

Sequential Graph Attention Learning for Predicting Dynamic Stock Trends (Student Abstract)

Tzu-Ya Lai¹, Wen Jung Cheng², Jun-En Ding^{3*}

¹ National Taipei University Master of Arts in Economics

² University of Connecticut Master in Financial Technology

³ National Yang Ming Chiao Tung University Institute of Hospital and Health Care Administration
s710961103@gm.ntpu.edu.tw, wen-jung.cheng@uconn.edu, m02040013@nycu.edu.tw

Abstract

The stock market is characterized by a complex relationship between companies and the market. This study combines a sequential graph structure with attention mechanisms to learn global and local information within temporal time. Specifically, our proposed “GAT-AGNN” module compares model performance across multiple industries as well as within single industries. The results show that the proposed framework outperforms the state-of-the-art methods in predicting stock trends across multiple industries on Taiwan Stock datasets.

Introduction

The stock market can be considered a dynamic system with numerous intricately connected parts that change over time. Graph Attention Networks (GATs) and Gated Recurrent Units (GRUs) have been used in some research to capture temporal information of the stock exchange due to the rapid growth of Graph Neural Networks (GNNs) (Hsu, Tsai, and Li, 2021). Each company is considered a node in the graph in these studies, and the correlation between two nodes at a certain time point determines the edge between them. In order to aggregate data on momentum spillover between companies, Cheng and Li (2021) previously proposed a model that integrates Graph Attention Network (GAT) and Attribute-Mattered Aggregator. In this paper, we argue that: a) the majority of research does not confirm the efficacy of the graph neural network approach, including whether it can be data-driven to capture company interrelationships; and b) the single-layer feed-forward neural network used as the Attribute-Mattered Aggregator in previous research (Cheng and Li, 2021) should be converted to a more methodologically sophisticated neural network module. We propose a novel model framework called the GAT-AGNN model to enhance prediction performance. We also suggest implementing the Corr-Cos model to confirm the efficacy of the graph neural network approach.

Problem Definition This study aims to forecast return on stocks across multiple industries based on their histories. We focus on two tasks founded in our analysis of return ratios from daily trading information: a) classification problems (classifying the future stock return from the closing price as a positive return or negative return); and b) regression problems (predicting future stock closing price return).

*Corresponding Author

Data Collection Our research uses daily data from *Taiwan Economic Journal* (TEJ) to predict the stock market trend. Every company has a highest/lowest price, opening/closing price and turnover rate. Data was collected from November 18, 2019, to June 20, 2022; a total of 630 days. The 30 companies studied were selected from the Yuanta/P-shares Taiwan Mid-Cap 100 ETF.

Proposed Method

Sequential Fusion Module with GAT and AGNN

Gated Recurrent Unit (GRU) Given the i -th company, and the independent variable for time series representation learning is $\mathcal{X}_i^{[t-T:t]} = [\mathbf{x}_i^{t-T}, \dots, \mathbf{x}_i^{t-1}]^\top$ which is the historical data for the first T days of the i -th firm, where T is the length of the rolling window. This research adopts GRU as the module for time series data extraction.

$$\mathbf{h}_i^t = GRU(\mathcal{X}_i^{[t-T:t]}). \quad (1)$$

The historical feature vector \mathbf{h}_i^t will be used to retrieve data for the previous T days in accordance with the approach described above.

Our approach utilizes the GRU module to create a graph-structured dynamic system (see Figure 1) that captures time-series dependency information for each company, and the nodes that constitute this dynamic system are derived from the fusion historical information vectors $\mathcal{H} = [\mathbf{h}_1^t, \dots, \mathbf{h}_N^t]$.

Graph Attention Network (GAT) The prospective link between companies i and j at time t is therefore captured by $U_{i,j}^t$.

$$U_{i,j}^t = u(\mathbf{h}_i^t, \mathbf{h}_j^t) = ELU(\mathbf{a}_u^\top \mathcal{W}_u [\mathbf{h}_i^t \parallel \mathbf{h}_j^t]), \quad (2)$$

To identify prospective company relationships via the shared attention mechanism, a single-layer neural network and an ELU activation function are used to train the correlation between nodes i and j . The softmax function is applied to normalize all pairwise connections, and the correlation between companies (i, j) , is normalized at time t , can be identified as follows:

$$\widetilde{Q}_{i,j}^t = softmax_j = \frac{\exp(U_{i,j}^t)}{\sum_{k \in N, k \neq i} \exp(U_{i,k}^t)}. \quad (3)$$

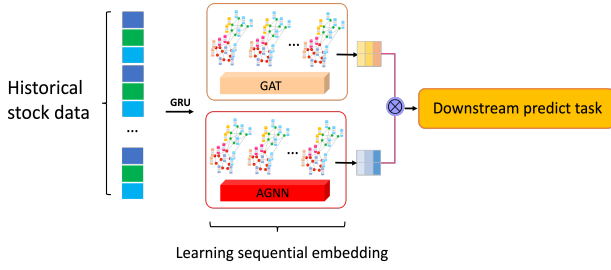


Figure 1: Diagram of GAT-AGNN’s architecture.

Attention-Based Graph Neural Network (AGNN) Using calculation $C_{i,j}^t$ can capture potentially connected organizations based on similarities between companies. By using a trainable α , the prospective connections between companies i and j at time t , can be identified as follows:

$$C_{i,j}^t = c(\mathbf{h}_i^t, \mathbf{h}_j^t) = \alpha \cos(\mathcal{W}_c[\mathbf{h}_i^{t\top}, \mathbf{h}_j^{t\top}]), \quad (4)$$

The softmax function is used to normalize all pairwise connections.

$$\widetilde{G}_{i,j}^t = \text{softmax}_j = \frac{\exp(C_{i,j}^t)}{\sum_{k \in N, k \neq i} \exp(C_{i,k}^t)}. \quad (5)$$

Sequential Relation Module The relational of the $i - th$ company at time t information from the GRU module vector is defined as $\mathcal{V} = [\mathbf{v}_1^t, \dots, \mathbf{v}_N^t]$, while \mathbf{v}_i^t is the correlation sequential vector calculated as follows:

$$\mathbf{v}_i^t = \tanh\left(\sum_{j, j \neq i}^N \widetilde{Q}_{i,j}^t W_s \mathbf{h}_j^{t\top} \otimes \widetilde{G}_{i,j}^t W_v \mathbf{h}_j^{t\top}\right), \quad (6)$$

In addition to the attention mechanism, we consider a multi-head option.

$$\mathbf{v}_i^t = \left\| \tanh\left(\sum_{j, j \neq i}^N \widetilde{Q}_{i,j}^{t,[m]} W_s^{[m]} \mathbf{h}_j^{t\top} \otimes \widetilde{G}_{i,j}^{t,[m]} W_v^{[m]} \mathbf{h}_j^{t\top}\right) \right\|_{m=1}^M, \quad (7)$$

The purpose of our prediction tasks are both Regression and Classification tasks. First, a single-layer neural network with the function $RELU(\mathcal{W}_i^T[\mathbf{h}_i^{t\top} \parallel \mathbf{v}_i^t] + \mathbf{b}_i)$ is used; then, the second uses a single-layer neural network with the function $Softmax(\mathcal{W}_i^T[\mathbf{h}_i^{t\top} \parallel \mathbf{v}_i^t] + \mathbf{b}_i)$.

Additionally, we replace the Sequential Relation Module of GAT and AGNN with the without training parameters module of Correlation coefficient and Cosine similarity (Corr-Cos model). With these settings, we can test which modules are more robust in their ability to learn complex information from industry relationships.

$$\mathbf{v}_i^t = \tanh\left(\sum_{j, j \neq i}^N \widetilde{Corr}_{i,j}^t W_s \mathbf{h}_j^{t\top} \otimes \widetilde{Cos}_{i,j}^t W_v \mathbf{h}_j^{t\top}\right). \quad (8)$$

Results and Discussion

In our research, we utilize different creating relationship functions (2) and (4) to enhance and combine the attention mechanisms of GAT and AGNN for the multi-head setting, which is superior to the traditional approach that relies only on an activate function with a single-layer neural network.

Problems Metrics	Classification		Regression	
	ACC	AUC	MSE	MAE
LSTM	0.5222	0.5154	1.7299	0.9591
LSTM+GAT	0.5283	0.5023	1.3874	0.8633
GRU	0.5139	0.5027	1.4963	0.8944
GRU+GAT	0.5150	0.4805	1.2083	0.8031
Corr-Cos	0.5328	0.5290	1.1850	0.7940
AD-GAT	0.5189	0.5056	1.2143	0.8183
GAT-AGNN	0.5350	0.5190	1.1585	0.7823

Table 1: Comparison of classification and regression prediction results across models.

	ACC	MSE
All Industry Industry6 ¹	0.5700	0.9851
Industry6 ¹	0.4767	1.2542

¹ Industry 6: shipping industry.

Table 2: Comparison of model performance across multiple industries and within single industries.

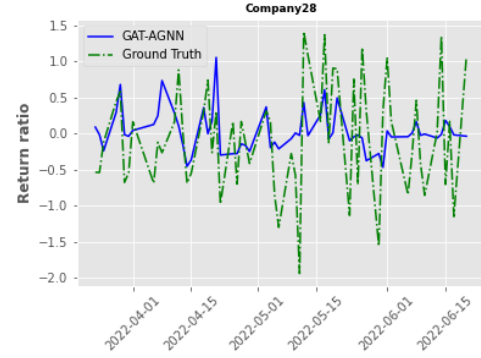


Figure 2: Stock time series forecasting visualization. Time-dependent pattern captured by GAT-AGNN.

The results on the classification and regression prediction tasks are displayed in Table1. According to the results below, the GAT-AGNN model outperforms the Corr-Cos model, indicating that graph neural networks are effective for such tasks. Moreover, the GAT-AGNN model lowers MSE by 5.5% over the AD-GAT model (Cheng and Li, 2021). Table1 shows, the GAT-AGNN model is not only significantly superior to GRU and GRU+GAT but also LSTM and LSTM+GAT.

It is evident in Table 2 that the MSE in the multiple-industry model decreases by 26.91%, while the ACC increases by 9.33% compared with the single-industry model. Thus, our study suggests that sequential graph structure in hierarchical information (see Figure 2) offers promise for the identification of industry relations and information.

References

- Cheng, R.; and Li, Q. 2021. Modeling the Momentum Spillover Effect for Stock Prediction via Attribute-Driven Graph Attention Networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 55–62.
- Hsu, Y.-L.; Tsai, Y.-C.; and Li, C.-T. 2021. FinGAT: Financial Graph Attention Networks for Recommending Top-K Profitable Stocks. *IEEE Transactions on Knowledge and Data Engineering*.