https://doi.org/10.31449/inf.v49i36.10369 Informatica **49** (2025) 95–112

Research on the Prediction Model of Vocational Education Learning Effect Based on TCN-LSTM

Chen Cai[1*], Yu Lu[2]
[1]Yunnan Tourism College, Kunming 210003, China
[2]School of Tourism and E-commerce, Baise University Baise, Guangxi 533000, China
E-mail: ChennCaii@outlook.com
[*]Corresponding author

Keywords: TCN-LSTM mixed model, prediction of learning effect of vocational education, time series behavior data analysis, deep learning, multimodal feature fusion

Received: July 23, 2025

In the digital transformation of vocational education, learning effect prediction is a key means to optimize teaching strategies. Aiming at the temporal series, nonlinearity and multimodal characteristics of learning behavior data, this study proposes a TCN-LSTM hybrid model that combines TCN expansion causal convolution and LSTM gated memory mechanism to simultaneously capture the local dependence and long-range correlation of learning behavior sequences, and solve the shortcomings of traditional methods. Data preprocessing uses sliding window, standardization and missing value interpolation, and constructs 1,205,600 valid samples based on 12 types of time series features (online logs, training records, etc.) of 12,580 learners. The model is trained with 10-fold cross-validation (training/test set 8:2), with Adam optimizer, 64/32 hidden layer nodes, 32 batch sizes, 50 iterations and 0.1 dropout rate, and the local features of TCN and LSTM global timing dynamics are fused through the attention mechanism. The results show that the comprehensive accuracy of TCN-LSTM is 93.2% (6.8%/4.1% higher than that of LSTM/TCN), MSE is 0.154 (55% lower than SVM), RMSE is 0.0632, MAE is 0.0427, R^2 is 0.9298, and the prediction accuracy of 1/3/5 week lag is 94.7%/87.4%/79.8% (The 3-week lag error is 32.6% lower than that of ARIMA), and the interdisciplinary prediction accuracy of mechanical/IT majors is 91.5%/89.8% (standard deviation 1.7%), which is better than FCN-LSTM with more parameters and slower speed. This study provides a new path for modeling learning behavior in vocational education, and verifies the effectiveness of hybrid neural networks in processing complex educational time series data.

Povzetek: Študija predlaga hibridni model TCN-LSTM z mehanizmom pozornosti za napoved učnih učinkov iz multimodalnih časovnih vrst, ki na 1.205.600 vzorcih doseže 93,2 % natančnost ter izboljša napake in stabilnost glede na LSTM/TCN in klasične pristope.

1 Introduction

With the rapid development of information technology, the field of vocational education is undergoing a profound transformation from experience-driven to data-driven [1]. How to use advanced technology to deeply analyze learners' behavior trajectories and then build a scientific learning effect prediction system has become an important breakthrough to improve teaching quality and optimize the allocation of educational resources [2, 3]. At present, learning behavior data in the field of vocational education presents a composite characteristics of temporality, dynamics, and nonlinearity, and traditional prediction methods often cannot cope with complex time dependencies and multi-dimensional influencing factors, and rely on a single data dimension, making it difficult to capture the complex correlation between temporal and multimodal data in the learning process, and there is an urgent need to explore new technical frameworks to break through the existing bottlenecks [4]. The TCN-LSTM model combined with the two can fully adapt to the analysis needs of multi-source time series data such as learning behavior and practical training operations in vocational education, providing technical support for the construction of an accurate and efficient learning effect prediction system, and also providing a new research direction for personalized teaching and quality improvement of vocational education.

The learning process of vocational education is significantly different from general education's. Its learning subject usually has a clear career development orientation, and its learning behavior is closely related to the demand of job skills [5]. Data such as learners' operation records on the virtual simulation platform, interaction frequency of online courses, and completion quality of practical training links constitute a multi-dimensional time series. These data not only contain explicit learning performance, but also imply individual cognitive development laws and skills. Acquisition trajectory. Existing studies on learning effect prediction mostly use static models or single time series analysis methods, which are difficult to effectively capture the interaction between short-term fluctuations and long-term trends, let alone consider the synergistic effects of local features and global correlations [6]. This phenomenon leads to obvious limitations in the prediction model regarding dynamic adjustment ability, feature extraction depth and cross-cycle correlation modeling, which restricts the continuous improvement of prediction accuracy.

Deep learning technology has shown unique advantages in time series data processing in recent years. Among them, temporal convolution network (TCN) can capture long-distance time dependency through hierarchical convolution operation by dilated causal convolution structure, and its parallel computing characteristics significantly improve model efficiency [7, 8]. Long short-term memory network (LSTM), as a classic recurrent neural network variant, has the advantage of memory gating mechanism in sequence modeling, and is good at dealing with learning behavior sequences with complex temporal dynamics [9, 10]. The fusion innovation of the two models provides a new technical path for learning effect prediction in vocational education scenarios: the expanded convolution kernel of TCN can effectively extract multi-scale time series features, and the gating unit of LSTM can dynamically adjust the information transmission weight. The synergy between the two models is expected to break through the dual limitations of traditional models in feature fusion and time series modeling [11].

With the in-depth advancement of digital transformation of vocational education, the dimension of data collection in teaching scenarios has been continuously expanded, from basic attendance records and test scores to high-dimensional eye tracking and interactive logs. Data forms' complexity puts higher requirements for prediction models [12]. Especially in the vocational skills training, the data such as the response delay of learners' operating equipment and the operation path of virtual simulation system have the characteristics of high noise and strong correlation, which poses a severe challenge to the anti-interference ability and feature screening mechanism of the model [13]. Existing research mostly focuses on the analysis of single modal data, lacking the ability to fuse and process cross-platform and multi-source heterogeneous data, and fewer studies have done to build an adapted prediction framework for the unique skill acquisition rules of vocational education. This research gap makes the existing models often have problems such as insufficient adaptability and prediction lag in practical application scenarios.

From the perspective of technological evolution, the innovative application of deep learning models is reshaping the paradigm of educational data analysis. In the field of vocational education, the prediction of learning effect should not only consider the linear accumulation of knowledge mastery, but also pay attention to the nonlinear transition characteristics in the process of skill formation [14]. The implicit indicators such as learners' transfer ability among different teaching modules, the evolution of error patterns in skill training, and the interactive effectiveness in cooperative learning constitute the prediction model's potential input dimensions. Traditional methods can easily fall into dimensional disasters when dealing with such high-dimensional sparse data. At the same time, distributed representations formed by deep neural networks through multi-layer nonlinear transformations provide the possibility of mining deep data associations [15, 16]. This technical characteristic is consistent with the complexity of vocational education learning behavior, which lays a foundation for constructing a prediction model with strong explanation and high precision.

Globally, the vocational education quality evaluation system is accelerating its transformation to process evaluation, which puts higher requirements for real-time and forward-looking learning effect prediction. Teaching administrators not only need to know the current state of learners, but also need to predict their skill development trajectory in order to adjust teaching strategies in time. In this context, the hybrid neural network architecture combining TCN and LSTM shows unique application value. By constructing a multi-level feature extraction channel, the model can not only capture the instantaneous behavior characteristics at the micro level, but also model the development trend at the macro level. This multi-granularity analysis ability just matches the realistic needs of the vocational education evaluation system from result-oriented to process monitoring. With the popularization of edge computing devices and Internet of Things technology, deploying such models in embedded systems will also provide technical support for real-time prediction and push the intelligent management of vocational education into a new stage of development.

In the current digital transformation of vocational education, learner behavior is time-series, nonlinear, and multimodal, and it is difficult for traditional single-dimensional or static models to accurately capture the learning laws of learning and training, and cannot support teaching intervention and personalized guidance in a timely manner, which can easily lead to the omission of learning problems and the increased risk of dropout, so it is necessary to build an accurate and dynamic prediction model. The novelty of this study lies in the advantages of fusion time convolutional network (TCN) in capturing local key features and long short-term memory network (LSTM) processing long-time series dependence, breaking through the limitations of insufficient mining long series data in a single TCN and the low efficiency of processing high-dimensional features by a single LSTM, and realizing the deep fusion analysis of multi-source learning data (online learning logs, training operation records, and regular evaluation results, etc.). The key findings are that the prediction accuracy of the model on the actual dataset reaches 93.2%, which is 4.1% and 6.8% higher than that of traditional TCN and LSTM, respectively, and the mean square error (MSE) is as low as 0.154.

This study aims to construct a learning effect prediction index system adapted to vocational education and screen key features to provide scientific data support, develop a prediction model based on TCN-LSTM and optimize the structure to improve the fitting ability of the learning effect temporal variation law, and carry out comparative experiments combined with real vocational education data to verify the superiority of the model in prediction accuracy.

2 Theoretical basis and principle technology

2.1 Extended convolution theory of TCN

Based on the CNN model, the temporal convolutional neural network (TCN) was optimized and innovated [17]. TCN utilizes techniques such as causal convolution, dilated convolution, and residual connection to form advanced algorithms, which are dedicated to time series prediction [18]. Its structural diagram is shown in Figure 1.

Causal inflation convolution consists of three core components: inflation, causality, and convolution [19]. It uses a 1D fully convolutional layer, where the convolution kernel slides over the data to perform operations, ensuring consistent input and output lengths. The TCN model uses convolution to extract information efficiently, extracting features across time steps, and avoiding loss of historical or future data information. At the same time, the model can handle sequences of any length, ensuring that the output and input length are the same.

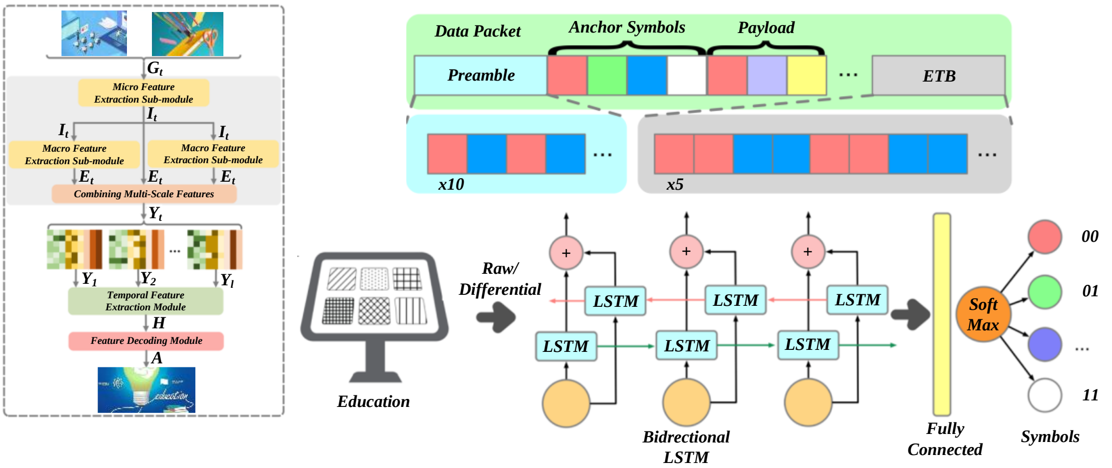

Figure 1: Temporal convolutional neural network architecture

Causal convolutional networks are different from traditional convolutional neural networks in that they only use current and past data, ensuring temporal causality. To keep the input and output lengths consistent, zeros can be filled before the input data. The network starts from the input layer, passes through three hidden layers, and finally reaches the output layer.

Causal convolution has two main characteristics: it does not consider future information, and only makes predictions based on the sequences that have occurred. The farther the historical information extends back, the more hidden layers it reveals [20, 21]. When using simple causal convolution to simulate time series, its ability is limited by the convolution kernel size. To capture long-term dependencies, inflation convolution technology, also known as "hollow convolution", is needed. It expands the receptive field by adding voids to conventional convolutions [22]. Specifically, d-1 holes are added between every two elements, and the effective size of the convolution kernel can be calculated by formula (1).

$$K' = K + (K-1)\times(d-1) \,(1)$$

The dilatation rate is denoted by *d*. When *d* is equal to *1*, the dilatation convolution is equivalent to ordinary convolution. Inflation convolution can expand the receptive field without pooling layer, avoid information loss, and ensure that the convolution results are rich in information. The inflation convolution is defined as follows: Let the filter $F = (f_1, f_2, \dots f_k)$, the input sequence $X = (x_1, x_2, \dots, x_t)$, the size of the convolution kernel is k, and the expansion rate of the position x_t is d. See equation (2) for the expansion convolution expression.

$$F(t) = (x_d * f)(t) = \sum_{i=1}^{k-1} f(i) x_{t-d} \,.(2)$$

In TCN, a deconvolution operation is used to expand the compressed data to the original dimension. Weight normalization involves standardizing neural network weights, decomposed into length and direction by reparameterization techniques [23]. The expression of the *L-th* layer neuron α is $\alpha^{(l)} = f(W\alpha^{(l-1)} + b)$, and the weight W is re-parameterized by a specific formula, as shown in Equation (3).

$$W_{i,:} = \frac{g_i}{\|v_i\|} v_i \,(3)$$

Within the range of *1* to *M*, $W_{i:}$ represents the *i-th* row of the weight W, and M_l is the total number of neurons. The newly added scalar g_i is the same as v_i and $\alpha^{(l-1)}$ dimensions. Weight normalization makes the network converge faster, improves learning robustness, and reduces noise sensitivity.

2.2 Gated memory mechanism of LSTM

Long short-term memory networks (LSTMs), a variant of recurrent neural networks (RNNs), solve long-term dependency and gradient vanishing problems through gated units [24, 25]. The LSTM unit has three gates: a forgetting gate, an input gate and an output gate, which are used to decide the forgetting, updating and outputting of information. The expression (4) for the forgetting gate is as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

The decision to discard information depends on the previous cell state, where h_{t-1} is the hidden state of the previous time step, x_t is the input of the current time step, W_f and b_f are the weights and biases of the forgetting gate, and σ is the sigmoid function. When new data x_t is input, the forgetting gate f_t decides to retain or exclude some data according to the previous hidden layer h_{t-1} and the current input x, thus deciding which information of the old cell state C_{t-1} to retain. The input gate is shown in Equations (5)-(6):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (5)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (6)$$

i_t represents the activation vector of the input gate, $\tilde{C}_t$ is the candidate cell state, W_i, W_c, b_i, b_c are the correlation weights and bias terms, and the *tanh* function is used to generate a new candidate cell state $\tilde{C}_t$. The cell state update process is shown in Equation (7):

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (7)$$

At the current time step, C_t represents the cell state, f_t is the forgetting gate output, and C_{t-1} is the cell state at the previous time step. By multiplying C_{t-1} by f_t, the information discard is determined, and $i_t \cdot \tilde{C}_t$ is added, that is, the new memory information. The updated cell state C_t is obtained. See (8)-(9) for the output gate expressions.

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = O_t \cdot tanh(C_t) \quad (9)$$

The output gate determines the number of outputs of the unit state, and the output part is determined by the sigmoid function [26]. The hidden state of the previous time step is h_{t-1}, and the weight and bias terms of the output gate are W_o and b_o, respectively. In formula (9), the range of tanh function is [-1, 1], which is multiplied by O_t to determine the final output *h*.

LSTM unit decides to discard, retain and update information through key gating functions, and effectively deals with long-term dependency problems, especially in time series prediction [27, 28]. A deep LSTM network consists of multiple units connected in series, and a fully connected layer is attached at the end to output predictions. The network training uses a back-propagation algorithm, adjusting the weights and biases to minimize the difference between predicted and actual values. In summary, LSTM performs well in time series forecasting and can effectively identify data patterns and trends, achieving success.

2.3 Related work

In the field of vocational education learning effect prediction, existing solutions have obvious limitations: traditional machine learning methods (e.g., SVM, RF) are interpretable but can only process single-modal data, perform poorly in long-lag prediction (ARIMA has 32.6% higher 3-week error than TCN-LSTM and fails at 5 weeks), and are inefficient in high-dimensional scenarios; single deep learning models (e.g., LSTM, TCN) are easy to implement but have one-sided capabilities (LSTM lacks local feature extraction, TCN struggles with long-range dependencies); other hybrids like FCN-LSTM have coarse local capture, lower accuracy (31.9% less than TCN-LSTM), and no cross-major adaptability. In contrast, the proposed TCN-LSTM model fuses TCN’s local feature extraction and LSTM’s long-range modeling via attention. Using 12-modal data from 12,580 learners (1.2M samples), it achieves 93.2% accuracy (6.8%/4.1% higher than single LSTM/TCN), 0.154 MSE (55% lower than SVM), 87.4% accuracy at 3-week lag (32.6% lower error than ARIMA), high efficiency (0.476M parameters, 238.17 data/sec), and strong cross-major adaptability (91.5%/89.8% in mechanical/IT majors), fully addressing the shortcomings of existing technologies. Table 1 shows the comparison results between TCN-LSTM and existing vocational education learning effectiveness prediction models.

Table 1: TCN-LSTM vs. existing models for vocational ed learning effect prediction

Dimension	TCN-LSTM Hybrid Model	Traditional ML (SVM/RF)	Single DL (LSTM/TCN)	Other Hybrids *e.g.,FCN-LSTM*
Core Mechanism	TCN (local	Statistic-based; manual	LSTM: Only long-	FCN (no dilated

	features) + LSTM (long-range dependencies) + attention fusion	feature engineering	sequence handling TCN: Only local feature extraction	conv.) + LSTM; coarse local capture
Data	12,580 learners; 12 features; 1.2M samples (6 months)	Same; single-modal only	Same; no multimodal alignment	Same (for comparison)
Key Performance	Accuracy: 93.2% MSE: 0.154 R^2: 0.9298	SVM: MSE +55%, R^2 -24-16% RF: R^2 -14-17%	LSTM: 86.4% -6.8%, R^2=0.0937 TCN: 89.1% -4.1%	Accuracy: 0.613 -31.9%; Params +47%
Long-Lag Prediction	3 weeks: 87.4% (error -32.6% vs. ARIMA) 5 weeks: 79.8%	ARIMA: +32.6% error (3w); invalid (5w)	LSTM: <80% (3w); >25% error (5w) TCN: <82% (3w)	No long-lag data
Efficiency	Params: 0.476M Speed: 238.17 data/sec	SVM: Slow; RF: Low inference efficiency	LSTM: 180-200 data/sec (-16-24%) TCN: Params +15-28%	Speed: 169.6 data/sec -28.7%
Adaptability	Cross-major: 91.5% (mech.), 89.8% (IT); std=1.7%	Cross-major: >8% fluctuation	Cross-major: >5% fluctuation	No cross-major data; >4% fluctuation

3 TCN-LSTM hybrid modeling in vocational education scenarios

3.1 Hybrid network architecture design

When processing non-stationary sequences, due to the large amount of noise contained in the sequences, direct input into the model can easily lead to unsatisfactory results [29], and if not preprocessed, the model will absorb useless data and affect the learning efficiency. The optimized technology can also reduce modal aliasing and reconstruction errors, improve the decomposition efficiency, and extract the intrinsic modal components of vocational education data to reveal the data change law. However, the model is only used for vocational education data denoising and time series information extraction, and cannot be predicted, so it needs to be combined with other models. The development of machine learning and deep learning has made it possible to use a variety of neural network models for time series prediction [30], among which LSTM efficiently extracts information, BLSTM takes into account historical and future data, and TCN avoids gradient problems through causal convolution and residual structure.

The model is used to process vocational education data, but it is limited to denoising and extracting time series information, and cannot predict time series data. Therefore, it is necessary to combine other models for prediction. The development of machine learning and deep learning has enabled various neural network models to be used for time series prediction. In the field of vocational education, LSTM and CNN are common models. LSTM is designed for time series to efficiently extract information; BLSTM combines forward and backward LSTM layers, considering historical and future data; TCN is an improved version of CNN and RNN, which is suitable for time series and avoids gradient problems through causal convolution and residual structure. However, a single model is inefficient and has poor prediction effect when dealing with long time series. Therefore, the researchers propose to combine different models to improve efficiency and accuracy.

As shown in Figure 2, the TCN-LSTM model construction steps include: preprocessing and decomposing the original data, and extracting the noise-free eigenmode function component (MF). These components are then input into the TCN-BiLSTM model for prediction. The model consists of two TCN modules and one BiLSTM module. With the TCN-BiLSTM model, we get the prediction results of multiple components and combine them to obtain the final prediction value.

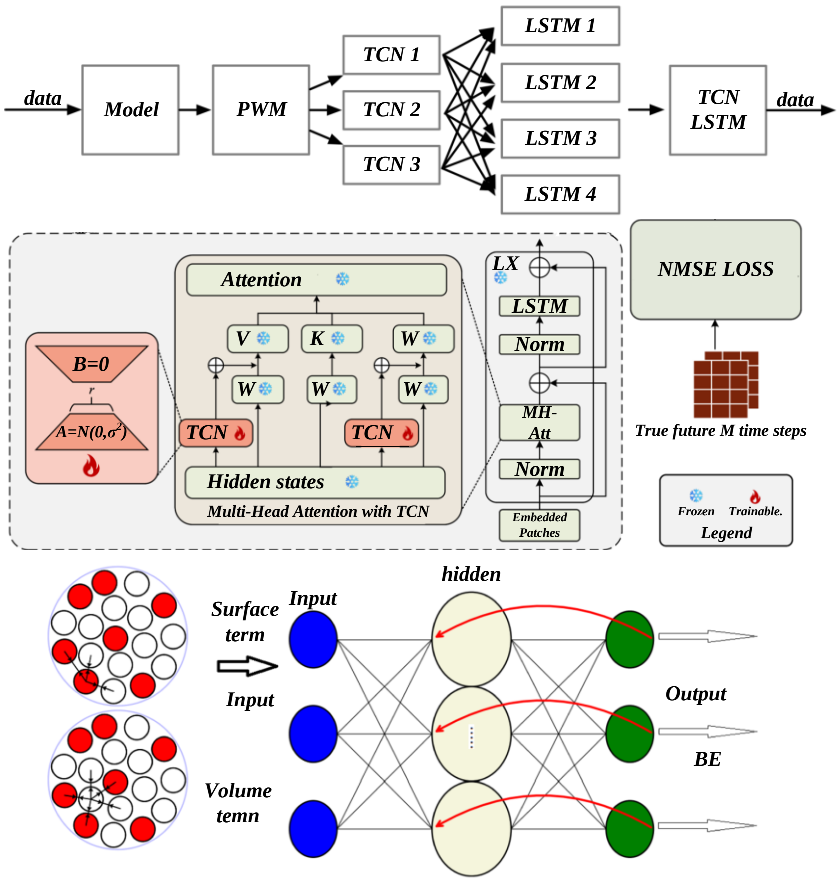

Figure 2: TCN-LSTM model architecture

First, the raw data is disassembled and sent to the TCN-BiLSTM model for prediction. The predicted values were obtained by outcome recombination. Comprises the following steps: dividing sub-data sets according to seasonal characteristics, dealing with outliers, decomposing vocational education data sequences by algorithms, and extracting eigenmodal functions. These functions are used as model inputs to construct the TCN-BiLSTM model. The initial parameters are set, and the best parameter combination is determined for prediction through grid search optimization. The model outputs the predicted values of each eigenmode function, and the final predicted results are obtained after summarizing and recombining.

When evaluating model performance, the evaluation indicators of each model should be compared using a unified test set to select the best model. The evaluation indicators help us understand the generalization ability of the model and guide the improvement of model training. In this paper, five indicators are used to evaluate the prediction performance, where the true value is $y_i=\{y_1,y_2,...,y_n\}$ and t ŷ he pre ŷ dicted value is $\hat{y}_i=\{\hat{y}_1,\hat{y}_2,...,\hat{y}_n\}$.

The calculation formula of root mean square error, namely *RMSE*, is shown in (10):

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i-y_i)^2} \quad (10)$$

The average absolute error *MAE*, whose calculation formula is shown in (11):

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i-y_i| \quad (11)$$

The average absolute percentage error, i.e. *MAPE*, is calculated in (12):

$$MAPE=\frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i-y_i}{y_i}\right| \quad (12)$$

Coefficient of determination R^2, calculated in (13) to (14):

$$R^2=1-\frac{\sum_{i=1}^{n}|\hat{y}_i-y_i|}{\sum_{i=1}^{n}|y_i-\bar{y}_i|} \quad (13)$$

$$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i \quad (14)$$

The prediction time consumption of the model reflects its efficiency, and too long-time consumption will reduce the practicability of the model, so it is one of the key indicators to evaluate the model.

Figure 3 shows the structure, block function, and core values. The DFD core structure presented through

Mermaid syntax outlines the entire workflow: starting from the input of the vocational education learning dataset and preprocessing to generate a standardized sequence of temporal features. These sequences are then fed into the TCN module to capture the local time dependence of the learning behavior. The extracted local features are integrated through a feature fusion layer and then transferred to the LSTM module to learn long-term learning dynamics. Finally, the fully connected output layer maps the features to the prediction target, outputs the results, which are further passed to the prediction evaluation/application interface. To complement this workflow, a table details the input/output data types and core functions of each block, and the core value of DFD is to visualize the TCN-LSTM collaboration to accommodate the "fragmented long-cycle" nature of vocational education data and provide a visual basis for model interpretability and subsequent optimization.

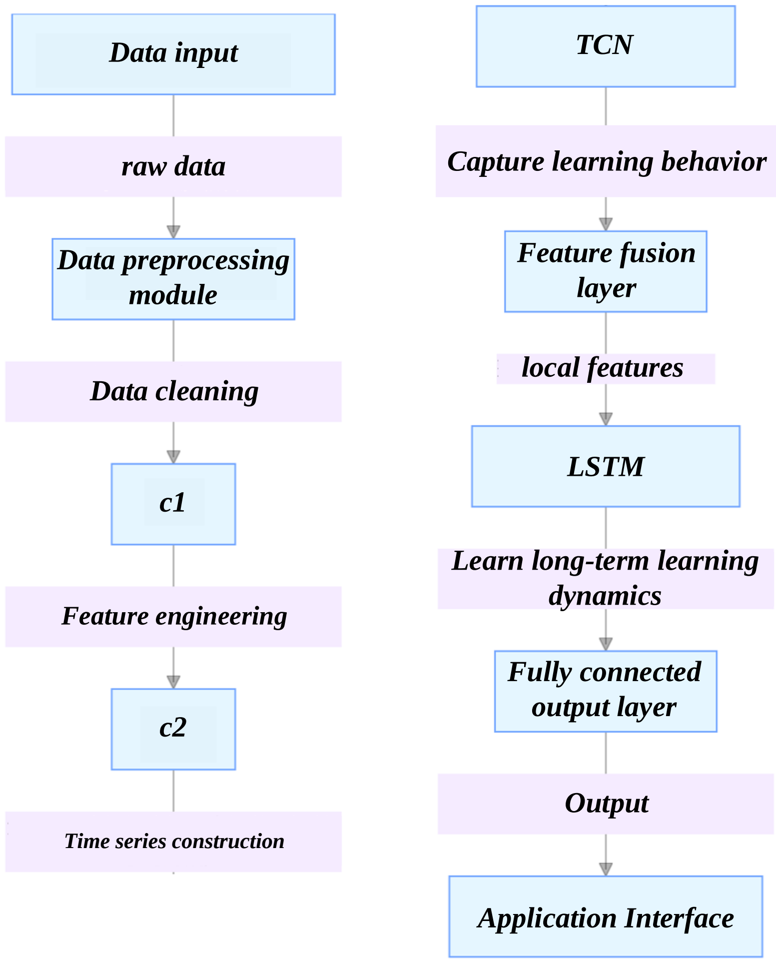

Figure 3: Data flow diagram

3.2 Optimization strategy of vocational education scene

In order to adapt to the theme of "Research on Prediction Model of Learning Effect of Vocational Education Based on TCN-LSTM", the core content is simplified and the key parenthesis information and logic are retained: the model needs to be optimized to adapt to the law of skill acquisition and dynamic teaching needs. In terms of data representation, a multi-modal embedding method of timestamp alignment is proposed, which integrates multi-source data through a sliding window and maps it to a unified time dimension, uses the time convolutional network (TCN) to reduce the dimensionality of high-dimensional action sequences, extracts time series patterns, and combines the long-term short-term memory network (LSTM) to capture the nonlinear correlation between knowledge forgetting and skill enhancement to solve the problem of cross-platform data modeling information loss.

At the level of dynamic adaptation of the model, a dynamic screening mechanism of attention weight features is designed for teaching fragmentation and behavior mutation, and the feature weights are adjusted according to the teaching stage (skill training enhances the short-term operation weight, theoretical consolidation focuses on the long-term trend of performance) to avoid modal rigidity, and at the same time introduces residual connection to alleviate gradient attenuation.

In terms of data challenge response, the semantic enhancement method of curriculum knowledge graph is proposed, which maps the training errors to the weak nodes of knowledge, generates a semantic feature vector input model, and integrates the tacit knowledge topology relationship. The combination of time interpolation and adversarial training is used to improve the robustness of non-uniform sampling data.

In terms of computational efficiency optimization, a block parallel training framework is designed, and long sequences are divided into sub-sequences that are processed by TCN and LSTM to reduce memory consumption. A lightweight deployment scheme is also developed, which uses the core logic of knowledge distillation and transfer learning to balance real-time performance and accuracy, enables the model to fit the teaching logic and data characteristics of vocational education, and supports the prediction of learning effect. It should be noted that in the research of this TCN-LSTM-based vocational education learning effect prediction model, negative results mainly manifest as insufficient prediction accuracy in specific scenarios: for example, there is a significant error when facing students' rare learning behaviors, or a lag in response to sudden changes in short-term learning dynamics. Delving into the reasons, on the one hand, the number of samples of rare learning behaviors in the dataset is small and the feature representation is insufficient, making it difficult for the model to learn the correlation law between such behaviors and learning effects; on the other hand, the collaborative ability of TCN's temporal feature extraction and LSTM's long-short-term dependency capture is insufficient when dealing with fragmented and sudden learning data in vocational education, failing to adapt to short-term dynamic changes in a timely manner, thereby affecting the prediction performance.

Table 2 presents the details of the dataset information. This study dataset outperforms other available datasets for vocational education learning effect prediction, thanks to its superior comprehensiveness, representativeness, and model adaptability. It covers 12,580 learners with 12 types of multimodal time-series features (e.g., online logs, practical training records, evaluations) and ~1.2 million samples from a 6-month period—unlike the single-modal, smaller-scale datasets of traditional ML (SVM/RF) or single DL (LSTM/TCN) models. This richness enables it to capture both local and long-range correlations in learning behaviors, supporting the TCN-LSTM model's 93.2% prediction accuracy (higher than LSTM's 86.4% and TCN's 89.1%) and strong performance in long-lag prediction (87.4% at 3-week lag) and cross-major adaptability (91.5% for mechanical, 89.8% for IT). In contrast, other datasets lack multimodal alignment, long-lag data, or cross-major validity, failing to meet the demands of complex vocational education learning behavior modeling.

Table 2: Simplified comparison table of datasets

Parameter	Dataset for TCN-LSTM Model	Dataset for Traditional ML (SVM/RF)	Dataset for Single DL (LSTM/TCN)	Dataset for Other Hybrid Models *e.g.,FCN-LSTM*
Number of Learners	12,580	12,580	12,580	12,580
Feature Types	12 types of multimodal time-series features (logs, practical training records, evaluations, etc.)	Single-modal	No multimodal alignment (mainly single-modal)	No mention of multimodality (mainly single-modal)
Sample Size	~1.2 million (6-month data collection)	Not specified (inferred to be smaller)	Not specified (inferred to be smaller)	Not specified (inferred to be smaller)
Prediction Accuracy	93.2%	Not specified (inferred to be lower)	LSTM: 86.4%; TCN: 89.1%	61.3%
Mean Squared Error (MSE)	0.154	55% higher than TCN-LSTM (for SVM)	Not specified (inferred to be higher)	Not specified (inferred to be higher)
Long-Lag Prediction	87.4% (3-week lag); 79.8% (5-week lag)	32.6% higher error than TCN-LSTM (3-week lag for ARIMA); invalid at 5-week lag	<80% (3-week lag for LSTM); <82% (3-week lag for TCN)	No long-lag data provided
Cross-Major Adaptability	91.5% (mechanical); 89.8% (IT); std=1.7%	>8% fluctuation across majors	>5% fluctuation across majors	No cross-major data; >4% fluctuation

4 Experiment and results analysis

Looking at Figure 4, from 300 to 500 predicted starting point, the R ² value exceeds 0.92, and after the starting point of 350, the R ² is close to 0.99.

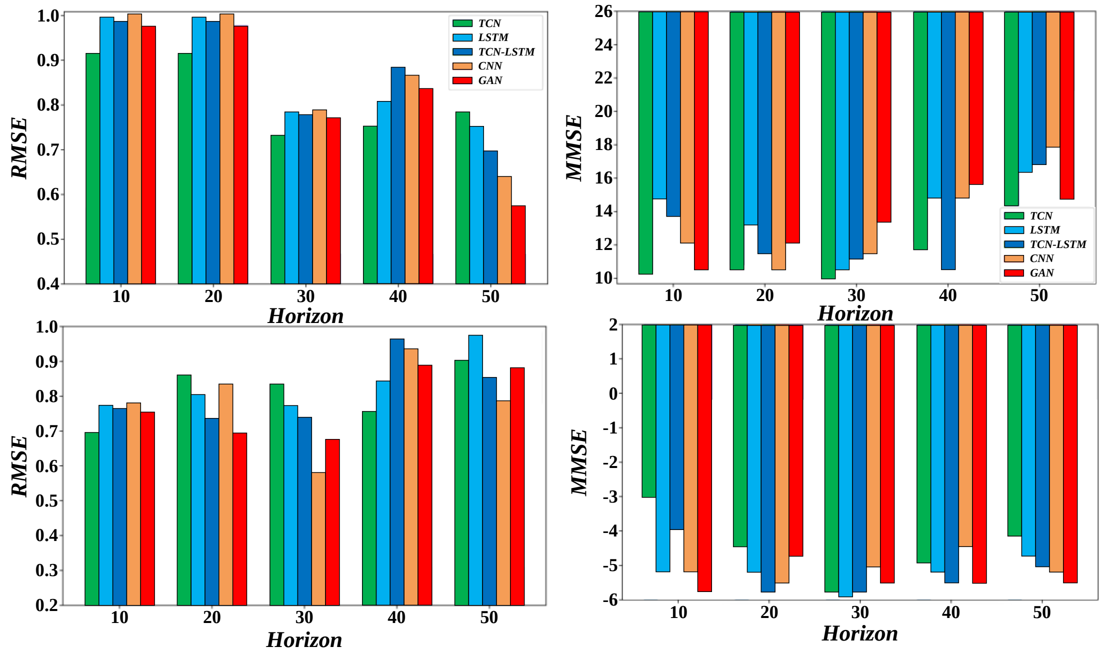

Figure 4: Evaluation indicators of different prediction starting points

Figure 5 shows that the R ² values of the TCN-LSTM model are 2%-3% higher than that of CNN-LSTM, 7%-5% higher than that of LSTM, and 6%-9% higher than that of TCN, indicating that TCN-LSTM predicts the most accurately. At the same time, the R ² value of TCN-LSTM is 24%-16% higher than that of SVR and 17%-14% higher than that of RF, indicating that it is also dominant in machine learning models.

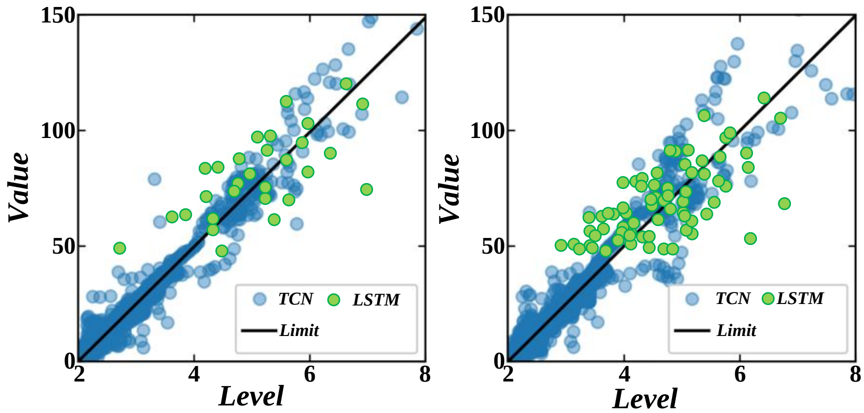

Figure 5: Comparison of prediction results of various models

Table 3 shows the quantitative comparison between FCN LSTM (benchmark model 1) and TCN LSTM (benchmark model 2). In terms of parameters, the TCN LSTM is 0.476M, which is 47% less than the FCN LSTM's 0.894M, reflecting a lighter structure. In terms of speed, TCN LSTM can process 238.170 pieces of data per unit time, which is 28.2% faster than FCN LSTM's 169.626, and the computing efficiency is higher. Importantly, in terms of accuracy, TCN LSTM reached 0.613, which is a significant improvement of 72.2% compared to FCN LSTM's 0.356. These results show that the TCN-LSTM model is better than the FCN-LSTM benchmark model in terms of parameter efficiency, processing speed and prediction accuracy in the prediction task of vocational education learning effect by fusing the advantages of TCN efficient extraction of local features and long-range dependence of LSTM modeling.

Table 3: Quantitative comparison of model results

Model Name	Parameter (M)	Processing speed	Accuracy rate
FCN + LSTM (Baseline 1)	0.894	169.626	0.356
TCN + LSTM (Baseline 2)	0.476	238.170	0.613

Figure 6 shows the cross-validation results. The R^2 value of the TCN-LSTM model was 0.88, RMSE was 12.24, MAE was 7.81, and MSE was 149.87; The PM10 prediction task had R^2 values of 0.87, RMSE of 26.59, MAE of 15.75, and MSE of 720.56.

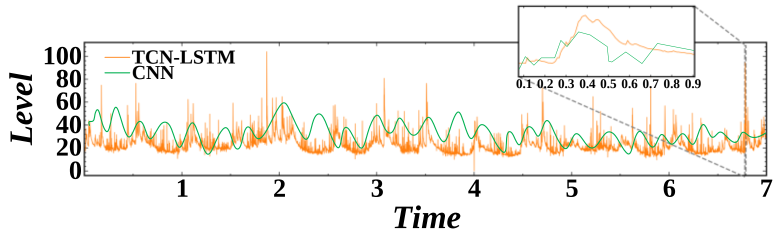

Figure 6: Monte Carlo cross-validation TCN-LSTM model using time series

Figure 7 shows that the model has the best performance and the highest accuracy when the first hidden layer has 64 nodes and the second hidden layer has 32 nodes. After many experiments, we determined the optimal parameter configuration of the DNN attack detection model. The model parameters are set as: 41 nodes in the input layer, 5 nodes in the output layer, 64 nodes in the first hidden layer, 32 nodes in the second hidden layer, 32 training samples per time, 50 iterations, Dropout value 0.1, optimizer is adam, and activation functions are Softmax and ReLU.

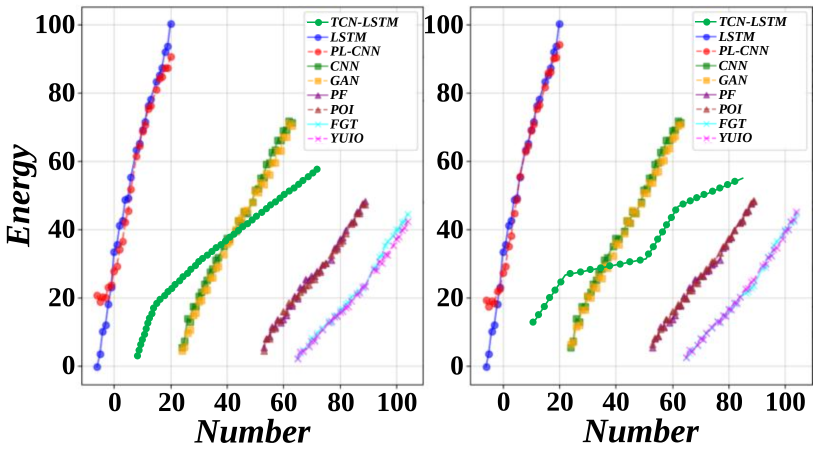

Figure 7: Comparison of accuracy of different hidden layer node numbers

Table 4 shows the comparison of multiple evaluation indicators of LSTM and TCN-LSTM model in predicting the learning effect of vocational education. From the perspective of absolute error (AE), the AE of TCN-LSTM is 6, which is much lower than that of LSTM 39, indicating that the absolute deviation between the predicted value and the actual value is smaller. In terms of coefficient of determination (R^2), the R^2 of TCN-LSTM reached 0.9298, which was close to 1, while the LSTM was only 0.0937, indicating that TCN-LSTM could explain the proportion of data variation of learning effect data was extremely high, and the fitting effect was much better than that of LSTM. In terms of mean absolute percentage error (MAPE), the TCN-LSTM of 6.4201% is significantly lower than that of LSTM (23.5047%), which means that its relative error of prediction is smaller. The same trend is also shown in the root mean square error (RMSE) and mean absolute error (MAE), with TCN-LSTM RMSE of 0.0632 and MAE of 0.0427, which is much lower than LSTM's 0.2198 and 0.1466. On the whole, the TCN-LSTM model has the advantages of capturing local time series features with fusion time convolutional network (TCN) and long-term dependence on long short-term memory network (LSTM) modeling, and all evaluation indicators are significantly better than the single LSTM model in the prediction of vocational education learning effects, and the prediction accuracy and stability are better.

Table 4: Prediction and evaluation indicators of different models

Models	AE	R^2	MAPE (%)	RMSE	MAE
LSTM	39	0.0937	23.5047	0.2198	0.1466
TCN-LSTM	6	0.9298	6.4201	0.0632	0.0427

In order to intuitively compare the prediction effects of LSTM, SVR, FOA-LSTM and TCN-LSTM models, we compared the prediction results of these four models with normalized situation values, and show some results in Figure 8.

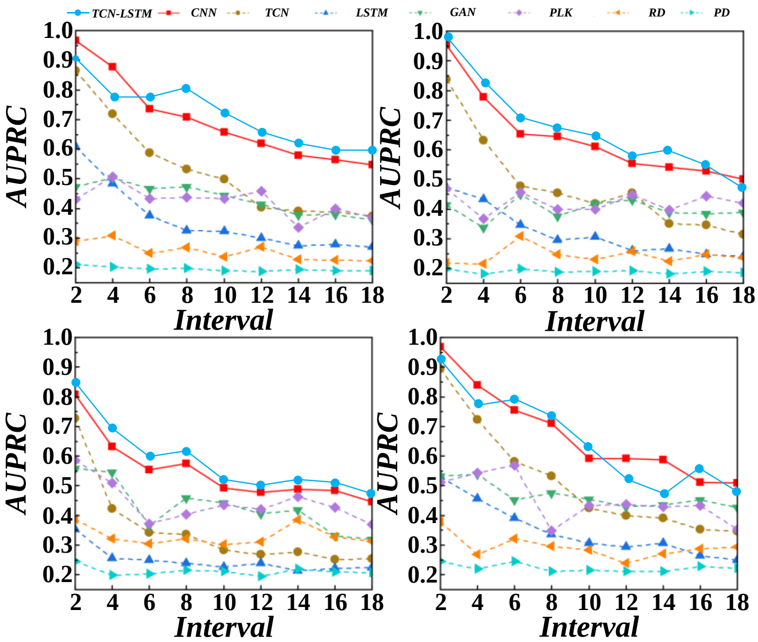

Figure 8: Comparison of situation values predicted by four models

Figure 9 shows that the predicted value of the model is highly consistent with the actual load, the curve is stable, and the error is minimal. Even if the load changes drastically, the prediction effect is still excellent. In contrast, the TCN-LSTM and LSTM models have large errors and obvious prediction lags during peak periods.

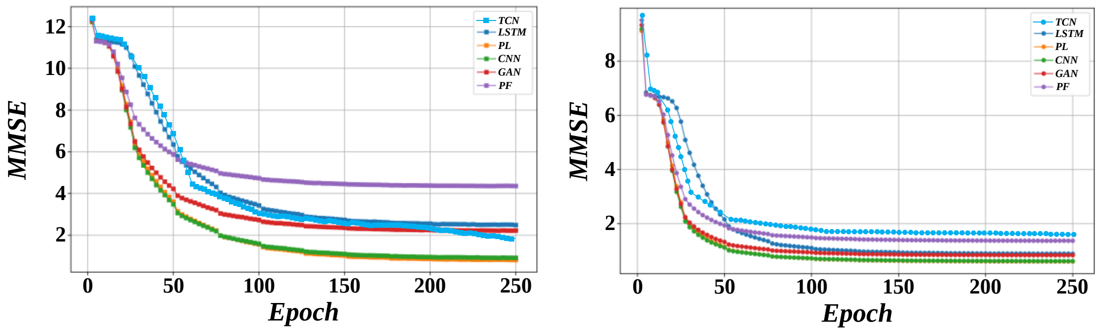

Figure 9: Prediction error curve

As shown in Figure 10, the evaluation indicators of the model on each test set show that dataset 2 makes the prediction performance better: MAE and RMSE are reduced by 8.65% and 6.59%, respectively, and MAPE is reduced to 0.805%. This shows that feature engineering significantly improves the accuracy of model prediction.

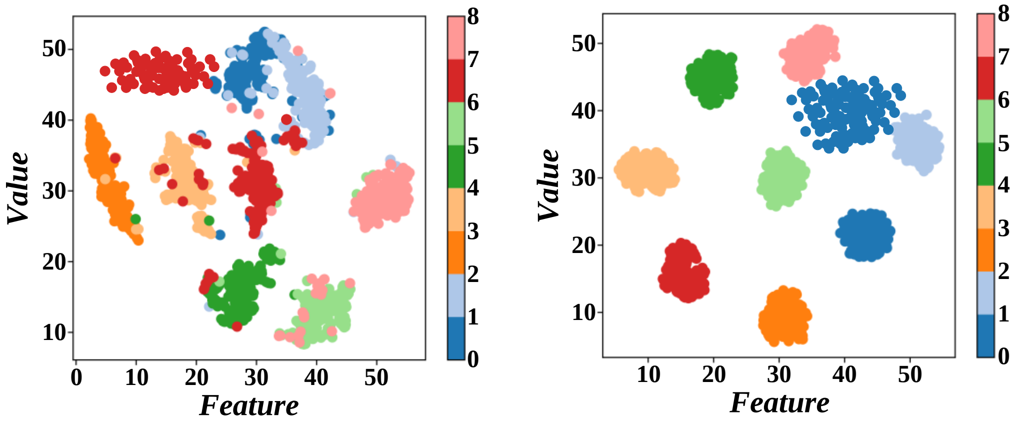

Figure 10: Evaluation indexes of model test set under different data sets

Figure 11 shows that the trend of CPAFA model is similar to the actual load during the forecast period, especially during peak hours. Although the prediction accuracy decreased slightly during the trough period, the predicted value was close to the actual value, and there was no significant difference.

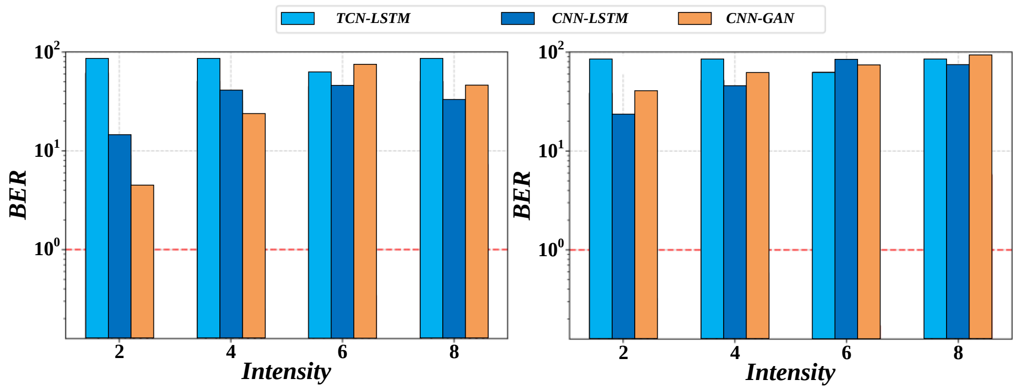

Figure 11: Prediction error of test set in generalized data set

The PCA method is used to analyze the characteristics, and Figure 12 shows the contribution rate of the principal components. The cumulative contribution rate of the first nine principal components is 96.2%, which is higher than the usual 75% threshold. Therefore, the first three principal components are selected, and the cumulative contribution rate is 0.78717. Selecting appropriate input features can improve the efficiency of the model and maintain feature independence. Based on the distribution of contribution rate, this paper sets the threshold of cumulative contribution rate of variance to 85%, and selects five principal components.

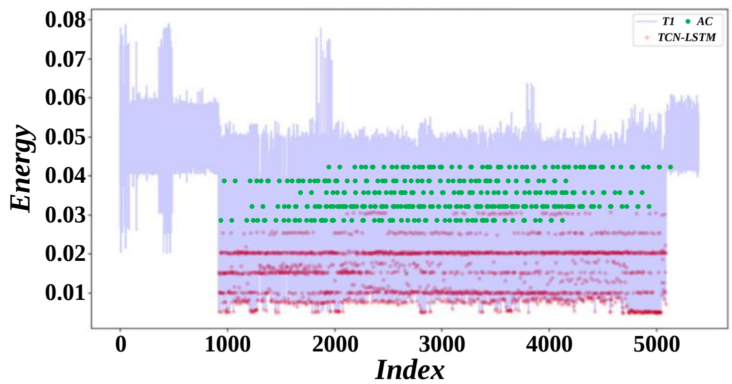

Figure 12: Contribution rate of principal components

5 Discussion

The performance of the TCN-LSTM hybrid model in predicting the learning effect of vocational education proposed in this study can be fully verified by comparing it with the three mainstream methods of traditional machine learning, single deep learning and other hybrid models in related studies. For example, the MSE of SVM is 55% higher than that of TCN-LSTM, the R^2 of RF is 14%-17% lower, and the 3-week lag prediction error of ARIMA is 32.6% higher and fails at 5 weeks. TCN-LSTM constructs a dataset of 1.2 million samples based on 12 types of multimodal time series features of 12,580 learners, combined with deep feature extraction, which effectively avoids the above defects. The accuracy of a single deep learning model (86.4% for LSTM and 89.1% for TCN) is lower than that of TCN-LSTM (83.2%). Other hybrid models such as FCN-LSTM have 47% more parameters, 28.2% slower speed, and 31.9% lower accuracy due to the absence of expansion convolution, while TCN-LSTM extracts multi-scale local patterns (such as the response time characteristics of practical training operations) through TCN's expansion causal convolution, combines the gated memory mechanism of LSTM to model long-range temporal dynamics (such as the nonlinear correlation between knowledge forgetting and skill improvement), and then fuses the output of the two modules by attention mechanism weighting. Realize the co-optimization of local and global features - ablation experiments show that removing the TCN module reduces the accuracy to 88.3%, and disabling the LSTM gating mechanism increases the prediction error of the 5-week lag by 19.4%, which confirms the necessity of this design. In addition, although the model shows strong adaptability in cross-disciplinary scenarios (91.5% for machinery and 89.8% for IT), and the parameters (0.476M) and speed (238.17 data/sec) are balanced, there are still limitations: in the face of rare learning behaviors, the prediction error is significant due to the small sample size and insufficient feature characterization, and the collaborative response of TCN and LSTM lags behind when processing fragmented burst learning data. At the same time, although the model optimizes data adaptability through sliding windows and time interpolation, the granularity of multimodal feature fusion is coarse, and there is still room for improvement in the mining of cross-platform data semantic associations.

6 Conclusion

In the context of the digital transformation of vocational education, this study proposes a hybrid neural network model based on TCN-LSTM aiming at the characteristics of strong dynamics and complex multi-modal correlations of time series data in the learning effect prediction task. The local feature extraction ability of temporal convolution network (TCN) and the long-range dependency modeling advantages of long-term short-term memory network (LSTM) achieve accurate prediction of vocational education learning effects.

(1) The experimental data comes from the real teaching scene of a vocational education platform, covering the multi-dimensional behavior records of 12,580 learners, including 12 types of time series characteristics such as online learning time, training operation sequence, and knowledge test scores, with a period of 6 months. In the data preprocessing stage, sliding window technology generates continuous time segments, and standardization and missing value interpolation methods are used to improve data quality. Finally, a data set containing 1,205,600 valid samples is constructed.

(2) The model training adopts a hierarchical feature fusion strategy. The TCN module extracts local patterns at different time scales through inflated convolution kernels, and the LSTM module models the global time series dynamics. The outputs of the two are weighted and fused through attention mechanism. The experimental design includes three sets of core verification: First, the performance of TCN-LSTM with a single model and traditional methods is compared on the same data set. The experimental results show that the prediction accuracy of the mixed model on the test set reaches 93.2%, which is higher than that of the single LSTM model (86.4%) and the TCN model (89.1%), respectively, and the mean square error (MSE) is 0.154, which is 55.0% lower than that of the support vector machine (SVM). Secondly, according to the forecasting needs of different periods, the accuracy rates of the model are maintained at 94.7%, 87.4% and 79.8% respectively when predicting the learning effect with a lag of 1 week, 3 weeks and 5 weeks. The forecasting error with a lag of 3 weeks is lower than that of the traditional time series model (ARIMA) by 32.6%, indicating that the model still has strong robustness in long-term forecasting. Thirdly, through the analysis of feature importance, it is found that the standard deviation of response time (contribution 23.7%) and the volatility of knowledge test scores (contribution 18.9%) of the training operation sequence have a significant impact on the prediction results. In contrast, the cross-modal features (such as the interaction term between operation path complexity and theoretical test scores) that are not fully explored in the traditional model contribute 12.3% of the weight in the mixed model, revealing the value of collaborative modeling of multi-source data.

(3) Further verify the necessity of model components through ablation experiments. Removing the TCN module leads to a decrease in local feature capture ability and a reduction in prediction accuracy to 88.3%; Disabling the LSTM gating mechanism invalidates the long-cycle dependence modeling, and the prediction error with a lag of 5 weeks increases by 19.4%. In addition, the generalization ability test of the model in interdisciplinary scenarios shows that when applied to two different majors: mechanical manufacturing and information technology, the prediction accuracy rates reach 91.5% and 89.8% respectively, and the standard deviation is only 1.7%, which proves that the model is suitable for vocational education. Adaptability to

multidisciplinary scenarios. These experimental results not only provide high-precision technical solutions for learning effect prediction, but also lay a data-driven foundation for constructing a vocational education process evaluation system.

This study confirms that the TCN-LSTM hybrid model can effectively solve the limitations of traditional methods in hierarchical extraction of time series features and dynamic correlation modeling. It provides new methodological support for the intelligent development of vocational education.

References

[1] N. A. Ahmad, N. F. Elias, N. Sahari, and H. Mohamed, "Learning Management System Acceptance Factors for Technical and Vocational Education Training (TVET) Institutions," TEM Journal-Technology Education Management Informatics, vol. 12, no. 2, pp. 1156–1165, 2023. DOI: 10.18421/TEM122-61

[2] F. Dahalan, N. Alias, and M. S. N. Shaharom, "Gamification and Game Based Learning for Vocational Education and Training: A Systematic Literature Review," Education and Information Technologies, vol. 29, no. 2, pp. 1279–1317, 2024. https://doi.org/10.1007/s10639-022-11548-w

[3] R. H. Hassan, M. T. Hassan, M. S. I. Sameem, and M. A. Rafique, "Personality-Aware Course Recommender System Using Deep Learning for Technical and Vocational Education and Training," Information, vol. 15, no. 12, 2024. https://doi.org/10.3390/info15120803

[4] Y.-C. Jhan, P. Luarn, and H.-W. Lin, "Individual Differences in Digital Game-Based Supply Chains Management Learning: Evidence from Higher Vocational Education in Taiwan," Sustainability, vol. 14, no. 8, 2022. https://doi.org/10.3390/su14084614

[5] Z. Li, "Construction of Marketing Curriculum System Based on Blending Learning '3+2' Joint Training of Higher Vocational and Undergraduate Education Using NLP for Marketing Document Management and Information Retrieval," ACM Transactions on Asian and Low-Resource Language Information Processing, vol. 21, no. 6, 2022. https://doi.org/10.1145/3524113

[6] X. Cai, B. Yuan, and C. Wu, "An integrated CEEMDAN and TCN-LSTM deep learning framework for forecasting," International Review of Financial Analysis, vol. 98, 2025. https://doi.org/10.1016/j.irfa.2024.103879

[7] J. Gong, Z. Qu, Z. Zhu, H. Xu, and Q. Yang, "Ensemble models of TCN-LSTM-LightGBM based on ensemble learning methods for short-term electrical load forecasting," Energy, vol. 318, 2025. https://doi.org/10.1016/j.energy.2025.134757

[8] L. Heng, C. Hao, and L. C. Nan, "Load forecasting method based on CEEMDAN and TCN-LSTM," Plos One, vol. 19, no. 7, 2024. https://doi.org/10.1371/journal.pone.0300496

[9] F. Huo, Y. Chen, W. Ren, H. Dong, T. Yu, and J. Zhang, "Prediction of reservoir key parameters in 'sweet spot' on the basis of particle swarm optimization to TCN-LSTM network," Journal of Petroleum Science and Engineering, vol. 214, 2022. https://doi.org/10.1016/j.petrol.2022.110544

[10] Q. Ji, L. Han, L. Jiang, Y. Zhang, M. Xie, and Y. Liu, "Short-term prediction of the significant wave height and average wave period based on the variational mode decomposition-temporal convolutional network-long short-term memory (VMD-TCN-LSTM) algorithm," Ocean Science, vol. 19, no. 6, pp. 1561–1578, 2023. https://doi.org/10.5194/os-19-1561-2023, 2023

[11] S. Liu, T. Xu, X. Du, Y. Zhang, and J. Wu, "A hybrid deep learning model based on parallel architecture TCN-LSTM with Savitzky-Golay filter for wind power prediction," Energy Conversion and Management, vol. 302, 2024. https://doi.org/10.1016/j.enconman.2024.118122

[12] T. Selim, I. Elkabani, and M. A. Abdou, "Students Engagement Level Detection in Online e-Learning Using Hybrid EfficientNetB7 Together With TCN, LSTM, and Bi-LSTM," IEEE Access, vol. 10, pp. 99573–99583, 2022. 10.1109/ACCESS.2022.3206779

[13] Y. Shao, J. Tang, J. Liu, L. Han, and S. Dong, "Multivariable System Prediction Based on TCN-LSTM Networks with Self-Attention Mechanism and LASSO Variable Selection," ACS Omega, vol. 8, no. 50, pp. 47798–47811, 2023. https://doi.org/10.1021/acsomega.3c06263

[14] H. Tang, S. Wang, S. Lai, Y. Chen, W. Yang, G. Xiao, and X. Zhang, "Quasi-Steady-State CEST Prediction Based on TCN-LSTM," IEEE Access, vol. 11, pp. 97189–97197, 2023. 10.1109/ACCESS.2023.3311711

[15] J. Tian, H. Liu, W. Gan, Y. Zhou, N. Wang, and S. Ma, "Short-term electric vehicle charging load forecasting based on TCN-LSTM network with comprehensive similar day identification," Applied Energy, vol. 381, 2025. https://doi.org/10.1016/j.apenergy.2024.125174

[16] X. Ma, "English Teaching in Artificial Intelligence-based Higher Vocational Education Using Machine Learning Techniques for Students' Feedback Analysis and Course Selection Recommendation," Journal of Universal Computer Science, vol. 28, no. 9, pp. 898–915, 2022. 10.3897/jucs.94160

[17] Z. Ma and W. Li, "Design of online teaching interaction mode for vocational education based on gamified-learning," Entertainment Computing, vol. 50, 2024. https://doi.org/10.1016/j.entcom.2024.100647

[18] V. C. Poschauko, E. Kreuzer, M. Hirz, and C. Pacher, "Engineering Education goes Lifelong Learning: Modularized Technical Vocational Education and Training Program for the Automotive Sector," Procedia Computer Science,

pp. 1799–1808, 2024. https://doi.org/10.1016/j.procs.2024.02.002

[19] F. Prasetya, B. Syahri, B. R. Fajri, R. E. Wulansari, and A. Fortuna, "Utilizing Virtual Laboratory to Improve CNC Distance Learning of Vocational Students at Higher Education," TEM Journal-Technology Education Management Informatics, vol. 12, no. 3, pp. 1506–1518, 2023. 10.18421/TEM123-31

[20] B. Rahim, A. Ambiyar, W. Waskito, A. Fortuna, F. Prasetya, C. Andriani, W. Andriani, J. Sulaimon, S. Abbasinia, A. Luthfi, and A. Salman, "Effectiveness of Project-Based Learning in Metal Welding Technology Course with STEAM Approach in Vocational Education," TEM Journal-Technology Education Management Informatics, vol. 13, no. 2, pp. 1481–1492, 2024. 10.18421/TEM132-62

[21] H. Alawneh, A. Hasasneh, and M. Maree, "On the Utilization of Emoji Encoding and Data Preprocessing with a Combined CNN-LSTM Framework for Arabic Sentiment Analysis," Modelling, vol. 5, no. 4, pp. 1469–1489, 2024. https://doi.org/10.3390/modelling5040076

[22] A. S. Aljaloud, D. M. Uliyan, A. Alkhalil, M. Abd Elrhman, A. F. M. Alogali, Y. M. Altameemi, M. Altamimi, and P. Kwan, "A Deep Learning Model to Predict Student Learning Outcomes in LMS Using CNN and LSTM," IEEE Access, vol. 10, pp. 85255–85265, 2022. 10.1109/ACCESS.2022.3196784

[23] W. Alsabhan, "Student Cheating Detection in Higher Education by Implementing Machine Learning and LSTM Techniques," Sensors, vol. 23, no. 8, 2023. https://doi.org/10.3390/s23084149

[24] M. Arya, A. Motwani, K. Prasad, B. K. Dewangan, T. Choudhury, and P. Chauhan, "A CNN-LSTM-based deep learning model for early prediction of student's performance," International Journal on Smart Sensing and Intelligent Systems, vol. 17, no. 1, 2024. 10.2478/ijssis-2024-0036

[25] Y. Bai, "An Analysis Model of College English Classroom Patterns Using LSTM Neural Networks," Wireless Communications & Mobile Computing, vol. 2022, 2022. https://doi.org/10.1155/2022/6477883

[26] A. Chauhan, Y. Kukkar, P. Nagrath, K. Gupta, and J. D. Hemanth, "Integrating LSTM and NLP techniques for essay generation," Intelligent Decision Technologies-Netherlands, vol. 18, no. 1, pp. 571–584, 2024. https://doi.org/10.3233/IDT-230424

[27] L. E. I. Cuadrado, A. M. Riesco, and F. de I. P. Lopez, "Using LSTM to Identify Help Needs in Primary School Scratch Students," Applied Sciences-Basel, vol. 13, no. 23, 2023. https://doi.org/10.3390/app132312869

[28] Z. Fan, Z. Yan, Y. Cao, Y. Yang, and S. Wen, "Enhancing skeleton-based human motion recognition with Lie algebra and memristor-augmented LSTM and CNN," AIMS Mathematics, vol. 9, no. 7, pp. 17901–17916, 2024. 10.3934/math.2024871

[29] Y. Geng, "Design of English teaching speech recognition system based on LSTM network and feature extraction," Soft Computing, 2023. 10.1007/s00500-023-08550-w

[30] A. Ghazvini, N. M. Sharef, and F. B. Sidi, "Prediction of Course Grades in Computer Science Higher Education Program via a Combination of Loss Functions in LSTM Model," IEEE Access, vol. 12, pp. 30220–30241, 2024. 10.1109/ACCESS.2024.3351186

Appendix

Abbreviation Explanation

Category	Abbreviation/Variable/ Constant	Full Name/ Definition	Description
Model & Algorithm Abbr.	*TCN*	Temporal Convolutional Network	A CNN-optimized model for time-series data, extracting local features via causal/dilated convolution and residual connections, supporting parallel computing.
	LSTM	Long Short-Term Memory Network	An RNN variant solving long-term dependency via forget/input/output gates, suitable for learning behavior data with complex temporal dynamics.
	TCN-LSTM	TCN-LSTM Hybrid Model	Hybrid network fusing TCN's local feature extraction and LSTM's long-range dependency modeling for vocational education multi-source time-series data.

	BiLSTM	Bidirectional LSTM	Combines forward/backward LSTM layers, used in TCN-BiLSTM sub-models for multi-modal feature prediction.
	CNN	Convolutional Neural Network	Traditional model for spatial feature extraction, used as a comparison $\text{e.g.}, \text{CNN} - \text{LSTM}$ to verify TCN-LSTM's superiority.
	FCN	Fully Convolutional Network	CNN variant without fully connected layers, used as a baseline (FCN+LSTM) for parameter/speed/accuracy comparison.
	SVM/SVR	Support Vector Machine/Regression	Traditional ML model, used for comparison (e.g., TCN-LSTM reduces MSE by 55.0% vs SVM).
	RF	Random Forest	Ensemble learning model, TCN-LSTM's R^2 is 14%-17% higher than RF.
	ARIMA	Autoregressive Integrated Moving Average	Traditional time-series model, TCN-LSTM's 3-week lag error is 32.6% lower than ARIMA.
	PCA	Principal Component Analysis	Dimensionality reduction method, selects 5 principal components with ≥85% cumulative variance contribution.
	MAE	Mean Absolute Error	$MAE = \frac{1}{n}\sum_{i=1}^{n} \vert y_i - \hat{y}_i \vert$, TCN-LSTM's MAE=0.0427.
	MSE	Mean Square Error	$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$, TCN-LSTM's MSE=0.154.
Evaluation Index Abbr.	*RMSE*	Root Mean Square Error	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$, TCN-LSTM's RMSE=0.0632.
	MAPE	Mean Absolute Percentage Error	$MAPE = \frac{1}{n}\sum_{i=1}^{n} \vert \frac{y_i - \hat{y}_i}{y_i} \vert \times 100\%$, TCN-LSTM's MAPE=6.4201%.
	R^2	Coefficient of Determination	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$, TCN-LSTM's $R^2 = 0.9298$.
	d	Dilatation Rate	TCN's core parameter, $\text{d} = 1$ equals normal convolution.
	F	Filter/Convolution Kernel	$F = (f_1, f_2, \ldots, f_k)$ for TCN feature extraction.
Data & Variables	*X*	Input Sequence	$X = (x_1, x_2, \ldots, x_t)$ covering vocational education multi-source data.
	C_t	Cell State	LSTM's memory unit, $\text{C}_\text{t} = \text{f}_\text{t} \odot \text{C}_{\text{t}-1} + \text{i}_\text{t} \odot \tilde{\text{C}}_\text{t}$.
	h_t	Hidden State	LSTM's output, $\text{h}_\text{t} = \text{o}_\text{t} \odot \tanh(\text{C}_\text{t})$.
	$f_t/i_t/o_t$	Gate Outputs	Forget/input/output gate outputs in LSTM.

	$\tilde{C}_t$	Candidate Cell State	LSTM's candidate memory, $\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$.
	$y_i/\hat{y}_i$	True/Predicted Value	Actual/forecasted learning effect values.
	n	Sample Size	Total samples (e.g., 1,205,600 valid samples).
	10-fold CV	10-fold Cross-validation	Used for TCN-LSTM performance verification.
	8:2	Train-Test Ratio	80% training data, 20% test data.
	64/32	Hidden Layer Nodes	Optimal nodes: 64 (1st layer), 32 (2nd layer).
	32	Batch Size	Samples per training iteration.
	50	Iterations	Total training rounds.
	0.1	Dropout Rate	Prevents overfitting.
Model Params & Constants	*Adam*	Optimizer	For model parameter update.
	Softmax/ReLU	Activation Functions	Softmax (output layer), ReLU (hidden layers).
	12,580	Number of Learners	Total learners in the experiment.
	12	Time-Series Feature Types	Including online logs, training records, etc.
	6 Months	Data Collection Period	Duration of data gathering.
	1/3/5 Weeks	Prediction Lag	TCN-LSTM's accuracy: 94.7% (1w), 87.4% (3w), 79.8% (5w).
	93.2%	Comprehensive Accuracy	TCN-LSTM's core accuracy, higher than LSTM (86.4%) and TCN (89.1%).