International Journal of Computational Intelligence Systems (2024) 17:12
https://doi.org/10.1007/s44196-023-00394-4

RESEARCH ARTICLE

GAN-Enhanced Nonlinear Fusion Model for Stock Price Prediction

Yingcheng Xu[1] · Yunfeng Zhang[1] · Peide Liu[2] · Qiuyue Zhang[2] · Yuqi Zuo[1]

Received: 22 October 2023 / Accepted: 11 December 2023

Abstract

Stock price prediction is a significant field of finance research for both academics and practitioners. Numerous studies have proved that the stock movement can be fully reflect various internal features of stock price including non-stationary behavior, high persistence in the conditional variance. The fusion of time-series prediction model such as Auto-Regressive Integrated Moving Average (ARIMA) and neural network is an availability but difficult approach for stock price prediction. Although the orientation has been studied through some methods in different research, there are still difficulties with the poor capture ability of time-series features and insufficient effectiveness of integrating temporal feature and frequency domain information. In this paper, we propose a Generative Adversarial Network (GAN) framework with the Convolution Neural Networks (CNN) as the discriminator and a hybrid model as the generator for forecasting the stock price. The hybrid model includes Attention-based Convolution Neural Networks (ACNN), Long Short-Term Memory (LSTM), and ARIMA model. Moreover, this proposed framework uses the Generative Adversarial patten and Attention Mechanism to achieve effective analysis and feature extraction for stock price movement. The extensive experiments in different history periods of dataset demonstrate an improvement in forecasting of stock price using our model as compared to the baseline models.

Keywords Stock prediction · Time-series prediction · Generative Adversarial Network · Hybrid model

1 Introduction

Stock market is one of the vital financial markets in current economic system which is gradually developing into rationality. The Efficient Market Hypothesis (EMH) [1] posits that stock price can comprehensively reflect relevant information of assets in a well-functioning stock market. However, the intricate nonlinear relationships among political factors, economic movement, social events, and investor behavior have a complex impact on stock prices, which makes stock market prediction a highly challenging task [2]. In the early research, many investors have focused on the direction of fundamental analysis [3] and technical analysis [4] to maximize stock profits. Fundamental analysis typically assumes that stock data are generated precisely by intricate financial systems [5]. Technical analysis often relies on the analyst's subjective judgment, disregarding other objective principle in stock market [6].Therefore, fundamental analysis and technical analysis tend to neglect the stock market sentiment and the real value of the company.

Statistical methods generally presume that stock data in sophisticated financial systems are generated by a linear and stationary process [7]. For example, Auto-Regressive Integrated Moving Average (ARIMA) [8] model and Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) [9] model have produced excellent effect in dealing with noise in data. With the flourishing of Machine Learning (ML) algorithm [10, 11], researchers have found superior prediction results by capturing nonlinear relationships in complex financial market data. The current popular ML techniques for predicting stock market data mainly include Genetic Algorithm (GA), fuzzy logic, and Support Vector Machine (SVM), which reduce prespecified data assumptions such as linearity, stationarity, homoscedasticity, or normality [12–14]. Although statistical methods have a certain effect on noise reduction, and machine learning models can fit the nonlinear characteristics of time series, they insufficiently analyze frequency domain information.

✉ Yunfeng Zhang
yfzhang@sdufe.edu.cn

1 School of Computer Science and Technology, Shandong University of Finance and Economics, Yaojia Sub-district Office, Jinan 250014, Shandong, China

2 School of Management Science and Engineering, Shandong University of Finance and Economics, Yaojia Sub-district Office, Jinan 250014, Shandong, China

Published online: 18 January 2024

In recent researches, Deep Learning (DL) models have been introduced to capture time and frequency domain features in stock data [15]. For example, Zahra and Mohamed [16] used Simple Recurrent Neural Network (SRNN) model to analyze and fit global stock price for a short term. Haiyao Wang et al. [17] focused on improving the ability to extract deep features that influence stock price in historical periods through Convolution Neural Networks (CNN). Ashish et al. [18] proposed a generalized Generative Adversarial Network (GAN) model which includes a generator and a discriminator for adversarial training to achieve good predictive performance. However, stock price data in different time periods reflect different market characteristics and behaviors [19]. The hybrid models integrating different methods in a nonlinear way have become a new research focus. Wenjie Lu [20] used a hybrid model integrating CNN model and Long Short-Term Memory (LSTM) model to extract the time features from historical data and make predictions for essentially improving the accuracy of stock prediction. Zhuangwei Shi et al. [21] have focused on improving the ability to capture key data features in historical periods, combining Attention Mechanisms with Convolution Neural Networks (CNN). Moreover, Generative Adversarial Network (GAN) has been successfully applied in many fields including image generation and conversion, which has led toward to conclude that GAN performs well in time-series prediction [18]. Domenico and Luca [22] suggest reducing the dimension of data in the GAN model so that it can generate low-dimensional sequence data. Eugene et al. [23] use multiple dimension key stock data to increase the data dimension for satisfying the prerequisite of the GAN model.

Most stock price predictive researches attempted to establish a direct connection between stock price and macro-external factors through time features. However, the nonlinear relationships among political policy, economic movement, social events, and investor behaviors led to the lag of prediction task. Analyzing the inherent law of stock price fluctuation is the key to improve the accuracy of the prediction model. The inherent law of stock price fluctuation is mainly reflected in the time characteristics and frequency domain information of the data. Moreover, noise in time-series data can affect that predictive model deficiently capture nonlinear features of data. Hence, existing researches are difficult to comprehensively consider the different problems in the stock price forecasting task.

The advantage of the hybrid model is that it combines the advantages of different models and solves the problem that single models insufficiently adapt to different stock market characteristics. Inspired by hybrid models and GAN, we propose a new Time-Series Generative Adversarial Network (TSGAN) model, including a generator and a discriminator. We integrate the time-series model ARIMA and a hybrid model integrating Attention-based CNN and LSTM in a nonlinear relationship. Under the premise of noise reduction of stock price data, integrated model can further extract the internal relationship between stock data in different periods. On this basis, we use CNN as a discriminator to deeply fuse time feature and frequency domain information of stock price. In summary, the contributions of our work include the following:

- A new Time-Series Generative Adversarial Network (TSGAN) based on ARIMA model, CNN–LSTM hybrid model, and CNN is proposed, which can effectively improve feature representation ability in different historical periods of stock price data.
- The fusion of LSTM and CNN as the generator is introduced for the first time for Generative Adversarial Network, which can effectively solve the long-term dependence problem of LSTM in adversarial training.
- In experiments, our proposed nonlinear fusion model is found to be significantly superior to single model and nonlinear integrated method, proving that our proposed method has stronger robustness and practicability in stock price prediction.

2 Related Work

In this section, this paper briefly reviews the related literatures on stock market prediction especially by applying time-series prediction approaches, Deep Learning methods, and hybrid models combining above two approaches on time-series analysis.

2.1 Traditional Time-Series Prediction Model

Stock price forecasting task is the use of historical data to predict a future period. In view of the high volatility and nonlinearity of stock price data, it is constructed by fitting a set of historical time-series $x_{t-1}, x_{t-2}, \ldots, x_{t-s}$. The purpose is to predict the future data $y_{t-1}, y_{t-2}, \ldots, y_{t-s}$ and describe the future stock price volatility trend. Many statistical models are currently used for stock price prediction, including ARIMA model [24]. The ARIMA model has achieved success to some degree in denoising and forecasting stationary time-series data. However, some studies showed that the nonlinear and non-stationary characteristics contained in stock price data are gradually increasing, which leads to poor performance of ARIMA model in stock price prediction [25].

2.2 Deep Learning Approach

As Deep Learning (DL) has achieved remarkable results in other fields, more and more researchers have begun to apply DL-based forecasting methods to the field of stock forecasting. Weiwei Jiang [26] introduced the application of different DL models and complex hybrid models in stock forecasting. Among them, neural network models such as CNN, LSTM, and various hybrid models were more commonly used in stock forecasting. Wenjie Lu et al. [20] proposed a CNN–LSTM hybrid model. First, CNN was used to extract the time features of various historical data including closing price, and then LSTM was used to predict stock price data. Kang Zhang et al. [27] proposed a GAN model with LSTM as the generator and Multi-Layer Perceptron as the discriminator, which forecasts the future periods of daily closing price through historical stock data in several past days. Domenico and Luca [22] used GAN to integrate different models with the aim of predicting stock price data for a future period. Its innovation is to use adversarial training to mine deeper data features and solve the problem of insufficient description of data nonlinear features. From the above research, the neural network model with strong nonlinear generalization ability has produced good application effect in the financial field.

2.3 The Hybrid Models

In recent years, more and more stock price forecasting researches were no longer limited to a single model. More concretely, studies and theory suggested that adopting hybrid models on stock price forecasting has gradually become an intense trend. As the relationship between external factors affecting stock price changes becomes more complex, it is increasingly difficult to extract the nonlinear and non-stationary characteristics of stock price data [28]. The single models poorly achieve comprehensive accurate results when applied to stochastic and complex time-series data. Meanwhile, the interpretability of time-series forecasting models fitting stock price data and the nonlinear generalization ability of neural network models in stock price data forecasting [29] are irreplaceable. Therefore, the hybrid models can excellently combine models with different advantages to make more accurate prediction targets [30].

The major way to improve stock market forecasting is to combine time-series forecasting models and neural network models in a nonlinear form, which can be used to help predict the nonlinear volatility of stock prices. The integrated model, namely, the combination of ML and DL approaches with time-series prediction models, has proved that the fitting effect of nonlinear fluctuation is greatly improved [29, 31]. Lihki and Keyla [32] proposed a hybrid model of ARIMA–SVR to reduce noise in financial time-series data for achieving more accurate volatility prediction. Zhuangwei Shi et al. [21] proposed a method to integrate the three models of ARIMA model, ACNN–LSTM hybrid model, and XGBoost model in a nonlinear way to improve the nonlinear feature capture ability of stock price data. Compared with other hybrid models, this model adopts the framework of nonlinear integration to enhance the robustness of prediction. The above research shows that the integrated model is more suitable for predicting the nonlinear fluctuation of the stock market than the single prediction model.

In summary, it is an effective way to improve the ability of capturing nonlinear characteristics of stock market by improving the DL framework of forecasting stock price data through nonlinear integration. Such a fusion framework that can not only effectively describe the volatility changes in the stock market, but also highlight the nonlinear characteristics of deep mining is needed to comprehensively improve the prediction accuracy and increase the suitability of the application. Indeed, the current DL framework still has much room for improvement in achieving a comprehensive analysis of the nonlinear volatility of the stock market. In this paper, we propose a new GAN framework TSGAN, which uses a nonlinear fusion hybrid model and CNN model as generator and discriminators, and combines the powerful nonlinear generalization ability of neural network to predict the stock price fluctuation trend in the financial market. Experiments show that the denoising ability of ARIMA model for historical time-series data has been verified. Therefore, we hope to design a fusion neural network structure that can process single-dimensional stock price data and extract nonlinear features from the corresponding stock price data for combination. Hence, we refer to the Generative Adversarial Network framework to fuse traditional time-series model and DL model. Especially, there are two advantages as follows: it is not limited by the data dimension, and the hybrid model in the network framework can process single-dimensional data, thus overcoming the limitations of traditional GAN network; it greatly improves the nonlinear generalization ability of predictive model, and enhances the fusion between time-series features and frequency domain information, thus avoiding the lag of data prediction. To sum up, TSGAN model provides a research basis for the combination of traditional models and neural networks, and achieves flexible and robust stock price prediction.

3 Materials and Methods

In this section, we first reveal the basic components of TSGAN, including the ARIMA model and several Attention Mechanisms. Then we present the overall architecture of TSGAN and investigate the internal process of the generator and discriminator.

3.1 Auto-Regressive Integrated Moving Average Model

The classical model in the time-series prediction problem mainly includes the Auto-Regressive Integrated Moving

Average ARIMA model. An ARIMA (p, q) model consists of an Auto-Regressive (AR) model and a Moving Average (MA) model.

An AR model is suitable for predicting movement related to its own historical data, and the structure is shown in Eq. (1). A MA model can accumulate the error terms in the AR model to effectively eliminate the random fluctuations in the prognostication, as shown in Eq. (2). When time-series $S = [s_1, s_2, \ldots, s_t]$ is stationary non-white sequence data, ARIMA(p, q) can use the historical data of the sequence to forecast a future period.

$$s_t = C_0 + \sum_{i=1}^{p} \gamma_i s_{t-i} + e_t \tag{1}$$

$$s_t = C_0 + \sum_{i=1}^{p} \beta_i s_{t-i} + e_t \tag{2}$$

An ARIMA model

$$s_t = C_0 + \sum_{i=1}^{p} \gamma_i s_{t-i} + \sum_{i=1}^{p} \beta_i s_{t-i} + e_t \tag{3}$$

where S_t is current value, γ_i and β_i are parameters, e_t is noise, i represents the number of historical data associated with the current value. In time-series prediction, ARIMA (p, q, d) is suitable for non-stationary sequence. Based on the feature of stock price data, difference method is applied to process the sequence. In the stock price prediction, the high-order difference can make the trend of the data more stable. However, if the first-order difference can make the sequence data stable, there is no need to use the high-order difference. Therefore, difference method can make the model have a good effect in denoising.

3.2 Attention Mechanism

Treisman et al. [33] in 1980 proposed the Attention Mechanism, which is a classical feature integration theory. The main idea of the theory is to simulate the process of human visual analysis. In stock price forecasting, the traditional neural network model usually ignores the difference in the influence weight of all sequence values of the input sequence on the context vector at each time point. Therefore, to overcome this limitation, we often add Attention Mechanism in the construction of neural networks. It filters key information from a large amount of feedforward information and gives them more weights by calculating the probability distribution of attention.

Traditional attention and multi-headed attention are the two basic divisions of the classical Attention Mechanism. Subsequently, traditional attention is primarily an addressing process. Given a task-related query vector for input $S = [s_1, s_2, \ldots, s_t]$ and index $Z = [1, 2, \ldots, N]$, Eq. (4) illustrates how to compute the attention value from the attention distribution with the relevant key and how to connect the determined value to the distribution, as shown in Eq. (5).

$$\alpha_i = p(Z = i | S, q) = \frac{exp(f(S_i, q))}{\sum_{j=1}^{N} exp(f(S_j, q))} \tag{4}$$

i.e.,

$$\alpha_i = \text{softmax}(f(S_i, q)) \tag{5}$$

$$f(S_i, q) = \frac{S_i^T q}{\sqrt{d_K}} \tag{6}$$

Here is attention score through scaled dot product, d is the dimension of input feature to avoid the problem of vanishing gradients.

Second, multi-head Attention Mechanism can easily capture the relationship between sequences. Suppose the input key-value pairs $(K, V) = [(k_1, v_1), \ldots, (k_N, v_N)]$, for given q, the specific process of the attention function as Eq. (7).

$$\begin{aligned} \text{att}((K, V), Q) &= \sum_{i=1}^{N} \alpha_i v_i \\ &= \sum_{i=1}^{N} \text{softmax}(f((K_i, V_i), q)) v_i \end{aligned} \tag{7}$$

The principle of the multi-head mechanism is to use multi-query $Q = [q_1, q_2, \ldots, q_M]$ as the key parameter of the attention function calculation.

$$\text{att}((K, V), Q) = (\text{att}((K, V), q_1) || \ldots || \text{att}((K, V), q_M)) \tag{8}$$

Here, || represents the concatenate operation. Equation (8) is commonly the description of multi-head attention (MHA). The precise structure may differ depending on the specific application. In our research, the Attention Mechanism is primarily employed to acquire the attention weights for each prediction step, while the procedure for each individual step essentially resembles the Attention Mechanism. Moreover, we have integrated the original model into the CNN, and the implementation process is elucidated in the specific components of the proposed model.

3.3 The Proposed Model

As presented in Sect. 2.3, recent studies have shown promising predictive results in stock price forecasting through hybrid model. Thus, this model aims to retain the advantages of the neural network method in nonlinear feature extraction,

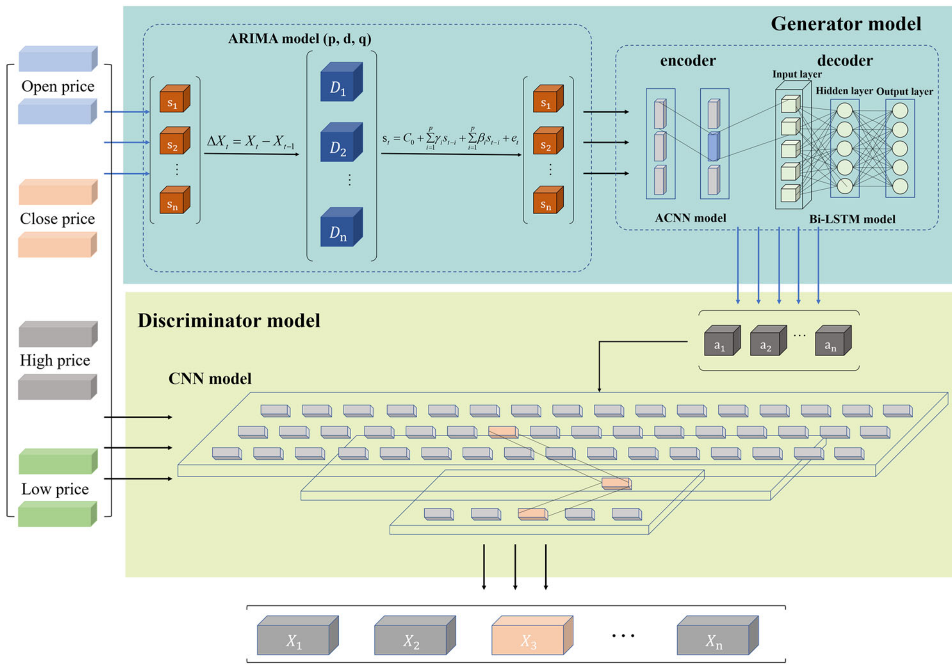


Fig. 1 The architecture of TSGAN

and at the same time, it has the advantages of traditional statistical model higher denoising ability. Our methodology is presented in Fig. 1. In this section, we first introduce the overall architecture of our proposed framework and then discuss details of the two main modules: (1) the Generator model produces prediction series, and (2) the Discriminator model is used to training data by adversarial iteration.

3.3.1 Overall Framework

The initial input data include open prices, close prices, highest price, and lowest price. First, we use ARIMA model to denoise the original data and obtain initial input vector. The proposed model can retain the denoising advantage of ARIMA model to the maximum extent. Subsequently, the Attention-based CNN (ACNN) model enables the derivation of the weight signal matrix pertaining to deep features. Furthermore, the Bidirectional LSTM serves as a decoder, adept at uncovering long-term time-series features. Finally, within the discriminator structure, the CNN model is deployed to assess whether the data are generated by the Generator model or is authentic.

To generate an origin price sequence with less random noise, we employ ARIMA ($p = 2, q = 0, d = 1$) model to pre-process the input data. Through the multi-head attention model, the CNN captures key signal that affects current prediction values. At the same time, the Bidirectional LSTM receives weight signal matrix, and the results are then distinguished into the real data and generated data by the CNN. The above process is inspired by the principle of GAN model, as shown in Fig. 1.

3.3.2 Generator Model

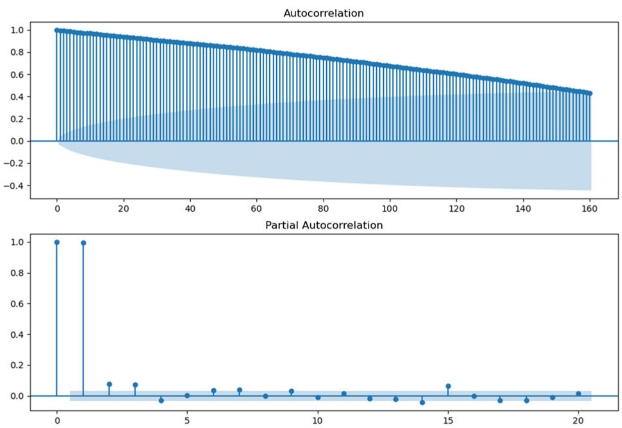
The Generator model is crafted as a hybrid model that combines ARIMA and ACNN–LSTM. We have elected to utilize the daily data from the previous 15 years, encompassing four financial factors, to forecast the future closing price. These four factors of the stock data for a given day include the high price, low price, open price, and close price.

Suppose our initial input is denoted as $S = [s_1, s_2, \ldots, s_t]$, representing the daily stock data over a period of t days. By applying the original stock market close price series $S = [s_1, s_2, \ldots, s_t]$ to the ARIMA model, we can generate a new series that effectively captures the underlying state. The results of the Augmented Dickey–Fuller (ADF) test, as presented in Table 1, indicate that the original sequence is non-stationary, while the first-order difference sequence exhibits stationarity. Having determined $d = 1$, we utilize the Autocorrelation Figure (ACF) and Partial Autocorrelation Figure (PACF) to deduce that adopting AR(2) and MA(0) is appropriate, as depicted in Fig. 2 respectively. Hence $p = 2$, $d = 1$, and $q = 0$.

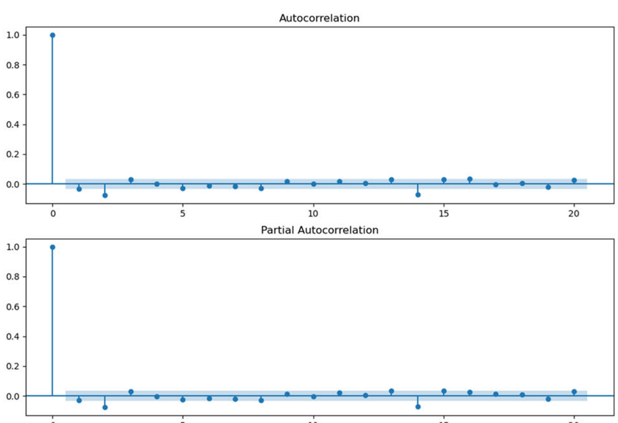
Then, DL architecture receives the new generated series $S' = [s'_1, s'_2, \ldots, s'_t]$. The DL architecture employed in this study is the Attention-based CNN-LSTM model, operating within a sequence-to-sequence framework. Within this architecture, the Attention-based CNN functions as the encoder, while the Bidirectional LSTM serves as the decoder. The model commences by utilizing convolutional operations to extract profound features from the original stock data. Subsequently, LSTM networks are employed to uncover long-term time-series features. Among them, ACNN encoder block

Table 1 ADF test for sequence

Metric	Value for the original sequence	Value for first-order sequence
Test statistic value	− 2.34702	− 14.7432
p-value	0.154220	2.48496
Lags used	18	18
Number of observations used	3504	3504
Critical value (1%)	− 3.446196	− 3.446196
Critical value (5%)	− 2.862747	− 2.862747
Critical value (10%)	− 2.557102	− 2.557102



(a) the original sequence.



(b) the first-order sequence.

Fig. 2 The ACF and PACF of the original and first-order sequence

comprises a multi-head attention layer and a CNN. The Q, K, V values are computed using Eq. (8) after the self-attention layer, and subsequently serve as inputs to the LSTM decoder block. Finally, the model generates synthetic data that closely resembles real data.

3.3.3 Discriminator Model

The primary objective of the discriminator is to establish a differentiable function D capable of classifying the generated data. The discriminator is designed to output 0 when presented with fake data and 1 when presented with real data, as described in Eq. (9). In this context, a CNN architecture was selected as the discriminative model D, consisting of five layers: an input layer, a convolution layer, an activation layer, a pooling layer, and an output layer, enabling convolutional operations on one-dimensional input. Furthermore, the activation function ReLU was chosen to mitigate the risk of gradient explosion. In particular, we determine if the data in the sequence $S' = [s'_1, s'_2, \ldots, s'_t]$ was fake.

$$D(S') = \begin{cases} 0, & \text{if} \quad d(S') = d(X_{\text{face}}) \\ 1, & \text{if} \quad d(S') = d(X_{\text{real}}) \end{cases} \tag{9}$$

where $d(\bullet)$ denotes the output of CNN. Both X_{real} and X_{fake} denote a complete data sequence. In our proposed GAN architecture, we maximize the discriminability of the discriminator and minimize the difference between the data generated by the generator and the real data. The loss function of the model is as follows:

$$\begin{aligned} \min_G \max_D O(G, D) = E[\log D(X_{\text{real}})] \\ +E[\log(1 - D(X_{\text{face}}))] \end{aligned} \tag{10}$$

4 Experiment

In this section, we provide a detailed description of the datasets utilized in our study. We then outline the experimental setup and elaborate on the modifications made to the model. Finally, we present a thorough account of the experiments and analyses conducted to assess the performance of our proposed TSGAN framework.

4.1 Data

The data utilized in this article are sourced from the open and freely available dataset provided by Tushare for conducting research on the Chinese stock market. This dataset offers rich data, user-friendly access, and convenient implementation. For our experiments, we specifically focus on the daily stock price data of Bank of China (601988.SH) in the Chinese stock market, covering the period from January 1, 2007, to June 22, 2021. The input vector comprises the high price, low price, open price, and close price.

The evaluation of prediction results is based on several criteria, including the mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and R^2. The calculation formulas for these metrics are as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|\hat{X}_t - X_t\right| \tag{11}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{X}_t - X_t)^2} \tag{12}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{X}_t - X_t}{X_t}\right| \tag{13}$$

$$R^2 = \frac{\sum_{t=1}^{n}\left\|\hat{X}_t - X_t\right\|^2}{\sum_{t=1}^{n}\left\|X_t - \bar{X}_t\right\|^2} \tag{14}$$

where i represents the predicted values, y_i represents the actual values, n is the total number of samples, and $\bar{y}$ is the mean of the actual values. These evaluation criteria provide insights into the accuracy, precision, and overall performance of the prediction models.

The experiments were conducted using an NVIDIA GTX 2080ti GPU with 12GB memory. The empirical study focused on analyzing the stock price of Bank of China (601988.SH) in the Chinese stock market. The stock price data were obtained from Tushare, a publicly available dataset. The selected data spanned from January 1, 2007, to March 31, 2022, with each day representing a point in the sequence. The train and test sets were divided on June 22, 2021, resulting in 3504 training samples and 180 testing samples.

4.2 Results

In this section, we present the experimental results of the ablation study and compare our model with other existing models. Initially, according to several criteria in Sect. 4.1, we give the comparison of the results of ablation experiments. Subsequently, we select a set of comparative experimental models, including machine learning (ML) models such as the hybrid ARIMA–SVR model [32]. In addition, we consider three DL models proposed by Zhuangwei Shi et al. [21], Ashish Kumar et al. [18], and Kang Zhang et al. [27]. To highlight the predictive capabilities of our model, we compare the errors of these integrated models and DL models with our own model.

4.2.1 Ablation Study

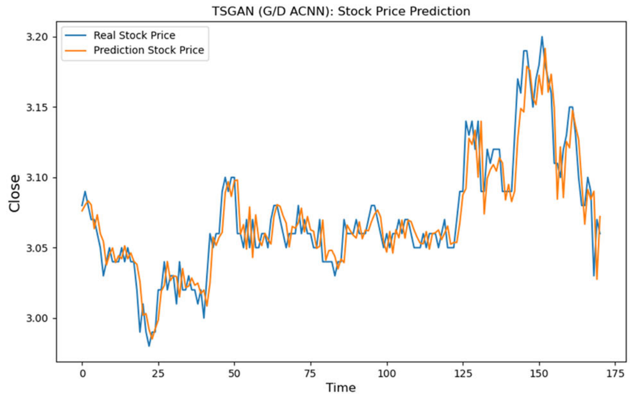
For a comprehensive analysis of the primary components within TSGAN, we have constructed three additional variants in addition to the fully equipped TSGAN model. One of these variants is referred to as TSGAN (G/D ARIMA), which indicates that the generator model does not incorporate the pre-process model. Moreover, we propose a variant without Attention Mechanism, named TSGAN (G/D Attention). Meanwhile, we propose a similar variant without ACNN in Generator model, named TSGAN (G/D ACNN), respectively.

- TSGAN (G/D ARIMA): TSGAN with the original input uses only ACNN–LSTM as Generator model.
- TSGAN (G/D Attention): TSGAN with the traditional CNN uses without Attention Mechanism.
- TSGAN (G/D ACNN): TSGAN with the original input uses only LSTM as a part of Generator model.

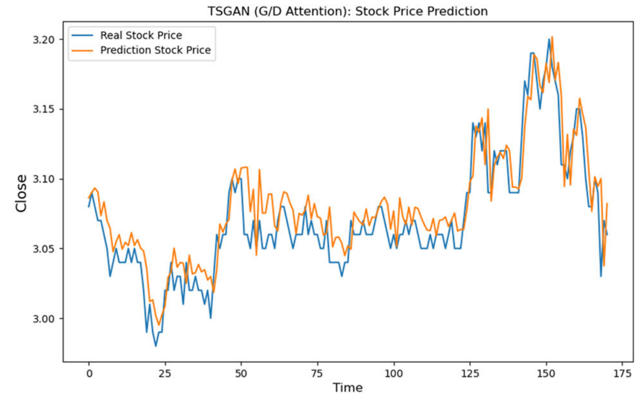
Figure 3 comprehensively depicts the predictive graphs of stock price on same dataset for the different TSGAN variants. Table 2 shows the performance of the different TSGAN variants. We can see that TSGAN (G/D ARIMA) yields the worst results, showing the necessity of the denoising effect of ARIMA model in Generator model. The performance of TSGAN (G/D ACNN) is unsatisfactory, demonstrating that mining nonlinear features in time sequence play an important role in Generator model. By contrast, TSGAN (G/D Attention) produces exceptionally competitive results, like those of TSGAN (G/D ARIMA). Consequently, the performance results of above variants confirm the positive effects of components of the proposed model, respectively, for stock price prediction.

4.2.2 Comparison with Other Models

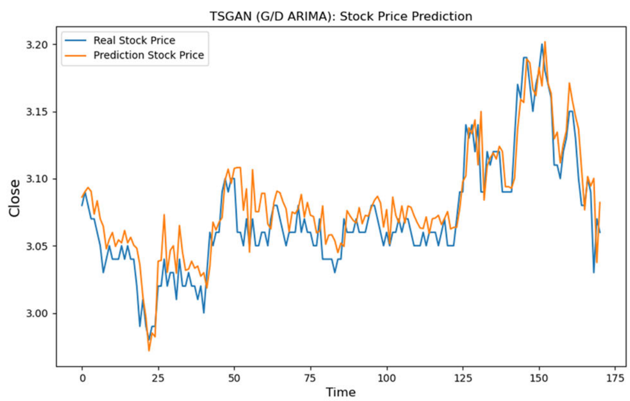
For the purpose of comparison, we have selected several classical and representative methods that share similarities with TSGAN and employ GAN architectures or similar approaches to address the stock market prediction problem. These chosen models serve as benchmarks for evaluating the performance of TSGAN.



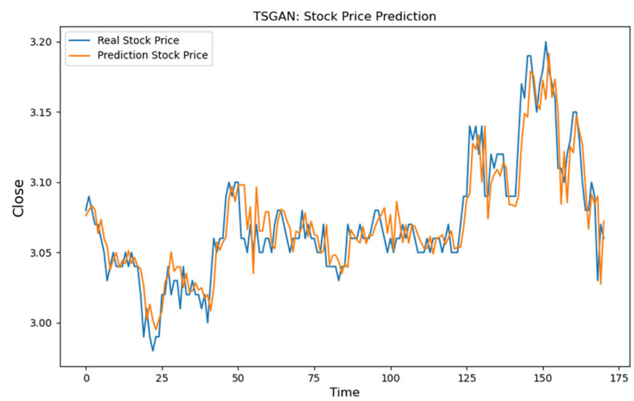
(a) TSGAN (G/D ACNN)



(b) TSGAN (G/D Attention)



(c) TSGAN (G/D ARIMA)



(d) TSGAN

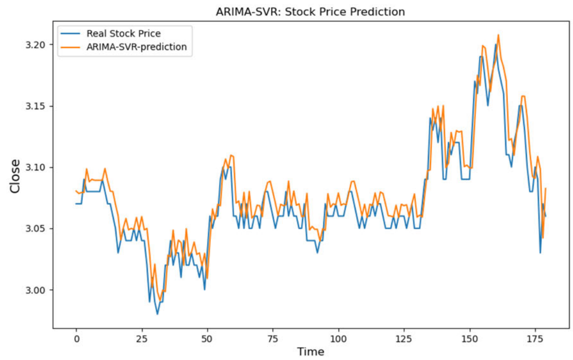
Fig. 3 Comparison of predictive results through different TSGAN variants: **a** TSGAN (G/D ACNN), **b** TSGAN (G/D Attention), **c** TSGAN (G/D ARIMA), **d** TSGAN

Table 2 Ablation experiment for four different models

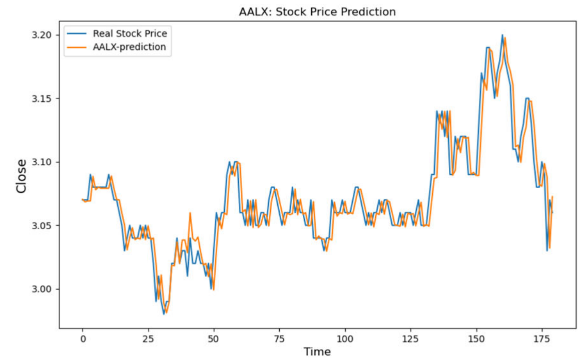
	MAE	RMSE	MAPE	R_2
TSGAN (G/D ACNN)	0.00029	0.01717	0.01287	0.83954
TSGAN (G/D Attention)	0.00040	0.02000	0.01577	0.78233
TSGAN (G/D ARIMA)	0.00041	0.01998	0.01572	0.78270
TSGAN	0.00018	0.01356	0.01033	0.89363

- *ARIMA–SVR* [32]: A novel hybrid time-series prediction model, which consists of two submodules: one makes use of the support vector regression model to forecast sequence for nonlinear stock data, and the other uses an ARIMA mo del to fit series data.
- *ARIMA–ACNN–LSTM–XGBoost* [21]: A novel NN prediction model with recurrent training. This model includes an ARIMA layer, an CNN layer, a LSTM layer, and a XGBoost fine-tuning layer.
- *GAN (G/D LSTM, CNN)* [18]: A novel NN prediction model based on Generative Adversarial Network, which demonstrates a high level of expressiveness when applied to sequential data. By employing these architectures, the model is able to extract deep features and effectively capture the dependencies present in sequential data.
- *GAN (G/D LSTM, DMLP)* [27]: A novel NN prediction model based on Generative Adversarial Network, which is highly expressive for sequential data. This model use LSTM and DMLP as Generator and Discriminator; it can capture the dependencies between sequences and has shown promising performance in the closing price prediction on real data.

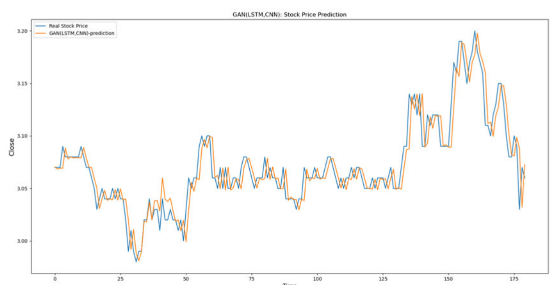
Figure 4 comprehensively depicts the predictive graphs of stock price on same dataset for the other competitive models. As presented in Table 3, the proposed TSGAN model demon-



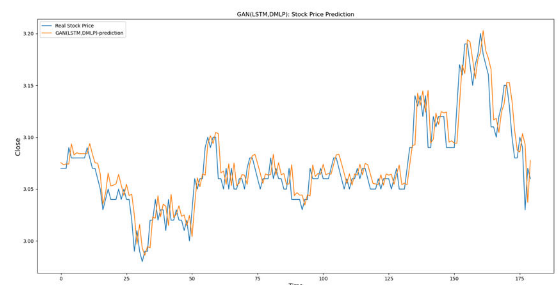
(a) ARIMA-SVR



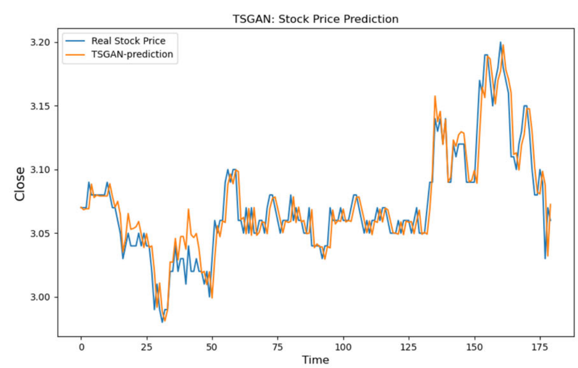
(b) ARIMA-ACNN-LSTM-XGBoost



(c) GAN(G/D LSTM, CNN)



(d) GAN(G/D LSTM, DMLP)



(e) TSGAN

Fig. 4 Comparison of predictive results through different models: **a** ARIMA–SVR [32], **b** AALX [21], **c** GAN (G/D LSTM, CNN) [18], **d** GAN (G/D LSTM, DMLP) [27], **e** ours

Table 3 Performance of other methods and TSGAN

	MAE	RMSE	MAPE	R_2
ARIMA–SVR [32]	0.00077	0.02456	0.02123	0.75957
GAN(G/D LSTM, CNN) [27]	0.00035	0.01635	0.01560	0.82134
GAN(G/D LSTM, DMLP) [23]	0.00027	0.01562	0.01327	0.85238
ARIMA–ACNN–LSTM–XGBoost [21]	0.00021	0.01498	0.01134	0.88356
Our	0.00018	0.1356	0.1033	0.89363

strates superior performance in most cases, as indicated by its low MAE, RMSE, and MAPE values. These metrics suggest that the predicted closing prices closely approximate the real data. The results show that our method achieves a competitive performance compared to other methods considered in the evaluation.

5 Conclusion

In this research paper, we introduce TSGAN, a novel DL model that integrates the Generator and Discriminator models. The concept of Generative Adversarial Network is applied to the stock market prediction problem. Through empirical evaluation, we substantiate the effectiveness of our model and assess its performance using diverse evaluation metrics. The comprehensive comparative analysis demonstrates that our model is superior to alternative approaches in various aspects. The model which employs Generative Adversarial Network framework enhances the accuracy of predictive results in stock price.

Through comprehensive comparisons with advanced methodologies, we have determined that our proposed method is suitable for practical applications. Moreover, our model has the potential to assist traders in mitigating financial risks and enhancing decision-making capabilities. Despite all this, there is ongoing work to further enhance the value of TSGAN in practical settings. One of the primary research objectives we will address in the future is reducing model training and pre-processing complexities, given the intricate and rapidly evolving nature of the stock market.

Author Contributions Conceptualization: YX, PL, and YZ; methodology: YX and QZ; visualization: YX and YqZ; writing—original draft: YX; writing—review and editing: YZ and YX; and software: YX.

Funding This work was supported in part by the Youth Innovation Team in Colleges and universities of Shandong Province (2022KJ185), in part by the Natural Science Foundation of Shandong Province (ZR2022MF245), and in part by Shandong Province Natural Science Foundation Youth Branch (ZR2023QF161).

Data Availability All data generated or analyzed during this study are included in this paper.

Declarations

Conflict of Interest The authors declare no competing interests.

References

1. Fama, E.F.: The behavior of stock-market prices. J Bus **38**(1), 34–105 (1965). http://www.jstor.org/stable/2350752
2. Enke, D., Thawornwong, S.: The use of data mining and neural networks for forecasting stock market returns. Expert Systems with applications. **29**(4), 927–940 (2005)
3. Fama, E.F.: Technical analysis of the financial markets: a comprehensive guide to trading methods and applications. J. Bus. **38**(1), 34–105 (1965). http://www.jstor.org/stable/2350752
4. Nti, I.K., Adekoya, A.F., Weyori, B.A.: A systematic review of fundamental and technical analysis of stock market predictions. Artificial Intelligence Review. **53**(4), 3007–3057 (2020). https://doi.org/10.1016/j.eswa.2011.03.001
5. Perols, J.: Financial statement fraud detection: An analysis of statistical and machine learning algorithms. Auditing: A Journal of Practice & Theory. 30(2), 19–50 (2011)
6. Fama, E.F.: Market efficiency, long-term returns, and behavioral finance. J. Financ. Econ. **49**(3), 283–306 (1998)
7. Kumar, D.A., Murugan, S.: Performance analysis of Indian stock market index using neural network time series model. In: 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, pp. 72–78 (2013)
8. Yan, W., Yuankai, G.: Forecasting method of stock market volatility in time series data based on mixed model of ARIMA and XGBoost. China Commun. **17**(3), 205–221 (2020). https://doi.org/10.23919/JCC.2020.03.017
9. Tim, B.: Generalized autoregressive conditional heteroskedasticity. J. Econom. **31**(3), 307–327 (1986). https://doi.org/10.1016/0304-4076(86)90063-1
10. Chen, Y., Hao, Y.: A feature weighted support vector machine and k-nearest neighbor algorithm for stock market indices prediction. Expert Syst. Appl. **80**(9), 340–355 (2017). https://doi.org/10.1016/j.eswa.2017.02.044
11. Sasan, B., Azadeh, A., Sergio, O.: Fusion of multiple diverse predictors in stock market. Inf. Fusion **36**, 90–102 (2017). https://doi.org/10.1016/j.inffus.2016.11.006
12. Suryoday, B., Saibal, K., Snehanshu, S., Luckyson, K., Sudeepa, R.D.: Predicting the direction of stock market prices using tree-based classifiers. N. Am. J. Econ. Finance **47**, 552–567 (2019). https://doi.org/10.1016/j.najef.2018.06.013
13. Qili, W., Wei, X., Han, Z.: Combining the wisdom of crowds and technical analysis for financial market prediction using deep random subspace ensembles. Neurocomputing **299**, 51–61 (2018). https://doi.org/10.1016/j.neucom.2018.02.095
14. Chung, H., Shin, K.-s: Genetic algorithm-optimized long short-term memory network for stock market prediction. Sustainability **10**(10), 3765 (2018)
15. Bao, W., Yue, J., Rao, Y.: A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLoS One **12**(7), 0180944 (2017)
16. Berradi, Z., Lazaar, M.: Integration of principal component analysis and recurrent neural network to forecast the stock price of Casablanca stock exchange. Procedia Comput. Sci. **148**, 55–61 (2019)
17. Wang, H., Wang, J., Cao, L., Li, Y., Sun, Q., Wang, J.: A stock closing price prediction model based on CNN-BiSLSTM. Complexity **2021**, 1–12 (2021)

18. Kumar, A., Alsadoon, A., Prasad, P.W.C., Abdullah, S., Rashid, T.A., Pham, D.T.H., Nguyen, T., Quoc, V.: Generative adversarial network (GAN) and enhanced root mean square error (ERMSE): deep learning for stock price movement prediction. Multimed. Tools Appl. **81**(3), 3995–4013 (2017). https://doi.org/10.1007/s11042-021-11670-w
19. Mahinda, M.K., Christoph, L., Pasi, L., Jari, P.: Machine learning techniques and data for stock market forecasting: a literature review. Expert Syst. Appl. **197**, 116659 (2022). https://doi.org/10.1016/j.eswa.2022.116659
20. Wenjie, L., Jiazheng, L., Yifan, L., Aijun, S., Jingyang, W.: A CNN-LSTM-based model to forecast stock prices. Complexity **2020**, 1–10 (2020)
21. Zhuangwei, S., Yang, H., Guangliang, M., Jianguo, W.: Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction. arXiv:2204.02623 (2022)
22. Santoro, D., Grilli, L.: Generative adversarial network to evaluate quantity of information in financial markets. Neural Comput. Appl. **34**, 17473–17490 (2022). https://doi.org/10.1007/s00521-022-07401-3
23. Lin, E., Lin, C.-H., Lane, H.-Y.: Relevant applications of generative adversarial networks in drug design and discovery: molecular de novo design, dimensionality reduction, and de novo peptide and protein design. Molecules **25**(14), 3250 (2020). https://doi.org/10.3390/molecules25143250
24. Jung-Hua, W., Jia-Yann, L.: Stock market trend prediction using ARIMA-based neural networks. In: Proceedings of International Conference on Neural Networks (ICNN'96), vol. 4, pp. 2160–2165 (1996)
25. Meher, B.K., Hawaldar, I.T., Spulbar, C.M., Birau, F.R.: Forecasting stock market prices using mixed ARIMA model: a case study of Indian pharmaceutical companies. Invest. Manag. Financ. Innov.. **18**(1), 42–54 (2021)
26. Jiang, W.: Applications of deep learning in stock market prediction: recent progress. Expert Syst. Appl. **184**, 115537 (2021)
27. Kang, Z., Guoqiang, Z., Junyu, D., Shengke, W., Yong, W.: Stock market prediction based on generative adversarial network. Procedia Comput. Sci. **147**, 400–406 (2019). https://doi.org/10.1016/j.procs.2019.01.256
28. Deepak, K., Pradeepta, K.S., Rajit, V.: A systematic review of stock market prediction using machine learning and statistical techniques. Mater. Today Proc. **49**, 3187–3191 (2022). https://doi.org/10.1016/j.matpr.2020.11.399
29. Chen, S., Zhou, C.: Stock prediction based on genetic algorithm feature selection and long short-term memory neural network. IEEE Access **9**, 9066–9072 (2020)
30. Mehdi, Z., Bahram, S.: Impact of foreign exchange rate on oil companies risk in stock market: a Markov-switching approach. J. Comput. Appl. Math. **317**, 274–289 (2017). https://doi.org/10.1016/j.cam.2016.10.012
31. Ken, L., Xiongwen, Q., Chen, J., Zhuangwei, S., Jinglong, Y.: An interpretable double-scale attention model for enzyme protein class prediction based on transformer encoders and multi-scale convolutions. Front. Genet. **13**, 885627 (2022). https://doi.org/10.3389/fgene.2022.885627
32. Rubio, L., Alba, K.: Forecasting selected Colombian shares using a hybrid ARIMA-SVR model. Mathematics **10**(13), 2181 (2022). https://doi.org/10.3390/math10132181
33. Anne, M.T., Garry, G.: A feature-integration theory of attention. Cogn. Psychol. **12**(1), 97–136 (1980). https://doi.org/10.1016/0010-0285(80)90005-5