International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 938-948

https://doi.org/10.70517/ijhsa46487

Modeling Teachers' Career Development Paths Based on Optimization of Artificial Intelligence Algorithms in Teacher Education

Mengting Guo[1,*] and Xiaopeng Liu[1]
[1] School of Education, Hanjiang Normal University, Shiyan, Hubei, 442000, China
Corresponding authors: (e-mail: mengtinguo@163.com).

Abstract With the continuous changes in the educational environment, the traditional teacher development path model is facing the challenges of multidimensionality and dynamism. The application of artificial intelligence technology, especially machine learning and optimization algorithms, provides new perspectives and methods for teacher development in education. Based on artificial intelligence algorithms, this study proposes a model for optimizing teachers' career development path. First, a mathematical model of the teacher career development path is constructed by combining the basic information of teachers, work performance and external environmental factors. Genetic algorithm and particle swarm optimization algorithm are used to optimize the path to maximize teachers' career potential and the utilization efficiency of school resources. In order to further improve the prediction accuracy, ARIMA and LSTM models were combined to model the linear and nonlinear parts of the time series data, respectively. The ARIMA model was used to obtain a smooth sequence through difference processing and to make preliminary predictions on the teacher career development data. Subsequently, the residual data from the ARIMA model was input into the LSTM model to capture the nonlinear trend and achieve a more accurate prediction of teacher career paths. The experimental results show that the ARIMA-LSTM hybrid model has a mean square error (MSE) of 1.0795, a root mean square error (RMSE) of 0.9456, and a mean absolute error (MAE) of 0.4516, which is significantly better than the traditional ARIMA and LSTM models. The optimization model provides new methods and ideas for the scientific planning of teachers' career development path.

Index Terms artificial intelligence, teachers' career, path optimization, ARIMA model, LSTM model, hybrid model

I. Introduction

In today's society, the roles and responsibilities of teachers are increasingly important, they are not only the transmitter of knowledge, but also the shaper of values and the explorer of students' potential [1]. For schools, teachers are the main force for high-quality development, talent cultivation and service to society, and their comprehensive quality is directly related to the effectiveness of school reform and development [2]-[4]. For students, teachers have an important guiding role in the formation of students' values, the development of thinking patterns, the mastery of professional knowledge and skills, and the cultivation of innovative and entrepreneurial spirit, and are the main support for schools to cultivate high-quality talents [5]-[7]. For teachers themselves, self-development is an intrinsic need to enhance professional ability and reflect professional value [8]. However, with the constant changes in the educational environment, teachers are facing more and more challenges, such as curriculum reform, technological innovation, and student diversity [9], [10]. All these challenges require teachers to continuously improve their professionalism and teaching ability to adapt to the new trend of educational development [11].

At this stage, most frontline teachers generally pay attention to improving students' academic level, but often neglect career planning centered on their own professional development [12], [13]. And suitable career planning can promote teachers to carry out education and teaching work more effectively and promote their professional development [14]. In particular, it should be pointed out that career planning is definitely not the patent of young teachers, and it is of great significance for classrooms of any teaching age to start planning their own careers [15], [16]. Therefore, with the assistance of artificial intelligence technology, with the goal of improving teachers' active learning, exploring ways to enhance teachers' career development paths can help to open up teachers' horizons and make them grow on a higher platform [17], [18].

This study combines two models, ARIMA and LSTM, to improve the prediction accuracy of teachers' career paths through an optimization algorithm. The ARIMA model first performs a time series analysis on the historical data of teachers' career development in order to predict its future development trend. However, the nonlinear prediction ability of the ARIMA model is weak, so this paper further applies the LSTM model to the residual processing of ARIMA to capture the fluctuation of the nonlinear part of the teachers' career path. This combination method can effectively improve the accuracy of prediction, which is especially practically important for the prediction of career paths containing complex dynamic factors. Specifically, firstly, a teacher career path optimization model adapted to different scenarios is constructed by analyzing and modeling multidimensional data such as teachers' basic information, job performance, and behavioral characteristics. Secondly, a combination of ARIMA and LSTM models is used to reduce the limitations of the traditional single model through step-by-step optimization and data fusion. Finally, the model is applied to actual data, and its effect in the prediction of teachers' career paths is verified through experiments, with a view to providing powerful support for education management.

II. Strategies for optimizing teachers' career development paths based on artificial intelligence

II. A. Career development path optimization model construction

The construction of the career development path optimization model based on artificial intelligence mainly relies on multidimensional data analysis and machine learning algorithms to achieve the scientific design and optimization of teachers' career development paths. In the model, the core task is to establish a mathematical model that can dynamically assess and optimize teachers' career development, combining data inputs, algorithmic reasoning and optimization goals to determine the best path.

The key variables input to the model include teachers' basic information (e.g., age, work experience, academic background, etc.), work performance data (e.g., performance ratings, experience, skill mastery, etc.), behavioral characteristics (e.g., career interests, work attitudes, etc.), and external environmental factors (e.g., industry trends, school development strategies, etc.). These variables will be used as model inputs through standardization.

When constructing the model, it is first necessary to define the objective function of career path optimization for teachers' career development in order to maximize teachers' career development potential and schools' resource utilization efficiency. Assuming that $f(x)$ denotes the career path optimization objective function, where x represents different career development paths of teachers, the function $f(x)$ should take into account the effects of teachers' performance and external conditions on career paths. The model can be expressed by the following equation:

$$f(x)=\sum_{i=1}^{n} w_i \cdot x_i \tag{1}$$

where w_i is the weight of each factor, x_i is the quantitative performance of the corresponding factor, and n is the number of factors affecting the path selection. The optimization objective is to maximize $f(x)$ so as to obtain the optimal career path.

For the objective function, genetic algorithm and particle swarm optimization algorithm (PSO) are used for path optimization. For example, genetic algorithms can select appropriate career paths by simulating the natural selection process and optimizing iteratively. Each generation of selection is based on a fitness function, which measures the quality of each path as a means of filtering out the optimal solution.

II. B. Data acquisition and processing

The data collection phase involves collecting teacher-related data from multiple channels and dimensions, including structured data and unstructured data. Structured data mainly include teachers' basic information (e.g., age, education, length of service, etc.), work performance data (e.g., performance ratings, project experience, etc.), and behavioral data (e.g., work attitudes, ability to work in a team, etc.). Unstructured data, on the other hand, include teachers' professional development goals, career interests, and affective tendencies. These data can be collected in a variety of ways, such as through performance evaluations and teacher questionnaires.

After the data collection is completed, the data preprocessing stage is entered to ensure the completeness and consistency of the data and to lay the foundation for the subsequent analysis and modeling. Data cleaning methods such as missing value filling, outlier detection, and duplicate data deletion are utilized for data cleaning to ensure data quality. For missing values, mean filling, interpolation or machine learning-based prediction filling methods are used; for outliers, they are identified and processed by statistical methods or model detection.

Feature selection and feature transformation are also required in the data processing process. The purpose of feature selection is to extract the most predictive and representative features from the original data to avoid redundant information interfering with the model. Feature transformation, on the other hand, includes methods such as normalization and normalization to make the data more suitable for use in machine learning models. For example, normalization uses the following formula:

$$x' = \frac{x-\mu}{\sigma} \tag{2}$$

where x' is the standardized data, x is the original data, μ is the mean of the data, and σ is the standard deviation of the data. The standardized data can eliminate the influence of different feature scales, making the model training more stable and efficient.

II. C.Artificial Intelligence Algorithm Selection and Model Optimization

II. C. 1) ARIMA time series modeling

The ARIMA model [19] is a widely used differential autoregressive moving average model for forecasting smooth time series. The model mainly focuses on capturing linear relationships, and is slightly weak in capturing nonlinear relationships. When dealing with non-smooth time series, the ARIMA model transforms the original series into a smooth series through the difference method, and then regresses the lagged values and random errors in the variables. As a result, the ARIMA model is able to utilize the past and current values of the time series to effectively forecast the future values. Therefore, based on the fact that the data in this paper are all time series data, the process of analyzing the data as time series and building an ARIMA time series model to capture the specific relationship between the level of teacher title and time is essential.

ARIMA time series model consists of autoregressive, difference and moving average three core components, specifically: AR represents autoregressive, I represents the number of single-integer orders, smoothness is crucial in the process of time series modeling, if the original time series is not smooth, it is necessary to use the difference method to transform it into a smooth series. The number of differences represents the number of single-integer orders, the number of differences that convert a non-smooth time series into a smooth series; MA represents the moving average.

The core of the AR model is to use the historical time data of the variables to describe and predict the current value, emphasizing the correlation between the historical value and the current value, but the time series to establish the AR model must satisfy the condition of smoothness, and the formula of the p-order autoregressive process is defined as follows:

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t \tag{3}$$

where y_t is the current value, μ is the constant term, p is the order, γ_i is the autocorrelation coefficient, and ε_i is the error term. The formula is expanded as:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots \alpha_p X_{t-p} + u_t \tag{4}$$

If the random perturbation term is a white noise, i.e., $u = 0$, the AR model is said to be a pure AR(p) process, denoted:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots \alpha_p X_{t-p} + \varepsilon_t \tag{5}$$

As can be seen from the formula, the prediction of the current value is based on historical data, and p is a key parameter in the model i.e. the order, which reflects how many periods of historical data are used to project the current value.

The following points need to be satisfied to use the autoregressive model for prediction:

(1) The autoregressive model is forecasting with its own data.

(2) The time series data must be smooth.

(3) It is suitable for the case of large correlation, when the autocorrelation coefficient is less than 0.5, it is not suitable to choose this method.

(4) Autoregression is only suitable for predicting phenomena that are correlated with their own prior periods.

The MA model can effectively avoid the influence of random fluctuations in forecasting through the moving average method. In the AR model, if u_t is not a white noise, it is usually considered to be a moving average of order q, denoted as:

$$u_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q} \tag{6}$$

where ε_i denotes the white noise sequence. In particular, when $X_t = u_t$, i.e., the current value of the time series is not related to the historical values but only to a linear combination of historical white noise.

$$X_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + ... + \beta_q \varepsilon_{t-q} \tag{7}$$

whereas the effect of historical white noise on the current predicted value in the AR model is indirect and cumulative, this is the key information that the MA model seeks to capture. The formula for the q-order MA model is defined as:

$$y_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{8}$$

The ARMA model is a comprehensive time series analysis model that combines the features of both AR(p) and MA(q) models to form the more general autoregressive moving average model ARMA(p, q), denoted as:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + ...\alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + ... + \beta_q \varepsilon_{t-q} \tag{9}$$

This equation shows that a stochastic time series can be represented and predicted by a combination of its own historical values (i.e., lagged values) as well as random perturbation terms, and that if this time series is smooth, i.e., its statistical properties do not change significantly over time, then the historical behavioral patterns of this series can be used to infer its future trends.

The general expression of ARIMA model is ARIMA (p, d, q), p is the number of autoregressive terms, q is the number of moving average terms, and d is the number of differencing needed to convert the original series into a smooth state. Difference as an important step in the modeling process, because non-smooth data is difficult to reveal the intrinsic laws of time series data, so data stability is crucial.ARIMA model is a combination of ARMA model and difference model, consisting of autoregressive model (AR) and moving average model (MA), for the non-smooth time series by difference to get a smooth sequence, and then the ARMA modeling.

The ARIMA model, or differential autoregressive moving average model, is a commonly used statistical model for forecasting smooth time series. The model is mainly used to capture linear relationships, and is not as powerful for capturing nonlinear relationships. When dealing with non-smooth time series, the ARIMA model transforms the raw data into a smooth series by differencing so that the dependent variable depends only on its lagged value and the current and lagged values of the random error term. This transformation enables more effective prediction of future values with the help of historical and present values of the time series in the ARIMA (p, d, q) model is structured as follows:

$$\begin{cases} \phi(B)\nabla^d x_t = \theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_s^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(\varepsilon_t \varepsilon_s) = 0, \forall s < t \end{cases} \tag{10}$$

where $\nabla^d = (1-B)^d B, \phi(B) = 1 - \phi_1 B - ... - \phi_p B^p$ is the autoregressive coefficient polynomial of the smooth reversible ARMA (p, q) model, $\theta(B) = 1 - \theta_1 B - ... - \theta_p B^p$ is the moving average coefficient polynomial of the smooth reversible ARMA (p, q) model.

Finally, the specific steps for building an ARIMA model for a set of time series data are shown in Figure 1.

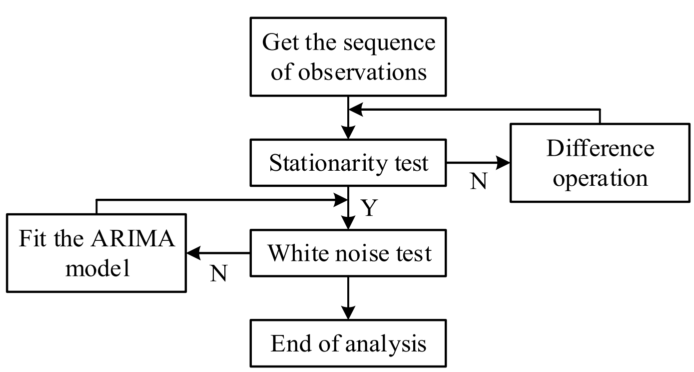

Figure 1: Modeling process of the ARIMA time series model

II. C. 2) LSTM model

LSTM (Long Short-Term Memory Network) [20] is an advanced form of Recurrent Neural Network (RNN) for efficiently processing time series data. Compared to traditional RNNs, LSTMs demonstrate enhanced performance in modeling and prediction of time series data by introducing complex gating mechanisms to solve the long-term dependency problem. A typical RNN structure consists of an input layer, an output layer, and a hidden layer. These hidden layers store and manage the sequential information in the time-series data through interconnected neurons. The LSTM structure is shown in Fig. 2, where x denotes the input value; h denotes the hidden layer; w_{zh} denotes the weight from the input layer to the hidden layer; w_{zh} denotes the weight from the hidden layer to the hidden layer; w_{zh} the hidden layer to the output layer weights; y denotes the output value.

LSTM improves long sequence learning by introducing gate control units, a mechanism that provides important improvements over RNNs. Compared with the traditional recurrent neural network RNN, the internal structure of long-short-term memory network LSTM is more complex. It not only sets up a hidden layer, but also introduces memory units, which include input gates, output gates, and forgetting gates. These gate control units utilize the Sigmoid activation function, which can selectively retain, forget, or add information, thus effectively solving problems such as gradient vanishing and gradient explosion. In this way, LSTM is able to better capture long-term dependencies and improve the learning ability of neural network on long sequence data.The structure of LSTM neural network is shown in Fig. 2.

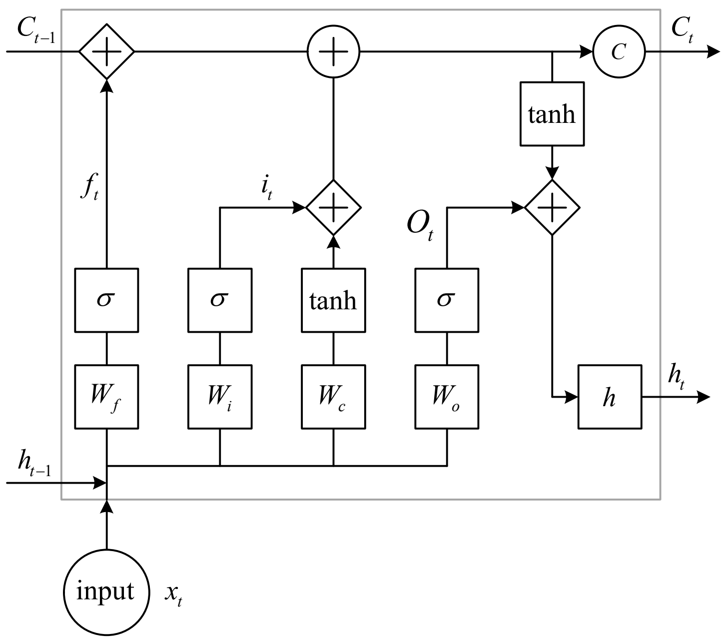

Figure 2: LSTM model

The input gate i_t, output gate o_t oblivion gate f_t, memory cell update state C_t, and the neural network output value h_t are computed as shown in Eqs. (11) to (16):

$$i_t = \sigma(W_i[h_{t-1}tx_t]+b_i) \tag{11}$$

$$f_t = \sigma(W_f[h_{t-1}tx_t]+b_f) \tag{12}$$

$$o_t = \sigma(W_o[h_{t-1}tx_t]+b_c) \tag{13}$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}tx_t]+b_c) \tag{14}$$

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \tag{15}$$

$$h_t = o_t \tanh(C_t) \tag{16}$$

where x_i denotes the variable input into the neural network unit at the moment t; and σ denotes the Sigmoid function, which produces vectors between [0, 1] based on the input; $\tilde{C}_i$ represents the candidate cell information; W_f, W_i, W_o, W_c denote the corresponding weights of different gates; The b_f, b_1, b_o, b_c denote the bias corresponding to the various gates.

II. C. 3) Model of ARIMA-LSTM Optimization Algorithm

In the prediction of teachers' career development path, there are both linear and nonlinear variations, which cause a single model can not be fitted well, thus leading to a large prediction error. On this basis, this study fuses the ARIMA model with the LSTM model to establish the ARIMA-LSTM optimization algorithm prediction model [21] to improve the prediction accuracy.

The ARIMA-LSTM optimization algorithm firstly predicts the collected data by ARIMA model, and then inputs the residuals into the ISTM model to get the predicted value of the residuals and corrects the ARIMA model. The steps are as follows: the optimization algorithm adopts the ARIMA model to predict the linear trend of the information related to teachers' career development paths, and the nonlinear part is retained in the residuals of the ARIMA model. On this basis, the residual data are input into the LSTM model to realize the prediction of the nonlinear trend error. Finally, the prediction results of the 2 models are integrated to obtain the prediction value of the best prediction model. The flow of the optimization algorithm constructed in this study is shown in Figure 3.

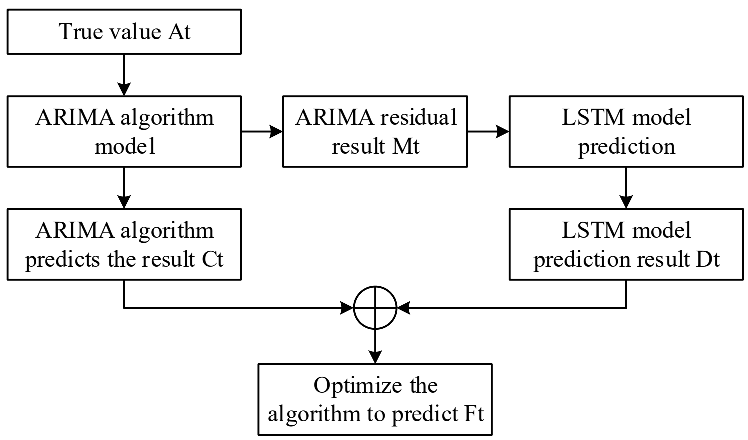

Figure 3: Model Structure of the ARIMA-LSTM optimization algorithm

In the ARIMA model, A_t is the true value at time t, C_t is the prediction result of the ARIMA algorithm, and the residuals of the ARIMA input to the LSTM are expressed as equation (17):

$$M_t = A_t - C_t \tag{17}$$

In the LSTM model, $\{M_t\}$ is modeled, and the nonlinear relationship of the residual values is described by LSTM to obtain the residual prediction value D_t:

$$D_t = f(e_{t-1}, e_{t-2}, \cdots, e_{t-m}) + \varepsilon_t \tag{18}$$

Finally, the outputs of ARIMA and LSTM models are combined to obtain the combined optimization model prediction results:

$$F_t = C_t + D_t \tag{19}$$

III. 3. Predictive analysis of models for optimizing teachers' career paths

III. A. Empirical analysis of the ARIMA model

ARIMA models were established for data in different sample periods, and the model prediction accuracy was verified with 2019 and 2023 data respectively. Before modeling, make a time series plot according to the original data Xt, Figure 4 shows the time series plot of the sequence Xt, which can be initially judged that the original data of teachers' career development path optimization is a non-stationary sequence.

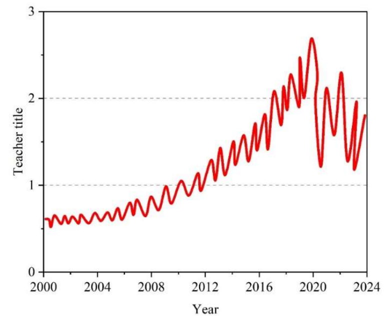

Figure 4: Sequence diagram

III. A. 1) Data smoothing

Before making predictions using the ARIMA(p,d,q) model, it is important to ensure that the data being processed are smooth and non-white noise series. Therefore, a series of preprocessing and tests were first performed on the raw data. By observing the time series plot of the sequence Xt, the sequence is initially judged to be an unsteady sequence. At the same time, it can be seen from the autocorrelation and partial correlation plots of the original sequence that the sequence is not a purely random sequence, indicating that there is a certain correlation between the level of teacher title in each time period. In order to eliminate the influence of subjective factors on the judgment of smoothness, ADF unit root test was carried out using EViews statistical software, and the test results are shown in Table 1. The value of the t-test statistic of the time series of the level of teachers' titles of the original data is -1.17045, which is greater than the critical value of the three confidence levels, so it can be judged that the series is non-smooth. Therefore, the original series needs to be processed by means of differencing.

Table 1: Adf unit root test

Adf statistic test	T statistic	P-value
1% confidence level	-1.17045	0.6889
5% confidence level	-3.45311	
10% confidence level	-2.87146	
Adf statistic test	-2.57278	

Next, the time series data of railroad passenger traffic is processed by first-order difference to obtain the first-order difference sequence graph, and the first-order difference sequence graph is shown in Fig. 5. Observation of the image shows that the first-order difference sequence Xt is no longer characterized by an obvious upward trend, and the linear trend of Xt has been more fully extracted.

The ADF unit root test is continued for the first-order difference series Yt. The value of its t-test statistic is -5.1045, which is smaller than the critical value of -3.4579 at the 1% level, rejecting the original hypothesis that the sequence has a unit root, and therefore the sequence after the first-order differencing process can be recognized as a smooth sequence.

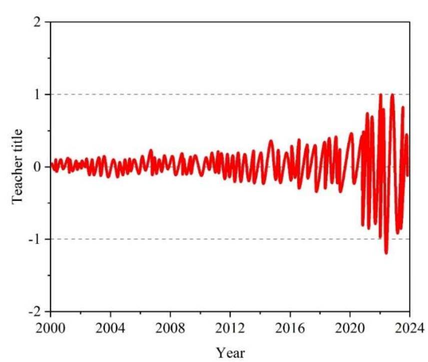

Figure 5: First order difference sequence diagram

Next, the autocorrelation and partial correlation plots of the first-order difference series Yt are plotted, and the statistical results of autocorrelation and partial correlation data are shown in Table 2. Through observation, it can be found that the white noise test shows that the P-value are less than 0.05, that is, under the 95% confidence interval, the original hypothesis can be rejected that the sequence is not a white noise sequence. The sequence Yt is able to perform the fitting of ARIMA (p,d,q) model with identified parameter 1.

Table 2: Independent correlation and partial correlation analysis

N	AC	PAC	Q-Stat	Prob
1	-0.175	-0.174	8.003	0.004
2	-0.136	-0.165	12.546	0.002
3	0.025	-0.038	12.988	0.004
4	-0.178	-0.215	21.456	0.000
5	-0.008	-0.097	21.568	0.001
6	0.034	-0.056	21.789	0.001
7	0.012	-0.024	21.799	0.002
8	-0.199	-0.267	32.564	0.000
9	0.133	0.007	37.456	0.000
10	-0.084	-0.166	39.487	0.000

III. A. 2) Model Identification and Establishment

The ARIMA model was fitted using EViews statistical software to determine the parameters p, q. Observing the sequence Yt, it was found that the sequence was suitable for ARMA(p, q). From the autocorrelation data, it can be seen that the autocorrelation coefficient is significantly not 0 at k=1~2. Similarly, the partial correlation data shows that the partial correlation coefficient is not 0 at k=1~2. Taking this information together, it can be determined that the maximum order for p and q is 2. Model fitting of the parameters by using EViews finally yields the model ARIMA(2,1,2), which not only conforms to the time-series characteristics, and the coefficients passed the significance test, the fitting effect is good. The parameter estimation results of the model are shown in Table 3.

Table 3: Parameter estimation of the model

Parameter	Coefficient	Standard error	T statistic	P-value
AR (1)	-0.4157	0.0724	-5.7397	0.0000
AR (2)	0.4958	0.0712	6.9445	0.0000
MA (1)	0.1162	0.0402	2.9012	0.0003
MA (2)	-0.8974	0.0392	-22.7617	0.0000

III. A. 3) Model Fit Tests

The sequence after information extraction should be converted into a white noise sequence by minimizing the amount of valid information. The residual series of ARIMA(2,1,2) model is subjected to the ADF unit root test, and its unit root statistic is -5.1997, which is smaller than the critical value of -3.4541 at the significant level of 1%.The autocorrelation and bias correlation plots of the residual series are further plotted, and the statistical results of the data are shown in Table 4. Observing the results of the white noise test, the P-value is greater than 0.05, indicating that the model is basically able to extract sufficiently for the time series information.

Table 4: Independent correlation and partial correlation analysis

N	AC	PAC	Q-Stat	Prob
1	-0.035	-0.035	0.246	0.615
2	0.051	0.047	0.845	0.654
3	0.074	0.078	2.051	0.577
4	-0.045	-0.046	2.522	0.645
5	-0.044	-0.052	2.915	0.715
6	-0.007	-0.012	2.924	0.745
7	-0.048	-0.037	3.514	0.874
8	-0.155	-0.154	9.456	0.341
9	0.098	0.088	11.451	0.256
10	-0.078	-0.056	12.648	0.212

III. A. 4) Prediction results and tests

The ARIMA(2,1,2) model was used to predict the time series of teachers' career development in 2024, and the model fitted well for teachers' career development as a whole. However, the nonlinear fitting ability of the model is poor, the fitting effect for the localization is not satisfactory, and the error gradually increases after 2021.The average absolute percentage error of the prediction result index in 2024 MAPE=0.1769, and the prediction value is as shown in Table 5, which shows that the overall prediction error is relatively large, and the prediction effect is average.

Table 5: The ARIMA model predicts the results (2024)

Time	APE
January	0.4944
February	0.1645
march	0.1105
April	0.2056
may	0.0894
June	0.0912
July	0.2456
August	0.1345
September	0.1754
October	0.1389
November	0.1875
December	0.0854
MAPE=0.1769	

In order to more intuitively see the fitting effect of the ARIMA model, the time series plot of the overall residuals is plotted. The time series plot of the overall residuals is shown in Figure 6, which shows that the residuals in the early stage are all within the error range, but the fluctuation of the residual series becomes more and more obvious with the extension of time, and the prediction error gradually increases. In particular, the actual value fluctuates more after experiencing the shock, and the model has a poor ability to identify and fit the outliers, leading to more inaccurate prediction data in the later period. The prediction results show that the ARIMA model is suitable for short-term prediction and has poor fitting ability for long-term data, and it only utilizes the historical data of teachers' career development, and it cannot learn the time series under the experienced shocks, and the prediction effect becomes worse with the great fluctuation of the actual values.

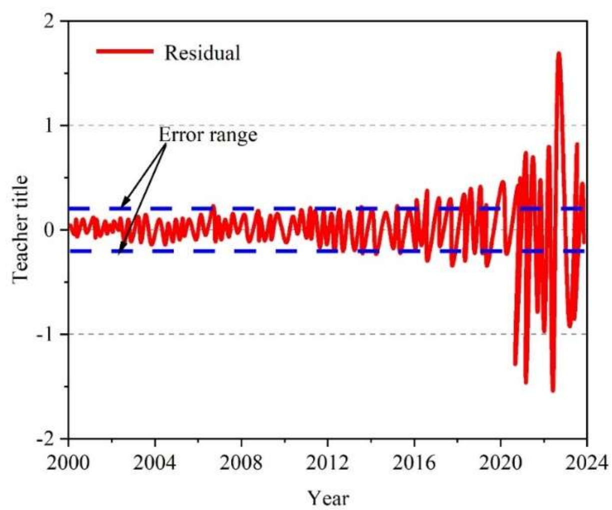

Figure 6: The residual differential sequence of the Arima model

III. B. ARIMA-LSTM hybrid model prediction

According to the prediction value obtained by adding LSTM neural network and the actual observation value, as Figure 7 shows the comparison results of the prediction made by the three models, which are the comparison of the prediction errors of the LSTM model, the ARIMA model and the ARIMA-LSTM model, respectively. From the figure, it can be seen that the LSTM model predicts the largest error, and the prediction error of the hybrid model is the smallest, and most of the error values oscillate above and below the zero value, and the prediction effect is the best. Table 6 shows a quantitative comparison of the mean square error, root mean square error and mean absolute error of the three models.

From the table, it can be seen that the prediction value after linear correction of the prediction error of the LSTM model using the ARIMA model is significantly better than the uncorrected prediction value, so the hybrid model of ARIMA-LSTM proposed in this paper is meaningful, and the prediction of the career development path of teachers in teacher education is more reasonable.

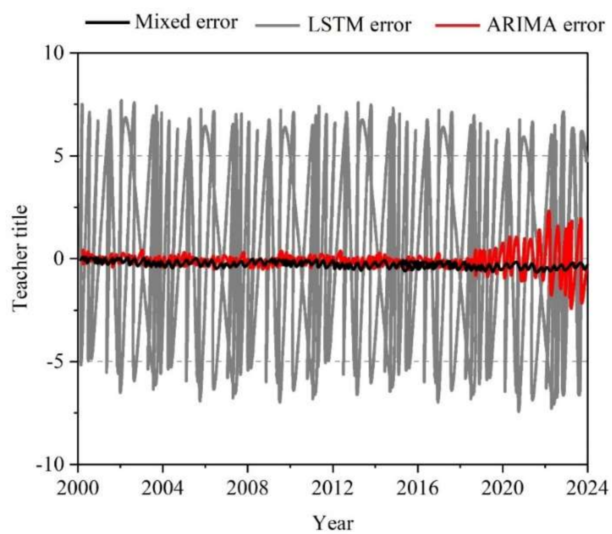

Figure 7: Error contrast diagram

Table 6: Comparison of three models

Model	MSE	RMSE	MAE
LSTM	26.1941	6.4961	4.4516
ARIMA	9.1245	3.1245	2.4125
ARIMA-LSTM	1.0795	0.9456	0.4516

IV. Conclusion

In this paper, a solution based on ARIMA-LSTM hybrid model is proposed through the study of optimization and prediction of teachers' career development path. In the comparison with the traditional ARIMA and LSTM models, the hybrid model is significantly better than the single model. Specifically, the mean square error (MSE) of the ARIMA-LSTM hybrid model is 1.0795, the root mean square error (RMSE) is 0.9456, and the mean absolute error (MAE) is 0.4516, which indicates that the model has a significant advantage in the prediction accuracy of the teachers' career development path. In contrast, the ARIMA model has a higher prediction error and lower accuracy especially when dealing with nonlinear data.The LSTM model has some limitations in overall trend prediction despite its ability to capture nonlinear trends better.

In addition, the validation results of the model show that the ARIMA-LSTM optimization model is able to better integrate the linear and nonlinear prediction advantages, providing more accurate predictions for the complex changes in teachers' career. This result not only provides new ideas for the scientific planning of teachers' career paths, but also provides references for other related studies in the field of education.

References

[1] Whalen, J., & Mouza, C. (2023). ChatGPT: Challenges, opportunities, and implications for teacher education. Contemporary Issues in Technology and Teacher Education, 23(1), 1-23.

[2] Toropova, A., Myrberg, E., & Johansson, S. (2021). Teacher job satisfaction: the importance of school working conditions and teacher characteristics. Educational review, 73(1), 71-97.

[3] Timm, J. M., & Barth, M. (2021). Making education for sustainable development happen in elementary schools: The role of teachers. Environmental Education Research, 27(1), 50-66.

[4] Ingersoll, R., Sirinides, P. M., & Dougherty, P. (2018). Leadership matters: Teachers' roles in decision making and school performance. American Educator, (543), 13-39.

[5] Keiler, L. S. (2018). Teachers' roles and identities in student-centered classrooms. International journal of STEM education, 5, 1-20.

[6] Cate, P. T., Ineke, M., Markova, M., Krischler, M., & Krolak-Schwerdt, S. (2018). Promoting Inclusive Education: The Role of Teachers' Competence and Attitudes. Insights into Learning Disabilities, 15(1), 49-63.

[7] Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: teacher assessment identity. Assessment in Education: Principles, Policy & Practice, 25(5), 442-467.

[8] Stevenson, R. B., Lasen, M., Ferreira, J. A., & Davis, J. (2017). Approaches to embedding sustainability in teacher education: A synthesis of the literature. Teaching and Teacher Education, 63, 405-417.

[9] Suyunov, A. G. (2023, August). DIRECTIONS FOR IMPROVING SELF-PROFESSIONAL DEVELOPMENT COMPETENCE OF SCHOOL TEACHERS. In International Scientific and Current Research Conferences (pp. 11-14).

[10] Almutarrafi, R. H. (2025). A Proposed Framework for Enabling Science Teachers to Implement Fourth Industrial Revolution Technologies in Science Education through Structured Self-Professional Development Using MOOCs. Journal of Scientific Research in Education, 26(2), 59-134.

[11] Seraji, F., & Khodaveisi, S. (2019). Teachers' professional development through online learning environment: A phenomenological study. Interdisciplinary Journal of Virtual Learning in Medical Sciences, 10(4), 40-53.

[12] Ince Aka, E., & Tasar, M. F. (2020). Prospective Science Teachers' Views on Career Planning and Their Self-Efficacy Levels for Career Decision-Making. International Online Journal of Education and Teaching, 7(3), 960-975.

[13] Zhang, J., Yuen, M., & Chen, G. (2018). Teacher support for career development: An integrative review and research agenda. Career Development International, 23(2), 122-144.

[14] McLennan, B., McIlveen, P., & Perera, H. N. (2017). Pre-service teachers' self-efficacy mediates the relationship between career adaptability and career optimism. Teaching and Teacher Education, 63, 176-185.

[15] Richardson, P. W., & Watt, H. M. (2018). Teacher professional identity and career motivation: A lifespan perspective. Research on teacher identity: Mapping challenges and innovations, 37-48.

[16] Eryılmaz, A., & Kara, A. (2021). Inhibitors of Teachers' Career Adaptability: Burnout and Loneliness in Work Life. International Journal of Educational Researchers, 12(3), 41-51.

[17] Martínez-Moreno, J., & Petko, D. (2024). What motivates future teachers? The influence of Artificial Intelligence on student teachers' career choice. Computers and Education: Artificial Intelligence, 7, 100296.

[18] Sadeghi, Z., & Shafiepour Motlagh, F. (2024). Educational career paths of future teachers based on artificial intelligence and model presentation. Research in Teaching, 12(2), 184-209.

[19] Reza Roshanpour,Amirreza Keyghobadi,Ali Abdi & Mohammad Ehsanbakhsh. (2025). Enhancing inflation forecasting across short- and long-term horizons in IRAN: a hybrid approach integrating machine learning, deep learning, ARIMA, and optimized nonlinear grey Bernoulli model. SN Business & Economics,5(6),63-63.

[20] Dou Hong,Fengze Li,Jieming Ma,Ka Lok Man,Huiqing Wen & Prudence Wong. (2025). Temporal environment informed photovoltaic performance prediction framework with multi-spatial attention LSTM. Solar Energy,296,113550-113550.

[21] Kamil Kashif & Robert Ślepaczuk. (2025). LSTM-ARIMA as a hybrid approach in algorithmic investment strategies. Knowledge-Based Systems,320,113563-113563.