Expert Systems With Applications 217 (2023) 119549

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

COVID19-MLSF: A multi-task learning-based stock market forecasting framework during the COVID-19 pandemic

Chenxun Yuan [a], Xiang Ma [a], Hua Wang [c], Caiming Zhang [a,b], Xuemei Li [a,*]

[a] School of Software, Shandong University, Jinan 250101, China
[b] Shandong Provincial Laboratory of Future Intelligence and Financial Engineering, Yantai 264005, China
[c] School of Information and Electrical Engineering, Ludong University, Yantai 264025, China

ARTICLE INFO

Keywords:
Stock market forecasting
COVID-19 pandemic
Multi-task learning
Feature fusion
K-nearest neighbor classifier

ABSTRACT

The sudden outbreak of COVID-19 has dramatically altered the state of the global economy, and the stock market has become more volatile and even fallen sharply as a result of its negative impact, heightening investors' apprehension regarding the correlation between unexpected events and stock market volatility. Additionally, internal and external characteristics coexist in the stock market. Existing research has struggled to extract more effective stock market features during the COVID-19 outbreak using a single time-series neural network model. This paper presents a framework for multitasking learning-based stock market forecasting (COVID-19-MLSF), which can extract the internal and external features of the stock market and their relationships effectively during COVID-19.The innovation comprises three components: designing a new market sentiment index (NMSI) and COVID-19 index to represent the external characteristics of the stock market during the COVID-19 pandemic. Besides, it introduces a multi-task learning framework to extract global and local features of the stock market. Moreover, a temporal convolutional neural network with a multi-scale attention mechanism is designed (MA-TCN) alongside a Multi-View Convolutional-Bidirectional Recurrent Neural Network with Temporal Attention (MVCNN-BiLSTM-Att), adjusting the model to account for the changing status of COVID-19 and its impact on the stock market. Experiments indicate that our model achieves superior performance both in terms of predicting the accuracy of the China CSI 300 Index during the COVID-19 period and in terms of sing market trading.

1. Introduction

As the stock market and artificial intelligence technology develop rapidly, a new generation of quantitative trading tools on the basis of machine learning has performed well in stock prediction tasks (Giudici, Polinesi, & Spelta, 2022; Ma, Zhao, Guo, Li, & Zhang, 2022; Shah, Bhatt, & Shah, 2022; Yan et al., 2020), and numerous quantitative stock trading researchers have gained huge profits from the stock market, which is prospering. Besides, COVID-19 has had a huge impact on the stock market, and Fig. 1 demonstrates a time series chart of the closing prices of the five major global stock indexes from January 2019 to May 2020. Consequently, the global stock market has experienced a significant decline, as is evident. The poor handling mechanism of most stock market forecasting models for unforeseen events limits the predictive power of the models during this period, which has prompted the investigation of stock market forecasting models that can handle COVID-19 pandemic events (Ronaghi, Salimibeni, Naderkhani, & Mohammadi, 2022; Štifanić et al., 2020). Based on this, we conducted in-depth investigation on two major aspects, namely, finding more data that could reflect external stock market features and further enhancing the performance of the time-series forecasting model.

In accordance to the efficient market hypothesis theory (Fama, 1970), the stock prices have fully reflected all valuable and pertinent information. And yet more research confirms the equal significance of information external to the stock market, including national economic policies, investors' investment sentiment, and the positive or negative effect of news, which can have a lasting or temporary effect on stock prices. Using stock prices and their derived technical indicators, and incorporating sentiment analysis of news and stock reviews for forecasting (Chen, Ma, Wang, Li, & Zhang, 2022; Jing, Wu, & Wang, 2021; Zhang et al., 2018), methods have been developed to quantify investor sentiment or news information.Despite being effective for stock market forecasting, these methods are not satisfactory due to the relatively homogeneous stock price information, the vast amount of information

* Corresponding author.
E-mail addresses: yuanchenxun@mail.sdu.edu.cn (C. Yuan), xiangma@mail.sdu.edu.cn (X. Ma), hua.wang@ldu.edu.cn (H. Wang), czhang@sdu.edu.cn (C. Zhang), xmli@sdu.edu.cn (X. Li).

https://doi.org/10.1016/j.eswa.2023.119549
Received 29 July 2022; Received in revised form 24 December 2022; Accepted 11 January 2023
Available online 16 January 2023

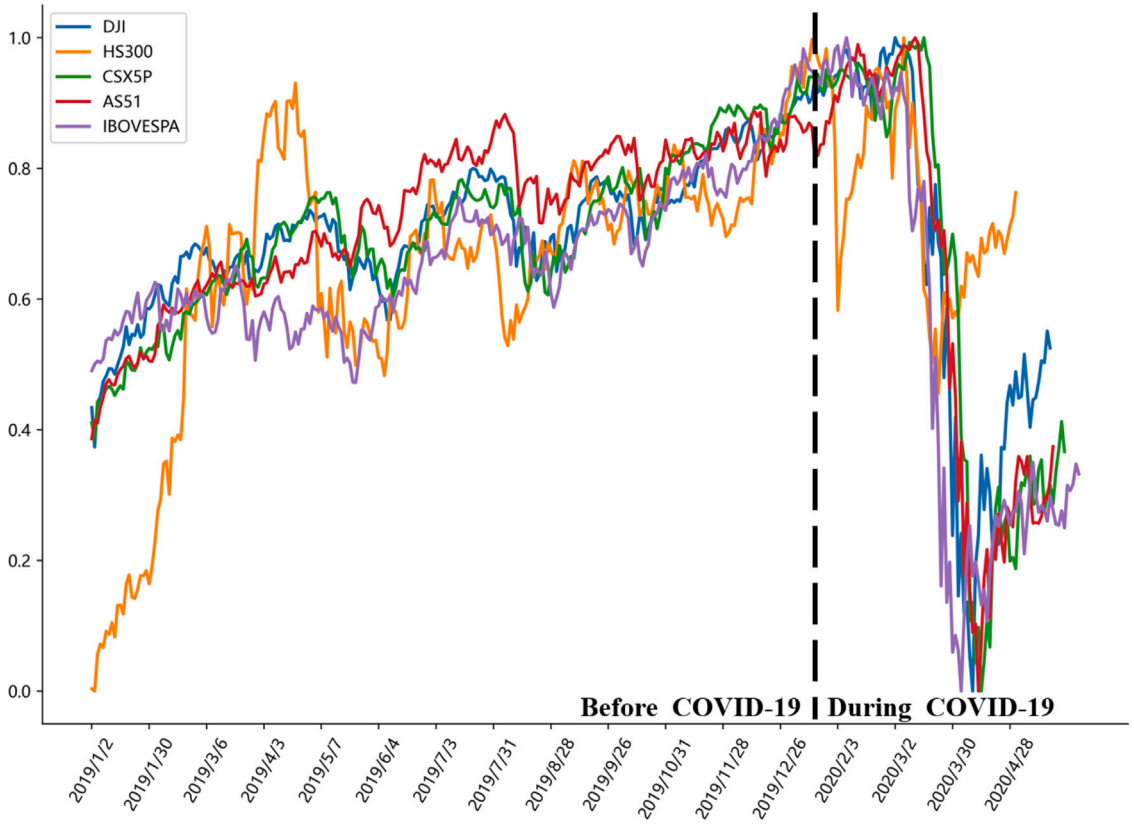

Fig. 1. Time-series chart of the closing prices of the five major global stock indexes between 2019.01 and 2020.05 (The closing price is normalized to the maximum and minimum).

available in textual information, and the fact that most of the textual information contains personal subjective opinions and with varying quality. This is particularly true during unpredictable occurrences such as the COVID-19 pandemic. Ahelegbey, Cerchiello, and Scaramozzino (2022) have confirmed in recent literature that COVID-19 can significantly affect financial markets, and using multi-characteristic data (such as market and sentiment data) to study the impact of COVID-19 on different industries, provides more effective evidence for studying the impact of COVID-19 on the stock market. In some studies, Twitter comments during COVID-19 (Ronaghi et al., 2022) and confirmed COVID-19 cases (Štifanić et al., 2020) have been employed to predict the stock market with positive results. Nonetheless, the polarity of the sentiment of Twitter comments is challenging to determine, either through manual annotation or natural language processing models. Furthermore, since different textual data, namely, news and stock reviews are intermingled with the stock market's short-term or long-term effects, it is difficult to classify and analyze the text data information. Consequently, we must identify other predictability-enhancing characteristics that are more efficient. It has been confirmed in some literature that Chinese stock market has features such as fast response to national policy regulations and parallel socioeconomic development with stock market (Los & Yu, 2008; Wang, 2010), therefore, macroeconomic data does have a longer-term impact on the stock market and can reflect the long-term market trend in advance, in comparison with stock commentaries or textual information. Additionally, due to the decrease in COVID-19 lethality, the growing number of cured cases, as well as the gradual development of new foreign trade policies in distinct countries, which have substantially aided the economic recovery it has become tough to reflect the development status of COVID-19 and the impact on the stock market employing only COVID-19 confirmed cases. In accordance to our study, the number of cured COVID-19 cases, the search volume of the epidemic in Baidu's index, and the media information statistics all reflect COVID-19's changing status to distinct degrees. Thus, we can utilize principal component analysis to downscale the aforementioned indicators and investigate the mechanism and mode of influence of an unexpected event including such COVID-19 on stock price fluctuations.

On the other hand, time-series prediction models including recurrent neural networks represented by long and short-term memory neural networks (Hochreiter & Schmidhuber, 1997), gated recurrent units (Cho et al., 2014), and temporal convolutional neural networks (Bai, Kolter, & Koltun, 2018) have become acknowledged as superior deep learning models in time-series analysis (Yan et al., 2020). Nonetheless, Recent studies have indicated that it has been challenging to extract more effective data with the aid of RNNs (Rezaei, Faaljou, & Mansourfar, 2021; Yan et al., 2020). Although the causal convolution structure proposed by TCN alleviates the issue arising from RNN, the traditional TCN is insensitive to the essential information in long time series. Additionally, this class of single-task prediction models lacks high-quality feature learning capabilities when dealing with nonlinear, highly noisy stock price series, particularly on the condition that there are multiple input sources and they may overlap more noise in the data, increasing the highly noisy nature of the data (Ko et al., 2021). The framework for multitask learning is an effective solution. Multitask learning refers to learning multiple related tasks at the same time, allowing these tasks to share knowledge in the learning process, and applying the correlation between multiple tasks to enhance the performance and generalization ability of the model in the task, which can serve a limited data enhancement function, and has been widely adopted and demonstrated excellent results in natural language processing (Collobert & Weston, 2008), computer vision (Girshick, 2015). Existing applications of multi-task learning in the stock market (Li, Song, & Tao, 2019; Ma & Tan, 2020, 2022; Zhang, Wu, & Li, 2022) have also largely enhanced the feature extraction capability of the models, and yet some current work (Li et al., 2019; Ma & Tan, 2020; Mootha, Sridhar, Seetharaman, & Chitrakala, 2020; Zhang et al., 2022)

only learns the stock price series to varying degrees. Consequently, the extracted individual features cannot adequately describe the state of the stock market. Many factors affect the rise and fall of stocks, and these factors usually do not exist in isolation, and stock prices are formed in such intertwined effects, which causes the multi-source nature of stock data, and the predominant advantage of multi-task learning, especially multi-task learning framework on the basis of soft sharing mechanism (Ma & Tan, 2022) is the capacity to learn diverse data sources by completing a variety of subtasks.

For the purpose of coping with the impact of stock price fluctuations caused by emergencies such as COVID-19 on stock market forecasting models. This paper uses the principal component analysis method to construct a New Market Sentiment Index (NMSI) owing to macroeconomic data and a COVID-19 index reflecting the development status of COVID-19 in China to characterize the external feature of the stock market in times of emergencies. We subsequently propose a multi-task learning-based forecasting model for COVID-19-MLSF, and the model framework is indicated in Fig. 2. In this model, we introduced the framework of multi-task learning and establish two forecasting subtasks whilst also combining the constructed external features of the stock market (NMSI and COVID-19 Index) with the low-frequency and high-frequency signals decomposed by stock price correspondingly. Besides, the subtasks extract global and local features of the stock price series, thereby effectively solving the problem of tough feature extraction and miscellaneous internal features of the stock price series. Due to the fact that multi-task learning reduces the impact of local parameters on the global and reduces overfitting of the main task, the main task's predictive ability is significantly enhanced. In addition, this paper adds attention mechanisms to dissimilar causal convolutional layers of TCNs, and the newly designed temporal convolutional neural network (MA-TCN) with multi-scale attention mechanism makes TCNs more sensitive to important features. Moreover, to address the issue that recurrent neural networks have a propensity to forget sequence features, this paper proposes a convolutional-bi-directional recurrent combining multiview convolutional neural network and temporal attention mechanism(MVCNN-BiLSTM-Att). In comparison with a single recurrent neural network, the feature extraction and memory capabilities of the network are improved. Experiments further reveal that our model accomplishes superior performance in predicting the China CSI 300 index and high returns on the bear market state during COVID-19.

The following are the major contributions of this paper:

- A new market sentiment index (NMSI) and COVID-19 index are constructed to reflect the external features of the stock market, introducing a novel method for studying the effects of unforeseen events, namely, COVID-19 on stock price volatility.
- Multitask learning forecasting framework is designed to handle multiple data sources of the stock market. Moreover, the decomposed stock price series and the constructed external features of the stock market are established as separate prediction subtasks, which alleviate the issue of complex features mingling in stock price data and likewise enhance the ability to make the neural network feature extraction module more capable of extracting effective features.
- Design a Multi-scale attention mechanism temporal convolutional neural network (MA-TCN), which effectively solves the traditional TCN's insensitivity to long sequences of significant information and thus calculates the weight parameters of distinct feature information better. Furthermore, design a multi-view convolutional-bi-directional recurrent neural network (MVCNN-BiLSTM-Att) to model how the COVID-19 state affects stock market volatility.

The remaining sections are organized as follows: Section 2 reviews some literature and novel methods related to this paper. Subsequently, Section 3 describes how the Market Sentiment Index (NMSI) and the COVID-19 Index are constructed. Section 4 presents the three task modules of the COVID-19-MLSF. Except for that, Section 5 presents experimental demonstrations, including comparison and ablation experiments as well as hyperparametric sensitivity tests. Ultimately, Section 6 concludes the entire paper and suggests future works.

2. Review of literature

In the last decade, stock market forecasting has gained a great deal of attention from financiers and computer scientists due to its unique features and wide range of potential applications.

2.1. Prediction based on machine learning and multi-task learning

Recurrent neural networks (RNNs) and their derivatives, long and short-term memory neural networks (LSTMs) and memory recurrent units (GRUs), have been proposed as alternative methods for extracting nonlinearity from stock price series in traditional time series models (Fu, Zhang, & Li, 2016; Hochreiter & Schmidhuber, 1997). In spite of this, these mentioned networks has serious problems, such as gradient disappearance and explosion, which adversely affect stock predictions for long time series. Bi-directional recurrent neural networks (Bi-RNN), such as the bi-directional long-term memory networks (Bi-LSTM), are better suited to predicting long- and short-term trends in stock data since they can capture information about both the past and future of the data (Althelaya, El-Alfy, & Mohammed, 2018; Siami-Namini, Tavakoli, & Namin, 2019; Yang & Wang, 2022). In addition to recurrent neural networks, temporal convolutional neural networks (TCNs) have been shown to outperform RNNs in a number of tasks (Bai et al., 2018; Cheng et al., 2021; Dai, An, & Long, 2022). However, TCNs are neither sensitive to the important information contained in the sequences nor to effectively extract it. The attention mechanism give new solution of the problem of feature attention in long-term sequences. Base on this process, recurrent neural networks incorporating the attention mechanism are widely used to predict stock prices. An example would be the two-stage attention recurrent neural network model (DA-RNN) proposed by Qin et al. (2017). This model through input attention mechanism to capture spatial features, while temporal features, and get better results of NASDAQ prediction.

It is common for researchers to resolve the time series data before inputting the model using the signal decomposition method (Lahmiri, 2016a; Lin, Lin, & Cao, 2021; Ma, Li, Zhou, & Zhang, 2021; Rezaei et al., 2021; Yan et al., 2020), with representative methods including Fourier transform (FT), wavelet transform (WT), empirical mode decomposition (EMD), ensemble empirical mode decomposition (EEMD), variable mode decomposition (VMD), etc. By decomposing the complex and high-noise financial time series into semaphores of different frequencies, it is possible to achieve a faster convergence during deep neural network learning and avoid overfitting. What is more, it is possible to separate the different valid information in the stock sequence. Nevertheless, the decomposition methods used in the aforementioned literatures need to decompose all data once, which means there may be leakage of future data, so that the prediction results will be prospectively biased. Liu, Ma, Li, Li, and Zhang (2022) decomposed the data selected through the sliding window in order to solve the problem of future data leaks, and stock price data used in this paper were also decomposed in this manner.

With the development of deep learning, deep multitasking learning has become one of the important subfields. Multi-task learning is to share the features extracted from different tasks by performing multiple prediction tasks at the same time. Multi-task learning not only has better feature extraction ability than a single model, but also has stronger generalization ability, and has been widely used in the prediction of stock data (Li et al., 2019; Ma & Tan, 2020, 2022; Zhang et al., 2022). Ma and Tan (2020) added attention mechanism to multitask learning to learn shared and private features from different tasks, and Zhang

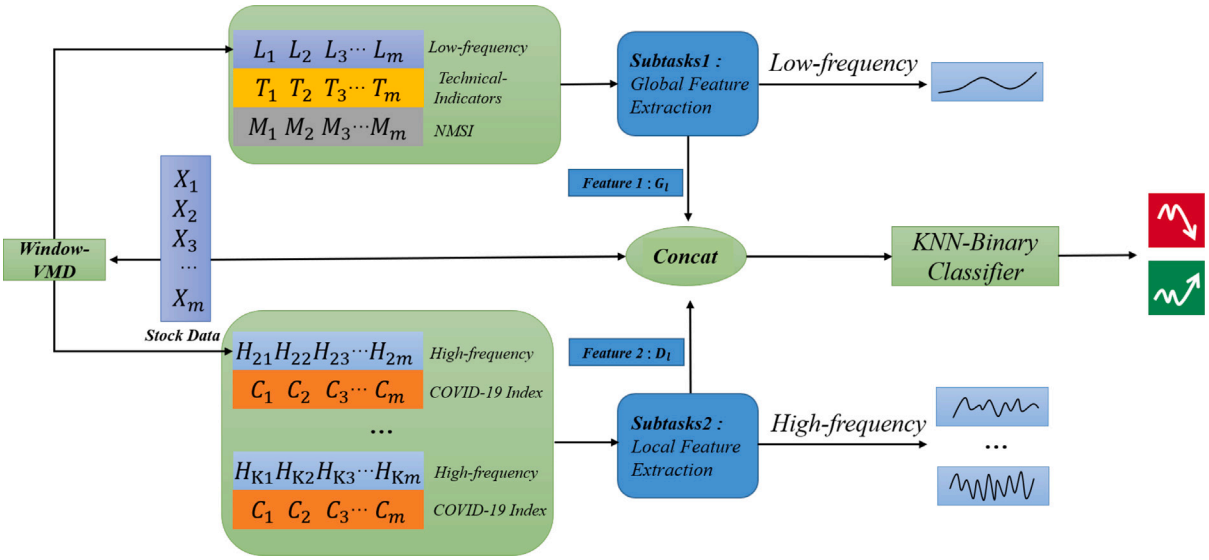

Fig. 2. COVID19-MLSF Framework.

et al. (2022) proposed an online multitask learning method with OMTL-LSSVR, which innovatively introduced online learning into multitask learning and significantly improved the prediction accuracy of the current task with continuously updated data. Nevertheless, entering the sequence of stock price into different neural network models and extracting features is the essence of the work. Then we complete the prediction after exchange these features. Because the use of stock price series only may result in similar features being extracted across different tasks, does not fully utilize the advantages of the multitask learning framework in handling multi-source data.

2.2. Prediction based on multiple features of the stock market

As the stock market becomes more prosperous, researchers are looking for more features to assist in market forecasting. It is easiest to use technical indicators, such as MA, MACD, RSI, etc., derived from stock prices, to assist in stock forecasting. These indicators are widely used because they can reflect long and short-term trends as well as phenomena such as overbought and oversold conditions (Agrawal, Khan, & Shukla, 2019; Gao & Chai, 2018). To alleviate the delaying prediction causing by single stock data, Baker and Wurgler (2006, 2007) used other data, such as macroeconomic data, to construct the BW sentiment indicators which reflect external information about the stock market. Additionally, multi-feature fusion forecasting is primarily based on news and investor sentiment; Jing et al. (2021) used stock forum text data from Oriental Fortune for sentiment analysis and CNN model for classification through Chinese sentiment corpus (ChnSentiCorp), combined with stock technical indicator information, then finally predicted by LSTM network, and obtained a low mean absolute percentage error; Zhang et al. (2018) performed a detailed analysis of the historical comments of experts in stock reviews and proposed a strategy to find a good stock reviewer by combining the trend of stock price changes and the opinion polarity (i.e., bullish or bearish) of stock reviewers to obtain high return returns in a backtest experiment. Hu, Liu, Bian, Liu, and Liu (2018) used economic news data to link specific stocks to news constructed a daily stock news corpus by linking it with stock price information, and then added stock price information to design a HAN hybrid network (a two-way GRU network incorporating Attention), which significantly improved annualized returns in real stock market simulations; with the negative impact of COVID-19 globally, some researchers have combined different perspectives on the COVID-19 and stock market linkages (Ronaghi et al., 2022; Štifanić et al., 2020). Ronaghi et al. (2022) extracted relevant data about COVID-19 from Twitter comments to build a COVID19-PRIMO Twitter dataset and then had stock prices for prediction, achieving a high accuracy rate. In order to effectively avoid the problem of data dimensional disaster, principal component analysis, as an important dimensionality reduction method in the data preprocessing stage, has shown good effectiveness in stock market prediction (Chen & Hao, 2018; Yan et al., 2020; Yue, Zhou, & Yuan, 2021; Zheng & He, 2021).

3. Prepare work

This section describes how China's Market Sentiment Index (NMSI) and COVID-19 Index are constructed, showing their relationship with the CSI 300 Index. The variational modal sliding window decomposition method used in the data pre-processing stage is also introduced.

3.1. New market sentiment index

As for choosing the original data used to construct the market sentiment index, in order to make the index highly similar(Pearson correlation coefficient)to CSI 300 (Gong, Zhang, Wang, & Wang, 2022; Yi & Mao, 2009), by continuously updating different basic data and calculating the similarity, 11 groups of basic data are selected as the basic data to create the index. Using neural networks to deal with them directly will result in redundancy of information. Therefore, we reduced dimensionality using principal component analysis. There is no need to denoise the raw data and remove outliers before doing so. This is due to the fact that these data are actual daily changes in market conditions and COVID-19. Noise reduction, particularly removing outliers, will decrease the correlation between the constructed index and the CSI 300 Index. Below is a brief description of what they mean.

Fund Discount Rate (DCEF): The extent to which the market price of a closed-end fund is below the net asset value.

Number of Initial Public Offerings (NIPO): The first time a company sells its shares to the public.

Revenue on the first day of IPO (RIPO): The yield on the first day a company sells its shares to the public for the first time.

Number of New Accounts (NA): The number of new natural persons investing, i.e. the number of new stockholders per month.

Number of New Investors (NewInvestors): The number of investors at the end of the current period minus the number of investors at the end of the previous period (the number of investors at the end of the period refers to the number of one-code accounts holding uncancelled, dormant A shares, B shares, credit accounts, and derivative contract accounts)

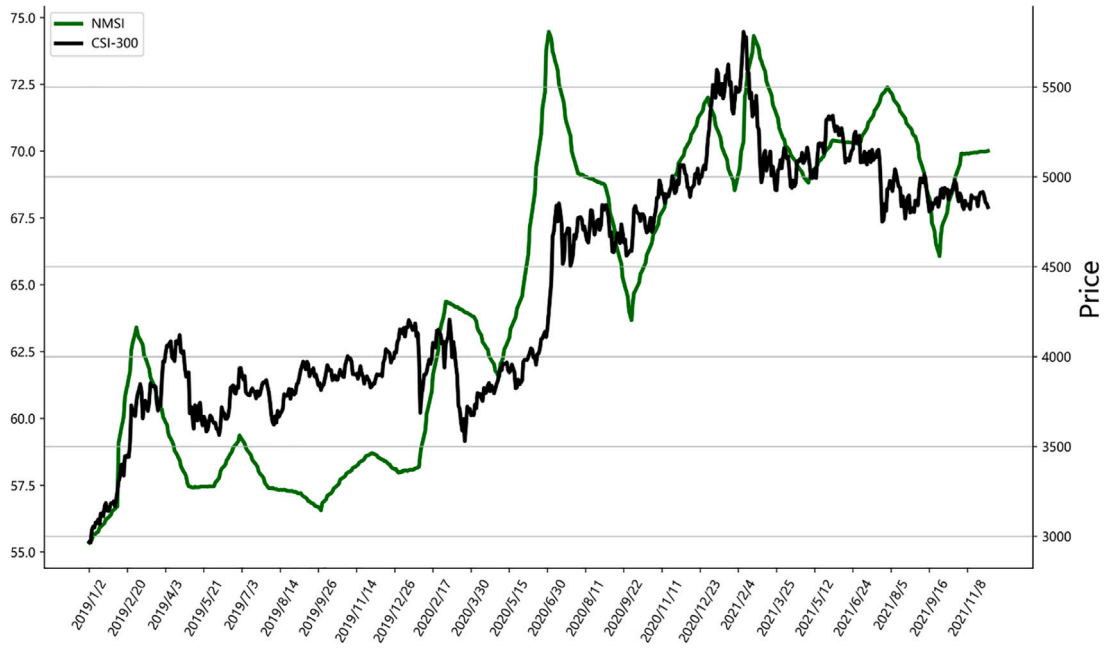

Fig. 3. NMSI and CSI 300 Trends.

Consumer Confidence Index (CCI): An indicator of the strength of consumer confidence, a comprehensive reflection and quantification of consumers' evaluation of the current economic situation and their subjective feelings about the economic outlook, income levels, income expectations and the psychological state of consumption, and a leading indicator for forecasting economic trends and consumption tendencies.

Consumer Price Index (CPI): It is a relative number reflecting the trend and degree of price changes of consumer goods and services purchased by urban and rural residents during a certain period of time.

Investor Confidence Index (ICI): The change in investment psychology and expectations of investors in the securities market under the current economic and market environment.

Social Financing Scale (AFRE): Social financing scale refers to the total amount of all funds obtained by the real economy from the financial system in a certain period of time. It reflects the relationship between finance and economy, as well as the aggregate indicator of financial support to the real economy.

Turnover Rate: It refers to the frequency of stocks changing hands in the market within a certain period of time, and is one of the indicators reflecting the strength of stock liquidity.

PE ratio: PE ratio is the ratio of stock price divided by earnings per share, and can be used as one of the indicators to assess whether the stock price level is reasonable.

For the above 11 indicators, most of them only provide monthly data because macroeconomic data indicators such as consumer confidence index and social financing scale are a level value to reflect a certain aspect of society in the current period, and their daily data have no specific reference significance. For a better correspondence with other daily data such as turnover rate and P/E ratio, and to facilitate the dimensionality reduction process, we first transform the monthly data into daily data using linear interpolation and eliminate non-trading days' data to ensure the overall trend of data distribution remained unchanged and achieve the data dimensionality correspondence. Here, we obtain the raw data matrix β of the market sentiment index (where $X, Y \ldots Z$ represent the 11 raw data and m denotes the number of rows), as shown in Eq. (1).

$$\beta = \begin{bmatrix} X_1 & Y_1 & & Z_1 \\ X_2 & Y_2 & \cdots & Z_2 \\ & \vdots & \ddots & \vdots \\ X_m & Y_m & \cdots & Z_m \end{bmatrix} \tag{1}$$

Among the many indicators mentioned above, there are many micro and macro factors that affect the stock market, which are not independent of each other. The use of neural network directly will cause information redundancy, so we use principal component analysis to reduce the dimensions. The specific steps are as follows. First, the original data matrix β is normalized using Eq. (2) to obtain the normalization matrix S.

$$S_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j}, i = 1, 2, \ldots, m; j = 1, 2, \ldots, p \tag{2}$$

where p denotes the number of columns, x_{ij} denotes the data matrix β values, and $\bar{x}_j$, σ_j denote the mean and standard deviation of each component j.

Then the initial eigenvalues of the matrix Z were extracted using Eq. (3).

$$\left| R - \lambda_j I \right| = 0 \tag{3}$$

where R denotes the correlation coefficient matrix of matrix S, and p eigenvalues λ_j are obtained. The principal components with eigenvalues greater than 1 are selected, and the score coefficient matrix A_{gj} (g is the number of principal components with eigenvalues greater than 1) and each principal component value Y_g are calculated for the original data, and then the final new market sentiment index (NMSI) M_m is calculated.

$$A_{gj} = U_g * \sqrt{Y_g} \tag{4}$$

$$Y_g = \sum \left(A_{gj} * \beta \right) \tag{5}$$

$$M_m = \sum \left(Y_g * \lambda_g \right) \tag{6}$$

where U_g is the loading of the principal component, Y_g represents the eigenvalue corresponding to each principal component, Y_g represents the gth principal component, and λ_g represents the eigenvalue corresponding to the gth principal component.

Fig. 3 illustrates the relationship between our constructed New Market Sentiment Index (NMSI) and the CSI 300 Index, and it can be observed that the trend of the constructed Market Sentiment Index is basically in line with the trend of the CSI 300 Index. In particular, the NMSI can reflect the trend of the CSI 300 index in advance, such as at

the beginning of March 2019, after the market sentiment index made a significant decline, while the CSI 300 index made a decline only in early April; between April–July 2020, the market sentiment index also rises ahead of the broad market index. We calculated the Pearson correlation coefficient of NMSI with CSI 300 using Eq. (7).

$$r = \frac{\sum_{i=1}^{m}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{m}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{m}\left(Y_i - \bar{Y}\right)^2}} \tag{7}$$

where X and Y denote the series values of NMSI and CSI 300, $\bar{X}$, $\bar{Y}$ denote the mean values of NMSI and CSI 300 series data, and m is the series length of both. the calculation results are in Table 1, and the correlation coefficients of both reach the value of strong correlation, which confirms that the constructed NMSI can reflect the trend of the broad stock market index well. NMSI does not use the stock price data, but can reflect the overall market trend, so NMSI is an important global external feature of the stock market.

3.2. COVID-19 index

Although the COVID-19 confirmed case data can reflect the changes of COVID-19, it still reflects the status of COVID-19 in an incomplete way. The data used to construct the COVID-19 Index are daily COVID-19 confirmed cases, new deaths, new cures, cumulative confirmed cases, cumulative deaths, cumulative cures, and the number of keywords "epidemic" in Baidu search and Baidu information in China. By observing the changes of COVID-19 and stock market related data and combining with news reports, we further summarize the mechanism and influence of COVID-19 on the stock market. According to our analysis, when the number of confirmed cases and deaths and the topic of "COVID-19" in online information increases, investors' investment sentiment is pessimistic, and the stock market will decline; When the topic of "COVID-19" decreases and the number of cured cases increases, investors' investment sentiment is optimistic, and the stock market will rise. Data also supports the above viewpoint, such as the Shanghai and Shenzhen 300 index rose nearly 5.8% after the Chinese government announced the free universal and widespread vaccination of the new crown vaccine on December 31, 2020. Another example is the World Health Organization (WHO) named the "Omicron" COVID-19 variant in November 2021 and confirmed a significant increase in the transmissibility of COVID-19, major stock markets in Europe and the US fell across the board, with all three major US stock indexes down more than 2%. Among them, the Dow Jones (DJI) fell 2.53%, the biggest drop of the year, while the Nasdaq (IXIC) and the S&P 500 (S&P500) fell 2.23% and 2.27%, respectively. In Europe, the major indexes also suffered heavy losses, with Germany's DAX closing down 4.15%, the largest one-day drop since 2021, and France's CAC 40 and the UK's FTSE 100 down 4.75% and 3.64%, respectively; in the Chinese market, the CSI 300 index fell 1.41% and the SSE index fell 1.47%.

Further, we use data on stock price volatility to more visually confirm the impact of COVID-19 on the stock market. As shown in Fig. 4, the relationship between monthly volatility and COVID-19 for the CSI 300 index for 2019–2021 is demonstrated. When there are concentrated outbreaks of COVID-19, such as between February–March 2020, February 2021 and July–August 2021, stock market volatility is significantly high in that month or adjacent months, indicating that the occurrence of COVID-19 exacerbates the volatile state of the stock market.

The method and procedure for constructing the COVID-19 Index are consistent with Section 3.1, and we omit the construction process here. For subsequent integration with the stock price series features, the COVID-19 Index for the period when COVID-19 did not occur is set to 0. In this way, the COVID-19 Index is expanded to a sequence C_m of the same length as the original stock price series. Fig. 5 shows the time series relationship plot of the constructed COVID-19 Index with the CSI 300 Index, and Table 1 shows the Pearson correlation coefficients of COVID-19 Index and CSI 300 Index.

Table 1
Correlation coefficients of NMSI, COVID-19 Index and CSI 300.

	Pearson correlation coefficient
NMSI-CSI 300	0.864
COVID-19 Index-CSI 300	−0.560

In Fig. 5, we observe that the trend of the whole epidemic changes from high to low, from a large national epidemic to a local epidemic, such as the epidemic in Shijiazhuang, Hebei Province in January 2021 and the epidemic in Nanjing Lukou International Airport in August 2021 are clearly reflected in the COVID-19 Index, indicating that the constructed COVID-19 Index can accurately describe the changes of the epidemic in China. In addition, the correlation coefficients in Table 1 also show that the COVID-19 Index has a significant negative correlation with the CSI 300 Index, indicating that the COVID-19 plays a certain inhibitory role on the development of the stock market. changes in the COVID-19 Index can cause short-term fluctuations in the stock market and can be used as a local feature external to the stock market, so we can use the COVID-19 19 Index's valid information for the study of short-term stock market volatility and reduce the impact of COVID-19 on stock market forecasting models.

3.3. Sliding window-VMD

Variational modal sliding window decomposition is used to decompose the stock price series with different frequency components. Numerous studies (Lahmiri, 2016b; Wu & Lin, 2019) have shown that variational modal decomposition (VMD) solves the serious endpoint effects and modal component mixing problems that occur in empirical modal decomposition, and has been widely used in biomedical signal processing (Lahmiri, 2014), time series prediction (Lahmiri, 2016b), mechanical fault diagnosis (Li, Liu, Wu, & Chen, 2020), and other fields. Due to the non-recursive signal processing method, VMD can determine the number of mode decompositions according to the actual situation. Therefore, VMD is used to decompose the stock series into low-frequency and high-frequency signal sequences by adjusting to the most appropriate number of decompositions.

For the VMD decomposition algorithm, the essence of the decomposition process is the variational problem. The existing VMD decomposition method decomposes the stock sequence $X(m) = \left[x_1, x_2, x_3 \ldots x_m\right]$.

$$\begin{gathered}\min_{(\omega_k, V_k)} \left\{ \sum_{k=1}^{n} \left\| \partial_m \left[(\delta(m) + j/\pi m) * V_k(m) \right] e^{-j\omega_i m} \right\|_2^2 \right\} \\ \sum_{k=1}^{n} V_k = X(m)\end{gathered} \tag{8}$$

where m denotes each moment of the sequence, $X(m)$ is the original sequence of stock prices, k is the number of modes, $\delta(m)$ is the Dirichlet function, $*$ denotes the convolution, $j = \sqrt{-1}$, and ∂_m is the partial derivative. After decomposition, k discrete modes are obtained, and the component of each mode k is V_k, and each V_k is concentrated around the center frequency ω_k of each eigenmodal function component. The components thus decomposed are a set of vector values; for example, the decomposition yields a low-frequency sequence that can be expressed as $V_{k=1} = \left[L_1, L_2, L_3 \ldots L_m\right]$. To better illustrate this, the first three values X_1, X_2, X_3 of the stock price series X are now decomposed to obtain the low-frequency series $V_{k=1}^{1} = \left[L_1^1, L_2^1, L_3^1\right]$, and obviously, the first three values of $V_{k=1}$ and $V_{k=1}^{1}$ are not equal, indicating that $V_{k=1}$ This set of components is obtained based on the decomposition of the whole sequence of X. The local values of $V_{k=1}$ still contain some information of the whole sequence. Therefore, using $V_{k=1}$ directly in the subsequent input to the neural network for prediction will result in future data leakage, causing the model to use the future data within the sliding window, resulting in forward-looking bias in the

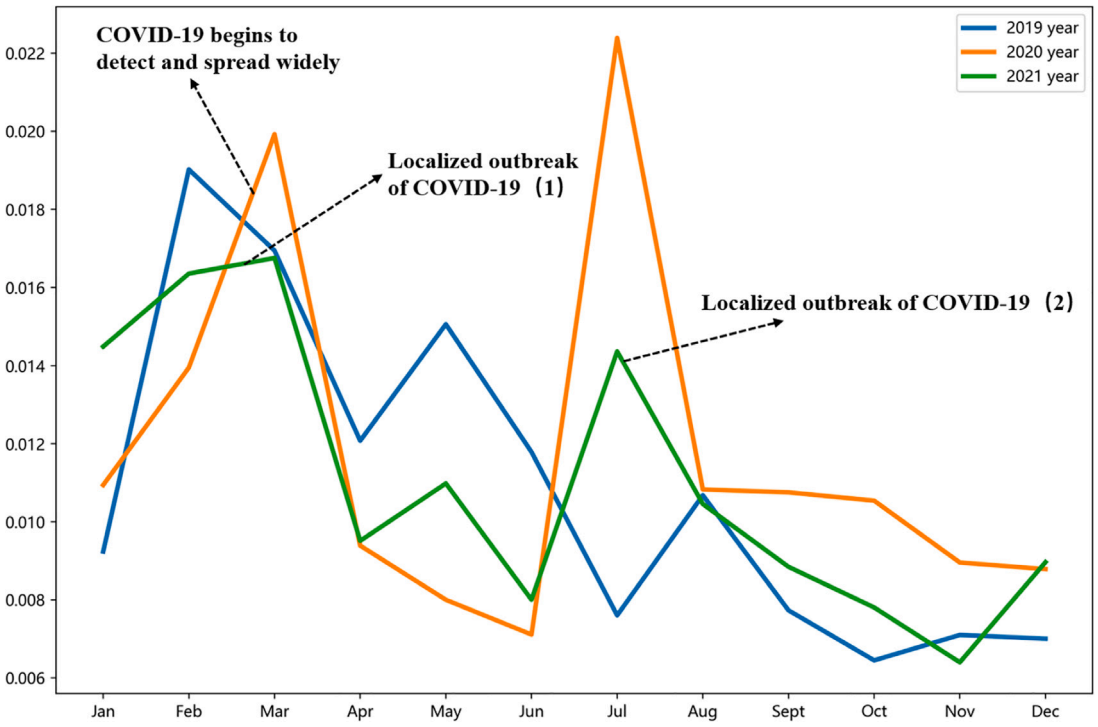

Fig. 4. CSI 300 2019–2021 Monthly Share Price Volatility and COVID-19 Relationship.

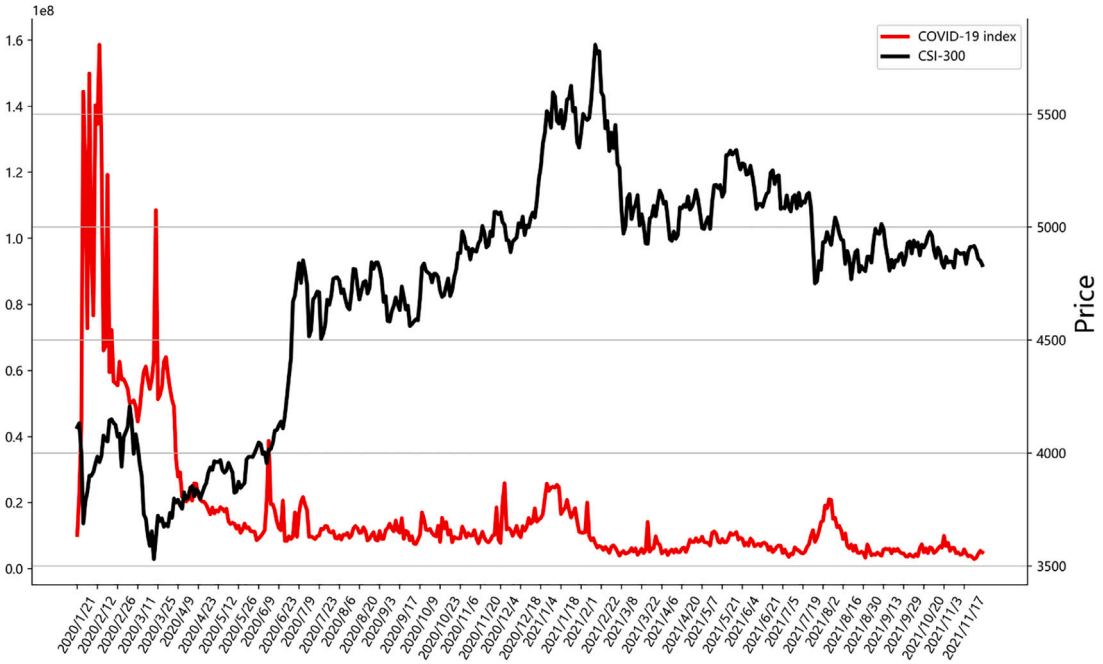

Fig. 5. COVID-19 Index and CSI 300 Trends.

model prediction, which cannot be decomposed in this way in practical applications.

The input stock price series is first sliding-window chunked, and then the series $V_{m1}, V_{m2}, V_{m3}, \ldots, V_{mk}$ is obtained after decomposing the data in each window of $X(m)$ using Eq. (8).

After the variational modal sliding window decomposition method, the low-frequency and high-frequency signal series are obtained. Among them, the high-frequency signal series reflect the short-term fluctuation features within the stock market due to the extraction of the sudden upward and downward trend of the stock price series; the low-frequency signal series reflect the long-term trend features within the stock market due to the extraction of the overall trend of the stock price series.

Stock price technical indicators are obtained from historical stock price series based on certain statistical methods, using certain mathematical formulas or quantitative models. They are a review of the historical market and can help us to a certain extent to make certain predictions about the future market. Since technical indicators such as MA, MACD, etc., are calculated based on historical stock prices over many consecutive days, they are essentially a smoothing of multi-day

Table 2
Correlation show among technical indicators.

	MA	MACD	RSI
MA	1	–	–
MACD	0.11	1	–
RSI	0.00	0.51	1

stock price data, reflecting the medium and long-term trend features within the stock market. In Table 2, we show the Pearson correlation coefficients for the three types of technical indicators included in formula (9). These three indicators have low correlations, or even no correlations at all, such as MA and RSI, etc.

$$T_m = \begin{bmatrix} \mathrm{MA5}_{\ 1} & \mathrm{MA15}_{\ 1} & & \mathrm{RSI}_{\ 1} \\ \mathrm{MA5}_{\ 2} & \mathrm{MA15}_{\ 2} & \cdots & \mathrm{RSI}_{\ 2} \\ & \vdots & \ddots & \vdots \\ MA5_m & \mathrm{MA15}_{\ m} & \cdots & \mathrm{RSI}_{\ m} \end{bmatrix} \tag{9}$$

Among them, the details of the indicators used are follow:

Moving Average(MA): The moving averages of different time windows can reflect the medium and long-term trend features of stock prices.

KDJ: Stochastic indicator. By calculating the proportional relationship between the highest price, the lowest price and the closing price of the last trading day that have occurred in a cycle, and then calculate the K value, D value and J value respectively, and draw a curve to reflect stock movement.

Convergence and Difference Moving Average(MACD): The dispersion and aggregation of the fast and slow moving averages represent the current market state and stock price trend.

Relative Strength Indicator(RSI): It is calculated based on the ratio of the rise and fall in a cycle, and reflects the degree of prosperity of the market in a certain period.

Up to this point, we have obtained multiple dimensional feature sequences for a stock, which are the low and high frequency sequences $V_{m1}, V_{m2}, V_{m3}, \ldots, V_{mk}$ of the decomposition of the stock price series, the market sentiment indicator sequence M_m, the COVID-19 Index C_m and the stock price technical indicator sequence T_m, as input data for the two subtask modules in the multi-task learning.

3.4. Summary

The non-stationary signal is decomposed using the decomposition method, and the relevant frequency signal is selected for analysis to extract the corresponding features. This paper uses the VMD method to decompose stock price series into low-frequency signals and high-frequency signals. According to the theory of VMD decomposition method, the low-frequency signal reflects the long-term trend, and the high-frequency signal reflects the short-term fluctuations. From the calculation methods of the market sentiment index(Section 3.1) and technical indicators, we can see that these two types of data are similar to low-frequency signals, reflecting the medium and long-term trend characteristics of stock price changes. We calculated the correlation of low-frequency signals, market sentiment index and technical indicators, as shown in Table 3. The low-frequency signal has a large correlation coefficient with the market sentiment index and the technical indicator MA. It can be seen that it is feasible to combine the three to extract global features. In the same way, the daily COVID-19 index has a corresponding impact on the stock market, as the variables that compose it change daily. As an example, when COVID-19 cases increased significantly, the stock market fell significantly. Therefore, COVID-19 can be used as a data feature supplement for high-frequency signals, and the designed neural network can be used to extract the features of stock price and COVID-19 index.

Table 3
Demonstration of correlation.

	NMSI	MA
Low frequency signal	0.85	0.94

4. Methodology

The COVID19-MLSF framework in this paper mainly consists of a global feature extraction module (subtask 1), a local feature extraction module (subtask 2), and a prediction result output module (main task). The global feature extraction module combines low-frequency signals, technical indicators, and NMSI to produce a feature matrix that reflects the long-term trends in the stock market, and the long-term trend features of the stock market are extracted using a newly designed multi-scale attention mechanism of the temporal convolutional neural network (MA-TCN); The MA-TCN uses a multi-layered attention mechanism that makes the TCN more responsive to important features. In the local feature extraction module, the high-frequency signals that reflect the short-term fluctuations of stock series and the constructed COVID-19 Index are jointly modeled to build a feature matrix that reflects the short-term fluctuations of the stock market, and the short-term fluctuations in the stock market, and these short-term fluctuation are extracted by our designed, multi-view convolutional-bidirectional recurrent neural network with temporal attention (MVCNN-BiLSTM-Att). The MVCNN-BiLSTM-Att employs multiple convolutional neural networks with different sensory fields, and incorporates a temporal attention mechanism in the recurrent neural network, which not only enhances the local feature extraction capability of the model, but also focuses on capturing the impact of the changing COVID-19 state on stock market volatility. In the above two modules, both TCN and BiLSTM structures are non-parametric machine learning models, which allow to combine continuous data (such as market data) with categorical ones(such as COVID-19 data). Finally, the features extracted from the two subtask modules and stock price sequences are fed into a KNN model for stock market trend prediction.

4.1. Subtasks1:Global feature extraction module

In the subtask 1 module, the low-frequency sequence V_{m1}, the market sentiment indicator sequence M_m and the stock price technical indicator sequence T_m 3-dimensional feature sequence obtained by decomposing the stock price sequence, their combination forms a multi-dimensional feature matrix $H\left[V_{m1}, M_m, T_m\right]$, and because the feature sequences of the three dimensions in H all reflect the long-term trend of stock prices, H can also be seen as a global feature matrix. This multidimensional global feature matrix is input to the global feature extraction module (structure shown in Fig. 6) for feature extraction. TCN can effectively alleviate the gradient disappearance problem of long time series prediction by causal convolution; its powerful multi-layer convolution kernel can efficiently extract some important information of the series. Based on these two advantages of TCN, we optimize the traditional temporal convolutional neural network and propose the multi-scale attention mechanism of temporal convolutional neural network (MA-TCN). The traditional temporal convolutional neural network performs the same level of convolution for the entire sequence, so it cannot extract the influence of different features. The MA-TCN introduces self-attention and attention mechanisms into different parts of the temporal convolutional neural network, which can dynamically adjust the weight parameter of each feature attribute, resulting in a model that is better suited to meet the actual situation and that fully explains the influence of the different attributes on the long-term trend.

Here the feature matrix H is input, and the objective is to extract the long-term trend features of the stock price changes. H with self-attention mechanism, assigning different weights to the information inside the feature matrix, and subsequently using the neural network

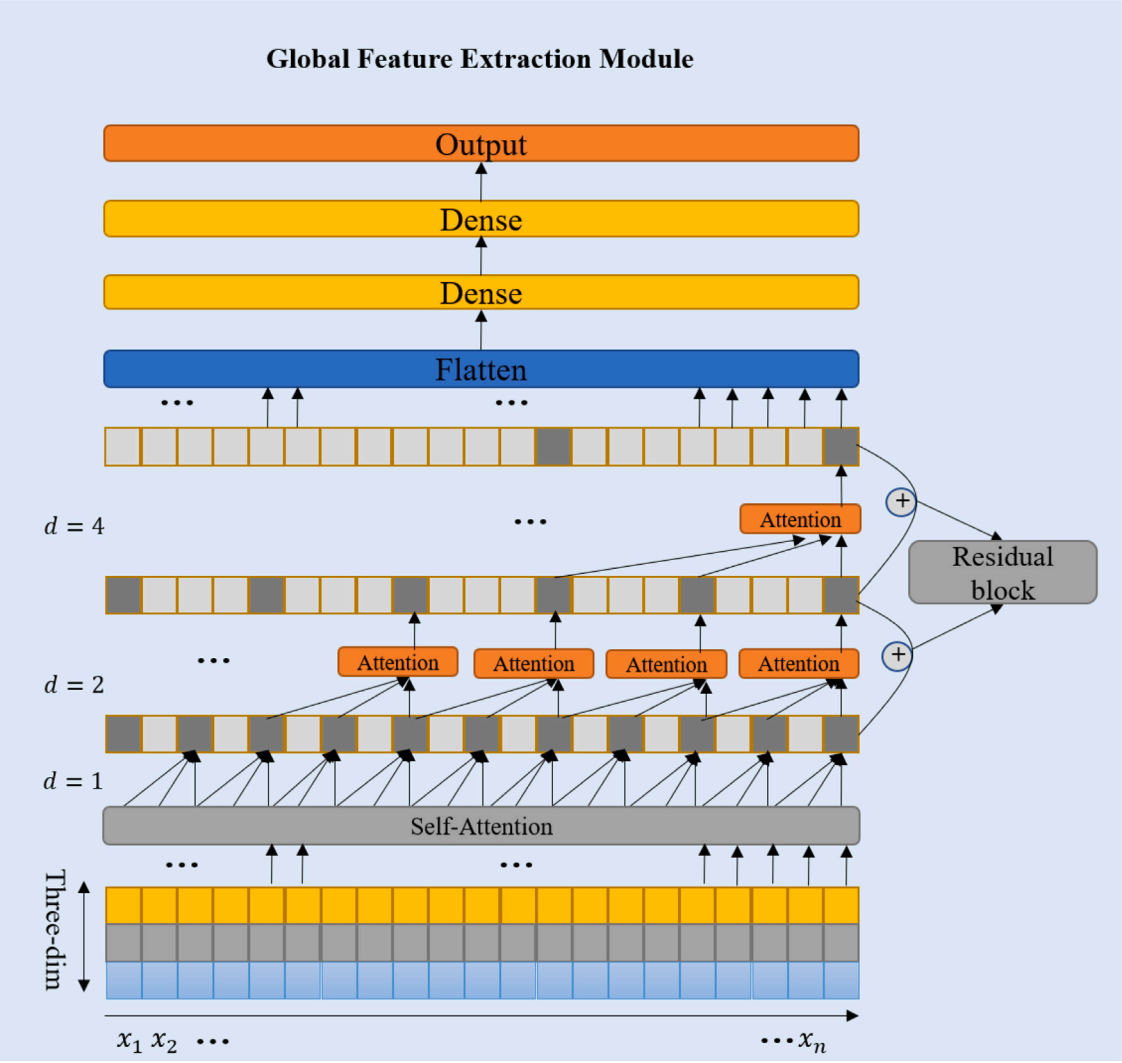

Fig. 6. Global feature extraction module structure.

for feature extraction and distance learning to improve the efficiency of feature extraction. The formula is:

$$H^1 = \text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \tag{10}$$

Where Q, K, V are the same tensor as H as shown in Eq. (11), d denotes the feature dimension of Q, and $softmax$ denotes the activation function.

$$\begin{aligned} Q &= \boldsymbol{W}^q H \\ K &= \boldsymbol{W}^k H \\ V &= \boldsymbol{W}^v H \end{aligned} \tag{11}$$

Where $\boldsymbol{W}^q, \boldsymbol{W}^k, \boldsymbol{W}^v$ denote the parameter matrix. Then the obtained H^1 feature matrix is input to the temporal convolutional neural network with the attention mechanism. Attention mechanism is added to each dilated convolutional layer, the purpose of which is to assign weights to the results of convolution operations in the process of feature extraction, which improves the previous traditional temporal convolutional neural network for single feature extraction from sequences. Which significantly improves the feature extraction capability and effectiveness. The feature extraction process of temporal convolutional neural network is as follows:

$$\mathcal{F}(t) = \sum_{i=1}^{k} f_i H^1_{t-d\cdot i} \tag{12}$$

where $f = (f_1, f_2, \ldots, f_k)$ denotes the convolutional kernel, k is the size of convolutional kernels, d is the expansion coefficient, and $H^1_{t-d\cdot i}$ represents the feature matrix before moment t. Where each residual module is subjected to two $\mathcal{F}(\cdot)$ transformations and the activation function uses the ReLU activation function,

$$\text{Residual block} = \text{Activation}\left(H^1 + \mathcal{F}\left(H^1\right)\right) \tag{13}$$

Using the temporal attention mechanism to get the final feature results,

$$G_l = \sum_{t=1}^{l} \frac{\exp\left(\beta^{\mathrm{T}} \tanh\left(\omega_r \mathcal{F}(t) + b_r\right)\right)}{\sum_{t=1}^{n} \exp\left(\beta^{\mathrm{T}} \tanh\left(\omega_r \mathcal{F}(t) + b_r\right)\right)} \mathcal{F}(t) \tag{14}$$

Where β^{T}, ω_r are the parameter matrices, b_r denotes the bias vector, $tanh$ is the activation function, and l is the training data length.

The features are extracted by a temporal convolutional neural network with a multiscale attention mechanism and then changed to one dimension by a decoding layer, and finally output by two fully connected layers. Using this module, extract the global features $G_l\left[G_1, G_2 \ldots G_l\right]$.

4.2. Subtasks2:Local feature extraction module

This subsection details the modules of subtask 2. At this point we have obtained k-1 high frequency signal sequences $V_{m2}, V_{m3}, \ldots, V_{mk}$ in

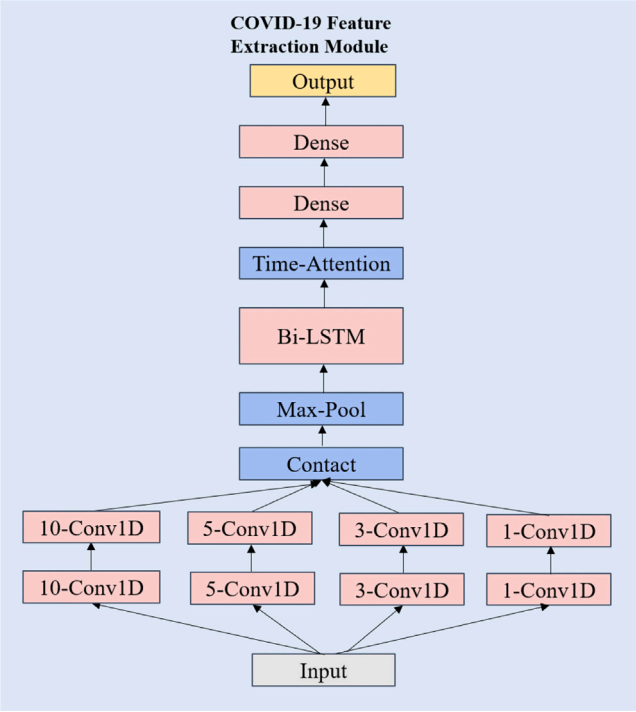

Fig. 7. COVID-19 feature extraction module structure.

addition to the first low frequency signal sequence, followed by feature fusion with the constructed COVID-19 Index to form local feature matrices

$S1\left[V_{m2}, C_m\right], S2\left[V_{m3}, C_m\right], \ldots,$ of number $k-1$ and length m The COVID-19 feature extraction module is used for feature extraction, and the specific structure of the COVID-19 feature extraction module is shown in Fig. 7.

In Fig. 7, we input the feature matrix S into the MVCNN-BiLSTM-Att, and first design two layers of one-dimensional convolutional neural network with different perceptual fields. Multiple parallel convolutional neural networks can extract more complete local information than a single one-dimensional convolutional neural network. CNN extracts features from the feature matrix for input into Bi-LSTM. By combining future and past information, Bi-LSTMs are capable of more effectively memorizing data features;and the constant change of COVID-19 has different degrees of influence on the stock market, we added the Time-attention mechanism to extract the degree of the effect of the changing COVID-19 status on the stock market. The structure of BiLSTM-Att is shown in Fig. 8. In Bi-LSTM, it is updated according to Eqs. (15)–(21).

$$i_t = \sigma\left(W_i\left[h_{t-1}, x_t\right] + b_i\right) \tag{15}$$

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right) \tag{16}$$

$$\tilde{C}_t = \tanh\left(W_C\left[h_{t-1}, x_t\right] + b_C\right) \tag{17}$$

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{18}$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t \tag{19}$$

$$h_t = o_t \circ \tanh\left(C_t\right) \tag{20}$$

$$B_t = \left[\overrightarrow{h_t}, \overleftarrow{h_t}\right] \tag{21}$$

Where Eq. (15)–Eq. (20) are the steps of LSTM and Eq. (21) is the step of Bi-LSTM. σ, $tanh$ denote the activation function, W_i, W_C, W_f, W_o denote the parameter matrix, b_i, b_c, b_f, b_o denote the bias vectors, where $\circ$ denotes the Hadamard product (element-wise multiplication), h_{t-1} denotes the hidden state value at the previous moment, x_t denotes the input at the current moment, $\tilde{C}_t$ denotes the temporary hidden variable at the current moment, $\overrightarrow{h_t}$ denotes the cell forward structural state, $\overleftarrow{h_t}$ denotes the cell backward structural state, and B_t is the output of the Bi-LSTM.

The formula for Time-Attention is as follows:

$$\zeta_t = a^{\mathrm{T}} \tanh\left(\omega B_t + b_v\right) \tag{22}$$

$$P_t = \frac{\exp\left(\zeta_t\right)}{\sum_{t=1}^{n} \exp\left(\zeta_t\right)} \tag{23}$$

$$D_t = \sum_{t=1}^{n} P_t B_t \tag{24}$$

where a, ω denote the parameter matrix, ζ calculates the attention weights for B_t. Finally, after calculating the probability P_t of the attention weight, perform a weighted summation,calculate output D_t

After the features are extracted in BiLSTM-Att, they are then output after two fully connected layers. Using this module,extracted $k-1$ local features $D_l\left[D_1, D_2, \ldots, D_l\right]$.

4.3. Maintasks:Prediction result output module

This section describes the main task module. In most of the literature (Rezaei et al., 2021; Ronaghi et al., 2022; Yan et al., 2020), after the neural network extracts the sequence features, the resultant output is performed using the Fully Connected Layer (FCL). The output of the fully connected layer depends heavily on the parameter settings of the anterior neural network, one parameter often has an important influence on the whole model's results, but our model separates the tasks of different modules, reducing the global impact of parameters in different tasks.

The output model of the main task is the K-nearest neighbor model. KNN is a simple algorithm with good classification effects and has been widely used in stock prediction (Chen & Hao, 2017; Nayak, Mishra, & Rath, 2015), and it does not require training sample data nor estimation of parameters, which makes it very suitable for further prediction based on the extracted features. The ablation experiment in Section 5.3.2 shows that the accuracy of the prediction result is about 13% higher than FCL when the features extracted from the network are used.

In implementing the KNN algorithm, the hyperparameter K and the distance measure have a great impact on performance. Regarding the distance measure, we calculate it by the reciprocal of the Euclidean distance,

$$d\left(x_j, x_i\right) = \frac{1}{\sqrt{\sum_{p=1}^{n}\left(x_p^* - x_p\right)^2}}, p = 1, 2, \ldots, l \tag{25}$$

where $x_j^*(j = 1, 2, \ldots, n)$ is the data to be classified in the test set and $x_i(i = 1, 2, \ldots, l)$ is the training set of known data. When the point to be predicted is closer to the sample point, the weight occupied will be larger, and vice versa, the weight will be smaller. Then we continue to fuse the stock price series X_l and the global features G_l (feature 1) and local features D_l (feature 2) obtained in Sections 4.1 and 4.2 into a feature matrix $W_l = \left[X_l, G_l, D_l\right]$, and set the stock price series X_t up or down coding using Eq. (26) as the label values for classification.

$$X_t = \begin{cases} 1 & X_{t+1} \geq X_t \\ 0 & X_{t+1} < X_t \end{cases} \tag{26}$$

At this point, the stock forecasting problem is described:

$$\hat{y}^{(t+1)} = F\left(W_l^{t-\partial}, W_l^{t-\partial+1}, \ldots, W_l^{t}\right) \tag{27}$$

where $\hat{y}^{(t+1)}$ is the prediction result, $F(\cdot)$ is the set KNN classification model, and ∂ is the number of days of input data.

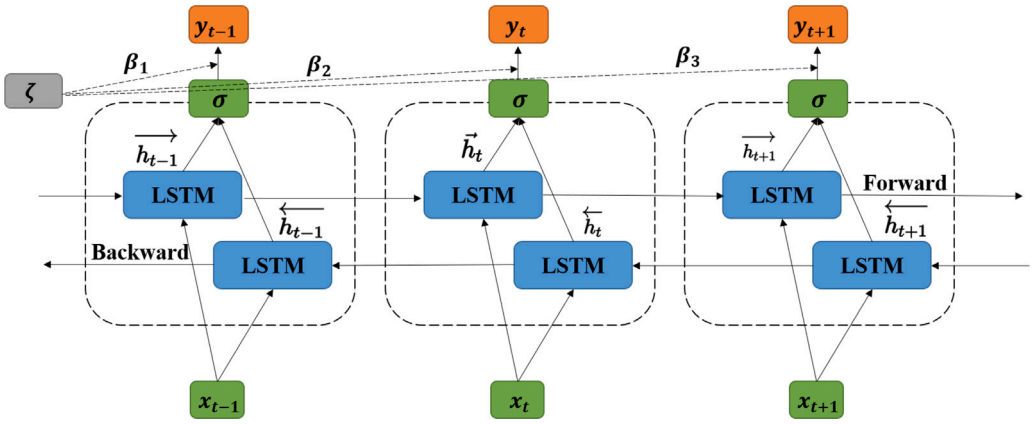

Fig. 8. BiLSTM-Att structure.

4.4. Network optimization and loss function

In COVID19-MLSF, the losses of both subtask modules are optimized using MSE:

$$\text{loss} = \frac{1}{l}\sum_{i=1}^{l}\left(\hat{y}_i - y_i\right)^2 \tag{28}$$

where $\hat{y}_i$ denotes the predicted value of the network, y_i denotes the true value, and the optimizer is Adam.

5. Analysis of experiments

Experimental results and comparisons are presented in this section to evaluate the proposed hybrid COVID19-MLSF framework for the task of stock market forecasting. Specifically, it includes the data introduction part, the evaluation index introduction part, the comparative experiment part, the ablation experiment part, and the parameter sensitivity test part.

5.1. Data and data preprocessing

This section describes the numerous data sources and data preprocessing methods used in this paper. The NMSI and COVID-19 Index constructed in Section 3 uses the DCEF, NIPO, RIPO, NA, NewInvestors, CCI, CPI, China daily COVID-19 new confirmed, new death, new cure, cumulative confirmed, cumulative death and cumulative cure cases from CSMAR database (https://www.gtarsc.com); Baidu search and information data from Baidu index (https://index.baidu.com); ICI data from (http://www.sipf.com.cn); AFRE data from (http://www.pbc.gov.cn), Turnover Rate, P/E Ratio, and CSI 300 Index data from Tushare (https://www.tushare.pro). The time period for all the above data is January 2019 to November 2021. Among them, stock price data, new confirmed cases of COVID-19, death cases, cured cases, cumulative confirmed cases, death cases, cured cases and Baidu Index data are all daily frequency data.

The stock data predicted by our model is China CSI 300 Index, and we divide 70% of the data as the training set and 30% as the test set. Since we construct the NMSI and COVID-19 index to reflect the Chinese stock market, we need to select an index that reflect the overall trend of the Chinese stock market, and the CSI 300 is the best choice.

This paper normalize the stock data and the feature date in the range [0,1],

$$x_* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{29}$$

Among of them, x is the original data, x_* is the normalized data.

5.2. Performance evaluation metric

In order to show the COVID19-MLSF performance, we use some classification metrics and stock return metrics for the presentation of the results. Among the evaluation indicators for classification are Accuracy (ACC), Mathews Correlation Coefficient (MCC), and F-score (F-score). The return indicators are Profit, Maximum Drawdown, the calculation formula of the classification evaluation index is as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{30}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \tag{31}$$

$$F - \text{score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{32}$$

where TP denotes the number of positive examples predicted by the classifier, which means are positive value, FP denotes the number of negative examples predicted by the classifier, which means are positive value, FN denotes the number of true positive examples predicted by the classifier as, which means are negative value, TN denotes the number of negative examples predicted by the classifier, which means are negative value. P_t is the net value of the product in a certain day, and P_y is the net value of the product in a certain day after t, k is the length of time. During the calculation of the profit, we simulate a real market transaction by setting the principal amount to RMB 1,000,000, but disregarding the transaction fees, which is incurred for buying and selling, etc. (Zhang et al., 2018). Since we predict the CSI 300 index, we consider the transaction as making a stock index futures.

$$\text{Profit} = \frac{P_t}{P_{t-k}} - 1 \tag{33}$$

$$\text{Maximum Drawdown} = \frac{\text{Max}\left(P_x - P_y\right)}{P_x} \tag{34}$$

where *Recall* denotes the recall rate, *Precision* denotes the precision rate,The calculation formula is:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{35}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{36}$$

5.3. Experimental results and performance comparison

This subsection presents the results of comparative experiments, ablation experiments and parameter sensitivity tests are presented, which further confirm the effectiveness of the features and models in Sections 3 and 4.

Table 4
Comparative display of experimental results.

	ACC	MCC	F-score	Profit	Maximum-Drawdown
COVID-19-MLSF	**58.85%**	**17.81%**	57.75%	**8.81%**	**7.93%**
BGVAR	49.01%	2.21%	52.09%	−7.25%	11.74%
MIDAS	53.47%	7.37%	57.39%	−10.84%	11.86%
Bayesian	52.08%	4.05%	**58.93%**	−7.97%	14.44%
Random Forest	54.17%	8.43%	52.69%	−2.82%	10.71%
LSTM	52.13%	6.17%	55.88%	−6.38%	10.56%
Adv-LSTM	55.73%	11.44%	56.41%	−2.12%	10.56%
SCINet	53.65%	7.22%	56.58%	2.10%	9.35%
Bi-LSTMSeq2Seq	52.60%	5.26%	51.85%	−3.76%	10.86%
Naive	43.23%	−13.58%	44.10%	−16.75%	18.59%
Buy and Hold	–	–	–	−14.61%	18.10%

Table 5
Ablation display of experimental results.

	ACC	MCC	F-score	Profit	Maximum-Drawdown
COVID-19-MLSF	**58.85%**	**17.81%**	**57.75%**	**8.81%**	7.93%
COVID-19-MLSF1	57.29%	14.66%	56.39%	2.14%	8.50%
COVID-19-MLSF2	56.25%	12.82%	52.81%	2.65%	**6.89%**
COVID-19-MLSF3	54.27%	8.66%	52.22%	1.45%	9.33%
COVID-19-MLSF4	56.77%	13.48%	53.84%	3.10%	8.54%
COVID-19-MLSF5	55.43%	12.22%	51.93%	2.44%	8.76%
COVID-19-MLSF6	51.22%	3.42%	48.88%	−6.31%	9.54%
COVID-19-MLSF7	53.48%	6.92%	49.07%	−5.97%	10.67%
KNN	50.52%	1.38%	43.11%	−5.36%	8.43%

5.3.1. Performance comparison experiment

First, we compared some traditional models for financial time series forecasting that use econometrics, including BGVAR (Ahelegbey, Billio, & Casarin, 2016; Ahelegbey et al., 2022), MIDAS (Gunay, Can, & Ocak, 2020). Then, we use normal machine learning classification models such as Random Forest, Bayesian, and LSTM. To further verify the superior performance of our model, we compared two advanced single-task models, Adv-ALSTM (Feng et al., 2018), SCINet (Liu, Zeng, Xu, Lai, & Xu, 2021), and multi-task model Bi-LSTM Seq2Seq (Mootha et al., 2020). In addition, we are also compare Naive algorithm (Cui, Xie, & Zheng, 2021) and the buy and hold algorithm. Here are brief descriptions of them:

BGVAR:The method Bayesian graph-based approach to identification in vector autoregressive (VAR) models.

MIDAS:The method allows different frequencies to be used in a regression model.

Random Forest:The method is to integrate many decision trees into a forest and use it to predict the final result.

Bayesian:The method calculates the probability that a classification object belongs to a certain class, and selects the class with the largest posterior probability as the class to which the object belongs.

LSTM:The method trains price features constructed based on historical sequence data, obtains sequence embeddings, and then uses a fully connected layer to predict.

Adv-ALSTM: This method proposes the use of adversarial training to train the model, which significantly improves the performance of the LSTM model.

SCINet: This method constructs the basic block SCI-Block, downsamples the input features into two subsequences, and then extracts each subsequence feature using different convolutional filters to retain the information of different features, and adds the learning of convolutional features between the two sequences in each SCI-Block, finally constructing a multilayer neural network framework for prediction.

Bi-LSTM Seq2Seq: This method constructs an encoder and decoder using Bi-LSTM to input the multidimensional price features of the stock and predicts multiple price series, such as closing price and opening price using a multitask learning framework.

Naive: This method use today's up and down as tomorrow's buying and selling signals.

In Table 4 shows the comparison results of our proposed COVID19-MLSF model with some single-task and multi-task models, therefore COVID19-MLSF model achieves good results in the field of stock prediction.

Table 4 compares our proposed COVID19-MLSF model with some single-task and multi-task models, thus COVID19-MLSF model achieves good results in the field of stock prediction. The bar chart in Fig. 9 shows how our model fares against other models for each evaluation metric, and it is obvious that our method fares better than other models for each evaluation metric. In Fig. 10, we show the change in cumulative returns between February 2021 to November 2021 for our model and the comparison model. This phenomenon was in a bear market according to the buy-and-hold return. Table 4, Fig. 9 and Fig. 10 demonstrate that using machine learning methods does increase the stock market returns. Among many models, our model has achieved the highest ACC, MCC, profit and low fallback rate, which indicates that our model not only provides benefits, but also reduces certain risks, including those caused by external macroeconomic policy adjustments in the stock market and COVID-19 factors.

5.3.2. Ablation experiment

To evaluate the construct external features of the stock market and the effectiveness of the design module, we transform COVID19-MLSF into the following six models, and conduct experiments on the same dataset.

(1) COVID19-MLSF1 means using NMSI alone, COVID19-MLSF2 means using COVID-19 Index alone for prediction.

(2) COVID19-MLSF3 shows that the variational modal sliding window decomposition method is not used.

(3) COVID19-MLSF4 shows that in the COVID-19 feature extraction module, the LSTM module is used to replace the MVCNN-BiLSTM-Att.

(4) COVID19-MLSF5 shows that in the global feature extraction module, the TCA module is used to replace the MA-TCN.

(5) COVID19-MLSF6 shows that the decomposed stock price sequence and constructed features are all have into a feature matrix for prediction.

(6)COVID19-MLSF7 shows that NMSI and COVID-19 Index were combined with the opposite frequency signal.

(7) KNN shows the prediction result of the main task module, which is used to both comparison and ablation experiments here.

In Table 5 the specific results of the ablation experiments are shown, and in Fig. 11, the bar chart shows the indicators between the model and each ablation experiment part.

Based on Table 5, we can draw the following conclusions can be got from the experimental results in Table 5. The COVID19-MLSF1 and COVID19-MLSF2 experiments proves that the constructed NMSI and COVID-19 Index features are effective. The COVID19-MLSF3 experiment proves that the decomposition of stock price series is very effective, and the decomposition of stock price series significantly improves the accuracy of stock forecasting. The COVID19-MLSF4 experiment proves that MVCNN-BiLSTM-Att is more effective than the traditional recurrent neural network model. The COVID19-MLSF5 experiment proves that MA-TCN is more effective than the traditional TCN. Multi-task learning with multiple data sources, which improves the feature extraction capability of complex data, is demonstrated by the experiments of the COVID19-MLSF6 and KNN experiments. COVID19-MLSF7 experiment proves NMSI is more effective when combined with low-frequency signals, and COVID-19 Index is combined with high-frequency signals.

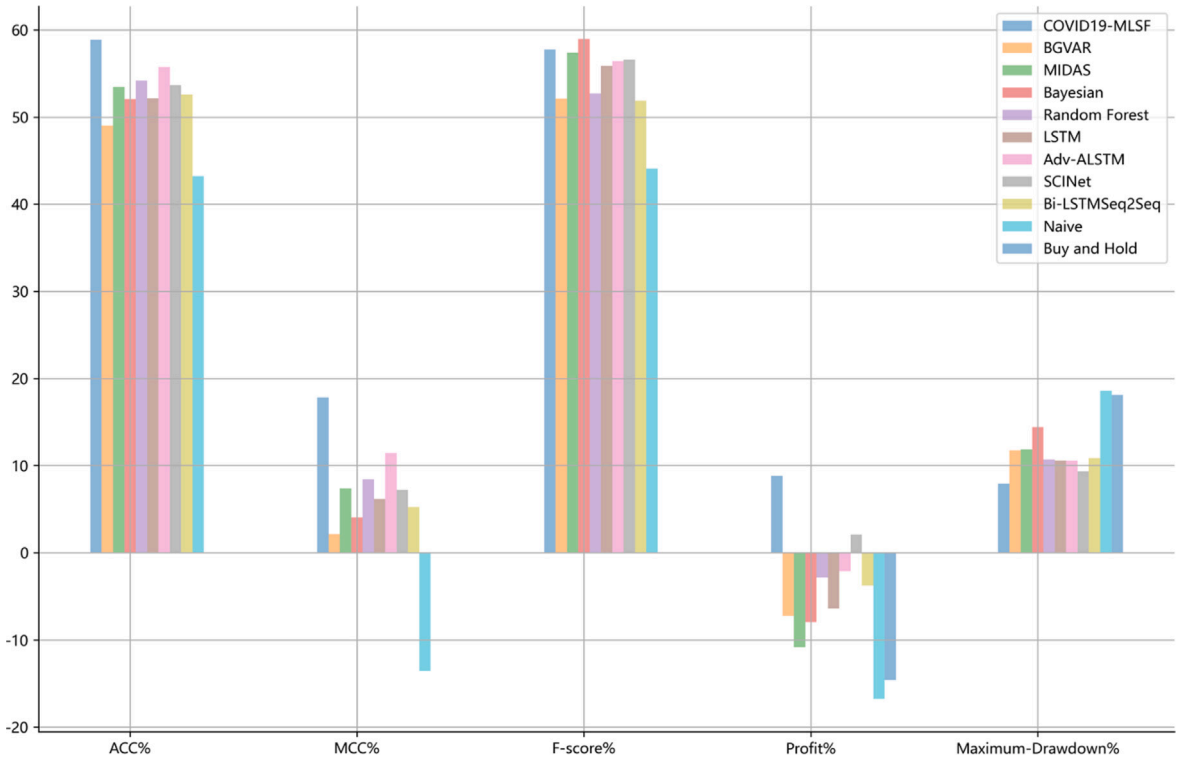

Fig. 9. Comparison of the metrics of COVID19-MLSF model and the comparison model.

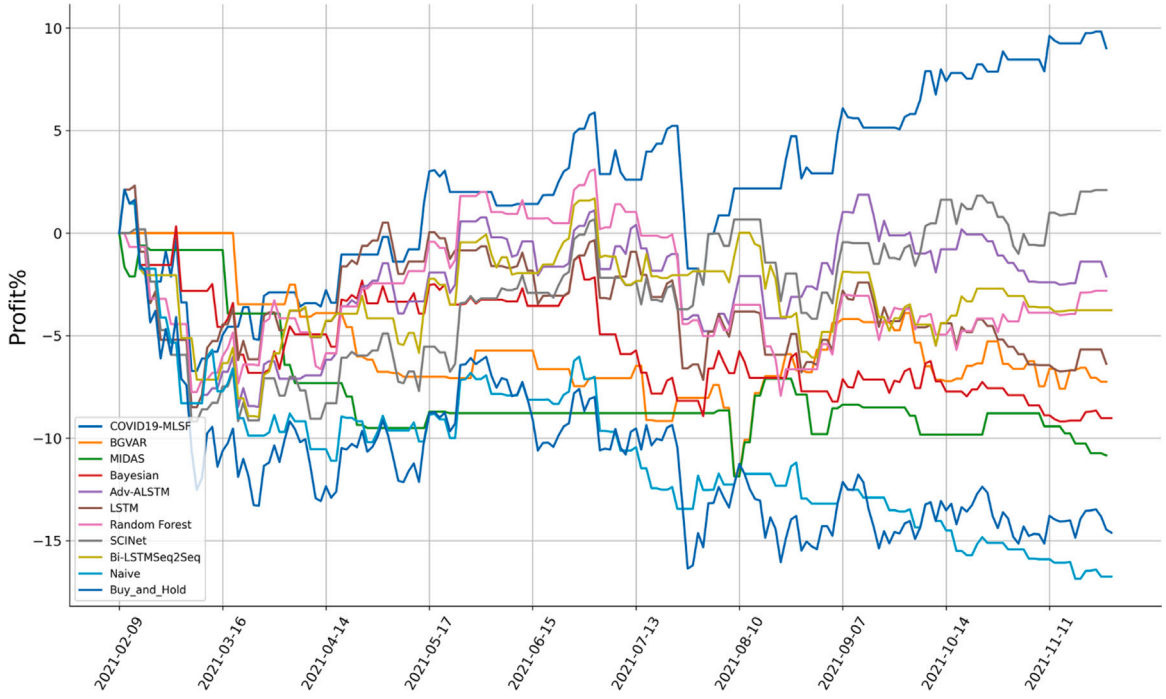

Fig. 10. Cumulative Return Comparison.

5.3.3. *Hyperparameter sensitivity experiment*

In KNN, the choice of K values is crucial for the final prediction results, so we performed a hyperparametric sensitivity analysis. Table 6 shows the prediction results based on the main task alone for different input days ∂, along with the results of joint prediction based on the features extracted from subtasks and closing prices. In Fig. 12, the data in Table 6 are presented as a line chart, allowing a more direct understanding of the data. It can be clearly that our predictions with the addition of subtask features are higher than the single main task, further demonstrating that our extracted features are effective when the

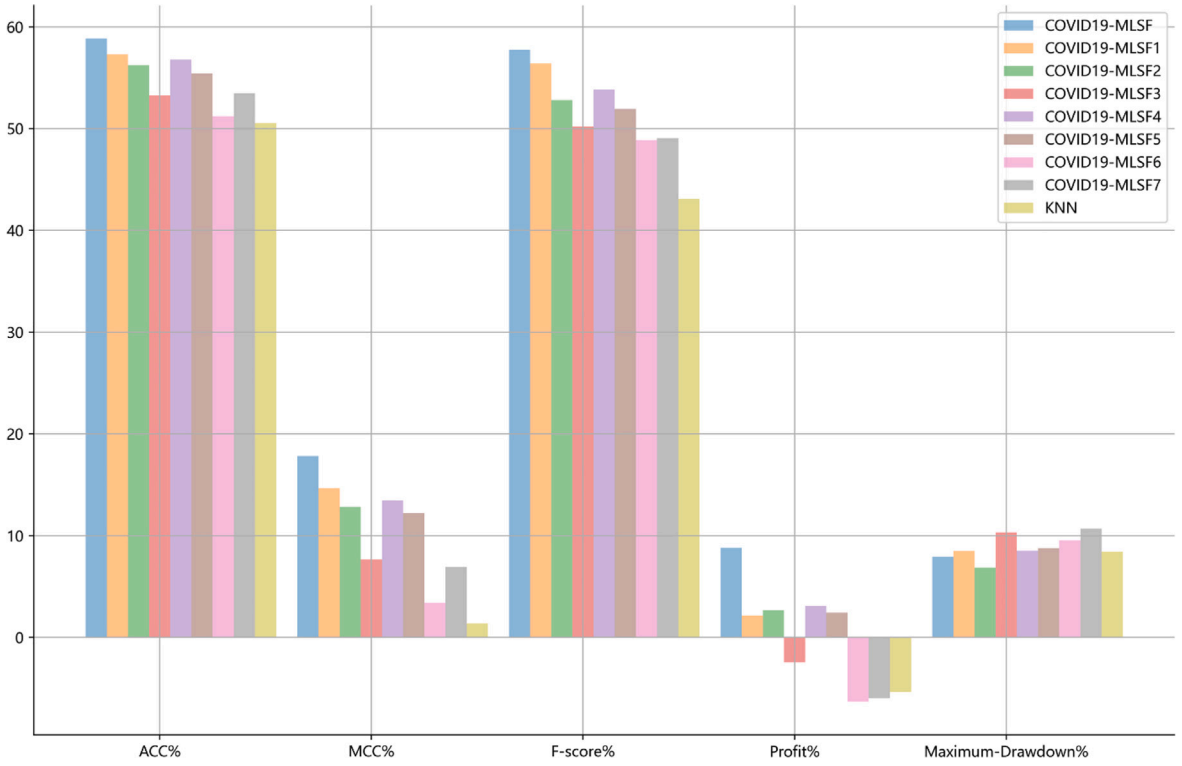

Fig. 11. Comparison of various indicators between the COVID19-MLSF model and the ablation experimental model.

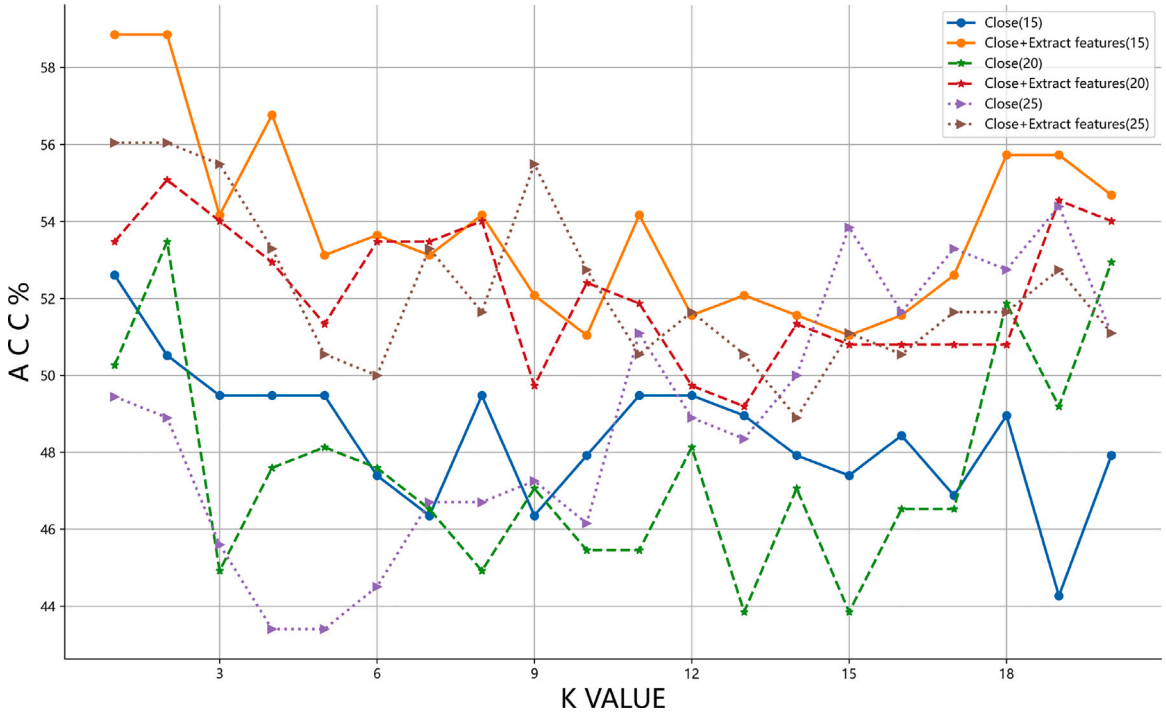

Fig. 12. The relationship between accuracy and hyperparameter K value.

input days ∂ and K values are the same. Furthermore, the multi-task model that we designed can improve significantly.

6. Conclusion and future work

By utilizing a multi-task learning framework, we propose a COVID19-MLSF stock forecasting framework that fully extracts the internal and external features of the Chinese stock market under COVID-19 pandemic. As a method of describing the external features of the stock market, this paper uses the newly constructed stock market sentiment index (NMSI) and the COVID-19 index. Two prediction subtask are established by combining them with decomposed stock price series separately, in which one uses MA-TCN to extract the globally important stock market features, while the other uses

Table 6
The relationship between prediction accuracy and hyperparameter K value under different input days ∂.

Enter days	∂=15		∂=20		∂=25	
Hyperparameter K	Close	Close + Extract features	Close	Close + Extract features	Close	Close + Extract features
1	52.60%	**58.85%**	50.27%	**53.48%**	49.45%	**56.04%**
2	50.52%	**58.85%**	53.48%	**55.08%**	48.90%	**56.04%**
3	49.48%	**54.17%**	44.92%	**54.01%**	45.60%	**55.49%**
4	49.48%	**56.77%**	47.59%	**52.94%**	43.41%	**53.30%**
5	49.48%	**53.13%**	48.13%	**51.34%**	43.41%	**50.55%**
6	47.40%	**53.65%**	47.59%	**53.48%**	44.51%	**50.00%**
7	46.35%	**53.13%**	46.52%	**53.48%**	46.70%	**53.30%**
8	49.48%	**54.17%**	44.92%	**54.01%**	46.70%	**51.65%**
9	46.35%	**52.08%**	47.06%	**49.73%**	47.25%	**55.49%**
10	47.92%	**51.04%**	45.45%	**52.41%**	46.15%	**52.75%**
11	49.48%	**54.17%**	45.45%	**51.87%**	51.10%	50.55%
12	49.48%	**51.56%**	48.13%	**49.73%**	48.90%	**51.65%**
13	48.96%	**52.08%**	43.85%	**49.20%**	48.35%	**50.55%**
14	47.92%	**51.56%**	47.06%	**51.34%**	50.00%	48.90%
15	47.40%	**51.04%**	43.85%	**50.80%**	53.85%	51.10%
16	48.44%	**51.56%**	46.52%	**50.80%**	51.65%	50.55%
17	46.88%	**52.60%**	46.52%	**50.80%**	53.30%	51.65%
18	48.96%	**55.73%**	51.87%	**50.80%**	52.75%	51.65%
19	44.27%	**55.73%**	49.20%	**54.55%**	54.40%	52.75%
20	47.92%	**54.69%**	52.94%	**54.01%**	51.10%	51.10%
Avg	48.44%	**53.83%**	47.57%	**52.19%**	48.87%	**52.25%**

MVCNN-BiLSTM-Att to extract other local stock market features and the degree of COVID-19 impact on the stock market. Our model achieves good results in predicting the Chinese stock market during COVID-19, reduces the impact of COVID-19 on the prediction model, particularly in constructing a more effective external feature of the stock market, which provides ideas for research into the impact of emergencies on the stock market.

With the subsequent unpredictable development of COVID-19, we can also look for new features, further study the impact of COVID-19 on individual stocks in different industries, and make use of machine learning and deep learning techniques to reduce the negative impact of COVID-19 on specific industries.

CRediT authorship contribution statement

Chenxun Yuan: Software, Investigation, Writing – original draft. **Xiang Ma:** Software, Conceptualization, Visualization. **Hua Wang:** Software, Writing & editing. **Caiming Zhang:** Visualization, Investigation. **Xuemei Li:** Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

We thank the anonymous reviewers for their constructive comments. This work was supported in part by the National Natural Science Foundation of China (Grant No. 62072281, No. 62007017)

References

Agrawal, M., Khan, A. U., & Shukla, P. K. (2019). Stock price prediction using technical indicators: a predictive model using optimal deep learning. *Learning*, *6*(2), 7.

Ahelegbey, D. F., Billio, M., & Casarin, R. (2016). Bayesian graphical models for structural vector autoregressive processes. *Journal of Applied Econometrics*, *31*(2), 357–386.

Ahelegbey, D. F., Cerchiello, P., & Scaramozzino, R. (2022). Network based evidence of the financial impact of Covid-19 pandemic. *International Review of Financial Analysis*, *81*, Article 102101.

Althelaya, K. A., El-Alfy, E.-S. M., & Mohammed, S. (2018). Evaluation of bidirectional LSTM for short-and long-term stock market prediction. In *2018 9th international conference on information and communication systems* (pp. 151–156). IEEE.

Bai, S., Kolter, J. Z., & Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, *61*(4), 1645–1680.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, *21*(2), 129–152.

Chen, Y., & Hao, Y. (2017). A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. *Expert Systems with Applications*, *80*, 340–355.

Chen, Y., & Hao, Y. (2018). Integrating principle component analysis and weighted support vector machine for stock trading signals prediction. *Neurocomputing*, *321*, 381–402.

Chen, X., Ma, X., Wang, H., Li, X., & Zhang, C. (2022). A hierarchical attention network for stock prediction based on attentive multi-view news learning. *Neurocomputing*, http://dx.doi.org/10.1016/j.neucom.2022.06.106.

Cheng, W., Wang, Y., Peng, Z., Ren, X., Shuai, Y., Zang, S., et al. (2021). High-efficiency chaotic time series prediction based on time convolution neural network. *Chaos, Solitons & Fractals*, *152*, Article 111304.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., et al. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.

Collobert, R., & Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on machine learning* (pp. 160–167).

Cui, Y., Xie, J., & Zheng, K. (2021). Historical inertia: A neglected but powerful baseline for long sequence time-series forecasting. In *Proceedings of the 30th ACM international conference on information & knowledge management* (pp. 2965–2969).

Dai, W., An, Y., & Long, W. (2022). Price change prediction of ultra high frequency financial data based on temporal convolutional network. *Procedia Computer Science*, *199*, 1177–1183.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417.

Feng, F., Chen, H., He, X., Ding, J., Sun, M., & Chua, T.-S. (2018). Enhancing stock movement prediction with adversarial training. arXiv preprint arXiv:1810.09936.

Fu, R., Zhang, Z., & Li, L. (2016). Using LSTM and GRU neural network methods for traffic flow prediction. In *2016 31st youth academic annual conference of Chinese Association of Automation* (pp. 324–328). IEEE.

Gao, T., & Chai, Y. (2018). Improving stock closing price prediction using recurrent neural network and technical indicators. *Neural Computation*, *30*(10), 2833–2854.

Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440–1448).

Giudici, P., Polinesi, G., & Spelta, A. (2022). Network models to improve robot advisory portfolios. *Annals of Operations Research*, *313*(2), 965–989.

Gong, X., Zhang, W., Wang, J., & Wang, C. (2022). Investor sentiment and stock volatility: New evidence. *International Review of Financial Analysis*, *80*, Article 102028.

Gunay, S., Can, G., & Ocak, M. (2020). Forecast of China's economic growth during the COVID-19 pandemic: a MIDAS regression analysis. *Journal of Chinese Economic and Foreign Trade Studies*.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Hu, Z., Liu, W., Bian, J., Liu, X., & Liu, T.-Y. (2018). Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction. In *Proceedings of the eleventh ACM international conference on web search and data mining* (pp. 261–269).

Jing, N., Wu, Z., & Wang, H. (2021). A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. *Expert Systems with Applications*, *178*, Article 115019.

Ko, J. U., Jung, J. H., Kim, M., Kong, H. B., Lee, J., & Youn, B. D. (2021). Multi-task learning of classification and denoising (MLCD) for noise-robust rotor system diagnosis. *Computers in Industry*, *125*, Article 103385.

Lahmiri, S. (2014). Comparative study of ECG signal denoising by wavelet thresholding in empirical and variational mode decomposition domains. *Healthcare Technology Letters*, *1*(3), 104–109.

Lahmiri, S. (2016a). Intraday stock price forecasting based on variational mode decomposition. *Journal of Computer Science*, *12*, 23–27.

Lahmiri, S. (2016b). A variational mode decompoisition approach for analysis and forecasting of economic and financial time series. *Expert Systems with Applications*, *55*, 268–273.

Li, H., Liu, T., Wu, X., & Chen, Q. (2020). An optimized VMD method and its applications in bearing fault diagnosis. *Measurement*, *166*, Article 108185.

Li, C., Song, D., & Tao, D. (2019). Multi-task recurrent neural networks and higher-order Markov random fields for stock price movement prediction: Multi-task RNN and higer-order MRFs for stock price classification. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 1141–1151).

Lin, G., Lin, A., & Cao, J. (2021). Multidimensional KNN algorithm based on EEMD and complexity measures in financial time series forecasting. *Expert Systems with Applications*, *168*, Article 114443.

Liu, T., Ma, X., Li, S., Li, X., & Zhang, C. (2022). A stock price prediction method based on meta-learning and variational mode decomposition. *Knowledge-Based Systems*, Article 109324. http://dx.doi.org/10.1016/j.knosys.2022.109324.

Liu, M., Zeng, A., Xu, Z., Lai, Q., & Xu, Q. (2021). Time series is a special sequence: Forecasting with sample convolution and interaction. arXiv preprint arXiv:2106.09305.

Los, C. A., & Yu, B. (2008). Persistence characteristics of the Chinese stock markets. *International Review of Financial Analysis*, *17*(1), 64–82.

Ma, X., Li, X., Zhou, Y., & Zhang, C. (2021). Image smoothing based on global sparsity decomposition and a variable parameter. *Computational Visual Media*, *7*(4), 483–497.

Ma, T., & Tan, Y. (2020). Multiple stock time series jointly forecasting with multi-task learning. In *2020 international joint conference on neural networks* (pp. 1–8). IEEE.

Ma, T., & Tan, Y. (2022). Stock ranking with multi-task learning. *Expert Systems with Applications*, *199*, Article 116886.

Ma, X., Zhao, T., Guo, Q., Li, X., & Zhang, C. (2022). Fuzzy hypergraph network for recommending top-K profitable stocks. *Information Sciences*, *613*, 239–255.

Mootha, S., Sridhar, S., Seetharaman, R., & Chitrakala, S. (2020). Stock price prediction using bi-directional LSTM based sequence to sequence modeling and multitask learning. In *2020 11th IEEE annual ubiquitous computing, electronics & mobile communication conference* (pp. 0078–0086). IEEE.

Nayak, R. K., Mishra, D., & Rath, A. K. (2015). A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices. *Applied Soft Computing*, *35*, 670–680.

Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G., & Cottrell, G. (2017). A dual-stage attention-based recurrent neural network for time series prediction. arXiv preprint arXiv:1704.02971.

Rezaei, H., Faaljou, H., & Mansourfar, G. (2021). Stock price prediction using deep learning and frequency decomposition. *Expert Systems with Applications*, *169*, Article 114332.

Ronaghi, F., Salimibeni, M., Naderkhani, F., & Mohammadi, A. (2022). COVID19-HPSMP: COVID-19 adopted hybrid and parallel deep information fusion framework for stock price movement prediction. *Expert Systems with Applications*, *187*, Article 115879.

Shah, H., Bhatt, V., & Shah, J. (2022). A neoteric technique using ARIMA-LSTM for time series analysis on stock market forecasting. In *Mathematical modeling, computational intelligence techniques and renewable energy* (pp. 381–392). Springer.

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019). The performance of LSTM and BiLSTM in forecasting time series. In *2019 IEEE international conference on big data* (pp. 3285–3292). IEEE.

Štifanić, D., Musulin, J., Miočević, A., Baressi Šegota, S., Šubić, R., & Car, Z. (2020). Impact of COVID-19 on forecasting stock prices: an integration of stationary wavelet transform and bidirectional long short-term memory. *Complexity*, *2020*.

Wang, L. (2010). *The effect of government policy on China's stock market* (Ph.D. thesis), Citeseer.

Wu, Q., & Lin, H. (2019). Short-term wind speed forecasting based on hybrid variational mode decomposition and least squares support vector machine optimized by bat algorithm model. *Sustainability*, *11*(3), 652.

Yan, B., Aasma, M., et al. (2020). A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM. *Expert Systems with Applications*, *159*, Article 113609.

Yang, M., & Wang, J. (2022). Adaptability of financial time series prediction based on BiLSTM. *Procedia Computer Science*, *199*, 18–25.

Yi, Z., & Mao, N. (2009). Research on investor sentiment measurement in Chinese stock market: construction of CICSI. *Financial Research*, *11*, 174–184.

Yue, X., Zhou, Y., & Yuan, C. (2021). Stock closing price prediction based on combined model of PCA-IMKNN. *International Journal of Modelling, Identification and Control*, *39*(3), 221–228.

Zhang, C., Wang, Y., Chen, C., Du, C., Yin, H., & Wang, H. (2018). Stockassistant: a stock ai assistant for reliability modeling of stock comments. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 2710–2719).

Zhang, H.-C., Wu, Q., & Li, F.-Y. (2022). Application of online multitask learning based on least squares support vector regression in the financial market. *Applied Soft Computing*, *121*, Article 108754.

Zheng, L., & He, H. (2021). Share price prediction of aerospace relevant companies with recurrent neural networks based on pca. *Expert Systems with Applications*, *183*, Article 115384.