Journal of Computer Science and Electrical Engineering
Print ISSN: 2663-1938
Online ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3161

TEMPORAL DEEP LEARNING FOR JOB COMPETENCY DEVIATION DETECTION IN COAL CHEMICAL VOCATIONAL TRAINING

Kai Fang[1*], Siga A[2]
[1]*School of Energy and Power Engineering, Xinjiang Vocational University of Technology, Kashgar 844004, Xinjiang, China.*
[2]*College of Intelligent Manufacturing, Xinjiang Vocational University of Technology, Kashgar 844004, Xinjiang, China.*
[*]*Corresponding Author: Kai Fang*

Abstract: Conventional assessment in coal chemical vocational training relies heavily on final scores and provides limited information about how operational competency evolves throughout the training process. To address this limitation, this study proposes a temporal deep-learning approach for detecting job-competency deviations from process-level training data. A multidimensional competency framework is established to characterize standardized operation, process regulation, abnormality recognition, fault handling, and safety and emergency response. Time-stamped operating events, action intervals, equipment states, and key process variables are organized into multivariate temporal sequences. A BiLSTM-Attention model is then developed to extract dynamic behavioral representations and estimate competency states, while competency deviations are quantified by comparing the predicted states with job-oriented reference levels. The proposed model outperforms Random Forest, XGBoost, LSTM, GRU, and Transformer baselines, achieving an R2 of 0.879 and an F1-score of 0.876. Dimension-specific analysis indicates relatively stable performance in standardized operation and safety-related tasks, whereas larger gaps occur in abnormality recognition and fault handling. Attention analysis further identifies alarm occurrence, critical process adjustment, and fault-handling stages as important contributors to competency estimation. The framework demonstrates the potential of process-level temporal data for fine-grained, data-driven competency diagnosis in coal chemical vocational training.
Keywords: Coal chemical vocational training; Job competency; Temporal deep learning; Competency deviation detection

1 INTRODUCTION

The coal chemical production consists of complicated process chains, continuous processes, interunit interaction, and strict safety requirements; frontline operators should not only understand the processes but also master skills in standard operation, process variable regulation, recognizing abnormalities, troubleshooting and responding emergencies. Practical training at vocational coal chemical programs shifts from isolated skill exercise towards integrated, job-oriented production tasks. With virtual simulation, distributed control systems (DCSs) and semi-physical training platforms, trainees can repetitively perform start-up, normal operation, process adjustment, shutdown, and fault handling under safe and controlled environment. Rich temporal data like action sequence, timestamp, equipment states, process variables, alarms, and task outcomes will be generated during these simulations. Yet most assessments rely on final scores, task completion, and instructors' observation. Outcome-based evaluations fail to reflect the development of competencies over an operating task. Trained individuals who achieve same final scores could show vastly different action sequences, response times, adjustment strategy, and dealing with abnormal conditions. We need to leverage temporal training data to uncover dynamic discrepancy between their actual performance and industrial job requirement.

Competency assessment and learning analytics research gradually shift towards process-oriented and data-driven approaches. Process-industry operator training uses various indicators: completion time, operational error, hint usage, event recognition, and equipment identification to describe performance, and simulation systems provide continuous behavioral traces [3, 7-8]; knowledge tracing research treat skill mastery as a latent state which evolves sequentially [1, 5]. Recurrent neural networks, long short-term memory network, and attention mechanism became popular tools for modeling learning behavior. Deep time series anomaly detection techniques showed capability to learn complex pattern from multivariate sequences and detect local or contextual deviation [2, 4, 10]. Artificial intelligence is being increasingly adopted in vocational competency assessment yet mostly focus on general performance prediction, knowledge-state estimation, or industrial equipment anomalies [6, 9]. There is limited work focusing on competency deviations in coal chemical vocational training connecting job requirements to continuous actions, changing process condition, and response to abnormal events.

In order to fill this gap, we propose a temporal deep-learning framework for detecting job-competency deviations in coal chemical vocational training. First, construct a job-competency representation for standardized operation, process

© By the Author(s) 2026, under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).

regulation, abnormality recognition, fault handling, and safety and emergency response. Then organize time-stamped actions, operating interval, equipment states, and key process variables into multivariate temporal sequences. Use BiLSTM-Attention model to capture the nonlinear relationship between operating behavior and evolving competency states. Quantify deviations by comparing the inferred states with stage-specific, job-oriented reference states. Evaluate our framework compared to conventional machine-learning and representative sequential models using attention analysis and ablation experiment to find out critical operating stages and source of information. Our approach changes assessment from one end-point score to process-level diagnosis, providing individual skill-gap identification and targeted training design.

2 DATA AND METHODOLOGY

2.1 Coal Chemical Vocational Training Scenario

In this work we consider typical operational training scenarios in vocational education for the coal chemical industry. In general industrial coal chemical operations consist of a series of tasks such as plant start-up, steady-state operation, adjusting process variables, shutdowns, or handling abnormalities with clear procedural dependencies and intensive interaction between operating units. A full operational trajectory holds much more competency related information compared to just a final task score.

The data considered in this framework are generated from a coal chemical virtual-simulation or DCS-based training platform. One complete training task performed by a trainee is treated as an individual training sample and consists of a sequence of time-stamped operating events together with the corresponding process states. For the i-th training sample, the raw temporal sequence is represented as

$$X_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,T_i}\} \tag{1}$$

where T_i is the number of operating events within a training session; x_i,t refers to the integrated system state on the t-th operating step. The state includes operating behavior, time information, equipment status, key process variables as well as alarm or abnormality information:

$$x_{i,t} = [A_{i,t}, \Delta t_{i,t}, P_{i,t}, T_{i,t}, F_{i,t}, L_{i,t}, S_{i,t}, E_{i,t}] \tag{2}$$

where A refers to the operating action, delta t means the time gap of two successive actions, P, T, F, and L refer to pressure, temperature, flow rate, and liquid level, respectively; S indicates the equipment/system operating state; E refers to alarm/abnormality/operation-feedback info. It allows us to describe together with which action a trainee operates, at what moment it occurs, as well as how the process responds.

2.2 Job Competency Framework

Detecting competency deviations requires an explicit job-oriented reference framework. Based on typical coal chemical operating tasks, practical-training assessment criteria, and industrial process-operation requirements, job competency is represented by five complementary dimensions: standardized operation, process regulation, abnormality recognition, fault handling, and safety and emergency response. As shown in Table 1.

Table 1 Job Competency Framework for Coal Chemical Vocational Training

Competency dimension	Main meaning	Representative observable behaviors
C1 Standardized operation	Execution of procedures according to standard requirements	Action sequence, omitted or incorrect steps, repeated actions, equipment start/stop procedures
C2 Process regulation	Regulation of key process variables according to operating conditions	Adjustment of temperature, pressure, flow rate and liquid level; stabilization time
C3 Abnormality recognition	Timely identification of deviations from normal operation	Alarm recognition, detection latency and diagnostic accuracy
C4 Fault handling	Appropriate response to abnormal conditions and equipment faults	Handling sequence, response time, recovery effectiveness and handling accuracy
C5 Safety and emergency response	Compliance with safety requirements and appropriate emergency actions	Interlock awareness, avoidance of unsafe actions, emergency shutdown and response

The competency state of trainee i at training stage k is represented as a five-dimensional vector:

$$C_{i,k} = \left[c_{i,k}^{(1)}, c_{i,k}^{(2)}, c_{i,k}^{(3)}, c_{i,k}^{(4)}, c_{i,k}^{(5)}\right] \tag{3}$$

where c_i,k^(j) in [0,1] denotes the normalized proficiency level of the j-th competency dimension. A stage-specific job reference state is further constructed from standard operating procedures and expert or high-performing operational trajectories:

$$C_k^* = \left[c_k^{*(1)}, c_k^{*(2)}, \dots, c_k^{*(5)}\right] \tag{4}$$

The use of stage-specific benchmarks instead of a single fixed threshold accounts for variations in task difficulty and competency requirements across different phases of the operating process.

2.3 Temporal Training Data Construction

Raw training logs contain both discrete operating events and continuous process variables and therefore require unified temporal preprocessing. All operating actions and process states are first ordered according to their timestamps to reconstruct the complete operational trajectory of each training session. Categorical variables, including valve operations, pump start/stop events, controller switching, and alarm acknowledgement, are transformed using discrete encoding or learnable embeddings. Continuous process variables are standardized using statistics calculated exclusively from the training set:

$$\tilde{x} = \frac{x - \mu_{train}}{\sigma_{train}} \tag{5}$$

where mu_train and sigma_train denote the mean and standard deviation of the corresponding variable in the training set. The interval between consecutive actions is retained as an independent feature to characterize response latency and decision time. To capture the dynamic evolution of competency during a training task, each complete operational sequence is divided into consecutive temporal windows of length W:

$$X_{i,k} = \{x_{i,k}, x_{i,k+1}, \dots, x_{i,k+W-1}\} \tag{6}$$

The window length and stride are determined using validation data. Because the number of operations varies across trainees and training sessions, padding and masking are applied to generate consistent input dimensions while preventing padded positions from contributing to model training or attention calculations. As shown in Table 2.

Table 2 Main Variables Used for Temporal Competency Modeling

Feature category	Representative variables	Information represented
Operating behavior	Valve operation, pump start/stop, controller adjustment	Action content and procedural sequence
Temporal behavior	Timestamp, action interval, response delay	Operating efficiency and response capability
Process variables	Temperature, pressure, flow rate, liquid level	Dynamic process conditions
Equipment state	Equipment status, control mode, interlock state	Operating and control conditions
Abnormal events	Alarm, process-variable exceedance, equipment fault	Abnormality recognition and handling context
Task information	Operating stage, task type	Operational context of individual actions

To avoid information leakage caused by records from the same trainee appearing in both training and test sets, dataset partitioning is performed at the trainee level rather than at the individual-event or sequence-window level. Accordingly, all records generated by a particular trainee are assigned exclusively to the training, validation, or test subset.

2.4 Temporal Deep Learning Model

Because practical operations in coal chemical training exhibit substantial sequential dependencies, a BiLSTM-Attention architecture is adopted as the principal temporal deep-learning model. The bidirectional long short-term memory network first extracts contextual representations of the operating trajectory, after which an attention mechanism identifies operating stages that contribute most strongly to competency estimation. The model finally produces simultaneous predictions for the five job-competency dimensions. For an input vector x_t, the LSTM controls the retention and updating of historical information through the forget, input, and output gates:

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \tag{7}$$

$$i_t = \sigma\left(W_i[h_{t-1}, x_t] + b_i\right) \tag{8}$$

$$\tilde{c}_t = \tanh\left(W_c[h_{t-1}, x_t] + b_c\right) \tag{9}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{10}$$

$$o_t = \sigma\left(W_o[h_{t-1}, x_t] + b_o\right), \quad h_t = o_t \odot \tanh\left(c_t\right) \tag{11}$$

The BiLSTM processes each operational sequence in both forward and backward directions, and the resulting hidden states are concatenated as

$$h_t^{bi} = \left[\overrightarrow{h_t}; \overleftarrow{h_t}\right] \tag{12}$$

allowing the model to represent the relationship between the current action and its surrounding operational context. An attention layer is subsequently applied to the BiLSTM outputs. The relevance score of the t-th operating state is computed as

$$e_t = v^T \tanh\left(W_a h_t^{bi} + b_a\right) \tag{13}$$

and the normalized attention weight is obtained as

$$\alpha_t = \frac{\exp\left(e_t\right)}{\sum_{s=1}^{T} \exp\left(e_s\right)} \tag{14}$$

The sequence-level representation is therefore

$$z = \sum_{t=1}^{T} \alpha_t h_t^{bi} \tag{15}$$

This allows us to assign higher weight to relevant actions (e.g., first response to an abnormal condition, adjustment of critical parameters, handling faults). This representation goes into a fully connected output layer for estimating the five competency dimensions:

$$\widehat{C}_{i,k} = \sigma\left(W_c z_{i,k} + b_c\right) \tag{16}$$

where the model output denotes the estimated dynamic competency state. Since all competency dimensions are represented as normalized continuous scores, a weighted multi-task mean squared error is employed as the primary training objective:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{5}\lambda_j\left(c_i^{(j)} - \hat{c}_i^{(j)}\right)^2 + \beta\|\Theta\|_2^2 \tag{17}$$

with λ_j being a loss weight for the j-th competency dimension, Theta are trainable model parameters, and beta is the regularization coefficient.

2.5 Competency Deviation Detection

The estimated competencies are then compared against a job-oriented reference state in order to identify competency deficiencies: The relative competency deviation for trainee i at training stage k on competency dimension j can be expressed by:

$$D_{i,k}^{(j)} = \max\left[0, \frac{c_k^{*(j)} - \hat{c}_{i,k}^{(j)}}{c_k^{*(j)} + \varepsilon}\right] \tag{18}$$

where c_k*^(j) denotes the reference competency level for the same operating stage and epsilon is a small positive constant introduced to avoid numerical instability. The formulation focuses specifically on competency shortages; performance exceeding the reference level is therefore not interpreted as a negative deviation. An overall job-competency deviation index is calculated as

$$D_i = \sum_{j=1}^{5}\omega_j\left(\frac{1}{K_i}\sum_{k=1}^{K_i}D_{i,k}^{(j)}\right) \tag{19}$$

where omega_j is the importance weight of competency dimension j and satisfies

$$\sum_{j=1}^{5}\omega_j = 1 \tag{20}$$

Dimension-specific decision thresholds tau_j are determined according to the distribution of competency deviations in the training data together with job-assessment criteria. A competency deviation is identified when

$$D_{i,k}^{(j)} > \tau_j \tag{21}$$

This formulation produces a dynamic diagnostic profile indicating both the type of competency deficiency and the operating stage at which it occurs.

2.6 Evaluation Metrics

Model performance is evaluated from two complementary perspectives: competency-state estimation and competency-deviation detection. For continuous competency-score estimation, mean absolute error (MAE), root mean squared error (RMSE), and the coefficient of determination (R2) are used:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{22}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{23}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{24}$$

For competency-deviation detection, Precision, Recall, and F1-score are calculated as

$$Precision = \frac{TP}{TP+FP} \tag{25}$$

$$Recall = \frac{TP}{TP+FN} \tag{26}$$

$$F1 = \frac{2\times Precision\times Recall}{Precision+Recall} \tag{27}$$

ROC-AUC and PR-AUC are additionally employed to evaluate threshold-independent detection performance. To examine whether temporal modeling provides additional information beyond conventional aggregated features, Random Forest and XGBoost are used as conventional machine-learning baselines. LSTM, GRU, and Transformer models are included as sequential-learning baselines and compared with the proposed BiLSTM-Attention architecture. Ablation experiments are also conducted by removing the attention mechanism, temporal-action features, and process-state variables.

3 RESULTS

3.1 Characteristics of Training Behavior and Job Competency

We initially analyzed the distribution of the five job-competency dimensions for an overview of the general competency structure among the trainees. The results are presented in Figure 1; there existed significant disparities on all five competency dimensions. Standardized operation (C1) had the highest overall level with a mean normalized score of

0.821 which indicates that most trainees can follow predefined operating procedures and perform routine equipment operations after basic practical training. Similarly safety and emergency response (C5) has a relatively high mean score of 0.793. Abnormality recognition (C3) and fault handling (C4) have significantly low scores of 0.712 and 0.684 respectively with more dispersed distributions. Process regulation (C2) with a mean score of 0.758 is situated between routine operation and complex abnormal-condition handling.

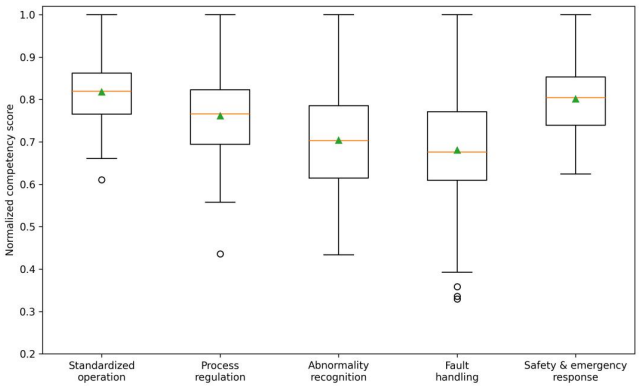


Figure 1 Distribution of Five Job-Competency Dimensions

These differences demonstrate that overall competency cannot be adequately represented by a single practical-training score. In particular, successful execution of routine procedures does not necessarily imply an equivalent capability to diagnose and handle complex abnormal operating conditions. Modeling the competency dimensions separately is therefore necessary for identifying more specific skill gaps.

3.2 Comparison of Different Models

We compare our BiLSTM-Attention model against Random Forest, XGBoost, LSTM, GRU and Transformer models for investigating how much temporal information is used on competency modeling. The performance comparison between these models on estimating competency state and detecting competency deviation are summarized in Table 3.

Table 3 Performance Comparison of Different Models

Model	MAE	RMSE	R2	Precision	Recall	F1-score	ROC-AUC	PR-AUC
Random Forest	0.091	0.118	0.742	0.781	0.748	0.764	0.842	0.801
XGBoost	0.083	0.109	0.781	0.804	0.775	0.789	0.861	0.823
LSTM	0.073	0.096	0.824	0.837	0.812	0.824	0.894	0.861
GRU	0.071	0.093	0.833	0.842	0.819	0.830	0.901	0.868
Transformer	0.068	0.090	0.846	0.857	0.831	0.844	0.912	0.883
BiLSTM-Attention	0.060	0.081	0.879	0.889	0.864	0.876	0.936	0.912

Traditional machine learning models learned valuable insights on aggregated behavior variables but they performed worse compared to the sequential models. Random Forest obtained an F1-score of 0.764 whereas XGBoost obtained 0.789. However, F1-scores for both LSTM and GRU rose up to 0.824 and 0.830 respectively indicating the added benefit of preserving the temporal order of operational events. Then, Transformer improved the F1-score to 0.844. Our proposed BiLSTM-Attention architecture had the highest performance in terms of MAE (0.060), RMSE (0.081) and R2 (0.879). Its F1-score, ROC-AUC, and PR-AUC were 0.876, 0.936, and 0.912 when it came to detecting competencies' deviation.

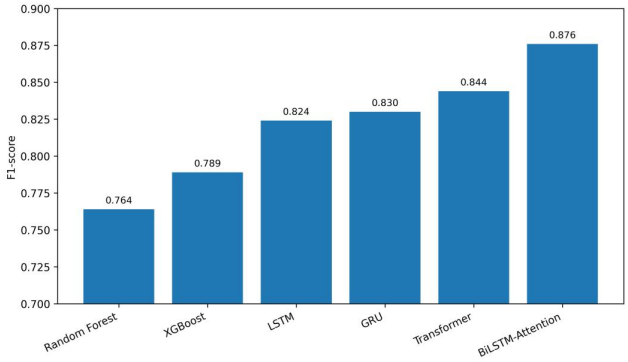


Figure 2 Competency-Deviation Detection Performance across Models

As shown by the overall trend in Figure 2 we observed an increase in detection when incorporating more temporal structure in our model. It appears clear that only counting the number of operations does not fully capture job competency; how actions are sequenced, their response time, and the change in state of processes provides important information for assessing operational competence.

3.3 Performance of Competency-Deviation Detection

We compare our estimated competency states against their respective job-oriented reference states by means of a deviation formulation as presented in Section 2.5. In Table 4 we present the mean competency score, mean deviation index, and dimension-specific detection performance.

Table 4 Dimension-Specific Competency and Deviation-Detection Results

Competency dimension	Mean competency score	Mean deviation index	Precision	Recall	F1-score
C1 Standardized operation	0.821	0.087	0.914	0.891	0.902
C2 Process regulation	0.758	0.148	0.884	0.861	0.872
C3 Abnormality recognition	0.712	0.216	0.873	0.850	0.861
C4 Fault handling	0.684	0.268	0.851	0.833	0.842
C5 Safety and emergency response	0.793	0.126	0.916	0.892	0.904

There is a clear trend in competency deviations among all five dimensions. Only 0.087 was the mean deviation index of standardized operation; however, the greatest deviation was found in fault handling (average: 0.268) then abnormality recognition (0.216). This demonstrates that routine procedural skills tended to be better developed compared to those which require interpreting abnormal process conditions. Accuracy detection varied between competency dimensions as well. Safety and emergency response had the highest F1-score (0.904), but fault handling had the lowest one (0.842); standardization safety procedures tend to have clearer behavior patterns so they can easily be distinguished by our model, while fault handling might contain multiple valid operating paths and depend heavily upon the underlying process state.

3.4 Multidimensional Competency-Deviation Profiles

In order to explore competency differences on an individual basis, we visualize in Figure 3 the five-dimensional deviation profiles for a set of 30 representative trainees. A large amount of variation is present not only between trainees but also with respect to the competency dimensions.

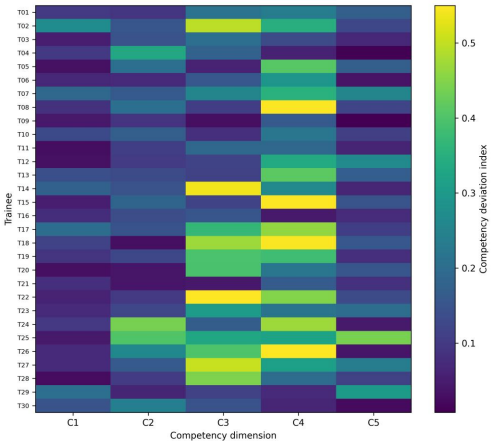


Figure 3 Trainee-Level Competency-Deviation Profiles

Several trainees exhibited little or no deviation in C1 and C5 but relatively high values in C3 and C4. These trainees were generally capable of completing standardized operating procedures but still showed deficiencies when required to recognize and respond to abnormal process conditions. Other trainees exhibited moderate deviations simultaneously in C2-C4, suggesting limitations in understanding dynamic process behavior rather than a single isolated operating step. Such differences cannot be reliably distinguished using a single final training score. The multidimensional deviation profile therefore transforms a global performance score into a diagnostic representation of specific competency deficiencies. The concentration of relatively high deviations in C3 and C4 further indicates that abnormality recognition and fault handling represent major competency bottlenecks and should receive greater emphasis in scenario-based training design.

3.5 Interpretation of Critical Operating Stages

In order to study the utilization of temporal information by our model we extract the attention weights for one typical abnormal condition handling sequence. The distribution of attention is shown in Figure 4 on 50 consecutive operating steps.

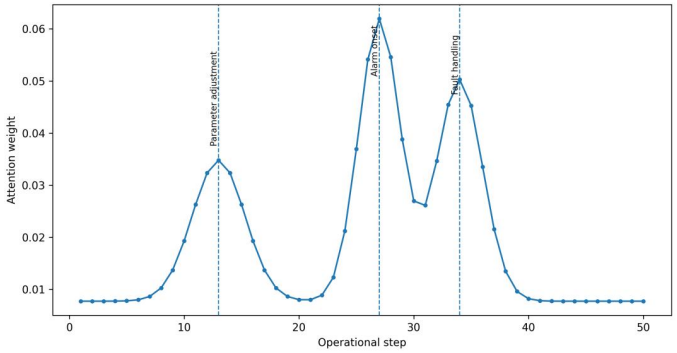


Figure 4 Attention Distribution over an Abnormal-Condition Handling Sequence

The attention weights were clearly non-uniform across the operational sequence. The first prominent increase occurred around Step 13, corresponding to a critical process-variable adjustment. The highest attention peak appeared near Step 27, coinciding with the onset of an alarm, while another substantial peak occurred around Step 34 during subsequent fault-handling actions. This pattern suggests that the model did not primarily rely on simple global statistics such as total task duration or the total number of actions. Instead, greater importance was assigned to the dynamic chain of abnormal-event occurrence, trainee response, and process recovery. Attention weights should not be interpreted as strict causal effects; in the present framework, they primarily provide stage-level interpretability and can be combined with feature-level methods such as SHAP to identify the specific process and behavioral variables contributing to a predicted deviation.

3.6 Ablation and Robustness Analysis

We conducted ablation experiments for quantifying contributions from various architectural components, as well as inputs. Results can be found in Table 5.

Table 5 Ablation Results of the Proposed Model

Configuration	R2	F1-score	PR-AUC
Full BiLSTM-Attention	0.879	0.876	0.912
Without attention	0.858	0.851	0.889
Without temporal interval features	0.847	0.838	0.876
Without process-state features	0.823	0.819	0.851
Single-direction LSTM	0.824	0.824	0.861

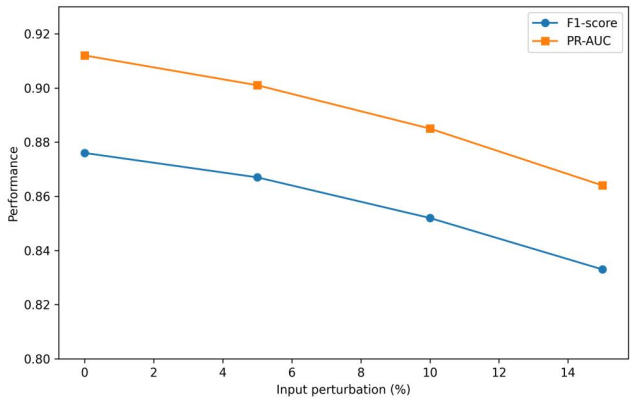


Figure 5 Robustness under Increasing Input Perturbation

Removing the attention mechanism reduced the F1-score from 0.876 to 0.851, confirming that adaptive weighting of different operational stages contributes to more accurate competency-deviation detection. Excluding temporal interval features reduced the F1-score to 0.838, demonstrating that response latency itself carries meaningful competency information. The largest performance deterioration occurred when process-state variables were removed: R2 decreased from 0.879 to 0.823 and the F1-score decreased to 0.819. This finding is particularly important for coal chemical training because knowing which action was performed is insufficient to determine whether it was effective; the resulting changes in temperature, pressure, flow rate, and other process variables must also be considered.

The model robustness is also tested through incrementally adding perturbations on the input variables which represent sensor noise, sampling uncertainty, and timing errors that are often present in practical training platforms. See Table 6.

Table 6 Robustness under Different Levels of Input Perturbation

Input perturbation	R2	F1-score	PR-AUC
0%	0.879	0.876	0.912
5%	0.868	0.867	0.901
10%	0.851	0.852	0.885
15%	0.832	0.833	0.864

As shown in Figure 5. Model performance decreased gradually with increasing perturbation but did not exhibit abrupt degradation. Under a 10% perturbation level, the F1-score remained at 0.852 and PR-AUC at 0.885. Even when the perturbation increased to 15%, the two metrics remained at 0.833 and 0.864, respectively. Overall, the comparative experiments, dimension-specific analysis, interpretability results, and ablation tests consistently demonstrate that competency deviations in coal chemical vocational training are both multidimensional and temporally dependent. By jointly incorporating operational behavior, temporal information, and evolving process states, the BiLSTM-Attention framework supports dynamic diagnosis at both the competency-dimension and operational-stage levels.

4 CONCLUSION

In this study we have constructed a temporal deep-learning approach to detect job-competency deviation during coal chemical vocational training which overcomes the limitations of traditional outcome-based assessment regarding dynamic difference of trainee performances. Using operational behavior, action intervals, equipment state and key process variable, our BiLSTM-Attention model captured five main competency dimensions and compared their estimated states against job-oriented reference levels. Our results show that temporal models are able to better utilize operational sequences, responses, and process-states than conventional machine-learning approaches mainly relying on aggregated features. Analysis specific to each dimension indicates that trainees tend to exhibit more consistent performance in standardized operations and safety related tasks; however, recognizing abnormalities and handling faults remain as the major competency weakness. This framework advances practical-training assessment beyond single-endpoint scores towards dynamic diagnosis at competency-dimension level and operational stage level. Through attention and ablation analyses, we found that critical operating stages, response timing and process-state information all significantly contribute to competency estimation indicating that it's important to evaluate not just if an operation has been done or not but how well the operating process was performed. In future research, validation of the framework will use larger datasets collected across various training units, job roles, and vocational institutions together with expert operational data establishing more robust competency benchmarks. Advanced architectures such as Transformer-based models, multimodal learning and explainable artificial intelligence might also be investigated through integration of operation logs with video, speech, and other behavioral information allowing fine-grain and real-time competency diagnosis and facilitating the use of data-driven assessment within real industrial training and occupational skill evaluation.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Abdelrahman G, Wang Q, Nunes B. Knowledge tracing: A survey. ACM Computing Surveys, 2023, 55(11): 1-37.
[2] Audibert J, Michiardi P, Guyard F, et al. Do deep neural networks contribute to multivariate time series anomaly detection? Pattern Recognition, 2022, 132: 108945.
[3] Fracaro S G, Glassey J, Bernaerts K, et al. Immersive technologies for the training of operators in the process industry: A systematic literature review. Computers & Chemical Engineering, 2022, 160: 107691.
[4] Li G, Jung J J. Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges. Information Fusion, 2023, 91: 93-102.
[5] Liang X, Liu M C, Hu J, et al. Deep knowledge tracing model by integrating problem difficulty and answering experience. Journal of South China Normal University (Natural Science Edition), 2023, 55(4): 81-86.
[6] Qin H, Yan J, Liu X, et al. An AI-driven learning analytics model for evaluating digital competence of vocational teachers. Array, 2026, 31: 100995.
[7] Shahab M A, Iqbal M U, Srinivasan B, et al. HMM-based models of control room operator's cognition during process abnormalities. 1. Formalism and model identification. Journal of Loss Prevention in the Process Industries, 2022, 76: 104748.
[8] Shahab M A, Iqbal M U, Srinivasan B, et al. HMM-based models of control room operator's cognition during process abnormalities. 2. Application to operator training. Journal of Loss Prevention in the Process Industries, 2022, 76: 104749.

[9] Yan J, Tian H, Sun X, et al. Role of artificial intelligence in enhancing competency assessment and transforming curriculum in higher vocational education. Frontiers in Education, 2025, 10: 1551596.
[10] Zamanzadeh Darban Z, Webb G I, Pan S, et al. Deep learning for time series anomaly detection: A survey. ACM Computing Surveys, 2024, 57(1): 1-42.