Stock Recommendations using Temporal Bipartite Graph Neural Network

CS224W
Fall 2019

Yuichi Kajiura, Rick Melucci, Jiushuang Guo {yuichik, rmelucci, jguo18}@Stanford.edu

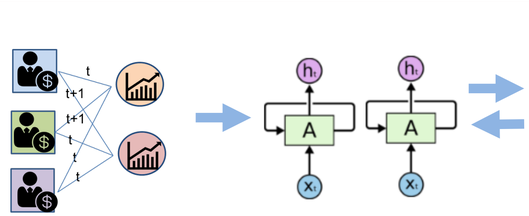
Motivation

It's hard for individuals with little financial knowledge or time to managing personal assets. But there exist groups of successful investors who focus on finding secure stocks that are undervalued compared to their intrinsic price: those investors are known as "**value investors**". Individual investors can follow the value investors' strategies.

Our contributions:

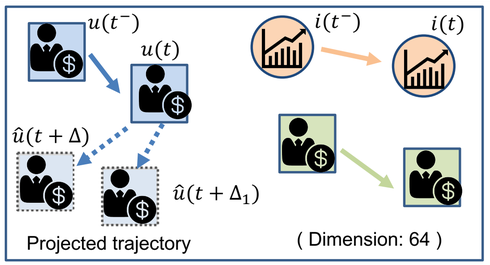
- Construct Investor-Stock Bipartite Graph and Directed Investor-Investor Graph
- Role detection by Motif, PageRank, HITs Analysis
- Detect secure investor communities by Louvain algorithm
- Implement a Temporal Graph Neural Network Model (JODIE) to predict future investment

Temporal Bipartite Graph Neural Network – JODIE



Temporal Interaction Network

Stock Embedding RNN & Investor Embedding RNN



Dynamic Embedding Trajectory

Features:
- price earnings
- price to earnings ratio
- price to book ratio
- current ratio
- dividend yield
- payout ratio

Symbol	Meaning
$\boldsymbol{u(t)}$ and $\boldsymbol{i(t)}$	Dynamic embedding of user u and item i at time t
$\boldsymbol{u(t^-)}$ and $\boldsymbol{i(t^-)}$	Dynamic embedding of user u and item i before time t
$\bar{\boldsymbol{u}}$ and $\bar{\boldsymbol{i}}$	Static embedding of user u and item i
$\tilde{\boldsymbol{u}}(\boldsymbol{t})$	Projected embedding of user u at time t
$\tilde{\boldsymbol{j}}(\boldsymbol{t})$	Predicted item j embedding

This Algorithm is Invented by Srijan Kumar, Xikun Zhang, Jure Leskovec

Edge weight :

$$w_b = \frac{\log V_{is}^t}{\sqrt{n_i^t \cdot n_s^t}}$$

Pruning edges with weight less than 0.2
- reduces size of dataset from 3M to 0.6M
- penalizes investor/stock which interact everywhere and enable us to focus on more informative transactions

⇒ eventually enhances prediction accuracy 10 times

Result :

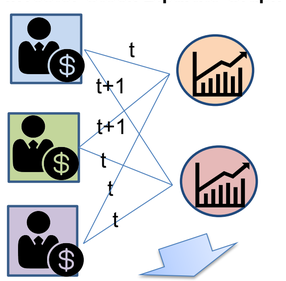
Architecture	Dataset Size(interactions)	Feature	MRR	Recall@10
JODIE	0.6million	No	0.040	0.087
JODIE	0.6million	Yes	0.042	0.096

JODIE with the features showed a 5% improvement in MRR and a 10% improvement in Recall@10 respectively

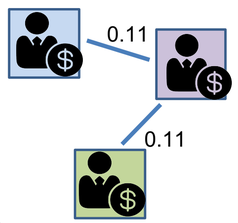
Data & Graph Construction

6 years of filings from Jan. 2013 quarterly. **23M reporting** from **6,454 investors** and **13,499 stocks** in total. From this data, we get **2M new investments** and **2M follow-on investments**. We also named **74 Value Investors**.

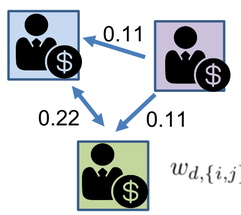
Investor-Stock Bipartite Graph



Investor-Investor Undirected Graph



Investor-Investor Directed Graph



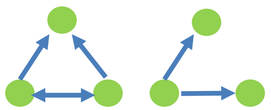
⇒Community detection

⇒Role detection

$$w_{d,\{i,j\}} = \sum_{t=1}^{T} \frac{E_{ij}^t}{(n_i^{t-1}+1)(n_j^t+1)}$$

Role Detection

Motif

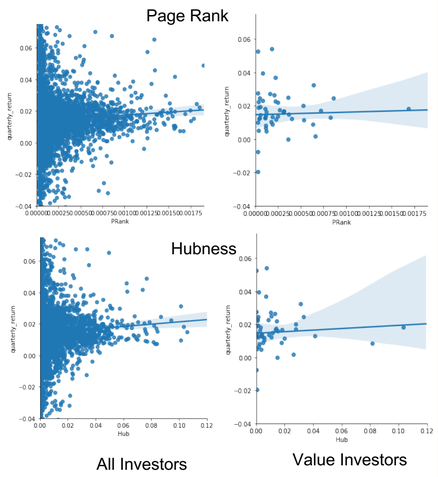


Two significant motifs.

Left motif implies the existence of investors who follow many other mutually connected investors (possible follower)

Right motif implies the existence of investors who are followed by many others (possible early mover)

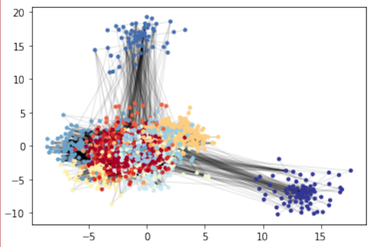
PageRank and Hubness



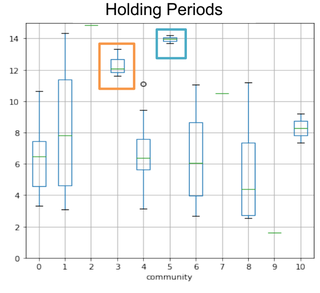
The results above show slight positive correlations between hubbiness/PageRank and return, which imply an "early mover" makes more profit than "followers".

Community Detection

Louvain Algorithm



Value investors in two of the communities play long term investments and enjoy moderate performance, implying potential investors that individual investors should follow.



MARKEL · HORIZON KINETICS · BERKSHIRE HATHAWAY · Boyar

Conclusion

1. Role Detection: identifies existence of leading investors for others to follow
2. Community Detection: shows clear differences among investor communities
3. JODIE: successfully predicts future investments of professional investor

⇒ Mixing those approaches will helps individual investors to greatly narrow down potential stocks to research before investing

Future Work

- Improve community detection, ex AGM
- Perform weighted PageRank on some particular communities detected by Louvain
- Improve JODIE by incorporating more features such as PageRank, HIT scores, and other financial indicators or by tuning parameters such as embedding dimension