Bimodal Stock Price Prediction Based on Holt-Winters Exponential Smoothing and PCA Whitening Transformation

Changsheng Zhu, Pengyun Lu, Wenfang Feng, Yanbo Wang

Abstract—Aiming at the nonlinear and high frequency characteristics of stock data, a hybrid stock price prediction model is proposed, which combines the Holt-Winters triple exponential smoothing method, PCA whitening transformation, PatchTST and BERT, called HW-PCAW-PB. The PCA whitening transformation is performed on the feature vector of investor stock reviews constructed by the BERT model, and the stock price features are processed using the Holt-Winters triple exponential smoothing algorithm and the PatchTST model. The cross-attention mechanism fuses the processed stock review feature information, stock price feature information, and multi-source features such as stock review reading volume and number of comments to predict stock prices. The prediction performance of the HW-PCAW-PB model is compared with benchmark models such as SVR, BiLSTM, TimesNet, KNN, ANN, LSTM, Autoformer and Informer using RMSE, MAE, MAPE and R^2 evaluation metrics. The experimental results show that the HW-PCAW-PB model achieves better scores than the benchmark models in all evaluation metrics.

Index Terms—Hybrid stock price prediction model, Holt-Winters triple exponential smoothing, PCA whitening transformation, PatchTST

I. INTORDUCTION

WITH the rapid development of the market economy, the scale of the stock market continues to expand, and more and more people are participating in stock investment. Predicting stock prices can help investors reduce investment risks and improve investment returns. However, as a kind of financial time series data, stock prices are affected by many factors and have the characteristics of high dimensionality, non-stationary and random fluctuations. All these factors work together, making stock prices unpredictable and challenging to forecast accurately [1].

In recent years, domestic and foreign scholars have achieved a lot of research results on stock price prediction methods. Conventional statistical methods, including the autoregressive integrated moving average (ARIMA) model [2] and the generalized autoregressive conditional heteroskedasticity (GARCH) model [3], are commonly applied in stock market analysis and forecasting. Ganesan and Kannan [4] used the ARIMA model to predict stock prices in the short term and achieved good results. Although traditional statistical methods have shown certain capabilities in stock price prediction, these models can only capture linear relationships and cannot capture nonlinear relationships. This means that statistical methods cannot effectively capture data on nonlinear factors such as investor comments and news, which may lead to a decrease in the precision of stock price forecasting. As these limitations of statistical learning methods in stock prediction are discovered, researchers have gradually tended to adopt machine learning algorithms to predict stock prices [5], [6]. Jayaswara et al. [7] applied support vector regression (SVR) for forecasting stock prices. Gao [8] applied the K-Nearest Neighbors (KNN) algorithm to estimate stock prices. Although traditional machine learning algorithms have advantages in processing nonlinear data, due to the time series correlation of stock data, these algorithms have limitations in feature extraction, which can easily lead to overfitting, weak generalization ability and falling into local optimal problems.

With the swift advancement of big data, deep learning has achieved remarkable advancements in feature extraction within the domain of time series [9]. For stock price prediction, deep learning methods, due to their complex network structure, can effectively uncover hidden patterns in stock price data and combine the relationship between time series information and features to provide more accurate stock price predictions. Recurrent Neural Networks (RNN) [10] have the ability to capture and memorize previous information of a sequence and are suitable for processing time series data. However, when processing long sequences, RNN is susceptible to problems such as gradient vanishing and gradient exploding. In order to solve these problems, Long Short-Term Memory (LSTM) [11] was introduced. LSTM introduces a gating mechanism to effectively regulate and update information flow, making it more suitable for processing long-term dependencies. The Transformer [12] model, initially designed for natural language processing, adopts an encoding-decoding framework. Its unique self-attention mechanism enables it to effectively extract time series features. The Autoformer model proposed by Wu et al.

Manuscript received June 18, 2024; revised November 28, 2024.

This work was financially supported by the National Natural Science Foundation of China (Grant Nos. 72161026).

Changsheng Zhu is a professor at the School of Computer and Communication, Lanzhou University of Technology, Lanzhou, 730050, China (Corresponding author;e-mail: 867320505@qq.com).

Pengyun Lu is a postgraduate student at the School of Computer and Communication, Lanzhou University of Technology, Lanzhou, 730050, China (e-mail: 1245400519@qq.com).

Wenfang Feng is a professor at the School of Economics and Management, Lanzhou University of Technology, Lanzhou, 730050, China (e-mail: 1036784024@qq.com).

Yanbo Wang is a postgraduate student at the School of Computer and Communication, Lanzhou University of Technology, Lanzhou, 730050, China (e-mail: 1663806213@qq.com).

[13] alternately optimizes prediction results and decomposes sequences. When predicting long sequences, its prediction accuracy exceeds that of the LSTM model. The PatchTST [14] model divides the time series into several time periods, which effectively prevents the overfitting problem. However, it is sensitive to outliers and noise, which may cause the performance of the model to deteriorate in some cases. This paper uses the quartile method and the Holt-Winters triple exponential smoothing algorithm [15] to reduce noise and outliers in the data, making the data more smooth and stable, and improving the predictive ability of stock prices.

Investor sentiment is considered a key factor in predicting stock trends [16]. However, earlier studies typically rely solely on historical stock price data for forecasting stock prices [17], which may lead to inaccurate predictions. To address this problem, researchers use bimodal data that combines text and numerical information to improve the accuracy of stock price prediction. Wu et al. [18] introduced the S_I_LSTM stock price prediction method, which incorporates various data sources and market sentiment. They discovered that its predictions align more closely with actual closing prices compared to using a individual data source. Zhuge et al. [19] combined sentiment data with stock price data as training data, and the results significantly improved the prediction accuracy.

Financial markets are complex, and stock prices are affected by a variety of inherent and human factors, such as public sentiment, political climate, and media events, all of which can introduce noise into stock data. Since an individual model is difficult to fully consider these factors and has inherent limitations, a hybrid approach of multiple models is used to solve the problem of stock price prediction [20]. Ma et al. [21] greatly enhanced the accuracy and stability of stock price prediction by combining the LSTM model with the ARIMA model. Zhao et al. [22] studied the multi-layer feature ablation method for stock review topic identification using the BERT model. Compared with the single BERT model, it further improved the performance of stock review topic identification. Daiya [23] used a fused convolution and Transformer model to analyze financial indicators and news data, boosting the precision of stock price forecasting.

Therefore, this paper establishes a hybrid prediction model that combines the Holt-Winters seasonal exponential smoothing algorithm, PCA whitening transformation, PatchTST model, and BERT model for stock price prediction. The key contributions of this paper are: (a) A new combined model is established to predict nonlinear and non-stationary stock price series with good accuracy and robustness. (b) The quartile method, Holt-Winters triple exponential smoothing algorithm, and PatchTST model are used to process time series data. (c) Natural language processing technology is introduced to extract stock review embedding vectors using the pre-trained BERT model, and then these stock review embedding vectors are processed using global average pooling and PCA whitening transformation. (d) Through the cross-attention mechanism, the processed stock review features, stock price features, and stock review reading volume and number of comments features are multi-feature fused to achieve multi-dimensional information integration, thus enhancing the accuracy of stock price forecasting.

The rest of this study is organized as follows: Section 2 outlines the developed model, Section 3 presents the experimental analysis and results, and Section 4 provides the conclusion.

II. Proposed Methodology

A. Framework of the Combined Model

The framework of the HW-PCAW-PB model is shown in Fig. 1. The HW-PCAW-PB prediction model mainly consists of the Holt-Winters triple exponential smoothing algorithm module, PatchTST module, Bert module, social engagement index normalization module, PCA whitening transformation module, and cross-attention mechanism module.

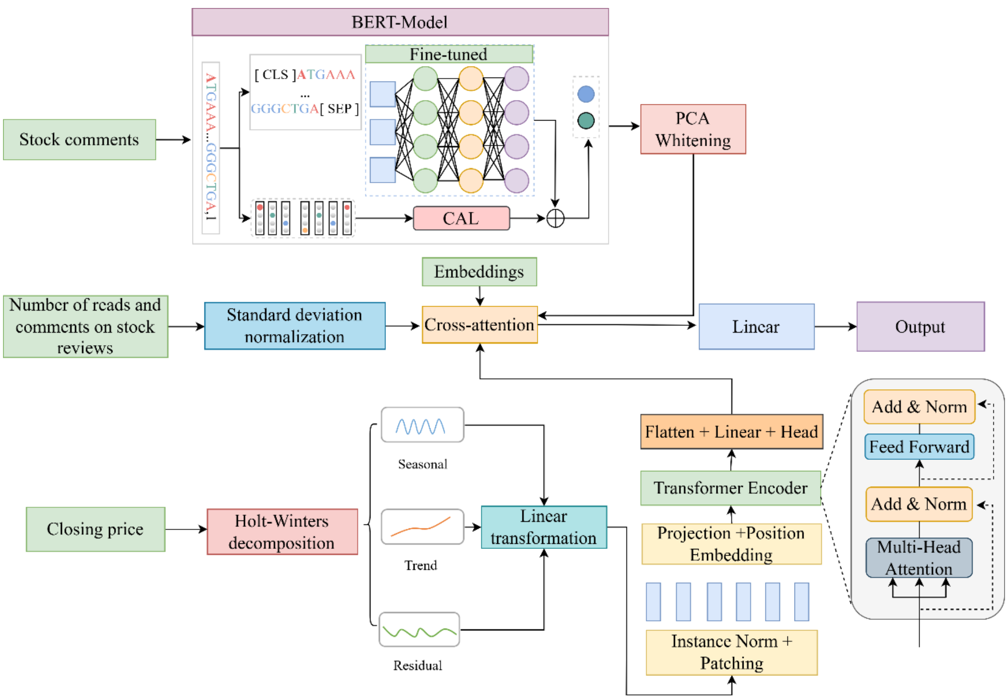


Fig. 1. Framework of the proposed model.

This paper introduces a hybrid model, HW-PCAW-PB, for stock price prediction. In order to improve the analysis and prediction capabilities of stock price data, its processing process includes multiple steps. First, the stock closing price data is downloaded using the Tushare financial interface package, and the quartile method and Holt-Winters triple exponential smoothing algorithm are used to clean and fit the data for outliers to obtain trend components, seasonal components, and residuals. These components are then normalized by standard deviation to ensure that their changes under different data ranges are comparable. Next, the normalized components are input into the PatchTST model for processing and extracting important features. Immediately after that, by crawling and preprocessing the stock reviews of Oriental Fortune Network, the pre-trained BERT model is used to extract the embedding vector of the stock price reviews, and the anisotropy in the stock review embedding is effectively eliminated through PCA whitening, and the semantic space formed by BERT is reasonably used to solve the representation degradation problem found in the traditional pre-trained language model, and enhance the precision of stock price forecasting. At the same time, the relevant information of each stock price review is normalized by standard deviation to provide more comprehensive information for the model. Embedding technology is used to map the features of each stock so that the model can better learn the characteristics of the stock. Finally, the cross-attention mechanism is used to fuse the stock price feature information processed by PatchTST with other relevant features and input into the model for stock price prediction.

B. Holt-Winters Triple Exponential Smoothing Algorithm

The stock price data has trend and rationality, so the Holt-Winters triple exponential smoothing method using the weighted average moving average method is selected to predict the time series, and the historical data, data trends and seasonal characteristics are weighted and summed. A weight value is assigned according to the distance between the data and the prediction time. Data at closer times are given a larger weight value, and data at farther times are given a smaller weight value, so as to achieve the time series prediction task.

The Holt-Winters triple exponential smoothing method consists of a prediction formula and three smoothing formulas:

$$L_i = \alpha(X_i - S_{i-s}) + (1-\alpha)(L_{i-1} + T_{i-1}) \tag{1}$$

$$T_i = \beta(L_i - L_{i-1}) + (1-\beta)T_{i-1} \tag{2}$$

$$S_i = \gamma(X_i - L_i) + (1-\gamma)S_{i-s} \tag{3}$$

$$F_{i+k} = L_i + kT_i + S_{i+k-s} \tag{4}$$

Where L_i、T_i、S_i are the level, trend and seasonal component of the original data at the ith moment respectively. α、β、γ are the smoothing coefficients of the level, trend and seasonality respectively, and their values vary between 0 and 1. X_i is the actual value of the data at the ith time point, F_{i+k} is the forecasted value of the next k cycles, and s is the length of the seasonal cycle.

Stock price data usually has random volatility, which is attributed to various random factors and noise in the stock market. Therefore, residuals are used to reflect the random fluctuations, noise, and other influencing factors in stock price data that are not explained by trend and seasonality. In the Holt-Winters triple exponential smoothing method, the residual is obtained by subtracting the trend and seasonal components from the original stock price data:

$$R_i = X_i - T_i - S_i \tag{5}$$

Where R_i is the residual of the ith observation.

C. PCA Whitening

Principal Component Analysis (PCA) is a popular method for analyzing data that is widely used to reduce high-dimensional data to a lower-dimensional representation to more effectively understand the structure and attributes of the data. The whitening operation converts vectors with a known covariance matrix into a series of new vectors, where the covariance matrix of each vector is a unit matrix, which means that the dimensions of the vector are uncorrelated, that is, the converted vector becomes a white noise vector. In this conversion, the mean of the vector needs to be adjusted to zero and the covariance matrix needs to be converted to a unit matrix. The transformation process is as follows:

$$\tilde{s_i} = (s_i - \mu)W \tag{6}$$

Among them, s_i represents the sentence vector, and the vector set is $\{s_i\}_{i=1}^{N}$. μ is the weighted mean value about s_i, and W is the transformation matrix to be solved. The method to calculate the covariance matrix of the original vector is as follows:

$$\sum = \frac{1}{N}\sum_{i=1}^{N}(s_i - \mu)^T (s_i - \mu) \tag{7}$$

After whitening, we can get the new covariance matrix of the vector $\tilde{\sum}$:

$$\tilde{\sum} = W \sum W^T \tag{8}$$

To achieve PCA whitening, i. e. , the new covariance matrix is a unit matrix:

$$W \sum W^T = I \tag{9}$$

According to (9), we get $\sum = W^{-1}(W^{-1})^T$, as the covariance matrix $\sum$ is a positive definite symmetric matrix, thus SVD (singular value decomposition) can be conducted:

$$\sum = U \wedge U^T \tag{10}$$

To obtain the PCA whitening transformation matrix by setting $W^{-1} = \sqrt{\wedge}\, U^T$, Where U stands for orthogonal matrix, $\wedge$ is a diagonal matrix with all positive diagonal elements:

$$W = U\sqrt{\wedge^{-1}} \tag{11}$$

By applying the PCA whitening transformation to the stock embedding vectors, more accurate and uniform semantic representations can be effectively extracted from the potentially biased distribution of the original data.

D. PatchTST Model

The PatchTST model is an independent prediction method specifically designed to process a single sequence of a time series dataset. Its main components include data processing layer, Transformer layer, and prediction head. Compared with traditional Transformer-based models, PatchTST treats each time point as a unit. PatchTST first normalizes the input sequence and then divides it into multiple time periods, which can have overlapping parts, and each time period is regarded as a unit. Next, each time segment is positionally encoded and directly input into the Transformer layer. Finally, the output vector of the Transformer is flattened and input into a prediction head to obtain a univariate output sequence. Given a multivariate time series dataset, a sequence of length L $(X_1, \cdots X_L)$ is input into the model, and the data processing layer divides the sequence into N patches, where N is calculated as:

$$N = \frac{(L-P)}{S} + 2 \tag{12}$$

Where L is the input sequence length, P is the patch length, S is the step size.

The Transformer layer maps a patch to a space of dimension D by multiplying it with $W_P \in R^{D\times P}$, obtaining the feature projection and summing it with the position encoding $W_{Pos} \in R^{D\times N}$ to obtain the time-ordered sequence $X_d^{(i)}$ containing the patch:

$$X_d^{(i)} = W_P X_P^{(i)} + W_{Pos} \tag{13}$$

Where W_{Pos} is the position encoding matrix, W_P is the mapping matrix, $X_P^{(i)}$ is the ith patch, and $X_d^{(i)}$ is the feature projection result of the ith patch.

Then using each head in the multi-head attention $h = 1,2,\ldots,H$ to convert it into Query vector (Q), Key vector (K), and Value vector (V):

$$Q_h^{(i)} = \left(X_d^{(i)}\right)^T W_h^Q \tag{14}$$

$$K_h^{(i)} = \left(X_d^{(i)}\right)^T W_h^K \tag{15}$$

$$V_h^{(i)} = \left(X_d^{(i)}\right)^T W_h^V \tag{16}$$

Where W_h^Q is the hth Q-transform matrix; W_h^K is the hth K-transform matrix; and W_h^V is the hth V-transform matrix.

Perform Softmax normalisation to obtain a weight matrix where each value of the matrix is greater than 0 and less than 1 and sums to 1. Multiply the weight matrix with V and calculate the weighted sum to obtain the output $O_h^{(i)}$:

$$\left(O_h^{(i)}\right)^T = Softmax\left(\frac{Q_h^{(i)} K_h^{(i)T}}{\sqrt{d_k}}\right) V_h^{(i)} \tag{17}$$

Where d_k is the scaling factor.

Finally, $O_h^{(i)}$ is flattened and input into the prediction head consisting of a dense layer to generate the final prediction.

E. Bert Model

Stock market-related commentary texts have an increasingly significant impact on the stock market. Investors tend to analyze information from them to adjust their investment strategies, thereby affecting stock prices. Although the traditional Word2Vec model can capture the contextual relationship between words, it cannot handle the polysemy of the same word in different contexts, resulting in the same vector for the same word in different contexts. In contrast, the BERT model can dynamically adjust the text representation according to the text context, so it is more suitable for handling polysemy, can better learn grammatical and semantic information, and conduct more accurate evaluation of the stock market's response.

This paper extracts text features from stock market related comments. The BERT model converts the input corpus into a feature vector representation through a multi-head self-attention mechanism and a feed-forward fully connected layer. First, we vectorize the stock market related comments and use the vocabulary of the BERT model to vectorize the words or phrases in each comment. In order to ensure the practicality of the extracted features, global average pooling is performed on each dimension of each vector. The vector after global average pooling can be used as the feature vector of the comment for subsequent data analysis and modeling.

III. Experimental Validation and Result Analysis

A. Experimental Data Sources and Processing

This paper selects SSE177 stock index as the research sample, covering more than 2 million daily transaction data from January 1, 2017 to December 31, 2017. The financial database Tushare is used to read public stock data, and the stock bar social media text data is obtained from Dongfang Fortune Network to add additional data sources.

The Python Spyder web crawler technology is used to collect unstructured text data, including post titles, number of reads, number of comments, author, latest update date, etc. In order to observe the prediction effect of different prediction methods on stock indexes, the last month's data of the overall closing day data set of each stock index, a total of 21 sets of data, are taken as the test set, and the first eleven months' data, a total of 223 sets of data, are taken as the training set.

This paper uses the quartile method to remove abnormal data. The principle of the quartile algorithm is to divide an ascending sequence $X = \{x_1, x_2, \cdots x_3\}$ into four parts according to three data points. Each part accounts for 25% of the total data sequence. The three data points are called the lower quantile Q_1, median Q_2, and upper quantile Q_3. These three data points are calculated separately and then Equation (18) is used to find the interquartile range I_{QR}.

$$I_{QR} = Q_3 - Q_1 \tag{18}$$

The inner range of outliers in sequence X is defined by the Whisker upper limit W_1 and lower limit W_2:

$$[W_1, W_2] = \left[Q_1 - 1.5I_{QR}, Q_3 + 1.5I_{QR}\right] \tag{19}$$

The outliers are the data in the sequence X that are out of the inner limit range. Since the value of I_{QR} is not significantly affected by individual outliers, the identification of outliers by I_{QR} is stable and reliable when the percentage of outliers is relatively small.

Considering the differences in the price fluctuation ranges of different stocks, it is necessary to eliminate the impact of different dimensions before making stock price predictions, that is, to standardize the data before making stock price predictionst:

$$N^* = \frac{N-\mu}{\alpha} \tag{20}$$

Where N is the original data, μ and α are the mean value and the standard deviation of the original data, and N^* is the normalized data.

B. Experimental Settings

Stock data is a typical time series data that changes over time. In order to capture the deep features in stock data, this paper uses the GridSearchCV method to optimize the hyperparameters of the HW-PCAW-PB model. The GridSearchCV method exhausts all possible parameter combinations within the specified parameter range to find the optimal parameters, thereby achieving the purpose of optimizing the model. In order to prevent the HW-PCAW-PB model from overfitting the training data during the training process, this paper uses the Early Stopping method to terminate the training before the model overfits, to improve the model's capacity for generalization. For the early convergence problem, the HW-PCAW-PB model selects a dynamic learning rate, whose main purpose is to effectively update the weights of each gradient descent when training the neural network to minimize the loss function and enhance the effectiveness of the HW-PCAW-PB model. The He Initialization method is used to reduce the deviation of the initial values of the model coefficients, thereby reducing the risk of falling into the local optimum. These optimized hyperparameters are set as the default parameters of the model. This method can ensure the model effect while reducing the complexity and uncertainty in the parameter selection process. Table I demonstrates some of the default hyperparameter settings for the experiment.

TABLE I
DEFAULT HYPERPARAMETER SETTINGS IN THE EXPERIMENT

Hyperparameter name	Specific settings
Optimization algorithm	Adam
Activation function	ReLU
Loss function	Mean Absolute Error
Learning rate	0.0001
Batch_size	256
Epochs	100

C. Evaluation Metrics

To evaluate the model's effectiveness in forecasting stock prices, metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and coefficient of determination (R^2) were used as evaluation indexes:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(x_i - \hat{x_i})^2} \tag{21}$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|x_i - \hat{x_i}| \tag{22}$$

$$MAPE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{(x_i-\hat{x_i})}{x_i}\right| \tag{23}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(x_i-\hat{x_i})^2}{\sum_{i=1}^{m}(x_i-\bar{x})^2} \tag{24}$$

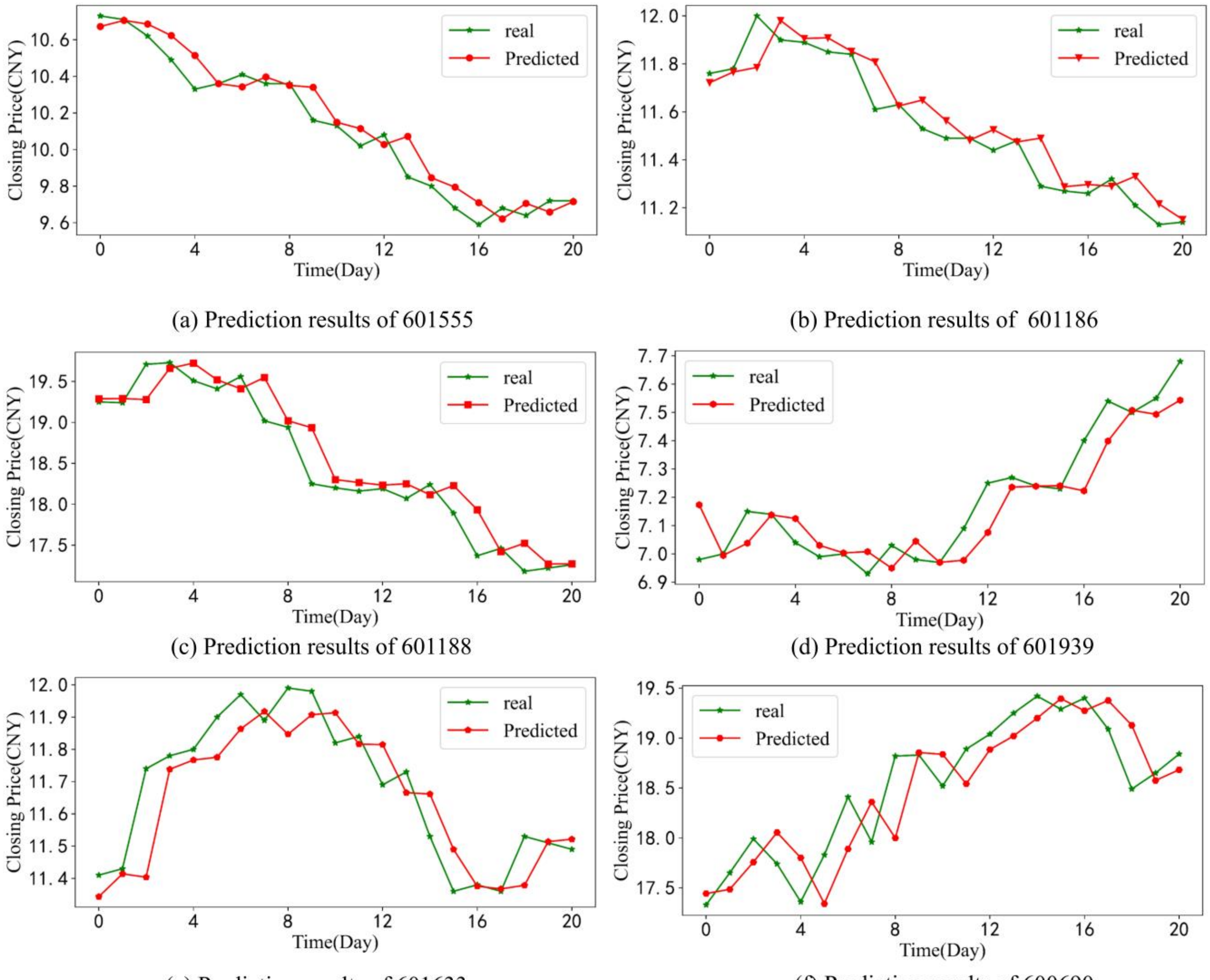
Where m refers to the total sample size, x_i denotes the actual closing price on day i, $\hat{x_i}$ stands for the predicted closing price on day i, and $\bar{x}$ is the average of the closing price of the stock. RMSE, MAE, and MAPE are employed to assess the discrepancy between the observed stock price and the forecasted price. The coefficient of determination, R^2, is often used to indicate the goodness of fit, which reflects the overall alignment between the predicted and observed values. R^2 reaches its maximum value of 1, with a higher value signifying a better model fit. However, since the predictions of the regression model may deviate from the true results, it leads to a negative value of R^2.

D. Results and Analysis

In the actual environment, the stock price of individual stocks fluctuates randomly and is determined by a range of variables, making it challenging to forecast. To evaluate the performance of the HW-PCAW-PB model, six representative stocks, including Dongwu Securities (601555), China Railway Construction (601186), Huatai Securities (601688), Construction Bank (601939), Great Wall Motors (601633), and Qingdao Haier (600690), are selected from the SSE177 stock index as prediction objects, and their stock price prediction curve is shown in Fig. 2.

It can be seen that the share price of Dongwu Securities in Fig. 2(a) slowly goes up after experiencing a wide range of decline in the previous period, and generally shows a downward trend. Fig. 2(b) China Railway Construction and Fig. 2(c) Huatai Securities share prices fall sharply after experiencing a small increase in the previous period, with an overall downward trend. Fig. 2(d) Construction Bank slowly declines after experiencing a small upward movement in the early,and then oscillates upward, showing an overall upward trend. Fig. 2(e) Great Wall Motor's share price climbs rapidly from a low level and then oscillates downward, showing a general downward trend. Fig. 2(f) Qingdao Haier stock price fluctuates with great ups and downs, with several peaks, and also showing an overall upward trend. Overall, the pattern of the forecasted and observed curves of the HW-PCAW-PB model shows good consistency, especially the stock price prediction curves are closely concave and convex to the real value curves, which illustrates the good accuracy and adaptability of the HW-PCAW-PB model in predicting future stock prices. However, there is still a degree of lag in the prediction curve, that is, there is a gap in time between the forecasted results and the actual value.

To provide a more comprehensive evaluation of the HW-PCAW-PB model's performance, its prediction outcomes were compared against those generated by several widely recognized models, including SVR, BiLSTM, TimesNet, KNN, ANN, LSTM, Autoformer, and Informer. The comparative analysis of the predictive accuracy and effectiveness across these models is visually represented in Fig. 3, highlighting the relative strengths and weaknesses of each approach.



(a) Prediction results of 601555

(b) Prediction results of 601186

(c) Prediction results of 601188

(d) Prediction results of 601939

(e) Prediction results of 601633

(f) Prediction results of 600690

Fig. 2. Plot of stock price prediction for the proposed.

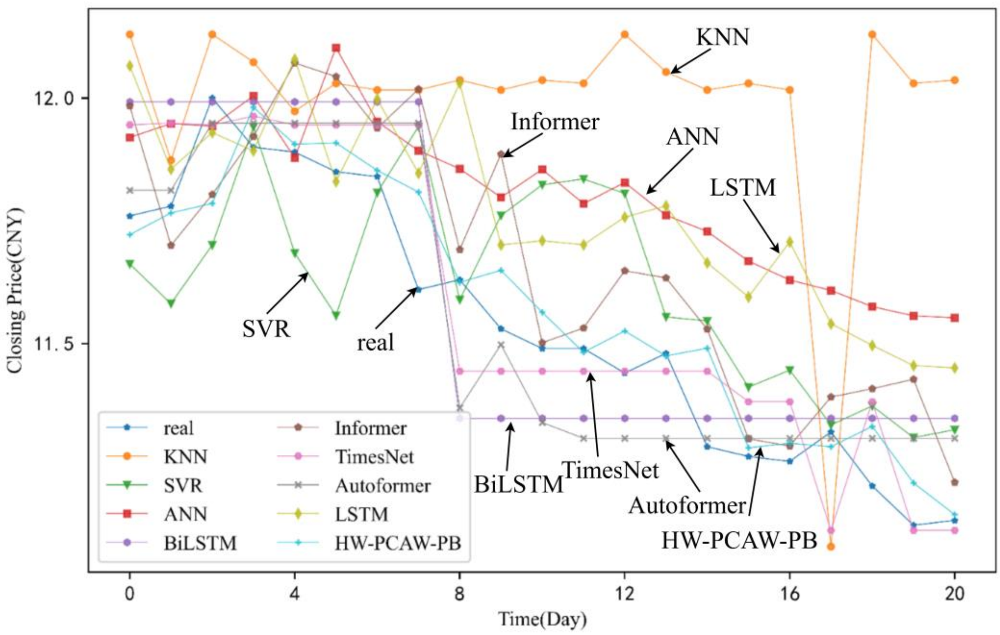


Fig. 3. Comparison of predictive model results.

The comparative analysis in Fig. 3 reveals that the HW-PCAW-PB model exhibits optimal prediction performance, with its prediction curve highly matching the actual stock price trend, showing excellent volatility tracking ability and stability. In contrast, the other models have different degrees of prediction bias: the ANN model generally overestimates the actual value and has systematic prediction bias; the KNN, SVR, Informer, and LSTM models show obvious prediction volatility, reflecting the lack of model stability; and the Autoformer, TimesNET, and

BiLSTM predictions are not as severe as KNN predictions that fluctuate up and down, but the predictions are also up and down and differ from the true values. In summary, among all models, the HW-PCAW-PB model shows the best prediction performance by virtue of its excellent fitting effect and outstanding fluctuation tracking ability.

To further verify the fitting effect of the HW-PCAW-PB model, we calculated the prediction error indicators of the model, including RMSE, MAE, MAPE, and R^2, and compared these indicators with the other Thirteen models. The results are detailed in Tables II-V.

As can be seen from Table II, for the six selected stocks, except for Huatai Securities, the HW-PCAW-PB model has the smallest RMSE and the best effect, and the prediction effect of the HW-PCAW-PB model is significantly better than the other Thirteen models. Compared with the machine learning models KNN, ANN, and SVR, the RMSE of the HW-PCAW-PB model is reduced by 87.65%, 66.24%, and 53.65%. Compared with the deep learning models BiLSTM and LSTM, the RMSE of the HW-PCAW-PB model is reduced by 37.60% and 53.28%. Compared to the time series models TimesNet, Autoformer, Informer, Transformer, TSMixer, Reformer, Dlinear, and ETSformer, the RMSE of the HW-PCAW-PB model is improved by 12.11%, 23.21%, 30.41%, 53.3%, 23.24%, 37.61%, 30.55%, and 12.84%. In general, from the perspective of RMSE evaluation indicator, the HW-PCAW-PB model performs well in stock price prediction.

From Table III, we can see that, except for Huatai Securities, the HW-PCAW-PB model has the best MAE and is significantly better than other models. Compared with the machine learning models KNN, ANN, and SVR, the MAE of the HW-PCAW-PB model is reduced by 88.83%, 70.64%, and 55.55%. Compared with the deep learning models BiLSTM and LSTM, the MAE of the HW-PCAW-PB model is reduced by 43.38% and 58.75%. Compared to the time series models TimesNet, Autoformer, Informer, Transformer, TSMixer, Reformer, Dlinear, and ETSformer, the MAE of the HW-PCAW-PB model is improved by 15.12%, 26.96%, 32.74%, 58.76%, 27.03%, 43.52%, 32.77%, and 15.22%. In general, from the perspective of MAE evaluation indicator, the HW-PCAW-PB model has the best stock price prediction performance compared with the other benchmark models.

As shown in Table IV, compared with the other Thirteen benchmark models, the HW-PCAW-PB model has the smallest MAPE for stock price prediction and has the best prediction effect. Compared with the machine learning models KNN, ANN, and SVR, the MAPE of the HW-PCAW-PB model is reduced by 89.97%, 74.94%, and 58.20% respectively. Compared with the deep learning models BiLSTM and LSTM, the MAPE of the HW-PCAW-PB model is reduced by 47.06% and 60.87% respectively. Compared to the time series models TimesNet, Autoformer, Informer, Transformer, TSMixer, Reformer, Dlinear, and ETSformer, the MAPE of the HW-PCAW-PB model is reduced by 20.22%, 30.35%, 35.86%, 60.87%, 32.55%, 47.35%, 35.86%, and 22.79% respectively. In general, from the perspective of MAPE evaluation indicator, the HW-PCAW-PB model has the best stock price prediction effect compared with the other Thirteen benchmark models.

As can be observed from Table V, the R^2 value of the HW-PCAW-PB model is closer to 1, and the prediction effect is significantly better than the other Thirteen models. Compared with the machine learning models KNN, ANN, and SVR, the R^2 of the HW-PCAW-PB model is increased by 20.29, 2.57, and 1.34 respectively. Compared with the deep learning models BiLSTM and LSTM, the R^2 of the HW-PCAW-PB model is increased by 0.42 and 1.07 respectively. Compared to the time series models TimesNet, Autoformer, Informer, Transformer, TSMixer, Reformer, Dlinear, and ETSformer, the R^2 of the HW-PCAW-PB model is improved by 0.13, 0.19, 0.45, 1.08, 0.19 , 0.42, 0.45, and 0.13. In general, compared with the other Thirteen benchmark models, the HW-PCAW-PB model has the best stock price prediction effect.

TABLE II
RMSE VALUES OF SIX STOCKS CORRESPONDING TO EACH MODEL

Model	601555	601186	601688	601939	601633	600690
SVR	0.1626	0.2201	0.4375	0.2687	0.4242	0.7193
BiLSTM	0.2417	0.1711	0.3894	0.1816	0.2014	0.4728
TimesNet	0.1054	0.1335	**0.2805**	0.1401	0.1621	0.3555
KNN	1.7845	0.5529	0.8895	1.1872	1.0453	2.9198
ANN	0.6802	0.2971	0.5201	0.4293	0.4159	0.7221
LSTM	0.3199	0.2649	0.4786	0.1644	0.3631	0.6234
Autoformer	0.1662	0.1371	0.3127	0.1412	0.1608	0.4293
Informer	0.1378	0.1877	0.2943	0.1382	0.2783	0.4504
Transformer	0.3206	0.2642	0.4792	0.1657	0.3641	0.6218
TSMixer	0.1663	0.1375	0.3127	0.1415	0.1606	0.4292
Reformer	0.2416	0.1713	0.3884	0.1817	0.2014	0.4738
DLinear	0.1378	0.1872	0.2952	0.1385	0.2785	0.4524
ETSformer	0.1063	0.1336	0.2875	0.1401	0.1632	0.3563
HW-PCAW-PB	**0.0976**	**0.0951**	0.2807	**0.096**	**0.1116**	**0.3536**

TABLE III
MAE VALUES OF SIX STOCKS CORRESPONDING TO EACH MODEL

Model	601555	601186	601688	601939	601633	600690
SVR	0.1362	0.1911	0.3682	0.22	0.3212	0.5546
BiLSTM	0.2088	0.1482	0.3247	0.1581	0.1519	0.4147
TimesNet	0.0871	0.1075	**0.2021**	0.1193	0.1178	0.3044
KNN	1.513	0.4786	0.7141	0.9533	0.831	2.6394
ANN	0.6628	0.27	0.461	0.3806	0.3244	0.6131
LSTM	0.3017	0.2368	0.421	0.1418	0.2801	0.5488
Autoformer	0.1309	0.1085	0.2364	0.1189	0.1232	0.3723
Informer	0.1069	0.1518	0.2303	0.1021	0.2229	0.3699
Transformer	0.3028	0.2352	0.4216	0.1437	0.2794	0.548
TSMixer	0.1308	0.1083	0.2366	0.1186	0.1236	0.3733
Reformer	0.2078	0.1488	0.3256	0.1589	0.1523	0.4165
DLinear	0.105	0.1522	0.2308	0.1022	0.2218	0.3724
ETSformer	0.0871	0.1076	0.2031	0.1193	0.1176	0.3046
HW-PCAW-PB	**0.076**	**0.0684**	0.2025	**0.0724**	**0.0828**	**0.2942**

TABLE IV
MAPE VALUES OF SIX STOCKS CORRESPONDING TO EACH MODEL

Model	601555	601186	601688	601939	601633	600690
SVR	0.0135	0.0166	0.02	0.0301	0.0275	0.0301
BiLSTM	0.0208	0.0128	0.0177	0.0221	0.013	0.0224
TimesNet	0.0086	0.0093	0.011	0.0167	0.0101	0.0165
KNN	0.1482	0.042	0.0383	0.1334	0.0714	0.1411
ANN	0.066	0.0236	0.0254	0.0537	0.0281	0.033
LSTM	0.0299	0.0207	0.0231	0.0197	0.0241	0.0297
Autoformer	0.0131	0.0094	0.0129	0.0165	0.0106	0.0202
Informer	0.0105	0.0132	0.0126	0.0142	0.0192	0.0201
Transformer	0.03	0.0206	0.0231	0.0199	0.024	0.0296
TSMixer	0.0130	0.0084	0.0126	0.0175	0.0116	0.0223
Reformer	0.0212	0.0132	0.0156	0.0232	0.0135	0.0227
DLinear	0.0103	0.0132	0.0126	0.0143	0.0191	0.0203
ETSformer	0.0088	0.0097	0.0123	0.0156	0.0113	0.0169
HW-PCAW-PB	**0.0076**	**0.0059**	**0.011**	**0.01**	**0.0071**	**0.016**

TABLE V
R^2 VALUES CORRESPONDING TO EACH MODEL FOR SIX STOCKS

Model	601555	601186	601688	601939	601633	600690
SVR	0.8028	0.3159	0.7451	-0.4793	-2.865	-0.2131
BiLSTM	0.564	0.5865	0.7981	0.3247	0.1287	0.476
TimesNet	0.9171	0.7481	0.8951	0.5976	0.4354	0.7037
KNN	-22.7578	-3.3117	-0.0536	-27.8743	-22.4689	-18.9855
ANN	-2.4513	-0.2465	0.6397	-2.775	-2.7154	-0.2224
LSTM	0.2367	0.0088	0.695	0.4466	-1.8315	0.0888
Autoformer	0.794	0.7346	0.8698	0.5918	0.4442	0.5679
Informer	0.8584	0.5022	0.8846	0.6086	-0.6636	0.5245
Transformer	0.2333	0.0141	0.6942	0.4375	-1.848	0.0935
TSMixer	0.798	0.7366	0.8689	0.5923	0.4456	0.5667
Reformer	0.5645	0.5859	0.7978	0.3239	0.1288	0.4769
DLinear	0.8583	0.5048	0.8839	0.6069	-0.6658	0.5203
ETSformer	0.9176	0.7488	0.8947	0.5969	0.4359	0.7040
HW-PCAW-PB	**0.929**	**0.8723**	**0.8951**	**0.8114**	**0.7323**	**0.7069**

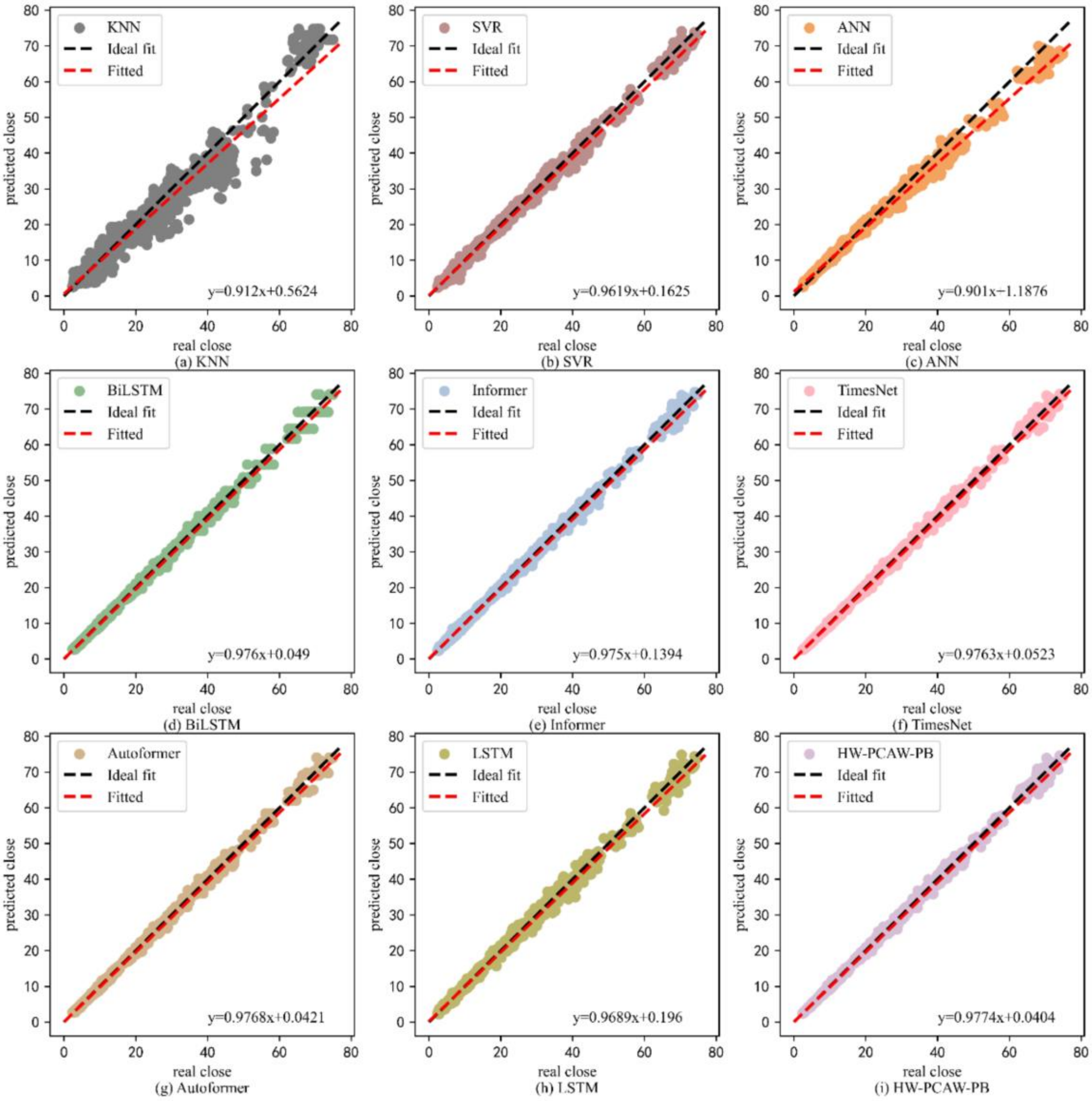


Fig. 4. Scatter plot of real and predicted values of stock prices on nine models.

Fig. 4 shows the scatter plots of the true value and forecasted value of the stock price on nine models: SVR, BiLSTM, TimesNet, KNN, ANN, LSTM, Autoformer, Informer, and HW-PCAW-PB. The scatter plots of the ANN model in Fig. 4(c) and the KNN model in Fig. 4(a) are relatively scattered, and the slopes of the fitted lines are 0.901 and 0.912, which shows that the forecasting performance of the ANN and KNN models are the worst. The scatter plots of the SVR model in Fig. 4(b) and the LSTM model in Fig. 4(h) are slightly clustered, and the slopes of the fitted lines are 0.9619 and 0.968, respectively, indicating that the prediction effects of the SVR and LSTM models are relatively good. In contrast, the scatter plots of the Fig. 4(d)BiLSTM, Fig. 4(e)Informer, Fig. 4(f)TimesNet, and Fig. 4(g)Autoformer models are relatively clustered, and the slopes of the fitted lines are 0.976, 0.975, 0.9763, and 0.9768, respectively, indicating that the BiLSTM, Informer, TimesNet, and Autoformer models have better stock price prediction effects. Among these nine models, the scatter plot of the HW-PCAW-PB model in Fig. 4(i) is the most clustered, and its slope of the fitted line is the largest, indicating that the HW-PCAW-PB model has the highest stock price prediction accuracy, the best fitting effect, and can more accurately capture the real trend and fluctuation of the data.

Fig. 5 shows the relative improvement percentage of the HW-PCAW-PB model compared with the SVR, BiLSTM, TimesNet, KNN, ANN, LSTM, Autoformer, and Informer models in different indicators. The RMSE, MAE, MAPE, and R^2 of the HW-PCAW-PB model are significantly better than the other eight benchmark models. Specifically, from Fig. 5(a) RMSE and Fig. 5(b) MAE, it can be seen that compared with the other eight prediction models, the RMSE and MAE of the HW-PCAW-PB model are improved by more than 10%, indicating that the HW-PCAW-PB model has significantly improved in stock price prediction accuracy. From Fig. 5(c) MAPE, it is noticeable that the MAPE of the HW-PCAW-PB model is improved by more than 20% compared with other benchmark models, further verifying the prediction accuracy of the HW-PCAW-PB model. From Fig. 5(d) R^2, it is clear that the R^2 of the HW-PCAW-PB model is markedly improved compared with other models. Overall, the HW-PCAW-PB model demonstrates superior performance in predicting stock indices, showing higher accuracy and better adaptability. It is more effective at capturing the intricate dynamics of the stock market.

Table VI compares the prediction performance of nine prediction models, including SVR, BiLSTM, TimesNet, KNN, ANN, LSTM, Autoformer, Informer, and HW-PCAW-PB, under three different decomposition methods (CEEMDAN, STL, and Holt-Winters). The RMSE, MAE, and MAPE of the HW-PCAW-PB model based on the Holt-Winters dataset are 0.4131, 0.2065, and 0.0130 respectively, which has the smallest prediction error compared with other prediction models. Compared with the single prediction models SVR, BiLSTM, TimesNet, KNN, ANN, LSTM, Autoformer, and Informer, the RMSE of the HW-PCAW-PB model based on the Holt-Winters dataset was improved by 39.97%, 17.93%, 5.21%, 83.82%, 66.23%, 39.67%, 6.64%, and 13.51% respectively, the MAE was improved by 50.75%, 29.79%, 10.57%, 86.83%, 73.41%, 49.07%, 17.26%, and 28.64% respectively, and the MAPE was improved by 62.28%, 46.67%, 20.56%, 88.93%, 82.22%, 56.61%, 34.41%, and 47.98% respectively. At the same time, the prediction results of each model based on the Holt-Winters dataset are significantly better than those based on the CEEMDAN dataset and the STL dataset. Compared with the CEEMDAN dataset and the STL dataset, the RMSE of the HW-PCAW-PB model based on the Holt-Winters dataset improved by 1.16% and 0.95%, the MAE improved by 0.39% and 4.41%, and the MAPE improved by 3.85% and 11.35%. Through the above analysis, the HW-PCAW-PB model based on the Holt-Winters dataset is very effective in stock price prediction.

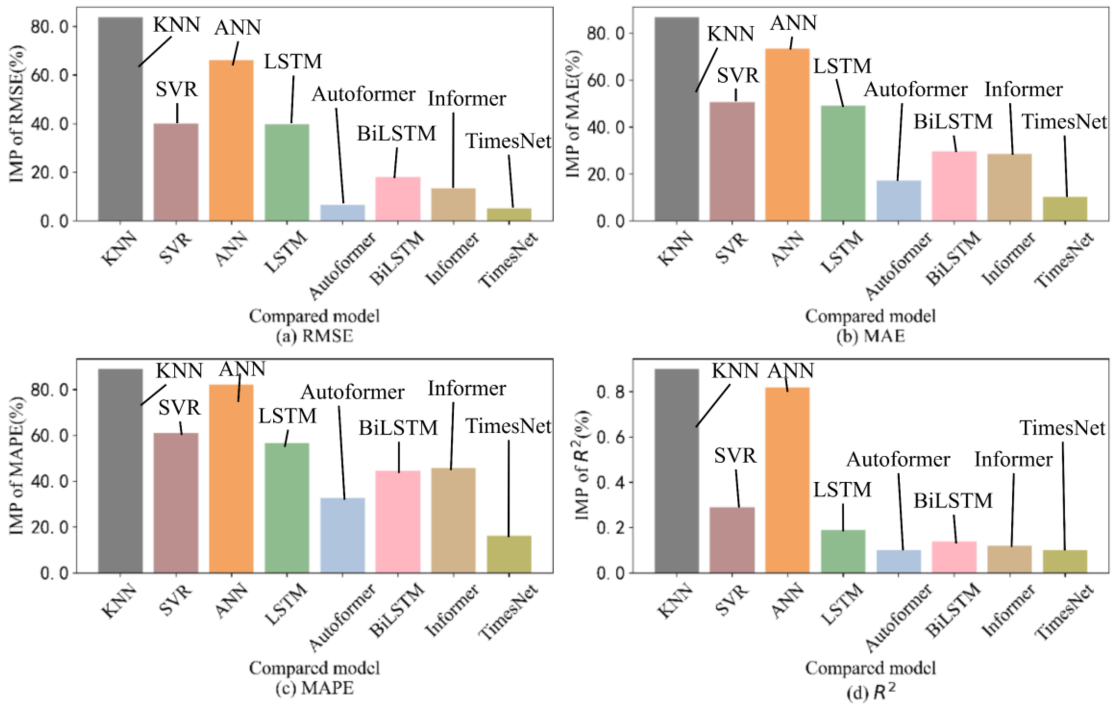


Fig. 5. Comparison of experimental predictors of HW-PCAW-PB model with other models.

TABLE VI
PERFORMANCE OF DIFFERENT MODELS UNDER THREE DECOMPOSITION METHODS

Model	CEEMDAN			STL			Holt-Winters		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
SVR	0.5556	0.3084	0.0245	0.7794	0.4748	0.0355	0.6800	0.4172	0.0321
BiLSTM	1.0882	0.3272	0.0232	0.5013	0.2957	0.0226	0.4979	0.2927	0.0225
TimesNet	0.4587	0.2346	0.0151	0.4309	0.2295	0.0151	0.4308	0.2294	0.0149
KNN	3.2871	1.6262	0.1094	2.5378	1.5683	0.1132	2.5241	1.5619	0.1130
ANN	1.1503	0.7092	0.0613	1.2452	0.7922	0.0714	1.2093	0.7735	0.0697
LSTM	0.9607	0.6902	0.6902	0.7841	0.496	0.0376	0.6769	0.4034	0.0289
Autoformer	0.4577	0.2529	0.0186	0.4384	0.2486	0.0184	0.4372	0.2483	0.0186
Informer	0.6811	0.449	0.0433	0.4765	0.2895	0.023	0.4721	0.2881	0.0231
HW-PCAW-PB	0.4131	0.2065	0.0130	0.4122	0.2152	0.0141	**0.4083**	**0.2057**	**0.0125**

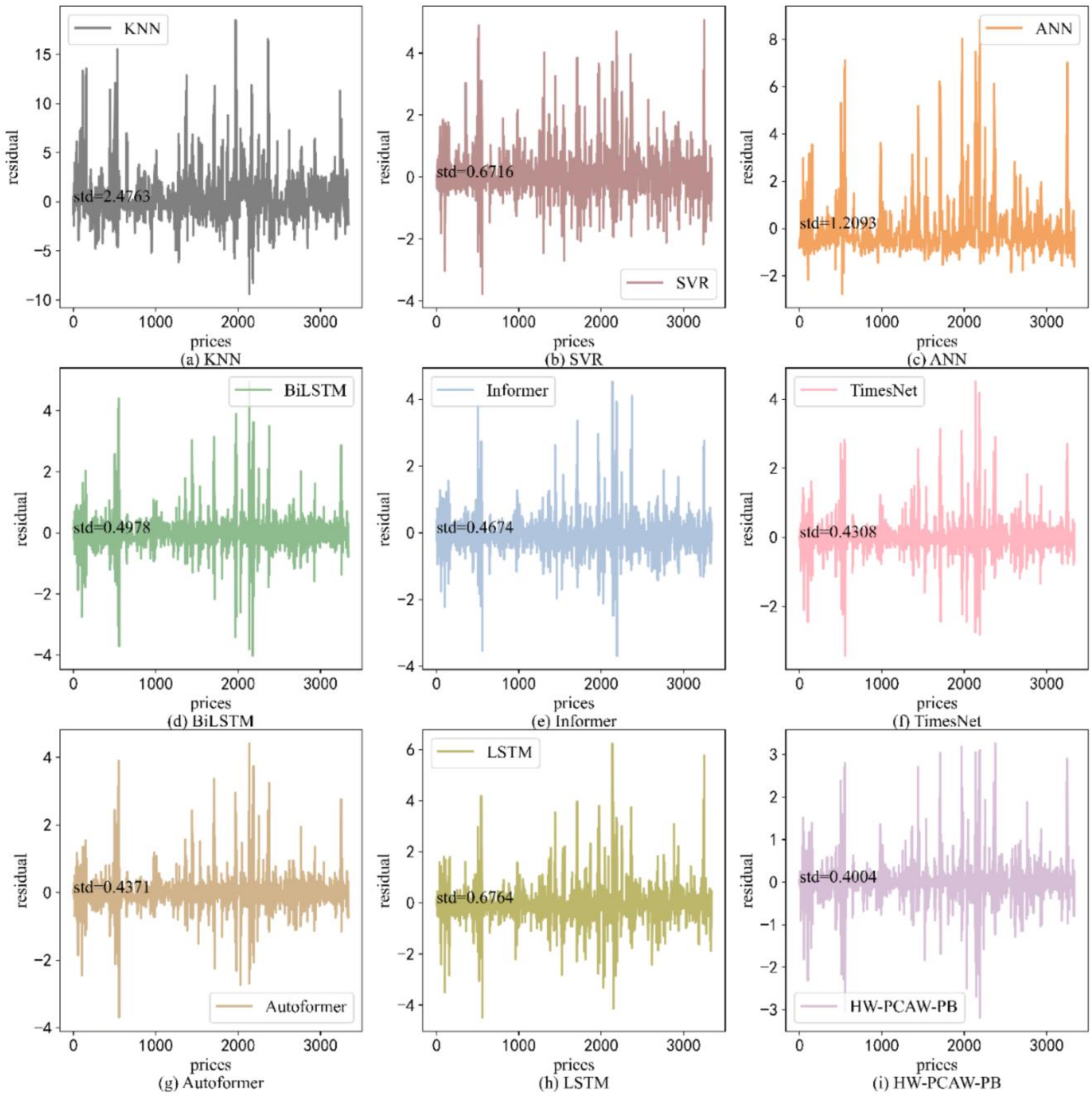


Fig. 6. Plot of predicted residuals for the nine models.

Fig. 6 shows the prediction residual graphs of nine models. The maximum residual value of the KNN model in Fig. 6(a) is 2.4763, indicating that the KNN model has the worst prediction effect. Secondly, the residual value of the ANN model in Fig. 6(c) is 1.2093, indicating that the fitting effect is better than KNN. The residuals of the SVR model in Fig. 6(b) and the LSTM model in Fig. 6(h) are relatively close, with 0.6716 and 0.6764. Compared with the ANN model, the predictions of the SVR and LSTM models are more accurate. Fig. 6(d) BiLSTM, Fig. 6(e) Informer, Fig. 6(f) TimesNet, and Fig. 6(g) Autoformer's residues are 0.4978, 0.4674, 0.4308, and 0.4371, indicating that the prediction accuracy of the BiLSTM, Informer, TimesNet, and Autoformer models is relatively high. The residuals of the HW-PCAW-PB model in Fig. 6(i) are closely distributed, and the smallest prediction residual among the nine models is 0.4004. This shows that the HW-PCAW-PB model is more accurate in predicting stock prices and has the best fitting effect on the data.

E. DM Test

In order to evaluate and compare whether there is a significant difference in prediction effectiveness of different stock price forecasting models, this study uses Diebold Mariano (DM) statistical test method. The DM test is a modified t-test used to test Whether there is a notable difference in the prediction errors of two prediction models. Its basic prerequisites or zero assumptions (H0) are the forecasting precision of the two models without significant differences. In this study, the test results of DM value and p value in Table VII show that at the 5% significance level, there is a significant inconsistency in the prediction efficiency between the proposed hybrid prediction model in this work and the benchmark model. This statistical difference shows that the hybrid forecasting model developed in this study shows significant superiority in stock price prediction over the benchmark model.

TABLE VII
DM TEST

Model	Metric	601555	601186	601688	601939	601633	600690
SVR	DM	5.31	3.47	3.76	3.61	2.34	3.07
	P	5.51E-8	2.55E-4	8.31E-5	1.52E-4	0	0
BiLSTM	DM	6.01	4.89	5.12	5.88	3.99	5.10
	P	9.05E-10	4.96E-7	1.47E-7	1.96E-9	3.18E-5	1.69E-7
TimesNet	DM	2.84	4.26	4.39	4.76	3.30	2.34
	P	0	9.93E-6	5.41E-6	9.39E-7	4.79E-4	0
KNN	DM	4.60	4.80	4.79	3.65	2.58	6.31
	P	2.16E-6	7.83E-7	8.01E-7	1.28E-4	0	1.37E-10
ANN	DM	11.11	6.15	5.54	5.84	4.27	4.76
	P	0	3.68E-10	1.47E-8	2.49E-9	9.73E-6	9.62E-7
LSTM	DM	6.17	4.86	5.55	3.84	2.03	3.88
	P	3.40E-10	5.69E-7	1.40E-8	6.13E-5	0.02	5.03E-5
Autoformer	DM	3.45	4.34	3.77	5.63	4.78	3.29
	P	2.83E-4	7.04E-6	7.86E-5	8.94E-9	8.39E-7	4.98E-4
Informer	DM	3.21	3.73	3.10	3.44	3.99	3.13
	P	6.48E-4	9.33E-5	9.46E-4	2.89E-4	3.30E-5	8.59E-4

F. Ablation Experiments

For validating the prediction effect of the proposed HW-PCAW-PB model, ablation experiments are performed on the key modules and modes. Different variant models are obtained by removing different modules and modes in the model. The variant model of the HW-PCAW-PB model (PCAW-PB) was obtained by removing the Holt-Winters analysis processing module and controlling that the rest of the HW-PCAW-PB model remained unchanged. The PCA whitening transformation module is removed and the other parts of the HW-PCAW-PB model are controlled to remain unchanged to obtain the variant model (HW-PB) of the HW-PCAW-PB model. In this paper, the HW-PCAW-PB model is trained based on unimodal data of social media text information (Text) and unimodal data of stock price series (Stock). The experimental comparison results are shown in Table VIII.

TABLE VIII
COMPARISON OF ABLATION RESULTS

Model	Metric	601555	601186	601688	601939	601633	600690
PCAW-PB	RMSE	0.1022	0.0952	0.2871	0.1061	0.1131	0.3561
	MAE	0.0806	0.0681	0.2119	0.0825	0.0864	0.3012
	MAPE	0.008	0.0059	0.0115	0.0114	0.0074	0.0164
	R^2	0.9221	0.872	0.8902	0.7695	0.725	0.7027
HW-PB	RMSE	0.1005	0.1123	0.3006	0.098	0.1265	0.3754
	MAE	0.0843	0.0795	0.229	0.0753	0.0966	0.3018
	MAPE	0.0084	0.0069	0.0124	0.0104	0.0083	0.0164
	R^2	0.9246	0.8219	0.8797	0.8033	0.6561	0.6696
Text	RMSE	0.1121	0.1127	0.3132	0.135	0.1427	0.3923
	MAE	0.0936	0.0915	0.297	0.0793	0.1252	0.3745
	MAPE	0.0097	0.0085	0.0165	0.0199	0.0132	0.0211
	R^2	0.9087	0.8168	0.8343	0.7889	0.6356	0.6389
Stock	RMSE	0.1011	0.0956	0.2789	0.0998	0.1127	0.3552
	MAE	0.0801	0.0681	0.2112	0.0812	0.0856	0.2998
	MAPE	0.0085	0.0061	0.0116	0.0112	0.0076	0.0176
	R^2	0.9231	0.8716	0.8905	0.8011	0.7178	0.7032
HW-PCAW-PB	RMSE	**0.0976**	**0.0951**	**0.2807**	**0.096**	**0.1116**	**0.3536**
	MAE	**0.076**	**0.0680**	**0.2025**	**0.0724**	**0.0828**	**0.2942**
	MAPE	**0.0076**	**0.0059**	**0.011**	**0.01**	**0.0071**	**0.016**
	R^2	**0.929**	**0.8723**	**0.8951**	**0.8114**	**0.7323**	**0.7069**

As shown in Table VIII, The HW-PCAW-PB model performed better than the variants in all respects, and compared to the variant model PCAW-PB, the RMSE, MAE, and MAPE of the HW-PCAW-PB model are reduced by an average of 2.38%, 4.19%, and 4.95%, and R^2 improved by 1.32% on average, compared to the variant model HW-PB, the RMSE, MAE, and MAPE of the HW-PCAW-PB model were reduced by 7.07%, 8.15%, and 8.28% on average, respectively, and R^2 improved by 3.88% on average. The effectiveness of incorporating Holt-Winters analytical treatment and PCA whitening transformation algorithm is verified. The Holt-Winters smoothing index methodology allows for a complete understanding of the cyclical features of stock price volatility by modelling trends, seasonality and smoothing, and the anisotropy problem existing in the embedding of stock comments can be effectively solved by applying the PCA whitening transformation to the stock comment embedding vector matrix, where the accuracy in predicting stock prices has been raised. This indicates that the addition of the Holt-Winters method and the PCA whitening transformation to the HW-PCAW-PB model significantly improves the prediction, which further verifies the superiority and effectiveness of the Holt-Winters method and PCA whitening transformation, and the capacity to predict future stock price trends has markedly improved. Compared to Text, the RMSE, MAE, and MAPE of the HW-PCAW-PB model are reduced by an average of 14.35%, 24.99%, and 35.21%, respectively, and the R^2 is improved by an average of 6.55%, and compared to Stock, the RMSE, MAE, and MAPE of the HW-PCAW-PB model are reduced by an average of 0.83%, 3.64%, and 7.99%, and R^2 is improved by 0.8% on average. Stock price series, as key time-series data containing trends and volatility of historical price movements, have a more direct impact on the forecasting model. Although social media text information can provide clues to market sentiment, it is still less valuable than stock price series. To improve forecasting performance, This paper combines multiple data sources and fully utilize the complementary characteristics of social media text and stock price series to achieve more accurate stock price forecasts.

G. Sensitivity Analysis

In this section, sensitivity analyses are performed for different hyperparameters, including the size of the sliding window and the number of encoder layers. The settings of each parameter are adjusted separately and the performance (MAPE and R^2) on the three datasets (601555, 601186, and 601688) is reported.

Fig. 7 shows the variation of MAPE and R^2 for HW-PCAW-PB on the three stock datasets under five different settings of the sliding window size $m \in \{5,10,15,20,25\}$. From Fig. 7(a), it can be seen that as the sliding window increases, the MAPE value of the model shows a trend of decreasing and then increasing. When the sliding window size is 10, the performance of the HW-PCAW-PB model reaches the optimum and its prediction is the most accurate. From Fig. 7(b), it can be seen that when the sliding window size is 10 or 15, the R^2 value of the HW-PCAW-PB model reaches the maximum, and the better the fitting effect of its HW-PCAW-PB model. In summary, this paper chooses the sliding window size of 10 as the optimal sliding window size.

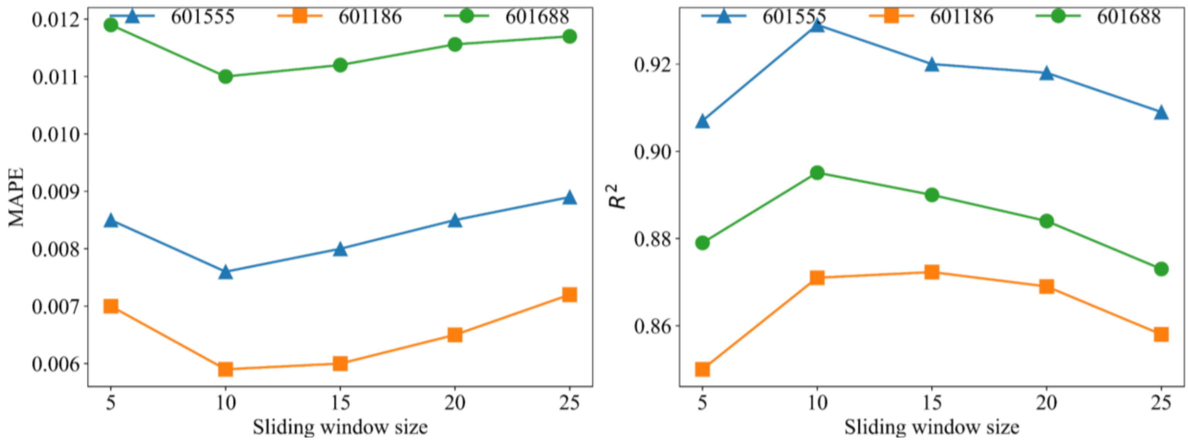


Fig. 7. Variation of model MAPE and R^2 with sliding window size on each stock dataset.

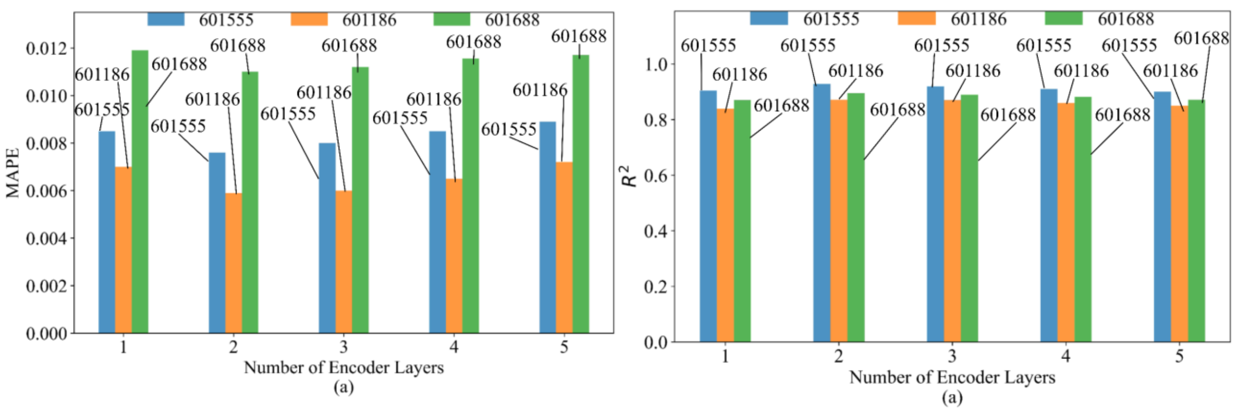


Fig. 8. Variation of model MAPE and R^2 with number of encoder layers on each stock dataset.

Fig. 8 demonstrates the variation of MAPE and R^2 of HW-PCAW-PB on the three stock datasets under five different choices of the number of encoder layers $n \in \{1,2,3,4,5\}$. From Fig. 8 (a) and Fig. 8 (b), it is apparent that the HW-PCAW-PB model has the smallest MAPE and the largest R^2 value at n=2, which indicates that the model reaches optimal performance with 2 encoder layers. Consequently, 2 encoder layers are chosen for this study.

IV. CONCLUSION

Forecasting stock time series poses a significant and intricate challenge within the field of financial, which is of great significance to both academic research and actual investment decisions. In this work, we propose a novel hybrid model named HW-PCAW-PB. The model uses the Holt-Winters triple exponential smoothing algorithm to boost the accuracy and robustness performance of prediction, introduces the PCA whitening transformation to eliminate the anisotropy problem in the embedding of stock reviews, and uses BERT for semantic embedding, which solves the representation degradation problem existing in traditional pre-trained language models, and further improves the prediction accuracy by using multi-source feature fusion. We compare the HW-PCAW-PB model with a variety of existing methods such as SVR, BiLSTM, TimesNet, KNN, ANN, LSTM, Autoformer, and Informer. The experimental results show that the HW-PCAW-PB model has significant improvements in RMSE, MAE, and MAPE indicators, with RMSE increased by more than 5%, MAE and MAPE increased by more than 10%, and R^2 also improved to a certain extent, which verifies the effectiveness of the hybrid model method and has good application prospects. This paper has some limitations. Due to time constraints in our experiments, the technique applied for extracting features from news was straightforward. Further research will explore hybrid models based on Transformer and other deformable technologies, as well as the automation of model hyperparameter optimization, with the aim of improving the accuracy and stability of stock price predictions.

REFERENCES

[1] H. Kim, "An empirical analysis of navigation behaviors across stock and cryptocurrency trading platforms: implications for targeting and segmentation strategies," *Electronic Commerce Research Journal*, pp. 1–29, Oct. 2022.

[2] D. Liang, "ARIMA Model-Based Research on Stock Price Prediction," *Science and Technology of Engineering, Chemistry and Environmental Protection*, vol. 1, no. 7, Jun. 2024.

[3] V. Vijayalakshmi, "Implementation of sentiment analysis in stock market prediction using variants of GARCH models," in *Computational Intelligence Methods for Sentiment Analysis in Natural Language Processing Applications*, pp. 227-249, Jan. 2024.

[4] Arvind Ganesan, and Adarsh Kannan, "Stock price prediction using ARIMA model," *International Research Journal of Engineering and Technology (IRJET)*, vol. 8, no. 08, pp226-234, 2021.

[5] L. N. Mintarya, J. N. Halim, C. Angie, S. Achmad, and A. Kurniawan, "Machine learning approaches in stock market prediction: A systematic literature review," *Procedia Computer Science*, vol. 216, pp. 96-102, 2023.

[6] Melina, Sukono, H. Napitupulu, A. Sambas, A. Murniati, and V. A. Kusumaningtyas, "Artificial Neural Network-Based Machine Learning Approach to Stock Market Prediction Model on the Indonesia Stock Exchange During the COVID-19, " *Engineering Letters*, vol. 30, no. 3, pp988–1000, 2022.

[7] D. G. Jayaswara, I. Slamet and Y. Susanti, "Prediction of Central Asia Bank's Stock Price using Support Vector Regression Method," in *Proceeding International Conference on Religion, Science and Education*, vol. 2, pp. 7-12, April. 2023.

[8] S. Gao, "Trend-Based K-Nearest Neighbor Algorithm in Stock Price Prediction," in *3rd International Conference on Digital Economy and Computer Application (DECA 2023)*, Atlantis Press, 2023, pp. 746–756.

[9] H. T. Liang, S. Liu, J. W. Du, Q. Hu, and X. Yu, "Review of deep learning applied to time series prediction," *J. Front. Comput. Sci. Technol*, vol. 17, no. 06, pp. 1285–1300, 2023.

[10] Y. Zhu, "Stock price prediction using the RNN model," in *Journal of Physics: Conference Series*, Vol. 1650, No. 3, p. 032103, Oct. 2020.

[11] Z. Li, H. Yu, J. Xu, J. Liu, and Y. Mo, "Stock market analysis and prediction using LSTM: A case study on technology stocks," *Innovations in Applied Engineering and Technology*, pp. 1-6, Nov. 2023.

[12] T. Muhammad *et al.* , "Transformer-based deep learning model for stock price prediction: A case study on bangladesh stock market," *International Journal of Computational Intelligence and Applications*, vol. 22, no. 03, p. 2350013, 2023.

[13] H. Wu, J. Xu, J. Wang, and M. Long, "Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting," *Advances in neural information processing systems*, vol. 34, pp. 22419-22430, Dec. 2021.

[14] X. Huang, J. Tang, and Y. Shen, "Long time series of ocean wave prediction based on PatchTST model," *Ocean Engineering*, vol. 301, p. 117572, Jun. 2024.

[15] M. Mohan, P. K. Raja, P. Velmurugan, and A. Kulothungan, "Holt-winters algorithm to predict the stock value using recurrent neural network," methods, vol. 8, no. 10, 2023.

[16] R. N. Paramanik, and V. Singhal, "Sentiment analysis of Indian stock market volatility," *Procedia Computer Science*, vol. 176, pp. 330-338, Jan. 2020.

[17] W. Lu, J. Li, Y. Li, A. Sun and J. Wang, "A CNN-LSTM-based model to forecast stock prices," *Complexity* , vol. 2020, no. 1, pp. 330-338, Jan. 2020.

[18] S. Wu, Y. Liu, Z. Zou, and T. H. Weng, "S_I_LSTM: stock price prediction based on multiple data sources and sentiment analysis," *Connection Science*, vol. 34, no. 1, pp. 44-62, 2022.

[19] Q. Zhuge, L. Xu, and G. Zhang, "LSTM neural network with emotional analysis for prediction of stock price, "*Engineering Letters*, vol. 25, no. 2, pp167–175, 2017.

[20] M. Jiang, L. Jia, Z. Chen,and W. Chen, "The two-stage machine learning ensemble models for stock price prediction by combining mode decomposition, extreme learning machine and improved harmony search algorithm," *Annals of Operations Research*, pp. 1-33, Feb. 2022.

[21] Chi Ma, Jie Wu, Hui Hu, Yue Nai Chen, and Jing Yan Li, "Predicting stock prices using hybrid LSTM and ARIMA model," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 3, pp. 424-432, 2024.

[22] F. Zhao *et al.* , "Multi-layer features ablation of BERT model and its application in stock trend prediction," *Expert Systems with Applications*, vol. 207, p. 117958, Nov. 2022.

[23] D. Daiya, and C. Lin, "Stock movement prediction and portfolio management via multimodal learning with transformer," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Jun 2021, pp. 3305-3309.

Changsheng Zhu is a distinguished professor and doctoral supervisor at Lanzhou University of Technology, where he plays a key role in advancing research and education. His primary research interests encompass high-performance computing, big data analytics, and engineering manufacturing information systems, among other areas. With extensive expertise in these fields, Professor Zhu has made significant contributions to both theoretical advancements and practical applications. Additionally, he serves as the director of the Traditional Chinese Medicine Big Data Engineering Research Center in Gansu Province, where he oversees projects aimed at leveraging big data technologies to promote innovation and development in the field of traditional Chinese medicine.