Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 214 (2022) 771–778

Procedia Computer Science

www.elsevier.com/locate/procedia

9th International Conference on Information Technology and Quantitative Management

Review of graph construction and graph learning in stock price prediction

Yunong Wang[a,b,*], Yi Qu[a,b], Zhensong Chen[c]

[a]Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences, Beijing 100190, China
[b]Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, Beijing 100190, China
[c]School of Management and Engineering, Capital University of Economics and Business, Beijing 100070, China

Abstract

Precise prediction of stock prices leads to more profits and more effective risk prevention, which is of great significance to both investors and regulators. Recent years, various kinds of information not directly-relevant with stock prices have received more attention, such as texts, images or connections. These external information has the potential of reflecting or influencing fluctuations, and thus, given the utilization of advanced analyzing techniques, the forecasting performance of stock prices could be promoted substantially. For instance, graph neural network models have expanded into many other disciplines including stock price prediction, and exhibited impressive representation learning ability. However, in stock markets, well-defined graphs are rarely seen and how to formulate the graph structures needed remains a challenging problem. Towards this end, this article presents a comprehensive overview of graph construction and graph learning in stock price prediction, by reviewing the existing studies, summarizing its general paradigm, special cases and proposing possible prospects. Our research not only systematically reveals the feasible ways of constructing graphs in financial markets, but also provides insights for further implementations of graph learning models into stock prediction tasks.

Keywords: Stock price prediction; Graph construction; Graph learning; Data structure.

1. Introduction

Prediction of stock price is an important research subject for investment decisions and quantitative management. As the stock market is influenced by a variety of factors, the inner mechanism and transmission path of stock price fluctuations are diverse and intricate. Besides the internal information such as daily trading indicators, external information should also be taken into consideration for mining patterns of stock price changes. External information such

* Corresponding author.
E-mail address: wangyunong20@mails.ucas.ac.cn

10.1016/j.procs.2022.11.240

as macro-economy, news, and social media commentary affects or reflects the stock price in various ways. To a certain degree, different external information represents the different causes and transmission paths of stock price changes. For example, stock market is a barometer of the macro-economy, so the macro-economy always influences the stock price through a top-down path. Meanwhile, social media comments reflect investor sentiment, and it may then translate into investor behavior and ultimately perform as stock price volatility. Thus, they affect the stock price through a chain transmission path. Traditional stock price forecasting studies, however, regard stocks as independent individuals and ignore the association between stock price of different firms. Therefore, besides the external information mentioned above, relational information is of increasing attention in the stock market. Different from top-down or chain transmission paths, this exploitation of stock-to-stock relational information follows a network-like transmission path of stock prices. Changes in the price of one stock may be simultaneously or subsequently shown to other stocks connected to it in the network. Roll[32] introduced stock price synchronicity in 1988, and subsequently, various measures and empirical studies[8, 17, 30, 47] have proved that cross-effects[24] indeed exist between different companies with certain relationships. Following the development of these theories as well as graph representation learning techniques, scholars have used graph structures to portray stock-related relationships in recent, improving stock price prediction.

In fact, the use of graphs to construct stock-to-stock relationships can be traced back to complex network analysis, which mainly uses graph properties, such as degree distribution, edge density, betweenness centrality, and so on to qualitatively analyze the network structure and the stock market it represents[6, 28, 35]. While with the booming of Graph Neural Network (GNN), graph contributes to stock price prediction through graph representation learning, which follows and portrays the network-like transmission path of stock price fluctuation mentioned above. The related work can be divided into semi-supervised learning, unsupervised learning and supervised learning, and supervised learning can be further divided into classification and regression, which refer to stock price trend and value prediction, respectively. Different from other relation-based problems such as traffic networks and citation networks, the relationships in the stock market are diverse, complex, and even implicit. Thus, appropriate graph construction may be more crucial than model selection and tuning. Only the relational network that conforms to realistic mechanisms can truly simulate the transmission path of stock price fluctuations and make correct predictions. Therefore, our review mainly focuses on graph construction based on the stock-to-stock relationships in a supervised forecasting task. As a supplement, we also briefly describe some graph constructions based on intra-stock relationships, which consider trading days or other factors of one stock as nodes.

The rest of this article will be arranged as follows: Section 2 will introduce the general paradigm of stock-to-stock graph construction. Section 3 will first present three special cases derived from the general paradigm and then introduce the graph construction of intra-stock relationships. Section 4 will describe the common models used to extract features and predict after the construction of graph. Section 5 will summarize and discuss the research trends in this area to provide guidance for better utilization of graph-based stock price forecasting.

2. Graph Construction: General Paradigm

According to the basic definition of graph data, a general graph includes at least nodes, node features, node labels, edges, edge features and other components. Nodes refer to entities and edges refer to relationships between entities. The connectivity of edges can be denoted by the adjacency matrix $\mathbf{A} = [a_{ij}]$ and a_{ij} represents the edge feature between node i and node j. It also illustrates the topology of the graph: the asymmetry and symmetry of the adjacency matrix $\mathbf{A}$ correspond to the directed and undirected graphs, respectively. The values of a_{ij} in the adjacency matrix can be either 0 or 1, or other values representing weights, corresponding to the unweighted and weighted graphs, respectively. This section will first introduce the node features with stocks as nodes, and then focus on classifying the common types of edge construction and edge attributes of the stock-to-stock relationships.

2.1. Node and Node Attributes

The main purpose of introducing graphs in the stock market is to portray and exploit the relationships of stocks, so it comes naturally to consider them as nodes. In this kind of literature, the raw data for node attributes are usually daily trading indicators of stocks [3, 14, 44] such as opening price, closing price, highest price, lowest price, trading volume and trading amount, or derived technical indicators [2, 12] like Moving Average (MA), Momentum (MOM), Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD) and so on. Furthermore, in order to leverage the time-series characters of these indicators, time-series modeling techniques are commonly used, which will be described in detail in 4.1. Some other feature extraction approaches are also available, for example, Chen et al. [2] convert technical indicators of stocks related to the target stock into image data and use Convolutional Neural Network (CNN) to obtain node features.

In addition to daily trading indicators, several studies use the embedding extracted from textual data such as news and events as features of stock nodes.[23] encodes the news headlines of stocks as node features through Long Short-Term Memory (LSTM), which has been employed with great success to encode the context information of the textual data. Xu et al.[43] utilize a type-specific multi-head attention mechanism and LSTM to obtain event information on past events of stocks, combined with the feedback of the stock to the events, i.e., the relative changes of the stock price and trading volume after the events. After feeding them into two different LSTMs, the final node features are obtained by concatenation.

2.2. *Edge Construction*

The existing stock-to-stock relationships can be divided into explicit and implicit, the first one refers to the relations that have real meaning while the second one can not be explained in reality but need to be mined artificially. These kinds of edge construction assume that stock prices of different companies fluctuate simultaneously or subsequently along the network formed with some kind of relationship. This section starts with two approaches to explicit relationship construction and then introduces implicit relationships calculated with similarity.

Relationships Constructed by Human Knowledge. Most manual relation construction can be divided into **coexistence** relationship and **connection** relationship. Coexistence relationship refers that two entities belonging together to the same concept, such as companies in the same industry[14, 33, 40], with the same business[23], with the same news topics[44]. Meanwhile connection relationship refers to the realistic association between companies, such as upstream and downstream companies in the supply chain[5], cross-shareholdings[23, 44], and competition, collaboration, strategic alliance[5, 20, 21].

Relationships Extracted via Knowledge Graph. Knowledge Graph is proposed by Google in 2012, and its main purpose is to support semantic search. Here we mainly discuss information extraction, including entity extraction, entity disambiguation, relationship extraction, attribute extraction and event extraction. Deng et al.[9] convert the news text of stocks from unstructured data to structured event tuples by OpenIE[10], and then disambiguate the entities by entity linking[36]. To make knowledge graph not so sparse, they link entities within one hop. Similarly, a number of studies[13, 51] build knowledge graphs from news data to assist in stock price prediction, which is not described in detail here. Other than stock relationship extraction based on news events, some articles extract stock-to-stock relationships through knowledge graph-based databases or documents. Feng et al.[11] describe a first-order relationship as subject and object stocks in the same statement, and a second-order relationship as stocks that have the same object in two different statements of Wikidata. Zhang et al.[46] use text data in the annual financial statements and prospectus documents, utilize Albert model for named entity recognition, and construct the industry chain knowledge graph of listed companies. Long et al.[25] find shareholding, relevance and affiliation relations from the companies' knowledge graphs.

Relationships Calculated with Similarity. In the implicit relationship construction, the most common is to calculate the similarity of the stock price series. After that, select the stock pairs with absolute values of correlation coefficients greater than a certain threshold or ranked in the top k to establish the edges, which is close to the idea of clustering. For example, the Spearman rank-order correlation coefficient, Pearson correlation coefficient, Detrended Cross-correlation coefficient (DCCA coefficient) are commonly used to calculate the similarity of opening price, closing price, historical revenue and so on [18, 29, 42]. In particular, some literature uses the method that specifically describe time series similarity, such as Dynamic Time Wraping (DTW), to establish connections between nodes [37]. In addition, the similarity of company accounting ratios is also available, such as Intangible Assets, BookValue Per Share, Long Term Debt, Debt in Current Liabilities and so on. Hou et al.[18] calculate the Euclidean distance between stock pairs by stock fundamental information extracted from variational autoencoder (VAE) and then construct edges by threshold filtering. In addition to sorting by absolute value, Chen et al.[2] distinguish the positive and negative correlation coefficients, and turn them into positive and negative relationship graphs respectively.

Along with the correlation coefficient, another method that can measure similarity to define the association between two variables is Normalized Mutual Information (NMI), which originated from Shannon's entropy theory and is widely prevalent in information theory. The nature of NMI is to calculate the probability distribution and joint probability distribution of variables to describe the amount of information that one variable contains in another variable. While in stock research, how to find this probability distribution becomes the key to constructing the edge. The related research discretizes the stock returns into several intervals, uses the frequency of the interval to estimate the probability, and then finds the correlation between different stocks[12, 16].

2.3. Edge Attributes

In terms of edge attributes, many studies process them as unweighted graphs with weights of 0 or 1. Instead of describing the weights of each relation in detail, this paper summarizes the weight calculations into the following types to facilitate reference for new relations construction. The first kind uses the similarity coefficient of a certain attribute to portray the closeness of the connection between nodes. The second kind uses the proportion of a certain attribute that a node contributes to another node to portray the importance of nodes to nodes. For example, the edge weight of node i to node j can be calculated by dividing the concept to which nodes i and j jointly belong by the concept to which j belongs. These two methods of calculating weights imply the undirected graph and directed graph respectively.

The stock market is dynamic and the strength of the relationship may change over time. Therefore, in order to obtain time-dependent relations, some works have introduced techniques such as attention mechanisms in the graph learning process after the initial construction described above. Xu[41] extracts dynamic spatial dependencies among nodes (stocks) based on transformed data with the kernel method. Feng et al.[11] use sequential embeddings to estimate the strength of relation by explicit modeling and implicit modeling. The explicit modeling mainly measures the similarity between two stocks at each timestep and models a nonlinear regression on the relations to define the relation strength. While implicit modeling relies on training parameters to capture the strength of the relationship.

3. Graph Construction: Special Cases

This section first briefly introduces three special graphs derived from the general graph, which are heterogeneous graph, hypergraph and subgraph. Then we present the graph construction based on intra-stock relations, which consider the trading days or other factors of one stock as nodes and try to find the pattern of price fluctuations from the connection.

3.1. Three graphs derived from general paradigm

Heterogeneous Graph. Heterogeneous graph refers to a graph with multiple types of nodes, which also implies multiple edge relationships, such as Knowledge Graph. Cheng et al.[4] build a heterogeneous graph with six kinds of nodes. The first three kinds represent companies, which are source nodes, target nodes and the bridge nodes connecting them. The rest three are attribute nodes, including news, event and market nodes, which are connected to the companies they describe. The advantage of heterogeneous graphs is to integrate multiple sources of information to provide a more comprehensive description of the stock relationship network.

Hypergraph. Hypergraph is a special kind of graph that can use one hyperedge to connect multiple stocks at the same time. Unlike the adjacency matrix or triple of a general graph, the hypergraph is represented by an incidence matrix, where each row represents a node and each column represents an edge. Sawhney et al.[34] construct the relationships between stocks in the same industry as a hypergraph. The hypergraph facilitates the depiction of multimodal data and more complex connections, compensating for the limitations of the pairwise relationship of traditional graph structures. The study of using hypergraphs in stock price forecasting is still relatively scarce, and with the diversification of data and relationships, this graph structure could be widely used in the future.

Subgraph. Subgraph is a partial graph cut out from the complete graph, mainly using four common filtering methods, which are Planar Maximally Filtered Graph (PMFG), Minimum Spanning Tree (MST), Asset Graph (AG) and Statistical Threshold Method (STM). Li et al.[22] first use Perceptually Important Points (PIP)[26] to find the important sequence from stock price series and convert them into a graph. Then they rank each node's importance score to select N important nodes and expand the subgraphs for these nodes by breadth-first search algorithm based on the preset minimum subgraph size. The advantage of subgraphs is to identify important nodes and relationships from a complete graph.

3.2. Graph based on intra-stock relations

Relationships Extracted from Time-series Relationships extracted from time-series refer to the relations between different trading days of one stock. Its nature is to turn time-series into a graph structure, with common methods being Visibility Graph (VG), Recurrence Networks (RNs), and Transition Networks (TNs). Taking VG as an example, its definition [19] is that the line connecting the values of two time points doesn't touch the value of their middle time point, then these two time points (nodes) have an edge. VG has many kinds, such as Horizontal Visibility Graph (HVG)[27] and Limited Penetrable Visibility Graph (LPVG)[38] and so on.

After the construction of the relationship based on the time series, we can use the graph directly to implement the stock price forecast, e.g., Zhang et al.[48] use VG to establish edges and forecast the future data through link prediction. Wu et al.[39] treat trading days as nodes and use graph embeddings to represent the association between time points as input to the prediction model. In addition, we can find stock-to-stock relationships based on the similarity of time-series graphs. Qi et al.[31] use VG to convert the time series of daily closing prices of each stock in a certain time window into a graph structure, and then calculate the correlation coefficients and distances of two stocks. By filtering graph method introduced in 3.1, to get the stock correlation network.

Relationships Based on Probability. The use of Bayesian Factor Graph (BFG) in stock price prediction is an early study[52, 53]. Its main idea is to pre-define a score function and find the highest score network structure. Common Bayesian statistics-based scoring functions are K2Metric[7], Bayesian Dirichlet equivalence uniforming (BDeu)[15], and search algorithms are K2 algorithm and hill-climbing algorithm. Bayesian networks usually use the trading days of stocks as nodes to establish the network and predict by maximizing the probability of stock price return based on this network. Also, nodes can be a set of economic factors, and BFG is used to select the key factors that correlate with the return. Further, some research introduces Dynamical Bayesian Factor Graph (DBFG) to depict dynamic relations among the nodes[49, 50].

In addition to BFG, another try that relies on the probability distribution is the combination of Hawkes Process and Maximum Likelihood Estimation (MLE). Yin et al.[45] process the time sequence of each stock into a Multi-dimensional Hawkes Process, then construct a correlation graph by Hawkes model.

4. After Graph Construction: Feature Extraction and Further Prediction

Since the final purpose of graph construction is stock price prediction, we will show the two main processes after graph construction, one is feature extraction and the other is final prediction.

4.1. Feature Extraction

Feature extraction usually consists of relational feature extraction and temporal feature extraction, which serve to capture patterns from the relationship and historical performance of the stock, respectively. One of the famous time-series modeling techniques is Recurrent Neural Network (RNN) and its variant LSTM and Gate Recurrent Unit (GRU), which introduce gate control to compensate for gradient vanishing or explosion of RNN. Also, attention mechanisms are widely combined with temporal modeling techniques to assign more attention to important nodes and edges, such as dual-stage attention-based RNN (DARNN) and attention LSTM. Certainly, some other deep learning models can also be employed as temporal feature extractors, as Chen et al.[2] use Dual Convolutional Neural Network (Dual-CNN) to capture individual trading features of each stock.

For the relational feature extraction, the common embedding methods for knowledge graphs are TransE, TransH, TransR, TransD. In addition, some graph embedding methods are also popular, such as node2vec, which is more concerned with local domain similarity, and struc2vec, which is more concerned with structural similarity. Moreover, Graph Convolutional Network (GCN) and Graph Attention Network (GAT) that obtain node representation through neighbor nodes aggregation have gained great attention in recent years. It is worth noting that there is no fixed order for relational feature extraction and temporal feature extraction, so they can be arranged and combined freely.

4.2. Further Prediction

Finally, a predictor is designed to perform a variety of tasks related to stock price prediction such as stock price movement forecasting, stock price forecasting, stock ranking, and investment trading portfolio. Some of the work completes the prediction task directly after feature extraction. While others connect to machine learning and deep learning models, such as GP[13], Artificial Neural Network (ANN)[49], Structural Support Vector Machines (SSVMs)[40].

5. Conclusion and Discussion

In this article, we present a brief overview of stock-related graph construction and learning, demonstrating the diversity and effectiveness of graph-related methods in stock market prediction. In a general graph, nodes are stocks, and node features can be daily trading indicators, text features, or indicators after model processing. The edges describing the stock-to-stock relationships can be classified into explicit and implicit constructions. Regarding the fixed weights of the edges, there are two main types of similarity coefficients and importance proportions. Moreover, the dynamics of the stock market lead some studies to use time-dependent weights rather than fixed weights. Subsequently, we introduce three graph structures derived from general graphs, namely heterogeneous graph, hypergraph and subgraph, which have the advantages of handling diverse data, describing multi-modal data and identifying important nodes, respectively. Also, we discuss the graph construction based on intra-stock relationships as a supplement. Finally, we describe some feature extraction techniques and predictors that support graph data to predict stock prices. Since we assume that changes in the price of a stock may be simultaneously or subsequently shown to other stocks connected to it in the network, it's vital to construct graphs based on proper stock-to-stock relationships. Thus, the more promising direction is to mine possible stock-related relations, and future work may be inspired by the following two perspectives.

Firstly, it may be a feasible attempt to construct a relationship by considering the inner mechanism behind the stock price fluctuations. The price of commodities is determined by supply and demand, and stocks, which are special financial products, also follow this rule. The difference in trading volume between buyers and sellers at the same time will lead to stock price fluctuations. In actual trading, the small amount of capital represented by individual investors will be quickly covered during the transaction, so its impact on the stock price is arguably negligible. The real trigger for major stock price fluctuations often originates from high-volume transactions, such as pension funds and social security funds. These large trades often focus on a group of stocks with common features, such as stocks in the same region, belonging to the same concept, and so on, which need to be explored. Stock relationships established after uncovering and understanding the underlying mechanisms of high-volume transactions may provide better performance on the stock price prediction.

Secondly, stock-related relationships can be discovered from qualitative, theoretical, and empirical research articles. For instance, ambiguity aversion causes people to prefer known uncertainties to completely unknown probability distributions, and as a result, investors exhibit a "home-loving tendency" when investing. Brennan et al.[1] find that fund managers over-invest in domestic stocks due to the psychological bias of ambiguity aversion by irrationally overvaluing the domestic stock market. Wang[30] verifies that companies headquartered in the same geographic region have a strong synergistic movement in stock returns. Therefore, it may be helpful to construct relationships through the company's geographical location to forecast the stock price. Relevant studies have laid the foundation, and we should make better advantage of this literature to build graphs.

Acknowledgements

This work has been supported by Key Projects (grants number 71932008, 72231010) and Youth Project (grant number 71901155) of National Natural Science Foundation of China.

References

[1] Brennan, M.J., Henry Cao, H., Strong, N., Xu, X., 2005. The dynamics of international equity market expectations. Journal of Financial Economics 77, 257–288.

[2] Chen, W., Jiang, M., Zhang, W.G., Chen, Z., 2021. A novel graph convolutional feature based convolutional neural network for stock trend prediction. Information Sciences 556, 67–94.

[3] Chen, Y., Wei, Z., Huang, X., 2018. Incorporating corporation relationship via graph convolutional neural networks for stock price prediction, in: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Association for Computing Machinery, New York, NY, USA. p. 1655–1658.

[4] Cheng, D., Yang, F., Xiang, S., Liu, J., 2022. Financial time series forecasting with multi-modality graph neural network. Pattern Recognition 121, 108218.

[5] Cheng, R., 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, pp. 55–62.

[6] Chmielewski, L., Amin, R., Wannaphaschaiyong, A., Zhu, X., 2020. Network analysis of technology stocks using market correlation, in: 2020 IEEE International Conference on Knowledge Graph (ICKG).

[7] Cooper, G.F., Herskovits, E., 2004. A bayesian method for the induction of probabilistic networks from data. Machine Learning 9, 309–347.

[8] Cziraki, P., Mondria, J., Wu, T., 2021. Asymmetric attention and stock returns. Management Science 67, 48–71.

[9] Deng, S., Zhang, N., Zhang, W., Chen, J., Pan, J.Z., Chen, H., 2019. Knowledge-driven stock trend prediction and explanation via temporal convolutional network, in: Companion Proceedings of The 2019 World Wide Web Conference, Association for Computing Machinery, New York, NY, USA. p. 678–685.

[10] Etzioni, O., Cafarella, M., Banko, M., 2014. Open information extraction.

[11] Feng, F., He, X., Wang, X., Luo, C., Liu, Y., Chua, T.S., 2019. Temporal relational ranking for stock prediction. ACM Transactions on Information Systems (TOIS) 37.

[12] Feng, S., Xu, C., Zuo, Y., Chen, G., Lin, F., XiaHou, J., 2022. Relation-aware dynamic attributed graph attention network for stocks recommendation. Pattern Recognition 121, 108119.

[13] Fu, X., Ren, X., Mengshoel, O.J., Wu, X., 2018. Stochastic optimization for market return prediction using financial knowledge graph, in: 2018 IEEE International Conference on Big Knowledge (ICBK), pp. 25–32.

[14] Gao, J., Ying, X., Xu, C., Wang, J., Zhang, S., Li, Z., 2021. Graph-based stock recommendation by time-aware relational attention network. ACM Transactions on Knowledge Discovery from Data (TKDD) 16.

[15] Heckerman, D., Geiger, D., Chickering, D.M., 2004. Learning bayesian networks: The combination of knowledge and statistical data. Machine Learning 20, 197–243.

[16] Hosseini, S.S., Wormald, N., Tian, T., 2021. A weight-based information filtration algorithm for stock-correlation networks. Physica A: Statistical Mechanics and its Applications 563, 125489.

[17] Hou, K., 2007. Industry information diffusion and the lead-lag effect in stock returns. The Review of Financial Studies 20, 1113–1138.

[18] Hou, X., Wang, K., Zhong, C., Wei, Z., 2021. St-trader: A spatial-temporal deep neural network for modeling stock market movement. IEEE/CAA Journal of Automatica Sinica 8, 1015–1024.

[19] Lacasa, L., Luque, B., Ballesteros, F.J., Luque, J., Nuño, J.C., 2008. From time series to complex networks: The visibility graph. Proceedings of the National Academy of Sciences 105, 4972 – 4975.

[20] Lai, L., Li, C., Long, W., 2017. A new method for stock price prediction based on mrfs and ssvm, in: 2017 IEEE International Conference on Data Mining Workshops (ICDMW), pp. 818–823.

[21] Leung, C.K.S., MacKinnon, R.K., Wang, Y., 2014. A machine learning approach for stock price prediction, in: Proceedings of the 18th International Database Engineering & Applications Symposium, New York, NY, USA. p. 274–277.

[22] Li, S., Wu, J., Jiang, X., Xu, K., 2022. Chart gcn: Learning chart information with a graph convolutional network for stock movement prediction. Knowledge-Based Systems 248, 108842.

[23] Li, W., Bao, R., Harimoto, K., Chen, D., Su, Q., 2020. Modeling the stock relation with graph network for overnight stock movement prediction, in: Twenty-Ninth International Joint Conference on Artificial Intelligence and Seventeenth Pacific Rim International Conference on Artificial Intelligence IJCAI-PRICAI-20.

[24] Lo, A.W., MacKinlay, A.C., 1990. When are contrarian profits due to stock market overreaction? The Review of Financial Studies 3, 175–205.

[25] Long, J., Chen, Z., He, W., Wu, T., Ren, J., 2020. An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in chinese stock exchange market. Appl. Soft Comput. 91, 106205.

[26] Lozano-Perez, T., Wesley, M.A., 1979. An algorithm for planning collision-free paths among polyhedral obstacles. Commun. ACM 22, 560–570.

[27] Luque, B., Lacasa, L., Ballesteros, F.J., Luque, J., 2009. Horizontal visibility graphs: exact results for random time series. Physical review. E, Statistical, nonlinear, and soft matter physics 80 4 Pt 2, 046103.

[28] Moghadam, H.E., Mohammadi, T., Kashani, M.F., Shakeri, A., 2019. Complex networks analysis in iran stock market: The application of centrality. Physica A: Statistical Mechanics and its Applications 531, 121800.

[29] Pillay, K., Moodley, D., 2022. Exploring graph neural networks for stock market prediction on the jse, in: Artificial Intelligence Research, Springer International Publishing, Cham. pp. 95–110.

[30] Pirinsky, C., Wang, Q., 2006. Does corporate headquarters location matter for stock returns? The Journal of Finance 61, 1991–2015.

[31] Qi, Z., Bu, Z., Xiong, X., Sun, H., Cao, J., Zhang, C., 2019. A stock index prediction framework: Integrating technical and topological mesoscale indicators, in: 2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI), pp. 23–30.

[32] Roll, R., 1988. R2. The Journal of Finance 43, 541–566.
[33] Sawhney, R., Agarwal, S., Wadhwa, A., Shah, R., 2021. Exploring the scale-free nature of stock markets: Hyperbolic graph learning for algorithmic trading, in: Proceedings of the Web Conference 2021, pp. 11–22.
[34] Sawhney, R., Agarwal, S., Wadhwa, A., Shah, R.R., 2020. Spatiotemporal hypergraph convolution network for stock movement forecasting, in: 2020 IEEE International Conference on Data Mining (ICDM), pp. 482–491.
[35] Shirokikh, O., Pastukhov, G., Boginski, V., Butenko, S., 2013. Computational study of the us stock market evolution: a rank correlation-based network model. Computational Management Science 10, 81–103.
[36] Sil, A., Yates, A., 2013. Re-ranking for joint named-entity recognition and linking, in: Proceedings of the 22nd ACM International Conference on Information & Knowledge Management, Association for Computing Machinery, New York, NY, USA. p. 2369–2374.
[37] Thitaweera, N., Sinthupinyo, S., 2021. Correlation network analysis in the stock exchange of thailand (set), in: 2021 6th International Conference on Machine Learning Technologies, Association for Computing Machinery, New York, NY, USA. p. 170–176.
[38] Tingting, Z., Ningde, J., Zhongke, G., Yuebin, L., 2012. Limited penetrable visibility graph for establishing complex network from time series. Acta Physica Sinica 61, 030506.
[39] Wu, J., Xu, K., Chen, X., Li, S., Zhao, J., 2022. Price graphs: Utilizing the structural information of financial time series for stock prediction. Information Sciences 588, 405–424.
[40] Xu, C., Huang, H., Ying, X., Gao, J., Li, Z., Zhang, P., Xiao, J., Zhang, J., Luo, J., 2022. Hgnn: Hierarchical graph neural network for predicting the classification of price-limit-hitting stocks. Information Sciences 607, 783–798.
[41] Xu, H., 2022. Using kernel method to include firm correlation for stock price prediction. Computational Intelligence and Neuroscience 2022.
[42] Xu, J., Zhou, J., Jia, Y., Li, J., Hui, X., 2020. An adaptive master-slave regularized model for unexpected revenue prediction enhanced with alternative data, in: 2020 IEEE 36th International Conference on Data Engineering (ICDE), pp. 601–612.
[43] Xu, W., Liu, W., Xu, C., Bian, J., Yin, J., Liu, T.Y., 2021. Rest: Relational event-driven stock trend forecasting, in: Proceedings of the Web Conference 2021, Association for Computing Machinery, New York, NY, USA. p. 1–10.
[44] Ye, J., Zhao, J., Ye, K., Xu, C., 2021. Multi-graph convolutional network for relationship-driven stock movement prediction, in: 2020 25th International Conference on Pattern Recognition (ICPR), pp. 6702–6709.
[45] Yin, T., Liu, C., Ding, F., Feng, Z., Yuan, B., Zhang, N., 2022. Graph-based stock correlation and prediction for high-frequency trading systems. Pattern Recognition 122, 108209.
[46] Zhang, B., Yang, C., Zhang, H., Wang, Z., Sun, J., Wang, L., Zhao, Y., Wang, Y., 2022. Graph representation learning for similarity stocks analysis. Journal of Signal Processing Systems , 1–10.
[47] Zhang, L., Aggarwal, C., Qi, G.J., 2017a. Stock price prediction via discovering multi-frequency trading patterns, in: Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, pp. 2141–2149.
[48] Zhang, R., Ashuri, B., Deng, Y., 2017b. A novel method for forecasting time series based on fuzzy logic and visibility graph. Advances in Data Analysis and Classification 11, 759–783.
[49] Zhao, S., Tong, Y., Meng, X., Yang, X., Tan, S., 2016a. Predicting return reversal through a two-stage method. 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS) , 341–344.
[50] Zhao, S., Tong, Y., Wang, Z., Tan, S., 2016b. Identifying key drivers of return reversal with dynamical bayesian factor graph. PLoS ONE 11.
[51] Zhu, P., Cheng, D., Luo, S., Xu, R., Liang, Y., Luo, Y., 2022. Leveraging enterprise knowledge graph to infer web events' influences via self-supervised learning. Journal of Web Semantics 74, 100722.
[52] Zuo, Y., Harada, M., Mizuno, T., Kita, E., 2011. Application of bayesian network for nikkei stock return prediction, in: 2011 International Conference on Technologies and Applications of Artificial Intelligence, pp. 194–199.
[53] Zuo, Y., Harada, M., Mizuno, T., Kita, E., 2012. Bayesian network based prediction algorithm of stock price return, in: Intelligent Decision Technologies, Springer. pp. 397–406.