



# Momentum portfolio selection based on learning-to-rank algorithms with heterogeneous knowledge graphs

Mei-Chen Wu<sup>1</sup> · Szu-Hao Huang<sup>2</sup> · An-Pin Chen<sup>3</sup>

Accepted: 7 March 2024 / Published online: 22 March 2024  
© The Author(s) 2024

## Abstract

Artificial intelligence techniques for financial time series analysis have been used to enhance momentum trading methods. However, most previous studies, which have treated stocks as independent entities, have overlooked the significance of correlations among individual stocks, thus compromising portfolio performance. To address this gap, a momentum trading framework is proposed that combines heterogeneous data, such as corporate governance factors and financial domain knowledge, to model the relationships between stocks. Our approach involves adopting a knowledge graph embedding approach to map relations among heterogeneous relationships in the data, which is then utilized to train a multitask supervised learning approach based on a learning-to-rank algorithm. This method culminates in a robust portfolio selection method on the basis of the framework. Experimental results using data from the Taiwan Stock Exchange demonstrate that our proposed method outperforms traditional linear models and other machine learning methods in predictive ability. The investment portfolio constructed serves as an invaluable aid to investment decision-making.

**Keywords** Portfolio selection · Learning-to-rank algorithms · Momentum trading strategy · Heterogeneous knowledge graph · Financial time series analysis · Artificial intelligence in finance

## 1 Introduction

The selection of financial assets marks the beginning of most decision-making processes in directional investment and trading, where portfolio selection centers on striking the best balance between risk and reward. The behavioral finance literature has extensively documented numerous financial uncertainties and market anomalies, such as calendar effects and momentum anomalies, which may lead to excess returns.

Financial markets harbor unpredictable or challenging-to-assess risks and volatility. This uncertainty can stem from various factors, including political, economic, monetary policy, market demand and supply dynamics, among others. Excessive financial uncertainty leads to market sentiment swings, unreasonable asset price fluctuations, and investor overreactions, ultimately impacting the efficient functioning of markets and capital allocation. To account for such aberrations in behavior, scholars have explored the use of modern artificial intelligence techniques, particularly in quantitative trading. This study proposes a momentum portfolio selection algorithm that incorporates machine learning-based ranking methods. The system models numerous factors pertaining to the state of the financial market and that of the wider economy using heterogeneous knowledge graph (KG) approaches, effectively integrating information on firm operations and the financial market.

Statistical analysis has traditionally been used for financial forecasting in investment. However, traditional linear regression models rest on an assumption of error term independence, which often fails to hold in financial time series data. To address this limitation, several approaches have been proposed [1–4]. Linear regression has the advantages of

---

✉ Szu-Hao Huang  
szuhaohuang@nycu.edu.tw

Mei-Chen Wu  
y10201y.mg02@nycu.edu.tw

An-Pin Chen  
apc888888@gmail.com

<sup>1</sup> Institute of Information Management, National Yang Ming Chiao Tung University, Daxue Rd., Hsinchu 300093, Taiwan

<sup>2</sup> Department of Information Management and Finance, National Yang Ming Chiao Tung University, Daxue Rd., Hsinchu 300093, Taiwan

<sup>3</sup> Department of Finance, Asia University, Lioufeng Rd., Taichung 413305, Taiwan

being easily interpretable, being able to mitigate overfitting through regularization, and allowing for efficient updating using stochastic gradient descent. Nonetheless, it cannot handle nonlinear relationships and lacks flexibility in complex model identification, which compromises its robustness.

Deep learning has become more accessible and powerful with advances in computer hardware. Deep learning (DL), a type of machine learning (ML), employs multiple layers of linear or nonlinear transformations for data feature extraction, which allows for automatic feature extraction. Conversely, conventional ML approaches depend on human-generated algorithms for feature creation, necessitating expert input—this process is termed feature engineering. DL's ability to automatically extract data features, model complex nonlinear systems, and improve the learning for the model substantially enhances the model's capabilities. DL's applications in finance have expanded considerably in scope [5–7].

Traditional linear regression models, which base their forecasts on time series stock prices, aim to predict whether future stock returns are positive or negative. However, these models primarily focus on the price fluctuations of individual stocks, overlooking the relationships between them. Historically, most stock-related studies have treated each stock as an independent entity, neglecting the interdependence between stocks. However, the stock prices of companies in the same industry or supply chain often move together. Feng et al. [8] demonstrated the intricate relationships and informational interconnections between stocks. They employed the temporal graph convolutional (TGC) method to rank stocks, effectively capturing time-sensitive stock relationships. Their approach, which incorporated industry attributes and supply chain relationships, yielded promising results.

Beyond predicting future stock price returns or trends, studies have also focused on the behavior of ongoing trends. For instance, Jegadeesh and Titman [9] proposed the concept of the momentum effect, where a stock's future returns are similar to its historical returns. Thus, investors can capitalize on this momentum effect by buying stocks with high historical returns and selling those with low returns, a tactic referred to as momentum investment. In this strategy, a portfolio contains a basket of strongly performing assets and thus has low risk and high yields.

Most studies on stock price forecasting have focused solely on individual stock prices, often neglecting a heterogeneous variety of information. With regard to corporate governance, each company has major institutional shareholders responsible for directing and managing the company's affairs. These shareholders, tasked with achieving the company's mission and maximizing profits, play a crucial role in corporate governance. Their influence extends to the relationships between different companies, especially when they hold appointments in multiple companies simultaneously.

The interconnections among these individuals can also influence stock price trends.

When major institutional shareholders hold positions across multiple companies, they often jointly influence the stock price trends of those companies. Despite this, research on utilizing information about major institutional shareholders for prediction purposes remains limited. In this study, a heterogeneous KG is constructed to represent the relationship chain among the major institutional shareholders of each company. This graph is combined with time series stock price data, using DL technology to extract relationship information between individual stocks, thereby enhancing the predictive accuracy of the model. Finally, stocks are ranked based on the model's predictions, and a portfolio is constructed from the top and bottom percentiles of stocks. The contributions of this study are as follows:

- The proposal of a novel multitask supervised learning approach based on the learning-to-rank algorithm for identifying investment opportunities.
- The implementation of a KG embedding approach that incorporates company governance factors and financial domain knowledge to model heterogeneous relationships, aiming to enhance the accuracy of financial decision models.
- The development of a momentum-based trading strategy that combines heterogeneous information and ranking results to create an optimal investment portfolio.
- An empirical test of the framework using real-world trading market data, focusing on achieving the best Sharpe ratio and return, thereby enabling more effective portfolio allocation for investors.

The remainder of this study is structured as follows: Section 2 discusses the use of artificial intelligence (AI) in finance, explores the combination of AI trading strategies with momentum, and reviews the literature on heterogeneous knowledge graphs. Section 3 summarizes the proposed system, discusses abnormal market situations, and reviews the application of ranking in finance. This section also describes the data used in this study. Section 4 presents the data set and experimental design, including a comparison with various baseline methods. Section 5 concludes the study with a summary of its findings.

## 2 Related work

This section reviews the literature on the application of AI in the financial sector, the use of AI in momentum trading strategies, and heterogeneous knowledge graphs.

## 2.1 Applications of AI in finance

Linear and nonlinear methods for forecasting stock prices and trends have been formulated. DL technology has gained prominence in stock price analysis, resulting in notable advancements. For example, various DL models have been employed, such as long short-term memory (LSTM), recurrent neural networks, convolutional neural networks (CNN), and stacked autoencoders (SAE). These models often incorporate sliding window techniques to forecast stock prices and capture the underlying data dynamics [10–16]. DL methods have been applied to a range of financial problems beyond stock price prediction [17, 18]. The advantages of DL models include their ability to identify patterns in data, extract relevant features, achieve high levels of accuracy, and generalize well outside the training data.

Multitask learning, which processes multiple targets simultaneously, improves model performance by leveraging shared features among tasks. Mahmoud et al. [19] addressed the problem of data deficiency in personalized time series prediction by proposing a multitask framework with a novel convolutional recurrent neural network. Their approach utilized transfer learning for multitask learning models. Similarly, a study conducted in the same year focused on multitask time series forecasting using shared attention mechanisms, aiming to enhance the accuracy of forecasting multiple time series concurrently [20]. Furthermore, high degree of success in hybrid adaptive and pre-training multitask learning have also been reported [21, 22].

In addition to predicting stock price movements to maximize investor returns, certain studies have concentrated on ranking the projected prices and recommending top-K stocks based on these rankings, thereby aiding in risk diversification. For instance, Hsu et al. [23] proposed a recommendation model named financial graph attention networks (FinGAT) that leverages time series data on stock prices and sector information to generate return ratios and suggest the top-K profitable stocks for investors. Their model outperformed other models in evaluation experiments. Another article combined the concept of momentum with graph-based techniques to capture inherent relationships and dynamics within financial markets [24].

Thus, DL brings numerous advantages in financial forecasting; it can handle vast, complex financial data sets and autonomously learn and extract features. This study aimed to utilize multitask DL techniques to comprehend both linear and nonlinear relationships within stock data, thereby capturing additional implicit market behaviors and financial trading patterns.

## 2.2 AI Trading in momentum application

In a momentous 1993 study, Jegadeesh and Titman [9] demonstrated that a simple trading strategy of buying stocks

that have recently outperformed (winners) and selling those that have recently underperformed (losers) yielded substantial returns on the US stock market. This study garnered widespread attention because it challenged the prevailing notion of stock market efficiency. The then-dominant hypothesis stated that the stock market is completely efficient; this means that all market information has already been priced in, leaving no room for profitable trading opportunities. However, Jegadeesh and Titman's findings suggested that this trading strategy still yielded robust returns, indicating that the stock market is not completely efficient.

Enhanced momentum trading strategies were later proposed by Takeuchi and Lee [25], who applied DL techniques to extract features from historical stock prices to predict future returns. Kim [26] also applied DL to enhance the effectiveness of a momentum-based investment strategy within the stock market. The traditional strategy is one where selecting 10% of the winners and losers are selected. Kim tested this strategy against another one where only 10% of the winners are selected. Additionally, Kim proposed a new strategy where, if the expected return for the holding period is positive, a purchase is made for stocks in the top 10 percentile, and conversely, a sale is made for stocks in the bottom 10 percentile if the expected return is negative. Subsequent research integrated the learning-to-rank method to enhance cross-sectional momentum portfolios [27]. The results demonstrated that the ranking model significantly enhanced the trading performance of cross-sectional strategies, outperforming traditional approaches.

Traditional momentum-based strategies often struggle to leverage the vast and intricate information within the market, resulting in insufficient responses to rapidly changing market conditions. However, DL can be used to uncover latent patterns in the market by analyzing the complex interconnections embedded in the data, and thus yield higher returns in momentum trading. Subsequent studies have continued to enhance the predictive efficiency of momentum trading approaches. This study contributes to this effort by using DL to understand the interaction between momentum trading behavior and high-heterogeneity data. A pre-trained model was used for portfolio selection. In the following section, the foundational concepts and existing literature on heterogeneous KGs are detailed.

## 2.3 Heterogeneous KG Embedding

KG embedding primarily involves converting data into low-dimensional vector representations. These vectors are designed to represent semantic similarity and correlations between words or phrases, thereby aiding tasks in ML and DL. KG embedding has been applied in diverse domains [28], including finance [29–32].

Given interdependencies and price comovements between stocks, numerous studies have incorporated rich, real-world information about the relationships between companies to obtain more accurate stock price predictions. This information pertains to factors such as supplier and customer relationships, industry affiliation, and geographical location.

To enhance stock price predictions, researchers have incorporated information about a target company and its relationships with other companies. They constructed a graph using real-world data, such as suppliers, customers, industry, and location. Subsequently, the model was trained to develop a distributed representation of the nodes in the graph. The results indicated that the model accurately captured the relationships between corporations [33, 34]. Kim et al. [35] proposed a hierarchical attention network for stock prediction, named HATS, which consolidates information from various types of relationships to derive valuable node representations. HATS integrates pertinent company information into its learning process, enabling predictions not only of individual stock prices but also of broader trends in financial market indexes. Other scholars have conducted related research, contributing further to this field [36–38].

Incorporating heterogeneous knowledge graph embedding techniques in this study models the relationships between stocks and relevant factors such as corporate governance and financial domain knowledge, where KG embedding positions associated knowledge points in close proximity in a low-dimensional space, enhancing the efficiency of queries and matches of related knowledge. The subsequent chapter details the methodology.

### 3 Methodology

This study introduces a momentum portfolio selection algorithm that integrates machine-learning-based ranking methods. The system is designed to effectively utilize company operations data and domain expertise in the financial market. It models numerous finance and economics-related factors through heterogeneous KG approaches. This framework employs deep neural network (DNN) and CNN technologies to extract data features. The extracted features are then used to rank stocks within a given stock pool. Based on these rankings, the top and bottom percentiles of stocks are selected to construct a portfolio, which is subsequently evaluated for performance.

#### 3.1 Primary

In this section, the primary baseline methods compared in this study are introduced, such as the dual-classifier method by Huang et al. and the relational stock ranking (RSR) model by Feng et al.

Huang et al. [39] proposed binary dual-classifier algorithms designed to accurately model overreactions and strengthen portfolio composition using contrarian trading strategies. The method combines financial knowledge with high-dimensional nonlinear models constructed through ML techniques, enabling the identification of financial time series patterns. Initially, the model selects the best and worst-performing 10% of stocks from the current investing stock pool. These stocks are then fed into separate classifiers, and the positive and negative labels of the winners and losers inputted into a learning algorithm. This process updates the classifiers for the winners and losers. Subsequently, the outcomes of these classifiers are used to update the winner and loser portfolios. This approach is termed the dual-classifier method, and its framework is depicted in Fig. 1

This study introduces a multitask framework model for portfolio selection, which incorporates the dual-classifier concept proposed by Huang et al. and employs a ranking approach to enhance the predictive capability of the model. Unlike traditional trend prediction tasks, this selection method prioritizes stocks that are likely to sustain momentum in the future.

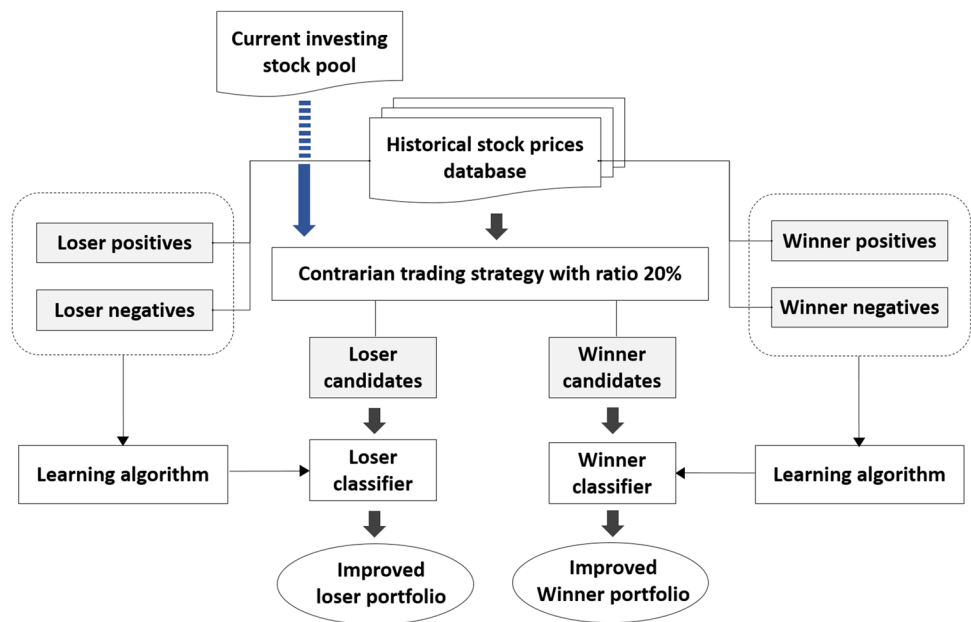
Feng et al., in 2019, [8] developed the RSR framework for predicting stock prices. This framework used the TGC method to capture the relationships between stocks. The RSR architecture is comprised of three layers: the sequential embedding layer, the relation embedding layer, and the prediction layer. Figure 2 illustrates the overall structure of the RSR framework.

The foundation of the RSR framework is the belief that past stock price changes significantly influence future ones. This belief is reflected in the use of a sequential model that identifies dependencies between each stock price sequence. The first layer of the model is an LSTM model, which is capable of retaining long-term sequence memory. The last hidden layer of the LSTM model serves as input for the subsequent layer. To capture the relationships between stocks, the model includes a relation embedding layer that incorporates industry attributes and supply chain relationships. The rationale is that if two companies are in the same industry or share a supply chain, their stock prices should exhibit similar trends or a transmission effect. The final prediction layer combines information from the sequential embedding layer and the relation embedding layer, concatenates them, and then outputs the predicted return ratio, which represents the ranking scores. These scores are then used to construct an investment portfolio.

The study used data from the NASDAQ and NYSE markets, covering the trading period from February 1, 2013 to August 12, 2017, and including 1026 and 1737 stocks, respectively. To evaluate the model's performance, three metrics were employed: mean square error (MSE), mean reciprocal rank (MRR), and cumulative investment return



Fig. 1 The Dual-classifier model framework



ratio (IRR). MSE is a widely used metric for evaluating regression tasks, whereas MRR is a metric for assessing ranking performance. The average reciprocal rank was computed for the stocks selected during the testing period. IRR, which directly reflects the effect of stock investment, was determined by summing the return ratios of the selected stocks. Smaller MSE values and larger MRR values indicate superior performance.

Table 1 presents the LSTM model as the baseline for RSR. Conversely, RankLSTM represents Feng et al.’s proposed RSR model. As indicated in the table, augmenting RankLSTM with industry relations yielded superior performance compared with the LSTM method alone. To explore the effect of integrating multiple relations, the RSR framework proposed by Feng et al. was augmented with multiple KG relations. Specifically, data from North American exchange-

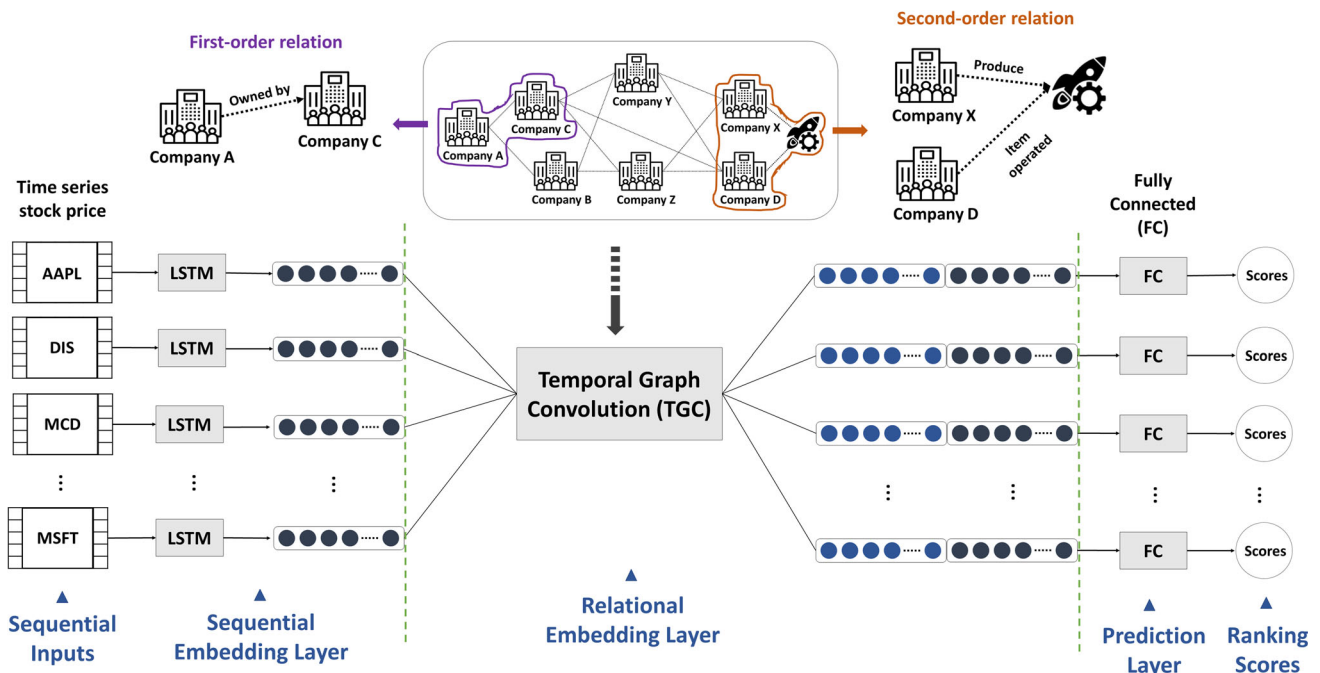


Fig. 2 Relational stock ranking framework

**Table 1** Results of the RSR model

	NASDAQ			NYSE		
	MSE	MRR	IRR	MSE	MRR	IRR
LSTM	0.000381	0.036400	0.130000	0.000231	0.027500	-0.900000
RankLSTM	0.000379	0.041700	0.680000	0.000228	0.037900	0.560000
RankLSTM+KG	0.000381	0.044700	1.490000	0.000228	0.044700	1.780000

traded funds (ETF) was introduced into the existing KG relation, comprising 1653 ETFs and 6620 stocks. ETFs are investment funds traded on exchanges, mirroring an index, commodity, bond, or a basket of assets. They offer diversified exposure, are bought and sold like stocks, providing investors access to various markets, sectors, or asset classes in a cost-effective and transparent manner. The outcomes, detailed in the RankLSTM+KG section of Table 1 indicate that the incorporation of multiple relations enhanced the correlation among stocks and significantly increased the IRR value.

Next, different market environments are examined. Taiwan's stock market is relatively small and volatile, which renders it susceptible to external factors. This financial uncertainty causes various fluctuations and unknown elements in the market, often influenced by political and economic news. As a result, unpredictable stock trends emerge, ultimately impacting the effective functioning of the market and capital allocation. Furthermore, the market is vulnerable to manipulation due to the substantial percentage of stocks held by major investors, some of whom may also hold positions within the companies in which they invest. Consequently, events such as elections or leadership changes within these companies can lead to substantial fluctuations in stock prices. Success in investing in Taiwan's stock market necessitates close monitoring of these external factors and an awareness of the potential for sudden and substantial changes.

In our study, the RSR model was employed to assess the performance of the Taiwan stock market using data from the Taiwan Economic Journal (TEJ) investment database. The same features as those employed by Feng et al were used, including closing prices and 5, 10, 20, and 30-day moving averages. These experimental tests are referred to as RankLSTM in Table 2. Additionally, fundamental corporate information for individual stocks was obtained and the effects of including industry relations were assessed, denoted as RankLSTM+KG (Industry) in the table. Data on

Board of directors, senior managers, and major institutional shareholders were also introduced for assessment, termed as RankLSTM+KG (Shareholder) in the table.

As illustrated in Table 2, the RankLSTM model proposed by Feng et al. achieved the best results in the top1, top5, and top10 evaluations. Its performance was further improved by augmenting the RankLSTM model with an industry KG beyond simply training on stock prices. Specifically, adding information on major institutional shareholders to the KG further enhanced the model's overall performance. This improvement is likely attributable to the model's enhanced ability to leverage intercompany information.

The experimental results clearly indicate that incorporating more meaningful information into the KG enhances the predictive capability of the model. The experiments also reveal that major institutional shareholders tend to influence market fluctuations in shallow-plate financial markets. Consequently, this study will expand upon this concept by employing heterogeneous KG approaches to learn the interrelationships among stocks, using corporate governance information relevant to major institutional shareholders as well as finance and economics-related factors. The trained KG embeddings are then integrated into the research framework to further augment its predictive abilities.

### 3.2 System overview

This study presents a novel and intelligent method for portfolio selection based on momentum portfolio selection, as illustrated in Fig. 3. The framework incorporates four types of data: time series data on individual stock prices, data on overall stock prices, information on stock correlations, and information on company operations represented by a KG. When extracting information for an individual stock, the system also gathers relevant data for other stocks. Each set of training data includes two individual stocks, their correlation with other stocks, the overall stock prices, and KG relations.

**Table 2** RSR model application in Taiwan's stock market

	MSE	MRR	IRR_1	IRR_5	IRR_10
RankLSTM	0.000742	0.081700	15.730000	5.030000	3.270000
RankLSTM+KG (Industry)	0.000723	0.085500	22.340000	6.390000	4.120000
RankLSTM+KG (Shareholder)	0.000715	0.093000	24.040000	6.770000	4.190000

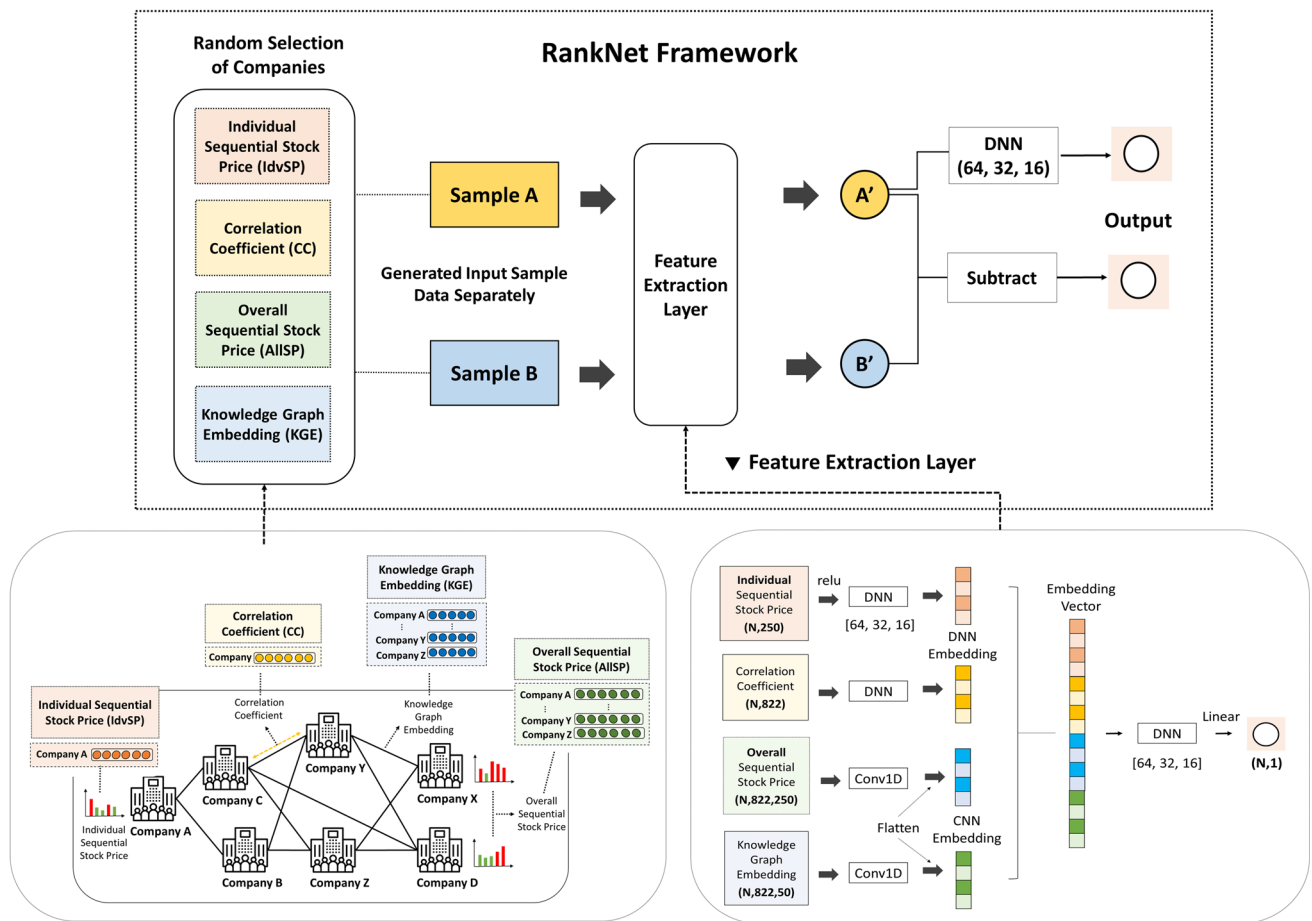


Fig. 3 System overview of the proposed model

To process this data, a DNN is employed for analyzing time series stock price and stock correlation data. A one-dimensional convolutional neural network (1DCNN) is used to reduce the dimensionality of overall stock price data and KG embeddings. The neural network training is based on the concept of RankNet. It processes the input data, and the final score is obtained by subtracting the values of the two samples at the subtraction stage. Finally, stocks are ranked according to these scoring results.

### 3.3 KG Construction

As previously mentioned, our framework employs KG embedding to encapsulate structured information about various entities and their interrelations. This section aims to present a comprehensive overview of the KG embedding techniques used in our proposed framework, encompassing definitions, notations, and fundamental concepts.

Knowledge graphs have evolved from semantic networks, which initially served to represent knowledge in a graph format. In semantic networks, nodes represent concepts and edges represent the relationships between these concepts.

The collection of all entities, denoted as  $E = \{e_1, \dots, e_E\}$ , and the relationships between these entities, identified as  $R = \{r_1, \dots, r_R\}$ , together constitute the knowledge base, represented as  $KG = \{E, R\}$ . Additionally, a set of triplets  $S = \{h, r, t\}$  was generated by combining the head entity, relation, and tail entity, respectively. In this study, attention was directed towards seven types of entities and eight types of relations. The seven types of entities are listed as follows:

1. **Company:** This study includes 822 companies listed on the Taiwan stock market.
2. **Shareholder code:** The identification codes of major institutional shareholder members for each of the 822 companies were obtained from the TEJ database.
3. **Shareholder group name:** The shareholder group code and name.
4. **Address:** The city where each company is located.
5. **Securities firm:** The firm entrusted by each target company to handle securities sales.
6. **Accounting firm:** The accounting firm tasked with managing the financial affairs of each target company.

7. **Industry category:** The category to which each target company belongs.

The eight types of relations can be divided into two categories: those between companies and those related to shareholders. These relations are defined as follows:

1. **Identity:** This relation between the company and shareholder code identifies the shareholders in the target company.
2. **Control category:** This relation describes the ownership classification of shareholders, which includes five categories: ultimate controllers, managers, group managers, friendly groups, and external parties.
3. **Entrusted entity of:** This relation between the company and accounting firm describes the accounting firm engaged by the target company.
4. **Sale of:** The relation between the company and securities firm indicates the securities firm that lists the target company's stock for sale.
5. **Located in:** This relation describes the city in which the company is located.
6. **Industry of:** This relation between the company and industry category denotes the industry to which the company belongs.
7. **Shareholder of:** This relation between the company and shareholder code denotes the shareholders of company.
8. **Represent:** This relation delineates the different shareholder codes within a shareholder group name.

The relationships within the proposed heterogeneous KG stem from the aforementioned seven entities and eight relationships. These relationships are depicted in Fig. 4.

Table 3 presents the abstract codifications of these entities and relations on the basis of their definitions. The first column describes the entities and relations as obtained from the data set, which are then mapped to corresponding triplets in the second column. These triplets help us learn the embeddings in the KG for each entity and relation, which is pivotal for the tasks discussed in the subsequent sections.

### 3.4 Training process and detail

Before the RankNet framework is applied, our strategy selects pools of stocks that have exhibited the best and worst returns over a certain period for constructing the investment portfolio. The process starts by gathering the monthly returns of all stocks for the past 24 months. From this data, the top and bottom 10% of stocks are identified based on their returns during the portfolio formation period. Operational and financial data for these companies is then acquired, which includes stock prices, correlation coefficients, and KG embeddings containing heterogeneous information.

The study utilizes data on the four following pieces of information: individual sequential stock prices, overall sequential stock prices, correlation coefficients, and KG embeddings. The RankNet method is applied to rank the stocks and to construct an optimal stock portfolio based on the model's ranking ability. The four types of input data are defined as follows:

- **Individual sequential stock price(IdxSP):** The daily stock prices of the target company over the past year are used to obtain the return:

$$r_t = \frac{SP_t - SP_{t-1}}{SP_{t-1}} \quad (1)$$

where  $t$  represents the current time,  $SP$  represents the stock price, and  $t - 1$  represents the time on the previous day.

- **Overall sequential stock price(AllSP):** The daily stock prices of the target company and other companies over the past year.
- **Correlation coefficient(CC):** The CC formula uses Pearson correlation coefficient to obtain the correlation between the target company and other companies, where the monthly return of the past 6 months is used as a comparison value:

$$CC = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}} \quad (2)$$

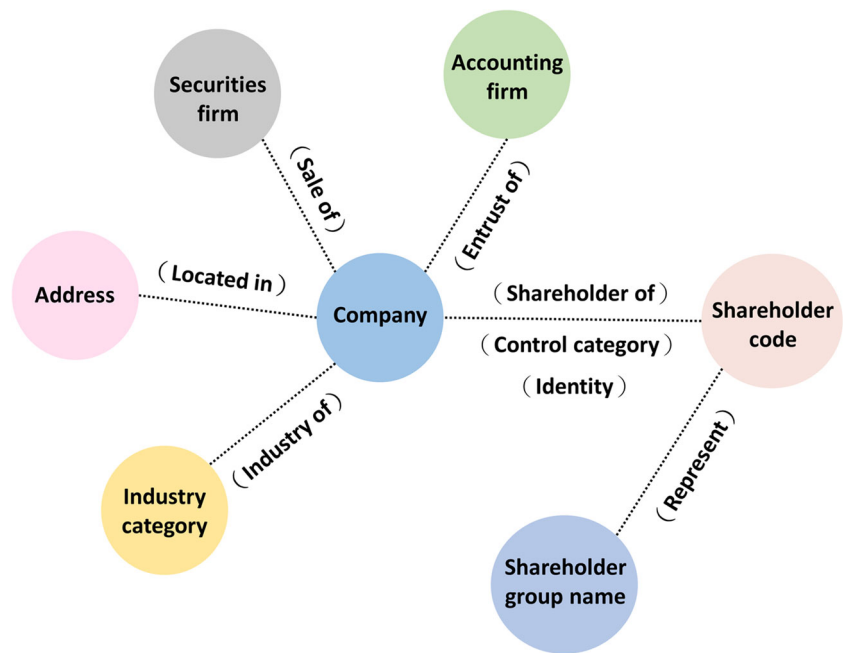
where  $n$  represents the number of pairs of stocks,  $\sum xy$  represents the sum of products of paired stocks.  $\sum x$  and  $\sum y$  represents the sums of the  $x$  and  $y$  scores.  $\sum x^2$  and  $\sum y^2$  represents the sums of the squared  $x$  and  $y$  scores.

- **Knowledge graph embedding(KGE):** This entails learning low-dimensional representations of entities and relations in a KG while preserving semantic meaning.
- **RankNet:** Our approach randomly samples two stocks and sorts them based on the size of the predicted output result.
- **Target:** This paper proposes a multi-task framework model that integrates a RankNet output for ranking two samples with an additional output for predicting future trends. Monthly returns serve as the target variable for both ranking and trend prediction tasks.

For IdvSP and CC, a three-layer DNN [64, 32, 16] is utilized for processing, whereas for AllSP and KGE, a three-layer 1DCNN [64, 32, 16] is employed for dimensionality reduction to obtain embeddings. During training, to ensure that data values are neither excessively high nor low, stock



**Fig. 4** Diagram of our heterogeneous knowledge graph



prices are standardized using (1) to scale values within [0, 1]. The model incorporates various heterogeneous data as input features, resulting in complex data characteristics prone to overfitting issues. To address this, two parameter settings were introduced: Dropout and Early Stopping. In order to mitigate overfitting during training, a Dropout rate of 0.5 is employed. Furthermore, "Early Stopping" is implemented to monitor the performance on the validation set and halt training prematurely in case of performance abnormalities. This measure also aids in preventing overfitting during training. In determining the learning rate value, rates of 0.1, 0.01, 0.001, and 0.0001 were experimented with. The results revealed that excessive learning rates led to unstable training, while rates that were too small resulted in the model's slow convergence. Based on these findings, a learning rate of 0.001 was chosen for subsequent experiments in this paper.

Regarding Batch Size configuration, sizes of 64, 128, 256, and 512 were tested. During the testing process, it was observed that smaller batch sizes led to slow convergence on

complex features, while larger ones required more memory and computational resources. The study also found that the accuracy of Batch Size 128 and 256 was very similar, but the computational resource consumption of Batch Size 512 was significant. Therefore, subsequent research adopted a batch size of 256 for consistency. In optimizer parameter settings, tests were conducted on "SGD", "Adagrad", "RMSprop", and "Adam". After multiple rounds of iterative testing, stable training results were consistently delivered by "Adam", surpassing "SGD" and "Adagrad". The latter exhibited unstable loss functions and poor accuracy predictions during training. While "RMSprop" mostly achieved good results, its performance fluctuated. The adaptive adjustment of learning rates for each parameter in "Adam" made it the optimal choice, consistently producing favorable outcomes after numerous tests.

Concerning the use of Activation functions, the model utilized "Sigmoid", "ReLU", and "Linear". "ReLU", which truncates negative values to zero, was suitable for the train-

**Table 3** The proposed method input KG triplets data

Description	Triples
Shareholder's identity	(Company, Identity, Shareholder code)
Shareholder's control category	(Company, Control category, Shareholder code)
Entrusted accounting firm	(Company, Entrust of, Accounting firm)
Entrusted securities firm	(Company, Sale of, Securities firm)
Company's location in a city	(Company, Located in, Address)
Company's industry category	(Company, Industry of, Industry category)
Company's shareholder	(Company, Shareholder of, Shareholder code)
Shareholder's group name	(Shareholder code, Represent, Shareholder group name)

ing data. ‘‘Sigmoid’’, mapping inputs to values between 0 and 1, was employed in the RankNet training framework. ‘‘Linear’’ was used for the final output in the framework. To prevent the problem of dying ReLU, avoiding a high learning rate and a larger negative bias, this paper adopted a learning rate of 0.001, which also helps mitigate the occurrence of this phenomenon. The remaining parameters are configured as follows: loss=‘‘binary crossentropy’’, and metrics=‘‘accuracy’’. The detailed architecture of the network, including layer numbers, configuration settings, and feature quantities are presented in Table 4 within this framework. Subsequent sections further elucidate the research procedure and provide detailed pseudocode, as presented in Algorithm 1.

After implementing the specified parameter settings and conducting model calibration on the necessary financial data, ready for recall experimental results. To ensure stability and reliability, several training iterations are typically performed. Additionally, certain financial data features incorporated into the model may require authorization from relevant institutions, potentially leading to some minor limitations in data acquisition. Following the acquisition of the required data features and parameter settings, the model undergoes multiple training iterations, typically around five times. The results generated from the execution closely align with the data presented in the paper.

### 3.5 Market anomalies

Market anomalies frequently occur in real financial markets and are typically categorized into three main types: time series anomalies, cross-sectional anomalies, and other market pricing anomalies. Time series anomalies refer to market volatility that fluctuates predictably over time. For example, the January effect, a calendar anomaly or the turn-of-the-year effect, is often associated with additional stock market rises in January. Weekly patterns also exist, such as the tendency for stocks to exhibit greater movement on Fridays than on

#### Algorithm 1 Proposed method to training process and detail

**Require:** The individual sequential daily stock price data,  $r_t$ ; The correlation coefficient data,  $CC$ ; The overall sequential daily stock price data,  $Allr_t$ ; The knowledge graph embedding,  $KGE$ ;  
**Ensure:** The RankNet predicted probabilities,  $P_{ij}$ ; The trend prediction,  $Pred_i$ ;  
**for** number of training epochs **do**  
    Obtain RankNet randomly Sample data  $U_i$  and  $U_j$ ,  
     $U_i$  and  $U_j$  contains randomly obtained data  $r_t, CC, Allr_t, KGE$ ;  
     $r_t$  and  $CC$  data are inputs DNN model respectively, and obtained DNN embedding  $\bar{r}_t$  and  $\bar{CC}$ .  
     $Allr_t, KGE$  data are inputs 1DCNN model respectively, and obtained 1DCNN embedding  $\overline{Allr}_t$  and  $\overline{KGE}$ .  
    Concat  $\bar{r}_t, \bar{CC}, \overline{Allr}_t$  and  $\overline{KGE}$  embedding are into DNN model, and obtain RankNet score  $s_i$  and  $s_j$ .  
    Concat  $\bar{r}_t, \bar{CC}, \overline{Allr}_t$  and  $\overline{KGE}$  embedding are into DNN model, and obtain trend prediction values  $Pred_i$ .  
**end for**

Mondays, with the market often closing higher on Fridays. Tung et al. [40] studied the constituent stocks of the S&P 500 index to explore the effect anomaly. They focused on the phenomenon where the daily return on Monday is typically lower than on the preceding Friday, known as the ‘‘week effect.’’ This study improves upon traditional trend forecasting and observation of buying and selling signals by applying AI methods and time series modeling to identify factors that accentuate daily effects. A predictive model for the occurrence of day-of-the-week effects is constructed through a five-stage experimental process. The results indicate that the constituents of the S&P 500 index exhibit a day-of-the-week effect.

Another anomaly is the momentum and overreaction anomaly, where the market tends to overreact, leading to stock prices continuing their prior trends. Jegadeesh and Titman [9] proposed the momentum effect, which suggests that stocks will persist in their recent directional trend. This observation suggests that stocks exhibiting higher recent returns are likely to maintain their performance going forward. Conversely, the reversal effect suggests that stocks will move

**Table 4** Number of layers, configuration settings, and feature quantities in the neural network

Category	Training algorithm	Number of neurons	Network structure	Activation function		Number of neurons
				Hidden layer	Output	
IdvSP	DNN	(N, 250)	64-32-16	relu	relu	(N, 16)
CC	DNN	(N, 164)	64-32-16	relu	relu	(N, 16)
AllSP	1DCNN	(N, 164, 250)	64-32-16	relu	relu	(N, 304)
KGE	1DCNN	(N, 164, 50)	64-32-16	relu	relu	(N, 304)
Sample A IdvSP+CC +AllSP+KGE	DNN	(N, 640)	64-32-16	relu	linear	(N, 1)
Sample B IdvSP+CC +AllSP+KGE	DNN	(N, 640)	64-32-16	relu	linear	(N, 1)
Sample A - Sample B	Subtract	(N, 1)-(N, 1)	X	X	sigmoid	(N, 1)
Trend prediction	DNN	(N, 640)	64-32-16	relu	linear	(N, 1)

opposite to their original trend in the future, marking turning point and a contrarian response.

In addition, anomalies in financial markets can be detected using financial time series data. Cheong et al. [41] introduced the spatiotemporal convolutional neural network-based relational network (STCNN-RN) to capture intricate correlations among multiple financial time series data sets and detect anomalies. To identify outliers in specific companies, genetic algorithms are employed. Studies have often fallen short in providing comprehensive explanations of these anomalies to investors. In this study, an interpretability model is employed to shed light on the timing of these company-specific anomalies and uncover the critical factors contributing to them. Empirical evidence from our experimental supports the efficacy of the proposed model in modeling various financial time series data sets and accurately detecting anomalies within related firms.

### 3.6 RankNet construction

RankNet is a neural network model proposed by Microsoft Research in 2005 for learning-to-rank. It transforms the ranking problem into a binary classification problem, aiming to predict the probability that one item is ranked higher than another. RankNet calculates the probability of the correct order between two documents and thereby simplifies the process for a sequence of documents by requiring probability calculations only for pairs of adjacent documents, reducing computational complexity. RankNet, which has been employed in various applications such as search engines, recommendation systems, and stock portfolio selection, consists of a multilayer perceptron (MLP) with two inputs and one output. The inputs are feature vectors representing two items to be compared, and the output is the probability that the first item ranks higher than the second.

The RankNet architecture requires defining two types of relevance probabilities for ranking: the predicted relevance probability and the true relevance probability. These probabilities are represented by the formula as illustrated in (3) and (4).

$$P_{ij} = P(U_i > U_j) = \frac{1}{1 + e^{-\sigma(s_i - s_j)}} \tag{3}$$

$$\overline{P}_{ij} = \frac{1}{2}(1 + (s_{ij})) \tag{4}$$

$U_i$  and  $U_j$  represent any sample pair. The parameter  $\sigma$  shapes the sigmoid function.  $s_i$  and  $s_j$  are the respective score output by the model.

RankNet evaluates the quality of a ranking by analyzing the relative relationships among multiple documents. The quality is higher when fewer pairs have incorrect relative relationships, which when a model incorrectly ranks

$U_i$  ahead of  $U_j$ , despite the true label indicating the opposite. RankNet aims to minimize the number of such incorrect pairs. When formulated as a cost function, RankNet incorporates the notion of probability, focusing on the likelihood  $P$  that  $U_i$  is ranked ahead of  $U_j$ , and it aims to minimize the difference between this predicted probability and the true probability. Finally, RankNet employs the cross-entropy cost function to quantify the level of fitting. The cost function formula is displayed in (5).

$$C = -\overline{P}_{ij} \log P_{ij} - (1 - \overline{P}_{ij}) \log(1 - \overline{P}_{ij}) \tag{5}$$

To maintain consistency, RankNet ensures that if  $U_i$  is more relevant than  $U_j$ , and  $U_j$  is more relevant than  $U_k$ , then  $U_i$  should also be more relevant than  $U_k$ . The true probability of  $U_i$  versus  $U_k$  is calculated using the true probabilities of  $U_i$  versus  $U_j$  and  $U_j$  versus  $U_k$ , as expressed in (6). The cross-entropy loss function is utilized during training by RankNet to minimize the difference between the predicted and actual rankings.

$$\overline{P}_{ik} = \frac{\overline{P}_{ij} \overline{P}_{jk}}{1 + 2\overline{P}_{ij} \overline{P}_{jk} - \overline{P}_{ij} - \overline{P}_{jk}} \tag{6}$$

## 4 Experimental Results

This chapter introduces the dataset and experimental design employed in this study. It then proceeds to compare various baseline methods, including traditional random choice, momentum-based, dual-classifier model, MLP, and the RSR model. Finally, the effectiveness of different methods is evaluated using performance indicators.

### 4.1 Dataset and experimental setup

Data from the TEJ investment database was analyzed, with a focus on companies listed on the Taiwan stock market. As of 2022, there were 971 listed companies in the Taiwan stock market. However, companies that delisted during the study period or had considerable missing data were excluded, leaving a total of 822 stocks for analysis. The period covered by the research data is from January 1, 2015, to December 31, 2022. A rolling window approach was employed, utilizing the past year's data as learning features to predict the monthly return for the subsequent month. This approach was applied in both RankNet and future trend forecasting. The data pertain to closing stock price, company operations, major institutional shareholders, and inter-company correlations. The closing price is used to calculate the return. The operational information and shareholder information includes the company ticker (Tic), shareholder code (denoting identity and control), shareholder group name, address,

securities firm, accounting firm, and industry category. This information contributes to construction of a KG, and the monthly returns of each stock over the past 6 months are used to determine the correlation between stocks.

The data are divided into four categories for input into the neural network, which utilizes the rectified linear unit (ReLU) activation function. The details of each type of input data are as follows:

- Individual sequential stock price (IdvSP): The return of the target company in the past year is obtained, and DNN [64, 32, 16] is used to extract features and generate embedding after dimensionality reduction.
- Correlation coefficient(CC): The correlation between the target company and other companies is obtained, using the monthly return in the past 6 months as a comparison value. DNN [64, 32, 16] is used to reduce dimensionality and generate embedding.
- Overall sequential stock price(AllSP): The return of the target company and other companies in the past year is obtained, and 1DCNN [64, 32, 16] is used to extract features, followed by generating embedding after dimensionality reduction.
- Knowledge graph embedding(KGE): Information such as Tic, shareholder code (denoting identity and control), shareholder group name, address, securities firm, accounting firm, and industry are used to construct a KG. This graph illustrates the relationship between all companies and yields embeddings that represent these relationships.

For the company with Taiwan stock market Tic 1101.tw, Table 5 lists the top 10 records used. “Shareholder code” is the identifier for major institutional shareholders. “Identity” denotes the role of these shareholders within the company. “Control” is categorized as follows: “A” for ultimate controllers, exerting the most significant influence over the company’s management and resource allocation decisions; “B” for managers, “C” for group manager, “L” for friendly groups, encompassing blood relations, marital relations, partnerships, mergers, divisions, or state ownership, and “X” for external individuals. “Shareholder group name” identifies the group to which the major institutional shareholders belong. “Address” denotes the city where the company is located. “Securities firm” identifies the entrusted securities firm, and “Accounting firm” refers to the appointed accounting firm handling the company’s financial affairs. “Industry” specifies the company’s business category.

Taking the company with Tic 1101.tw as an example, the shareholder code “P000160025” is the code used to indicate independent directors and supervisors in 1101.tw. For the same period, the same code “P000160025” also indicates the same for 2535.tw and 2617.tw. This study

incorporates these relationships into the training model to enhance predictive performance by leveraging the inter-linkages among individual stocks.

After obtaining the four mentioned input data types, samples are randomly generated for input into the RankNet architecture. RankNet aims to sort two randomly selected stocks from the samples based on the magnitude of the prediction outcome. The target of this study is to predict the monthly return for the next month.

## 4.2 Performance metrics

Several key indicators, including the Sharpe ratio, Calmar ratio, maximum drawdown (MDD), and Sortino ratio, are utilized in this study to evaluate performance.

The Sharpe ratio is a measure of the amount of excess return earned for each unit of risk. It is commonly used to evaluate the performance of long-term investments such as funds and asset allocations. The Sharpe ratio is defined as follows:

$$\text{SharpeRatio} = \frac{R_p - R_f}{\sigma_p} \quad (7)$$

$R_p$  and  $R_f$  represent the portfolio’s return and risk-free rate, respectively.  $\sigma_p$  represents the standard deviation of the portfolio’s excess return.

The Calmar ratio quantifies the relationship between returns and the maximum drawdown, which is calculated by dividing the annualized rate of return by the historical maximum drawdown. A higher Calmar ratio indicates more favorable fund performance. The formula for the Calmar ratio is as follows:

$$\text{CalmarRatio} = \frac{R_p - R_f}{MDD} \quad (8)$$

$R_p$  and  $R_f$  represent the portfolio’s return and risk-free rate, respectively.  $MDD$  represents the maximum drawdown.

The maximum drawdown measures the largest decline in an account’s net value from its peak. It represents the worst possible scenario for an investment entered at any point in time.

The Sortino ratio evaluates a portfolio’s performance and stability in fund investment or asset allocation. It quantifies the amount of return achievable per unit of risk taken.

$$\text{SortinoRatio} = \frac{R_p - R_f}{\sigma_d} \quad (9)$$

$R_p$  and  $R_f$  represent the portfolio’s return and risk-free rate, respectively.  $\sigma_d$  represents the standard deviation of negative returns.

**Table 5** The company and director-supervisor information for 1101.tw

Tic	Shareholder code	Identity	Control	Shareholder group name	Address	Securities firm	Accounting firm	Industry
1	1101 C000005671	Major shareholder	A	G1101	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
2	1101 C000000275	Major shareholder	X	G2883	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
3	1101 P000160025	Independent board directors	X	nan	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
4	1101 C000000322	Major shareholder	L	G2891	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
5	1101 P000225583	Independent board directors	X	nan	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
6	1101 P038753241	Independent board directors	X	nan	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
7	1101 P035630059	Independent board directors	X	nan	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
8	1101 C000012386	Board of directors and supervisors	A	G1101	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
9	1101 C000000249	Major shareholder	X	G2882	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
10	1101 C000007180	Major shareholder	L	G2504	Taipei	ChinaTrust Commercial Bank	Deloitte & Touche	M11A
7153	2535 P000160025	Board of directors and supervisors	X	nan	Taipei	JihSun Holdings	Deloitte & Touche	M25B
7606	2617 P000160025	Board of directors and supervisors	X	nan	Taipei	Taishin Securities	Deloitte & Touche	M26A



The main difference between the Sortino ratio and the Sharpe ratio lies in how they account for risk. The Sortino ratio focuses solely on downside risk, measuring the excess return per unit of downside risk (represented by the standard deviation of negative returns). This reflects the return earned for each unit of risk resulting in a loss. Conversely, the Sharpe ratio considers both the downside and upside risk and quantifies the additional return per unit of total risk, as indicated by the standard deviation of all returns.

### 4.3 Performance evaluation

In this section, various baseline methods are compared over a rolling test period from 2018 to 2021. These methods include the random choice method, traditional momentum-based method, dual-classifier model, MLP, RSR, and our proposed method.

Before comparing these methods, stocks exhibiting momentum are identified. Tests at various time intervals, ranging from the 6th to the 24th month at 6-month intervals, are conducted for the formation and holding periods in momentum trading. The results reveal that the optimal profitability has 24-month formation and holding periods; thus, these periods were used in subsequent experiments. Next, the concepts and practices of the different baseline methods are outlined. In the random choice method, 10% of stocks are randomly selected from either the positive or negative monthly return stock pools. Decisions to assume long or short positions, are based on positive or negative returns, respectively, and performance is evaluated based on daily returns over a 24-month holding period. In the traditional momentum-based method, with a 24-month formation period used as a reference, individual stock returns are ranked. Subsequently, the top and bottom 10% of returns are selected to form stock pools. Long positions are taken based on positive returns, whereas short positions are based on negative returns for trading. Finally, the performance of the investment portfolio is then evaluated using daily returns over a 24-month holding period. The distinction between this method and the random choice method lies in the selection criteria: the random choice method involves selecting stocks randomly from the stock pool, whereas the momentum-based method involves selecting the best and worst-performing stocks from the pool to create long and short investment portfolios.

The dual-classifier model involves constructing a dual-classifier through the application of the momentum contrarian strategy. The procedure for selecting the top and bottom performing stocks mirrors that of the previous baseline methods. After the initial long-short investment portfolio is established, candidates for stock selection are chosen based on past contrarian trading results and future trend persistence as defined by binary targets using monthly return as a criterion. The classifier, which utilizes support vector machines

(SVM), is tested. After the dual-classifier is trained, its results inform adjustments to the investment portfolio. In the MLP method, the SVM approach of the dual-classifier model is replaced with a four-layer neural network [32, 16, 8, 1]. The sigmoid activation function is used for the first layer, and the monthly return is the target. Stocks are sorted based on the predicted results, and the top and bottom 10% are selected to construct the investment portfolio.

The final baseline method, RSR, involves inputting sequential historical prices into LSTM and incorporating KG embedding similar to the proposed method. The experimental settings follow the guidelines of the original study, with a historical sequence length of 4 for features, 64 hidden units in LSTM, and “leaky\_relu” as the final layer’s output activation, with a learning rate of 0.001. After the model makes its predictions, the best and worst-performing stock pools are selected for portfolio evaluation. For a conceptual understanding of the RSR model, please refer to Chapter 3.1.

As for the proposed method, consistent with the intervals used in the baseline methods, this experiment extracts monthly returns from  $T$  to  $T-24$  as the formation period to screen the top and bottom 10% of stock pools and identify candidate stocks. The training data for these candidate include data on daily returns for the past 250 days, KG embedding after dimensionality reduction, similarity calculated from monthly returns over the past 6 months between two stocks, and the daily return of all candidate stocks. All these data are inputted into the training model. These heterogeneous data are processed through CNN and DNN for dimensionality reduction and embedding generation to map different data types to the same dimension, achieving the goal of combining heterogeneous data. For detailed network parameter settings, refer to Chapter 3.4.

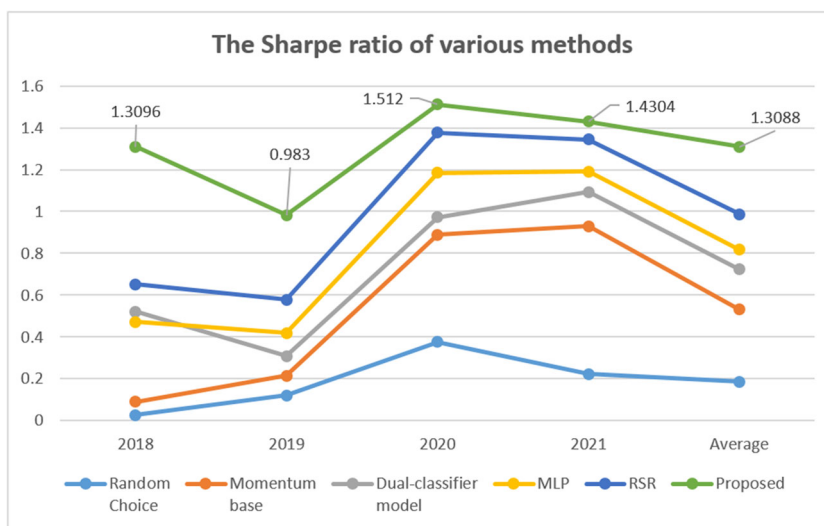
In the experiment of this study, data were obtained from the TEJ database. Stocks with missing data or significant anomalies were excluded. In cases where minor missing or anomalous data were encountered, zero values were substituted. Table 4 presents experiments conducted over different time intervals from 2018 to 2021. Each interval included two years of training samples and one year of defined targets. To prevent overfitting of the training data set due to the limited number of training samples, an extensive number of samples were generated from the candidate stocks. For instance, when employing a long or short strategy with 160 candidate stocks,  $160 \times 159 = 25440$  samples were generated, each comprising the four types of heterogeneous data previously mentioned.

The experimental results, as illustrated in Table 6, indicate that the random choice method, which involves randomly selecting stocks based on positive and negative returns, consistently exhibited the poorest performance each year, as indicated by the lowest Sharpe ratio values. The traditional momentum-based method, which selects 10% of

**Table 6** Results of different testing years in 2018-2021

	Model	2018	2019	2020	2021	Average
Random Choice	Cumulative Returns	0.0007	0.0069	0.0236	0.0176	0.0122
	Annual Return	0.0004	0.0035	0.0122	0.0090	0.0063
	Annual Volatility	0.0346	0.0343	0.0336	0.0446	0.0368
	Sharpe Ratio	0.0280	0.1202	0.3759	0.2234	0.1869
	Calmar Ratio	0.0116	0.0971	0.1964	0.1691	0.1186
	MDD	-0.0323	-0.0365	-0.0619	-0.0533	-0.0460
	Sortino Ratio	0.0402	0.1720	0.5679	0.3357	0.2790
Momentum base(1993)	Cumulative Returns	0.0041	0.0108	0.0591	0.0584	0.0331
	Annual Return	0.0021	0.0056	0.0302	0.0297	0.0169
	Annual Volatility	0.0272	0.0276	0.0341	0.0320	0.0302
	Sharpe Ratio	0.0904	0.2146	0.8886	0.9290	0.5307
	Calmar Ratio	0.0388	0.1311	0.8737	1.0830	0.5317
	MDD	-0.0539	-0.0424	-0.0345	-0.0274	-0.0396
	Sortino Ratio	0.1304	0.3074	1.2624	1.3893	0.7724
Dual-classifier model(2011)	Cumulative Returns	0.1257	0.0651	0.0798	0.0917	0.0906
	Annual Return	0.0624	0.0330	0.0405	0.0463	0.0456
	Annual Volatility	0.1332	0.1346	0.0418	0.0422	0.0880
	Sharpe Ratio	0.5219	0.3095	0.9721	1.0919	0.7239
	Calmar Ratio	0.2720	0.1446	0.6906	0.8251	0.4831
	MDD	-0.2295	-0.2284	-0.0587	-0.0561	-0.1432
	Sortino Ratio	0.6562	0.3893	1.3548	1.6058	1.0015
MLP	Cumulative Returns	0.0483	0.0239	0.0745	0.0807	0.0569
	Annual Return	0.0244	0.0123	0.0379	0.0408	0.0289
	Annual Volatility	0.0541	0.0301	0.0318	0.0341	0.0375
	Sharpe Ratio	0.4724	0.4196	1.1848	1.1904	0.8168
	Calmar Ratio	0.2494	0.2717	1.1025	1.4703	0.7735
	MDD	-0.0978	-0.0451	-0.0344	-0.0278	-0.0513
	Sortino Ratio	0.6647	0.6311	1.7499	1.7722	1.2045
RSR(2019)	Cumulative Returns	0.0675	0.1076	0.3422	0.5693	0.2716
	Annual Return	0.0339	0.0541	0.1645	0.2614	0.1285
	Annual Volatility	0.0535	0.0998	0.1155	0.1859	0.1137
	Sharpe Ratio	0.6512	0.5782	1.3766	1.3435	0.9874
	Calmar Ratio	0.4464	0.3046	0.6614	0.9211	0.5834
	MDD	-0.0760	-0.1775	-0.2487	-0.2838	-0.1965
	Sortino Ratio	0.8809	0.7140	1.7780	1.7581	1.2827
Proposed	Cumulative Returns	0.1367	0.1116	0.2049	0.2295	0.1707
	Annual Return	0.0677	0.0561	0.1012	0.1123	0.0843
	Annual Volatility	0.0510	0.0571	0.0652	0.0765	0.0625
	Sharpe Ratio	<b>1.3096</b>	<b>0.9830</b>	<b>1.5120</b>	<b>1.4304</b>	<b>1.3088</b>
	Calmar Ratio	1.0705	0.8829	0.9522	0.9950	0.9752
	MDD	-0.0632	-0.0635	-0.1063	-0.1129	-0.0865
	Sortino Ratio	1.9046	1.3694	2.1228	2.0825	1.8698

**Fig. 5** Comparing the Sharpe ratio results of different periods between the baseline and the proposed method



stocks from positive and negative return pools to construct investment portfolios, yielded a superior Sharpe ratio than the random choice method. The dual-classifier model refined the traditional momentum-based stock selection method by applying binary classification for readjusting the selected stock group’s portfolio, resulting in a higher Sharpe ratio compared with the traditional approach.

Using the MLP method, the test initially selected the top 10% of stocks using the traditional momentum-based approach and then applied a multilayer neural network to further adjust the investment portfolio. Except for a slight decrease in the Sharpe ratio in 2018, the performance from 2019 to 2021 surpassed that of the dual-classifier model. The RSR method, which incorporated the graph relationships between stocks, enhanced the model’s predictive capabilities. The experimental findings indicate that the RSR method outperformed the previous methods in terms of Sharpe ratio from 2018 to 2021.

Figure 5 illustrates a comparison of Sharpe ratio results across different periods between baseline methods and the proposed method. The figure demonstrates that the random choice and traditional momentum-based methods yield inferior Sharpe ratio results. Conversely, our proposed method, which incorporates heterogeneous data such as company operations data and financial domain knowledge, not only

enhances the model’s ranking capability but also achieves the highest performance in terms of Sharpe ratio values.

#### 4.4 Ablation study

This section presents an ablation analysis of our proposed method. Table 7 presents the results for the test year 2020. Four inputs, namely IdvSP, CC, AllSP, and KGE, were analyzed.

Initially, when using IdvSP and AllSP as inputs for the same period, the Sharpe ratio value is positive at 0.9703, indicating robust forecasting ability from stock price data alone. Adding the correlation CC between individual stocks reveals its effect on the model, improving predictive ability and yielding a Sharpe ratio of 1.0405. Incorporating company operations information KGE further improves the Sharpe ratio to 1.1772, indicating enhanced predictive power with additional stock information. When the model incorporates IdvSP, CC, and KGE, the Sharpe ratio reaches 1.2955, indicating improved predictive ability. Including all data types, the model achieves a Sharpe ratio of 1.5120, indicating improved predictive ability with comprehensive knowledge.

Applying the same method, Table 8 presents the test results for the year 2021. Initially, with only IdvSP and AllSP as

**Table 7** Results of the ablation study in 2020

Model	Cumulative Returns	Annual Return	Annual Volatility	Sharpe Ratio	Calmar Ratio	MDD	Sortino Ratio
IdvSP+AllSP	0.1240	0.0623	0.0645	0.9703	0.6250	-0.0997	1.3850
IdvSP+AllSP+CC	0.1198	0.0603	0.0579	1.0405	0.8722	-0.0691	1.4574
IdvSP+AllSP+KGE	0.1303	0.0654	0.0551	1.1772	0.9647	-0.0678	1.6650
IdvSP+KGE+CC	0.1785	0.0887	0.0674	1.2955	0.7995	-0.1110	1.8009
IdvSP+AllSP+KGE+CC	0.2049	0.1012	0.0652	1.5120	0.9522	-0.1063	2.1228

**Table 8** Results of the ablation study in 2021

Model	Cumulative Returns	Annual Return	Annual Volatility	Sharpe Ratio	Calmar Ratio	MDD	Sortino Ratio
IdvSP+AllSP	0.1373	0.0685	0.0760	0.9105	0.5657	-0.1211	1.2728
IdvSP+AllSP+CC	0.2060	0.1014	0.0878	1.1442	0.7018	-0.1444	1.6224
IdvSP+AllSP+KGE	0.2844	0.1377	0.1069	1.2598	0.7820	-0.1760	1.7525
IdvSP+KGE+CC	0.3682	0.1753	0.1286	1.3213	0.8199	-0.2139	1.8090
IdvSP+AllSP+KGE+CC	0.2295	0.1123	0.0765	1.4304	0.9950	-0.1129	2.0825

inputs, the Sharpe ratio is 0.9105. However, including CC significantly enhances predictive power, resulting in a Sharpe ratio of 1.1442. Furthermore, adding KGE further improves the model, increasing the Sharpe ratio to 1.2598. Incorporating IdvSP, CC, and KGE together leads to a strong Sharpe ratio of 1.3213. Finally, with all data types inputted, the model's learning and predictive abilities further improve, as evidenced by our proposed method's effective stock price prediction.

Figure 6 visually represents the results of the ablation study for 2020 and 2021. The graph depicts that all Sharpe ratios are positive, with the highest value achieved when all four types of information are included. This implies that the model's learning ability and predictive capabilities improve with the incorporation of more comprehensive information.

#### 4.5 Top-N percentage evaluation

Next, the predictive ability of our model across different top-N percentages is discussed. Experiments are conducted using 10%, 15%, and 20% of stocks for the years 2020 and 2021. The experimental results are presented in Table 9. The experimental results using 2020 data indicate that with 10%, 15%, and 20% of stock included, the obtained Sharpe ratios are positive, amounting to 1.5120, 1.3991, and 1.2636, respectively. These findings indicate that our model has strong predictive ability at any selected percentage of stocks. Specif-

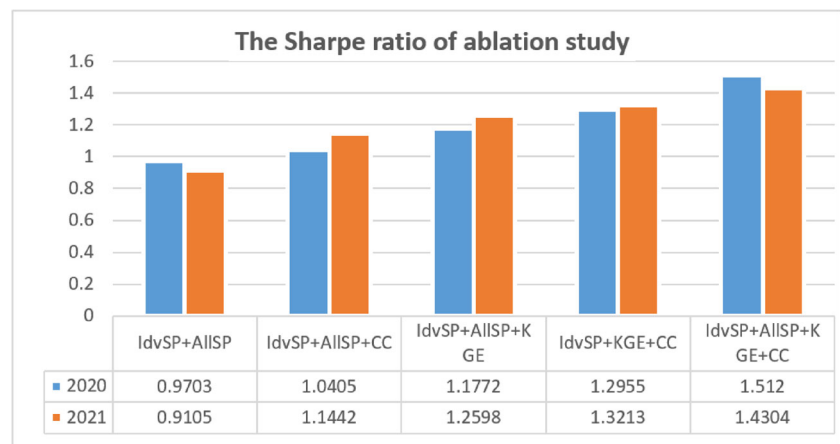
ically, the best experimental results are obtained when using 10% of the stocks for prediction.

The experimental results using 2021 data indicate that with 10%, 15%, and 20% of stock included, the obtained Sharpe ratios are positive, amounting to 1.4304, 0.0926, and 1.1973, respectively. Again, the best experimental results are obtained when using 10% of the stocks for prediction. Based on these results, it is concluded that the experiments using 10% of the stocks consistently delivered the best performance, establishing 10% as the baseline for all experiments in this study.

#### 4.6 Different periods of formation and holding period results

This section examines the performance of our model over different formation and holding periods. The results are presented in Table 10. Profit performance for formation and holding periods of 6, 12, 18, and 24 months was tested, with these tests conducted for both 2020 and 2021. For 2020, the Sharpe ratio for the short-term 6-month period is negative, with a value of -0.9522. However, extending the holding period to 12 months results in a positive Sharpe ratio of 0.8575. With holding periods of 18 and 24 months, the Sharpe ratios are positive, at 0.6845 and 1.3949, respectively. Applying the same experimental design to test 2021, the results indicate that the short-term 6-month Sharpe ratio is negative

**Fig. 6** Illustrates the comparison of Sharpe ratio outcomes derived from the proposed method's ablation study



**Table 9** Results of extract top-N percent stock

Year	Top-N (%)	Cumulative Returns	Annual Return	Annual Volatility	Sharpe Ratio	Calmar Ratio	MDD	Sortino Ratio
2020	10%	0.2049	0.1012	0.0652	1.5120	0.9522	-0.1063	2.1228
	15%	0.1708	0.0850	0.0596	1.3991	1.4457	-0.0588	2.0384
	20%	0.1050	0.0530	0.0416	1.2636	1.6088	-0.0330	1.9857
2021	10%	0.2295	0.1123	0.0765	1.4304	0.9950	-0.1129	2.0825
	15%	0.0080	0.0041	0.0734	0.0926	0.0422	-0.0976	0.1285
	20%	0.1723	0.0854	0.0705	1.1973	0.7753	-0.1101	1.6745

at -0.8322. For holding periods of 12, 18, and 24 months, the Sharpe ratios are positive, with values of 0.6261, 0.7938, and 1.4304, respectively.

Based on these findings, a limited backtesting period of only 6 months for the formation and holding periods leads to unstable profit performance. For formation and holding periods of 12, 18, and 24 months, the Sharpe ratios are consistently positive. Notably, the highest and most positive Sharpe ratio is observed when the period is set to 24 months. Therefore, for all subsequent experiments, this study uses 24 months as the baseline period.

#### 4.7 Testing different trading markets

This study selected NASDAQ as the test market to verify the performance of different trading markets. NASDAQ is a prominent U.S. stock exchange known for its technology and biotech listings. It offers a wide range of securities, from traditional stocks to innovative offerings. The research data covers the period from January 1, 2015 to December 31, 2022, divided into four time intervals for testing. The input features mirror those used in testing the Taiwan stock market, comprising daily returns from the past year, monthly returns from the previous 6 months, daily returns of other stocks over the past year, and financial information-based KG. Notably, there's a difference in the KG information due to difficulties

in obtaining board of directors and supervisors' information in the U.S. stock market. Hence, data was obtained from 158 ETFs, totaling 1,652 constituent stocks. Each ETF includes multiple constituent stocks, establishing associations across multiple ETFs akin to Taiwan's board of directors and supervisors' information.

The training dataset exhibits a significant correlation with the experimental results. When the dataset is segmented into different time intervals or influenced by the interactive effects of the prevailing environment, uncertainty factors emerge, leading to noticeable variations in the training outcomes. Examining Table 10 across different testing periods of 6, 12, 18, and 24 months, the experimental results reveal that 24 months yield the optimal outcomes for different testing years. Conversely, the shortest period of 6 months produces the poorest results, indicating that different segmentation intervals impact the experimental outcomes.

Tables 6 and 11 present similar findings, indicating that the selection of different data features results in varying impacts. RSR model, as shown in the tables, utilizes only stock prices, industry classification, and relationships with upstream and downstream sectors as input features. Despite achieving positive returns, its profitability slightly lags behind the Proposed model. In this study incorporates diverse data features, allowing it to learn more information and exhibit higher sensitivity to environmental changes. Coupled with our care-

**Table 10** Results of Different Period

Year	Different Period	Cumulative Returns	Annual Return	Annual Volatility	Sharpe Ratio	Calmar Ratio	MDD	Sortino Ratio
2020	6	-0.1468	-0.0789	0.0827	-0.9522	-0.2796	-0.2820	-1.4778
	12	0.0857	0.0435	0.0512	0.8575	0.4253	-0.1022	1.2311
	18	0.0756	0.0384	0.0575	0.6845	0.3591	-0.1071	0.9850
	24	0.1674	0.0834	0.0587	1.3949	0.9953	-0.0838	1.9695
2021	6	-0.1783	-0.0963	0.1139	-0.8322	-0.2710	-0.3552	-1.2596
	12	0.1086	0.0546	0.0916	0.6261	0.3557	-0.1534	0.8415
	18	0.1123	0.0564	0.0724	0.7938	0.5960	-0.0946	1.1031
	24	0.2295	0.1123	0.0765	1.4304	0.9950	-0.1129	2.0825



**Table 11** Results of NASDAQ testing across different years from 2018 to 2021

Model	Metrics	2018	2019	2020	2021	Average
RSR(2019)	Cumulative Returns	0.0020	0.0043	0.0058	0.0019	0.0035
	Annual Return	0.0010	0.0021	0.0029	0.0009	0.0017
	Annual Volatility	0.0013	0.0024	0.0034	0.0012	0.0021
	Sharpe Ratio	0.8086	0.8732	0.8472	0.8156	0.8362
	Calmar Ratio	0.5448	0.5898	0.6128	0.7266	0.6185
	MDD	-0.0019	-0.0036	-0.0047	-0.0013	-0.0029
	Sortino Ratio	1.1946	1.2671	1.2291	1.1709	1.2154
Proposed	Cumulative Returns	0.0033	0.0040	0.0031	0.0035	0.0035
	Annual Return	0.0189	0.0306	0.0419	0.0341	0.0314
	Annual Volatility	0.0149	0.0224	0.0345	0.0200	0.0229
	Sharpe Ratio	<b>1.2656</b>	<b>1.3588</b>	<b>1.2074</b>	<b>1.6898</b>	<b>1.3804</b>
	Calmar Ratio	3.4082	2.1316	3.0697	8.4455	4.2638
	MDD	-0.0056	-0.0143	-0.0137	-0.0040	-0.0094
	Sortino Ratio	2.2960	1.8468	1.5982	2.6250	2.0915

fully selected parameter configurations, it achieves superior predictive capabilities. The proposed model yields positive returns in different intervals, specifically 1.2656, 1.3588, 1.2074, and 1.6898, compared to the RSR model. This highlights that with more comprehensive input features, the proposed model provides more assistance during training, resulting in favorable outcomes.

The experimental framework, parameter settings, and network layers mirror those used in testing the Taiwan stock market. In addition, this study includes the RSR model for comparison, utilizing the same feature parameters as described in Section 3.1. As shown in Table 11, this demonstrates commendable predictive capabilities when tested in different time intervals within the NASDAQ trading markets.

## 5 Conclusion

This study proposes a novel multitask supervised learning approach using a learning-to-rank algorithm to construct a portfolio selection framework in momentum trading. This framework integrates heterogeneous data, including company operations data and major institutional shareholder details, forming a relationship chain graph. This integration, coupled with time series stock price data, aims to address a gap in the literature on intelligent trading strategies, where the focus has been on individual stock price changes.

The experimental results demonstrate the strong performance of our model in ablation analysis when all stock information is included, allowing the model to obtain more

knowledge and improve its predictive ability. The obtained Sharpe ratios are favorable. In the top-N percent experiment, the Sharpe ratios for 10%, 15%, and 20% are all positive, confirming the model's robust predictive ability at any selected percentage. The most effective results occur with 10%. Our model, when its parameters are optimized, outperforms other baseline methods for the 2018-2021 period. This model can efficiently select the top-N percent of stocks demonstrating momentum behavior from a massive stock pool and recommend these stocks for portfolio inclusion. It also offers valuable portfolio adjustment suggestions to investors.

Future studies can (1) explore different KG approaches to construct embeddings using various node measurement methods and (2) investigate the use of different learning-to-rank algorithms beyond RankNet for testing.

Our belief is that combining heterogeneous data to construct momentum trading strategies, coupled with learning-to-rank algorithms for ranking, can further enhance

**Author Contributions** All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Mei-Chen Wu, Szu-Hao Huang and An-Pin Chen. The first draft of the manuscript was written by Mei-Chen Wu and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** Open Access funding enabled and organized by National Yang Ming Chiao Tung University. This work was supported in part by the Ministry of Science and Technology, Taiwan, under Contract MOST 111-2221-E-A49-125-MY3; and in part by the Financial Technology (FinTech) Innovation Research Center, National Yang Ming Chiao Tung University.

**Data availability and access** The data that support the findings of this study are available from the TEJ database (<https://www.tejwin.com/en/>)

but restrictions apply to the availability of these data, which were used under licence for the current study, and so are not publicly available.

## Declarations

**Ethical and informed consent for data used** This study did not involve any ethical and informed consent issues.

**Competing Interests** All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

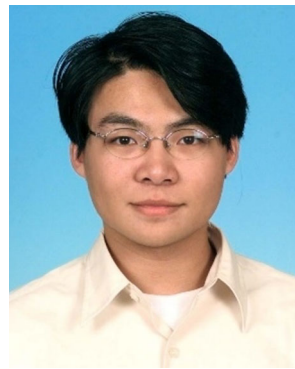
- Box GE, Jenkins GM (1976) Time series analysis: forecasting and control. Holden Day, San Francisco
- Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica J Econometric Soc* pp 987–1007
- Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics* 31(3):307–327
- Nelson DB (1991) Conditional heteroskedasticity in asset returns: A new approach. *Econometrica J Econometric Soc* pp 347–370
- Dixon M, Klabjan D, Bang JH (2015, November) Implementing deep neural networks for financial market prediction on the Intel Xeon Phi. In: Proceedings of the 8th workshop on high performance computational finance pp 1–6
- Peng Y, Jiang H (2015) Leverage financial news to predict stock price movements using word embeddings and deep neural networks. arXiv preprint [arXiv:1506.07220](https://arxiv.org/abs/1506.07220)
- Chen YY, Chen CT, Sang CY, Yang YC, Huang SH (2021) Adversarial attacks against reinforcement learning-based portfolio management strategy. *IEEE Access* 9:50667–50685
- Feng F, He X, Wang X, Luo C, Liu Y, Chua TS (2019) Temporal relational ranking for stock prediction. *ACM Trans Inform Syst (TOIS)* 37(2):1–30
- Jegadeesh N, Titman S (1993) Returns to Buying Winners and Selling Losers: Implications for. *J Financ* 48(1):65–91
- Selvin S, Vinayakumar R, Gopalakrishnan EA, Menon VK, Soman KP (2017, September) Stock price prediction using LSTM, RNN and CNN-sliding window model. In: 2017 International conference on advances in computing, communications and informatics (icacci), IEEE, pp 1643–1647
- Bao W, Yue J, Rao Y (2017) A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS ONE* 12(7):e0180944
- Hossain MA, Karim R, Thulasiram R, Bruce ND, Wang Y (2018, November) Hybrid deep learning model for stock price prediction. In: 2018 IEEE Symposium series on computational intelligence (ssci), IEEE, pp 1837–1844
- Chen L, Chi Y, Guan Y, Fan J (2019, May) A hybrid attention-based EMD-LSTM model for financial time series prediction. In: 2019 2nd International conference on artificial intelligence and big data (ICAIBD), IEEE, pp 113–118
- Nguyen DHD, Tran LP, Nguyen V (2019) Predicting stock prices using dynamic LSTM models. In: Applied informatics: second international conference, ICAI 2019, Madrid, Spain, Proceedings 2 pp 199–212. Springer International Publishing. Accessed 7–9 Nov 2019
- Yu P, Yan X (2020) Stock price prediction based on deep neural networks. *Neural Comput Appl* 32:1609–1628
- Mehtab S, Sen J (2020) A time series analysis-based stock price prediction using machine learning and deep learning models. *Int J Bus Forecast Mark Intell* 6(4):272–335
- Wu ME, Syu JH, Lin JCW, Ho JM (2021) Portfolio management system in equity market neutral using reinforcement learning. *Appl Intell* pp 1–13
- Hung TY, Huang SH (2022) Addressing the cold-start problem of recommendation systems for financial products by using few-shot deep learning. *Appl Intell* 52(13):15529–15546
- Mahmoud RA, Hajj H, Karamah FN (2020) A systematic approach to multi-task learning from time-series data. *Appl Soft Comput* 96:106586
- Chen Z, Jiawe E, Zhang X, Sheng H, Cheng, X (2020) Multi-task time series forecasting with shared attention. In: 2020 International conference on data mining workshops (ICDMW), IEEE, pp 917–925
- McDermott M, Nestor B, Kim E, Zhang W, Goldenberg A, Szolovits P, Ghassemi M (2020) A comprehensive evaluation of multi-task learning and multi-task pre-training on ehr time-series data. arXiv preprint [arXiv:2007.10185](https://arxiv.org/abs/2007.10185)
- Mirfallah Lialestani SP, Parcerisa D, Himi M, Abbaszadeh Shahri A (2022) Generating 3D geothermal maps in Catalonia, Spain using a hybrid adaptive multitask deep learning procedure. *Energies* 15(13):4602
- Hsu YL, Tsai YC, Li CT (2021) FinGAT: Financial Graph Attention Networks for Recommending Top-K Profitable Stocks. *IEEE Trans Knowl Data Eng* 35(1):469–481
- Zhang S, Ma X, Fang Z, Pan H, Yang G, Arce GR (2023) Financial time series forecasting based on momentum-driven graph signal processing. *Appl Intell* pp 1–17
- Takeuchi L, Lee YYA (2013) Applying deep learning to enhance momentum trading strategies in stocks. In Technical Report. Stanford University, Stanford, CA, USA
- Kim S (2019) Enhancing the momentum strategy through deep regression. *Quantitative Finance* 19(7):1121–1133
- Poh D, Lim B, Zohren S, Roberts S (2021) Building cross-sectional systematic strategies by learning to rank. *The Journal of Financial Data Science* 3(2):70–86
- Yang YC, Chen CT, Lu TY, Huang SH (2023) Hierarchical Reinforcement Learning for Conversational Recommendation With Knowledge Graph Reasoning and Heterogeneous Questions. *IEEE Trans Serv Comput*
- Zehra S, Mohsin SFM, Wasi S, Jami SI, Siddiqui MS, Syed MKURR (2021) Financial knowledge graph based financial report query system. *IEEE Access* 9:69766–69782
- Hsu PY, Chen CT, Chou C, Huang SH (2022) Explainable mutual fund recommendation system developed based on knowledge graph embeddings. *Appl Intell* pp 1–26
- Chou YC, Chen CT, Huang SH (2022) Modeling behavior sequence for personalized fund recommendation with graphical deep collaborative filtering. *Expert Syst Appl* 192:116311

32. Zheng J, Li Y (2023) Machine learning model of tax arrears prediction based on knowledge graph. *Electronic Research Archive* 31(7):4057–4076
33. Chen Y, Wei Z, Huang X (2018, October) 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* pp 1655–1658
34. Matsunaga D, Suzumura T, Takahashi T (2019) Exploring graph neural networks for stock market predictions with rolling window analysis. *arXiv preprint arXiv:1909.10660*
35. Kim R, So CH, Jeong M, Lee S, Kim J, Kang J (2019) Hats: A hierarchical graph attention network for stock movement prediction. *arXiv preprint arXiv:1908.07999*
36. 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
37. Chen W, Jiang M, Zhang WG, Chen Z (2021) A novel graph convolutional feature based convolutional neural network for stock trend prediction. *Inf Sci* 556:67–94
38. Feng S, Xu C, Zuo Y, Chen G, Lin F, Xiahou J (2022) Relationship-aware dynamic attributed graph attention network for stocks recommendation. *Pattern Recogn* 121:108119
39. Huang SH, Lai SH, Tai SH (2011) A learning-based contrarian trading strategy via a dual-classifier model. *ACM Trans Intell Syst Technol (TIST)* 2(3):1–20
40. Tung HH, Cheng CC, Chen YY, Chen YF, Huang SH, Chen AP. Binary classification and data analysis for modeling calendar anomalies in financial markets. In: *2016 7th International conference on cloud computing and big data (CCBD)*, IEEE, pp 116–121
41. Cheong MS, Wu MC, Huang SH (2021) Interpretable stock anomaly detection based on spatio-temporal relation networks with genetic algorithm. *IEEE Access* 9:68302–68319

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Mei-Chen Wu** received the B.S. and M.S. degrees in information management from Yu Da University of Science and Technology, Miaoli, Taiwan, in 2012 and 2013, respectively. She is currently pursuing the Ph.D. degree in information management at National Yang Ming Chiao Tung University, Hsinchu, Taiwan. Her research interests include information hiding, digital watermarking, deep learning, artificial intelligence, and financial technology.



**Szu-Hao Huang** received the B.E. and Ph.D. degrees in computer science from National Tsing Hua University, Hsinchu, Taiwan, in 2001 and 2009, respectively. He is currently an Associate Professor with the Department of Information Management and Finance and the Chief Director of the Financial Technology (FinTech) Innovation Research Center, National Yang Ming Chiao Tung University. He is also the Principal Investigator of the Financial Technology Innovation

Industrial–Academic Alliance and several cooperation projects with leading companies in Taiwan. He has authored more than 70 papers published in the related international journals and conferences. His research interests include trustworthy AI, recommendation systems, computer vision, information security, and reinforcement learning for trading.



**An-Pin Chen** received his Ph.D. in Industrial & Systems Engineering from the University of Southern California. He currently serves as a Chair Professor in the Department of Finance at Asia University, Taiwan. His research focuses on finance decision support systems, data mining, financial engineering, knowledge management, and computational intelligence applications in finance. In 2020, he was honored as a Fellow of the Chinese Society for Management of Technology.