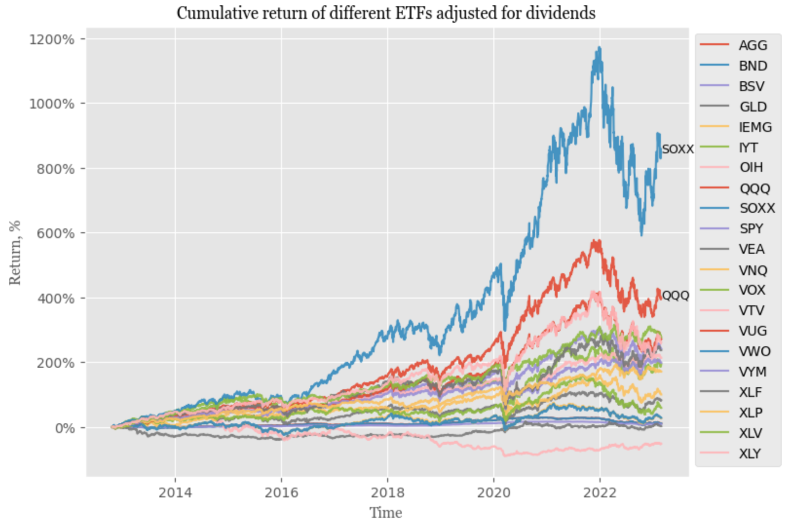
I. *By asset class*: Stocks (SPY), Bonds (AGG), Real Estate (VNQ), Commodities (GLD)

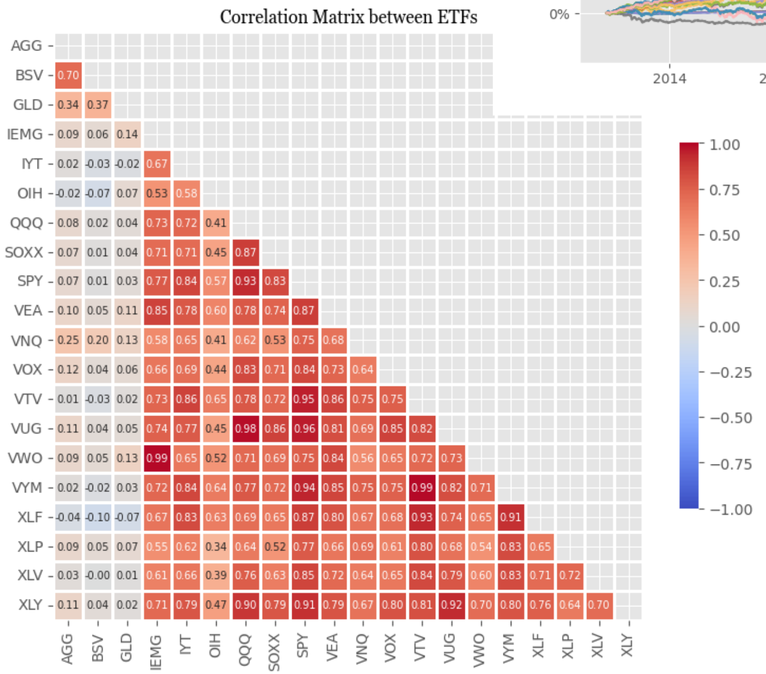
II. *By investment style*: Growth (VUG), Value (VTV), Dividends (VYM)

III. *By geography*: Developed (VEA), Developing markets (IEMG)

IV. *By sector*: Finance (XLF), Technology (QQQ), Health Care (XLV) etc.

Data collected from Yahoo Finance





Features generation

To generate features I used a collection of technical indicators and cumulative past returns for periods [1, 3, 7, 14, 30, 40, 50, 60, 70, 80, 90].

In total each node vector consisted of 23 features at each point in time.

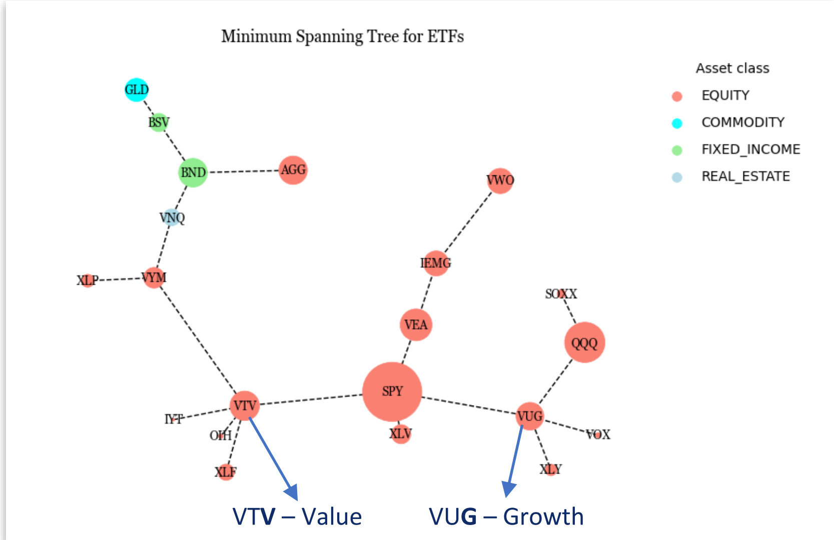
Name	Type	Parameters Specification
RSI	Momentum	Periods: 14, 28
MACD	Trend	Periods_slow: 26, 36; Periods_fast: 12, 18; periods_sign: 9, 12
Vortex Indicator	Trend	Periods: 14, 28
Stochastic Oscillator	Momentum	Periods: 14, 28
Williams Indicator	Momentum	Periods: 14, 28
Ulcer Index	Volatility	Periods: 14, 28

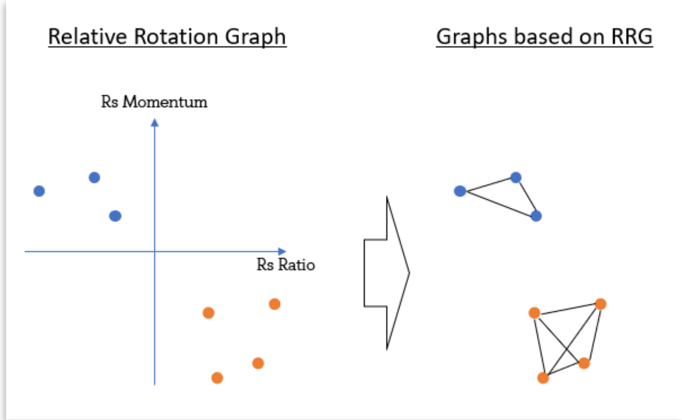
Graphs generation

To create graphs I used several approaches.

1. I calculated correlation between price changes(returns) as well as for volume traded changes with period of 90 days.
2. A used Gowel distance to normalize correlation $$d_{i,j} = \sqrt{\frac{1}{2}(1-\rho_{i,j})}$$
3. Calculate Minimum Spanning Tree (MST) and Planar Maximally Filtered Graph (PMFG) for graph representation.
4. In addition, I developed an approach based on Relative Rotation Graph (RRG).

№	Factor	Type of graph
1	Daily return	MST
2	Daily return	PMFG
3	Daily trading volume change	MST
4	Daily trading volume change	PMFG
5	Daily Return	Clustering based on RRG



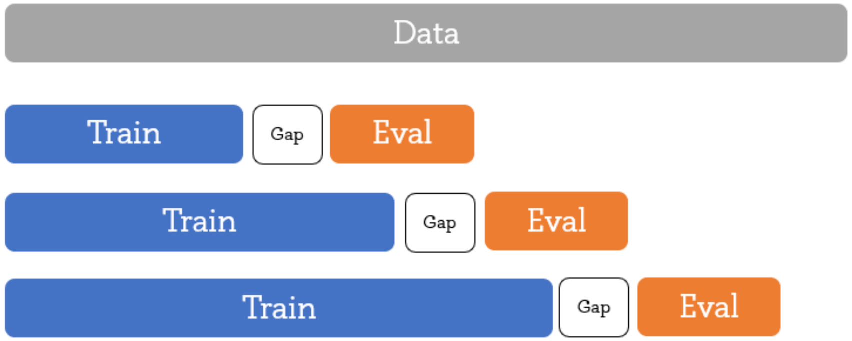


Experiments on market data

To test the model I completed a set of experiments changing different hyperparameters, including:

- PER buffer size
- Weights constraints (ability to short stocks)
- Number of attention heads
- Gamma γ parameter (Risk-appetite)
- Comparison with traditional approaches

Equal-weighted portfolio was chosen as a benchmark to compare results against to. To test this approach, I used expanding window with 4 consequent evaluation sets

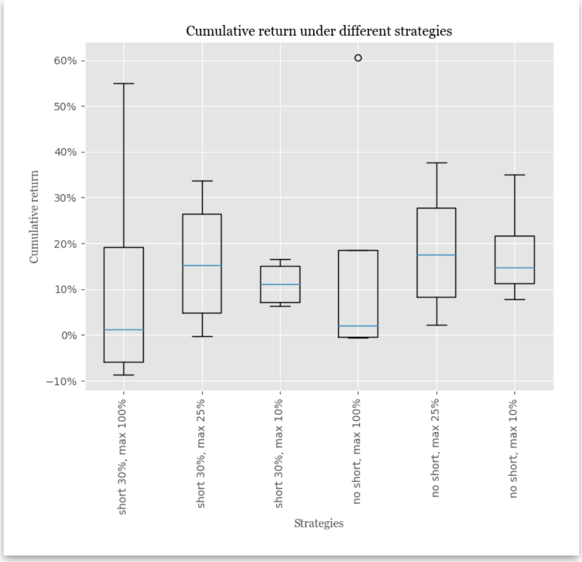


Train period	Evaluation period
2013-03-08 : 2015-07-24	2015-12-02 : 2017-07-05
2013-03-08 : 2017-02-24	2017-07-06 : 2019-02-06
2013-03-08 : 2018-09-26	2019-02-07 : 2020-09-08
2013-03-08 : 2020-04-30	2020-09-09 : 2022-04-08

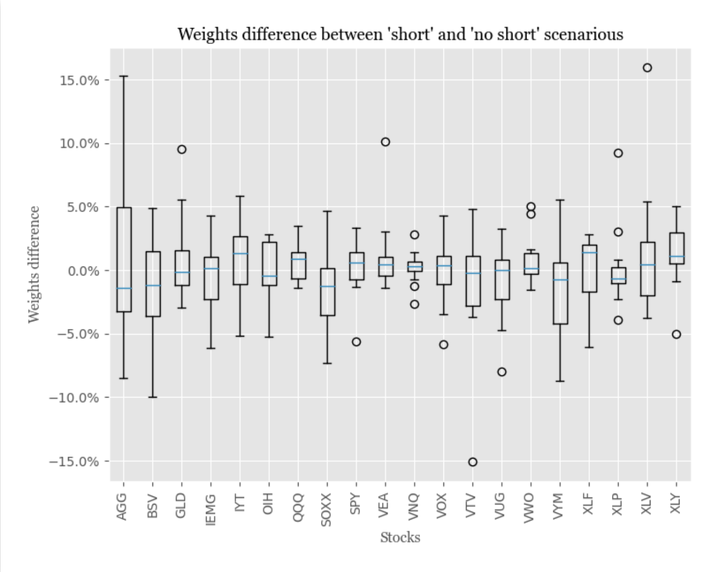
Experiments on market data

Weights constraints & Ability to short

No statistical difference in performance across scenarios…

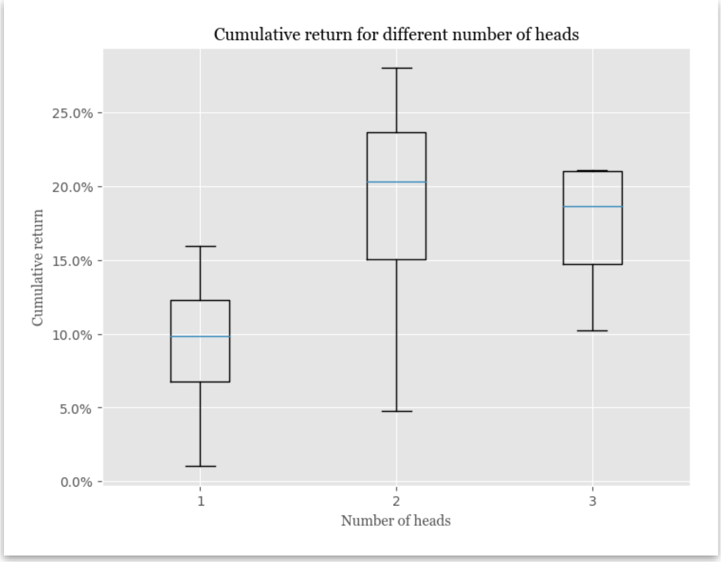


…and no indication that model "prefers" to short particular tickers across evaluation sets.



Number of heads

No significant difference in results between 3 scenarios of 1,2 and 3 heads per relation type.

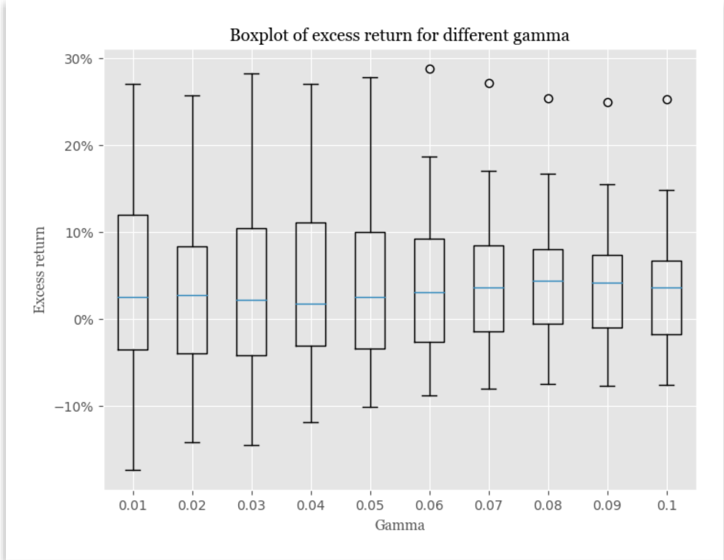


Number of heads	Mean performance	Information ratio (IR)
1	9,1%	0.47
2	18,3%	1,11
3	17,1%	0.85

Experiments on market data

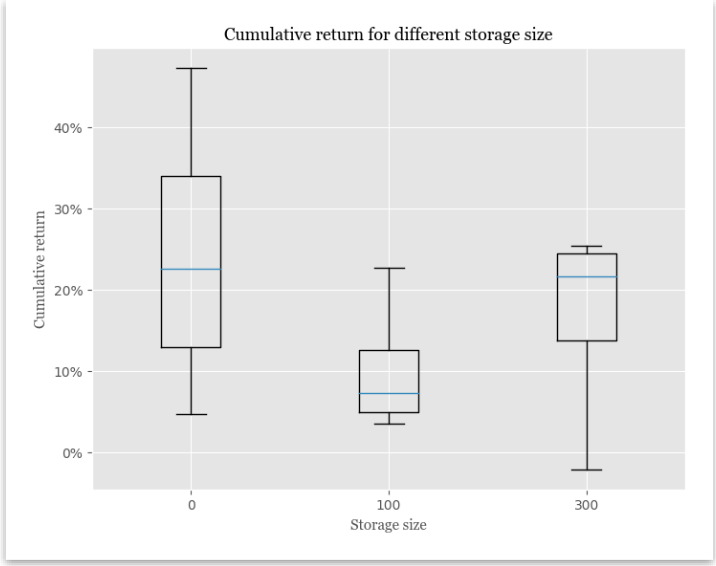
Different values of gamma

- a higher gamma value appears to yield more significant results compared to lower values



Gamma	T-statistics	P-value
0.01	1.18	0.12
0.02	1.32	0.10
0.03	1.36	0.09
0.04	1.62	0.06
0.05	1.89	0.03
0.06	2.13	0.02
0.07	2.36	0.01
0.08	2.40	0.01
0.09	2.27	0.01
0.1	2.16	0.02

Different storage size (PER buffer)



Evaluation Period	Best model			Benchmark return
	Storage size	Return	Information ratio	
1	300	24%	1,7	20%
2	300	25%	1,6	7%
3	0	29%	0,9	15%
4	0	47%	2,9	33%

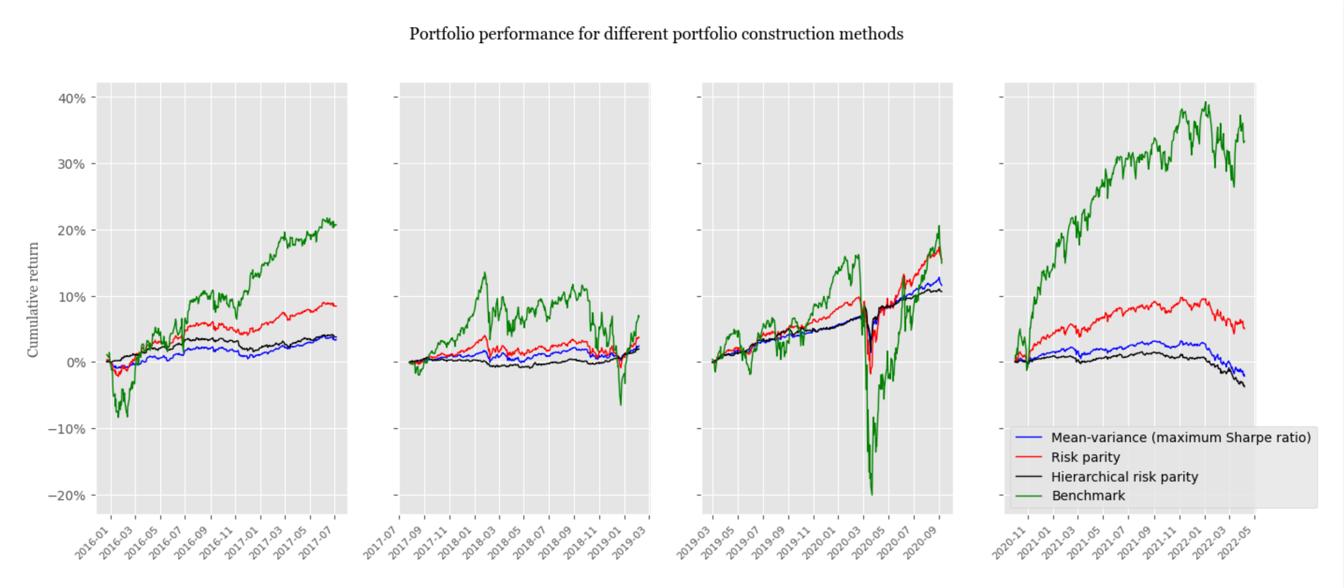
Comparison with traditional approaches

Results obtained from cross-evaluation testing of 200 models

Train period	Evaluation period	Average return	Benchmark return	Excess return	T-statistics	P-value
2013-03-08 : 2015-07-24	2015-12-02 : 2017-07-05	21,7%	20,7%	0,9%	0,75	0,237
2013-03-08 : 2017-02-24	2017-07-06 : 2019-02-06	8,5%	6,8%	1,7%	2,77	0,003
2013-03-08 : 2018-09-26	2019-02-07 : 2020-09-08	25,2%	14,9%	10,3%	6,92	<0,001
2013-03-08 : 2020-04-30	2020-09-09 : 2022-04-08	35,6%	33,3%	2,3%	1,92	0,03

Comparison of popular portfolio optimization approaches:

- *Equal-weight (benchmark) (green)*
- *Mean-variance (blue)*
- *Risk-parity (red)*
- *Hierarchical Risk Parity (black)*



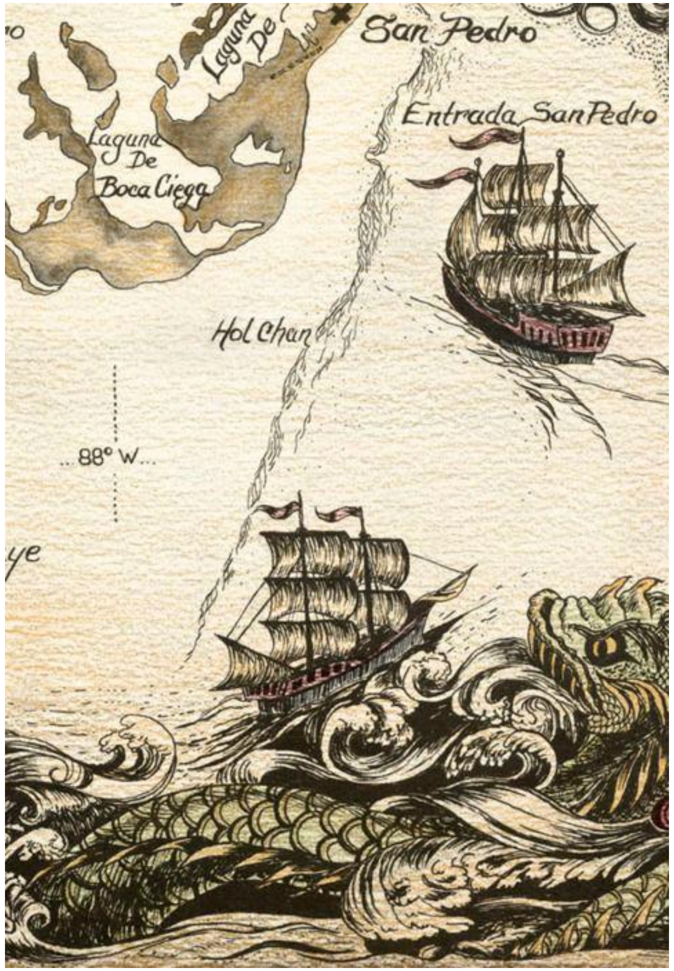
Conclusion

I. Graph Neural Networks represent a promising approach to predicting and analyzing financial markets.

II. It was a first known attempt to use GNN in end-to-end optimization task, and a first attempt to use PER in nominal loss calculation. Presented model is a rather simple and a lot can be improved in its behavior.

III. While the results are mixed, it is still worth noting that they are statistically significant.

IV. A great area of future research lies in building and testing more sophisticated graph models.



Painting by Savanna Redman

HSE UNIVERSITY • EST 1992