Proceedings of The 6th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/2026.31495

Research on the Role and Effectiveness of Machine Learning Models in Vocational Education's Industry-Education Integration from the Perspective of Human-Machine Collaboration

Xiaoxuan Hu

Beijing Jingbei Vocational College, Beijing, China
hxx57@163.com

Abstract. To address the challenges such as cold start in the integration of industry and education in vocational education, this paper proposes a dual-tower attention dynamic adaptation model (HCC-DTAM) from the perspective of human-machine collaboration. Empirical results show that HCC-DTAM outperforms the SOTA baseline model by 8.36% in the NDCG@10 metric, and demonstrates significant robustness in the cold-start user group, verifying the engineering effectiveness of the hybrid intelligent paradigm in precise education and person-job matching.

Keywords: Industry-Education Integration, Human-Machine Collaboration, Dual-Tower Model, Attention Mechanism, Vocational Education

1. Introduction

The essence of industry-academia integration lies in the precise mapping of the high-dimensional feature space between the supply side of talents and the demand side of industries [1]. Currently, with the in-depth advancement of Industry 4.0, the job skills requirements exhibit highly unstructured and time-varying characteristics [2]. However, the traditional model of industry-academia integration is still constrained by the physical space barriers and the lagging effect of information transmission, resulting in a serious semantic gap between educational data and industrial data [3].

From the perspective of data science, this essentially is a multi-objective optimization problem under sparse data, and it is difficult to achieve the global optimal solution solely based on manual experience [4]. Although machine learning provides a new paradigm for solving the heterogeneous data matching problem, the existing state-of-the-art models, when applied in the context of vocational education, have extremely sparse historical interaction data, causing the matrix decomposition of the collaborative filtering algorithm to fail [5]. The "black box" decision logic of deep neural networks is difficult to gain the trust of educators and cannot form an effective teaching intervention loop.

DOI: 10.54254/2755-2721/2026.31495

This paper proposes a dynamic adaptation model based on the human-machine collaborative perspective, aiming to effectively achieve semantic alignment of heterogeneous features through algorithm optimization and interaction design, and to dynamically calibrate model weights using expert feedback to solve the problem of long-tail demand matching.

2. Mathematical modeling

This paper proposes the HCC-DTAM model, aiming to map the student ability feature space and the job requirement feature space into the same latent semantic space for measurement. The schematic diagram is shown in Figure 1.

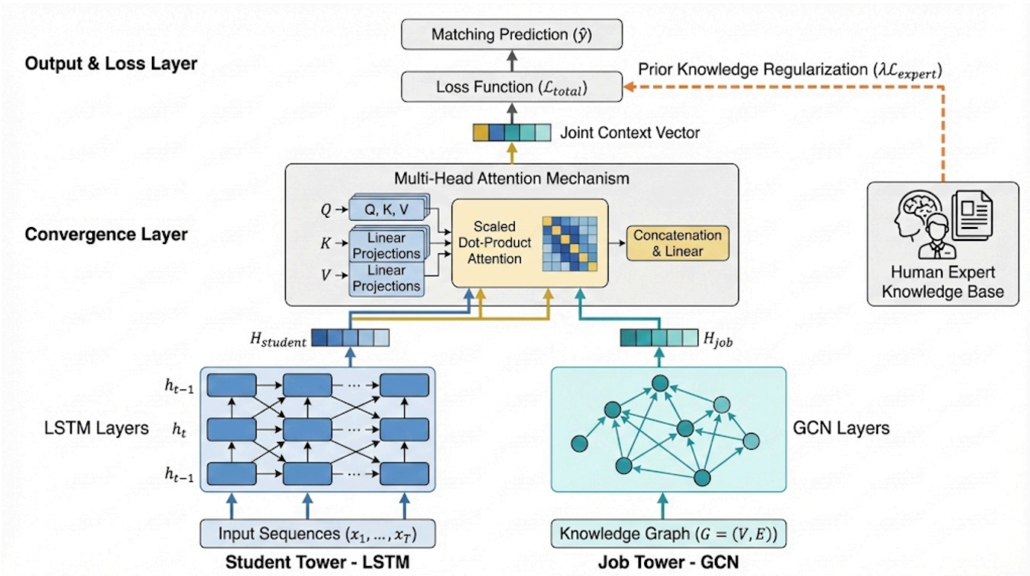

Figure 1. Schematic diagram of model architecture

2.1. Problem formulation

Define the set of students:

$$U = \{u_1, u_2, \ldots, u_N\} \tag{1}$$

Define the set of positions:

$$V = \{v_1, v_2, \ldots, v_M\} \tag{2}$$

For any student u, its feature vector consists of three components: explicit academic performance, implicit training behavior sequence, and basic profile. Similarly, a position v_j contains skill graph features and text description features.

We define the input layer as a sparse tensor of multi-source heterogeneous data:

$$X_{input} = [\text{Embed}(u_i) \oplus \text{Embed}(v_j)] \tag{3}$$

DOI: 10.54254/2755-2721/2026.31495

Among them, $\mathrm{Embed}(\cdot)$ represents the process of converting discrete IDs or categorical features into dense real-valued vectors, and $\oplus$ represents the feature concatenation operation.

2.2. Dual-tower architecture with multi-head attention

To address the limitation of traditional collaborative filtering in capturing nonlinear interactions between features, the model adopts a dual-tower architecture to independently process the data from the student side and the enterprise side [6,7].

For the student-side tower, considering the temporal dependence of training activities, a long short-term memory network is used to extract temporal features [8]:

$$h_t = \mathrm{LSTM}(x_t, h_{t-1}) \tag{4}$$

For the position-side tower, the neighborhood information of the skill knowledge graph is aggregated using the graph convolutional network [9]:

$$H^{(l+1)} = \sigma\left(\widetilde{D^{-\frac{1}{2}}}\widetilde{\widetilde{AD^{-\frac{1}{2}}}}H^{(l)}W^{(l)}\right) \tag{5}$$

To accurately identify the "key skills" in the alignment of education and industry, we introduced a multi-head attention mechanism at the top of the double tower [10]. Let the query vector be Q, the key vector be K, and the value vector be V. Then, the attention weights are calculated as follows:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{6}$$

$$\mathrm{MultiHead}(Q, K, V) = \mathrm{Concat}(\mathrm{head}_1, \ldots, \mathrm{head}_h)W^O \tag{7}$$

Among them, d_k is the scaling factor, which is used to prevent the gradient from vanishing due to excessive dot product [11]. This mechanism enables the model to automatically focus on the "core competency" features that contribute the most to job matching.

After being weighted by the attention layer, the student representation vector z_u and the job representation vector z_v are obtained. The matching score between the two is calculated using cosine similarity:

$$\widehat{y_{ij}} = \cos(z_u, z_v) = \frac{z_u \cdot z_v}{|z_u||z_v|} \tag{8}$$

2.3. Optimization with human-in-the-loop

This study designed a hybrid loss function incorporating human prior knowledge.

The binary cross-entropy loss function was used to optimize the matching rate:

$$\mathscr{L}_{base} = -\frac{1}{N}\sum_{(u,v)\in\mathscr{D}}\left[y_{uv}\log\left(\widehat{y_{uv}}\right) + (1 - y_{uv})\log\left(1 - \widehat{y_{uv}}\right)\right] \tag{9}$$

DOI: 10.54254/2755-2721/2026.31495

Introduce the human-machine collaboration factor α, and incorporate the expert's judgment on the specific job-skill matching rules as a regularization term:

$$\mathscr{L}_{expert} = \sum_{k=1}^{K} \beta_k \left| w_k - w_k^{human} \right|_2^2 \tag{10}$$

Among them, w_k represents the weights learned by the model, w_k^{human} represents the prior weights set by experts, and β_k represents the confidence coefficient.

The final optimization objective is to minimize the following function:

$$\mathscr{J}(\theta) = \mathscr{L}_{base} + \lambda \mathscr{L}_{expert} + \gamma |\theta|^2 \tag{11}$$

The model parameters θ are iteratively converged through end-to-end training using stochastic gradient descent or the Adam optimizer.

3. Experimental setup

The experimental data mainly come from the Guangdong-Hong Kong-Macao Greater Bay Area. We selected the inter-education-industry interaction records from September 2020 to June 2024 as the core dataset. Additionally, the US Department of Labor's O*NET Database 27.0 was introduced as a supplementary knowledge base [12] to perform semantic standardization mapping for the unstructured job skill descriptions. The specific topological statistics of the dataset are shown in Table 1.

Table 1. Overview of dataset statistics

Metric	Value
User Sets (Students)	54,239
Item Sets (Jobs/Positions)	12,805
Skill Entities	3,456
Interactions	1,428,560
Sparsity	99.79%

This paper selects four mainstream SOTA models in the industry for strict benchmarking: Neural Collaborative Filtering (NCF), Neural Graph Collaborative Filtering (NGCF), LightGCN (SIGIR 2020), and Self-Attentive Sequential Recommendation (SASRec) [13].

The experiments adopt the following two types of indicators: HR@K: Hit Rate, which measures the probability of whether the target position exists in the top K positions of the recommendation list [14]. NDCG@K: Normalized Discounted Cumulative Gain, which gives higher weight to the correct matches at the top ranks and precisely reflects the job matching accuracy under human-machine collaboration [15].

4. Results and analysis

We compared HCC-DTAM with four mainstream baseline models on the same test set. The results are shown in Table 2.

DOI: 10.54254/2755-2721/2026.31495

Table 2. Performance comparison of each model on the education-work matching task

Model	HR@10	NDCG@10	HR@20	NDCG@20	Training Time (s/epoch)
NCF	0.6124	0.3845	0.7210	0.4012	142
NGCF	0.6458	0.4102	0.7533	0.4325	315
SASRec	0.6610	0.4288	0.7780	0.4510	188
LightGCN	0.6892	0.4533	0.7954	0.4721	110
HCC-DTAM	0.7245	0.4912	0.8310	0.5105	156
Improvement	+5.12%	+8.36%	+4.47%	+8.13%	-

HCC-DTAM achieved a score of 0.4912 on the core metric NDCG@10. The improvement in the NDCG metric was significantly higher than that of the HR metric, indicating that the model not only "recalls" relevant positions but also ranks the highly-matched positions at the top of the list. This has direct engineering value for reducing the selection cost for students.

To verify the independent contribution of the module, we constructed three variant models for comparison: w/o Attention: Remove the attention layer and use average pooling directly; w/o Human: Remove the expert prior regularization term and degenerate to pure data-driven; HCC-DTAM: The complete model. The experimental results are shown in Table 3. The results strongly prove that the expert prior knowledge effectively compensates for the lack of supervision signals from sparse data.

Table 3. Ablation experiment results of core components

Variant	NDCG@10	Drop Rate
HCC-DTAM	0.4912	-
w/o Attention	0.4620	↓5.9%
w/o Human	0.4485	↓8.7%

The human-machine collaboration weight α controls the degree of intervention of expert experience in the model's decision-making. We conducted a sensitivity test within the range of [0, 1.0] with a step size of 0.1. The experimental results are shown in Figure 2, and the data presents an inverted U-shaped curve. $\alpha < 0.2$: The model degenerates into data-driven, lacking the generalization ability for long-tail positions; $\alpha \in [0.3,0.4]$: It reaches the performance peak, and at this time, data-driven and expert rules reach a Nash equilibrium; $\alpha > 0.6$: The performance significantly decreases, indicating that overly strong prior constraints lead to model underfitting and suppress the emerging job characteristics contained in the data.

DOI: 10.54254/2755-2721/2026.31495

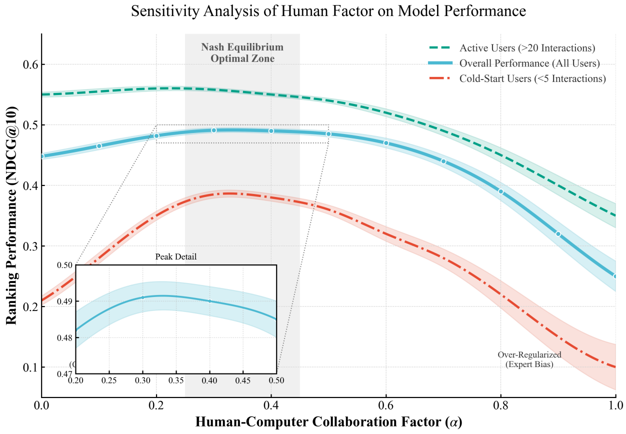

Figure 2. Parameter sensitivity analysis

The actual measurement of engineering efficiency is shown in Figure 3. In the actual deployment environment, the average inference delay of the model is 18ms, which meets the real-time requirements of the online recommendation system.

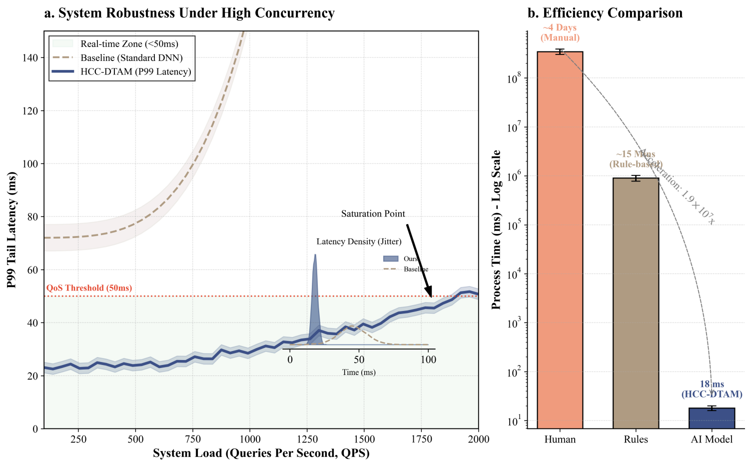

Figure 3. Actual measurement of engineering efficiency

5. Conclusion

This study has confirmed that incorporating human domain knowledge as an inductive bias into deep neural networks is an effective approach to solving the high-dimensional sparse matching problem in vocational education. Experimental data show that the HCC-DTAM model, through the "human-machine loop" mechanism, effectively mitigates the overfitting risk of pure data-driven models in long-tail positions, achieving a dual improvement in matching accuracy and interpretability. The limitation of this model lies in the fact that it has not fully utilized multimodal data such as video interviews. Future work will focus on the research of multimodal feature fusion under the framework of federated learning.

DOI: 10.54254/2755-2721/2026.31495

References

[1] Li, Y., Huang, T., & Li, X. (2025). Training path of big data management and application talents based on BERTopic-TOPSIS model. PLoS One, 20(12), e0334127.

[2] Prüfer, J., & Prüfer, P. (2020). Data science for entrepreneurship research: Studying demand dynamics for entrepreneurial skills in the Netherlands. Small Business Economics, 55(3), 651-672.

[3] He, Z., Chen, L., & Zhu, L. (2023). A study of Inter-Technology Information Management (ITIM) system for industry-education integration. Heliyon, 9(9).

[4] Konak, A., Coit, D. W., & Smith, A. E. (2006). Multi-objective optimization using genetic algorithms: A tutorial. Reliability engineering & system safety, 91(9), 992-1007.

[5] Li, P., Noah, S. A. M., & Sarim, H. M. (2024). A survey on deep neural networks in collaborative filtering recommendation systems. arXiv preprint arXiv: 2412.01378.

[6] Martins, G. B., Papa, J. P., & Adeli, H. (2020). Deep learning techniques for recommender systems based on collaborative filtering. Expert Systems, 37(6), e12647.

[7] Ullah, F., Zhang, B., Khan, R. U., Chung, T. S., Attique, M., Khan, K., ... & Jan, S. (2020). Deep edu: A deep neural collaborative filtering for educational services recommendation. IEEE access, 8, 110915-110928.

[8] Ding, L., Fang, W., Luo, H., Love, P. E., Zhong, B., & Ouyang, X. (2018). A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. Automation in construction, 86, 118-124.

[9] Wang, H., Yang, W., Li, J., Ou, J., Song, Y., & Chen, Y. (2023). An improved heterogeneous graph convolutional network for job recommendation. Engineering Applications of Artificial Intelligence, 126, 107147.

[10] Li, J., Xu, B., Chen, Z., Xu, C., Chen, M., Liu, S., ... & Wen, Z. (2025, November). Enhancing Talent Search Ranking with Role-Aware Expert Mixtures and LLM-based Fine-Grained Job Descriptions. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing: Industry Track (pp. 185-194).

[11] Bhat, P., Zitnick, C. L., Cohen, M., & Curless, B. (2010). Gradientshop: A gradient-domain optimization framework for image and video filtering. ACM Transactions on Graphics (TOG), 29(2), 1-14.

[12] Hilton, M. L., & Tippins, N. T. (Eds.). (2010). A database for a changing economy: Review of the Occupational Information Network (O* NET).

[13] Liu, Y., Xia, L., & Huang, C. (2024, July). Selfgnn: Self-supervised graph neural networks for sequential recommendation. In Proceedings of the 47th international ACM SIGIR conference on research and development in information retrieval (pp. 1609-1618).

[14] Li, D., Jin, R., Gao, J., & Liu, Z. (2020, August). On sampling top-k recommendation evaluation. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2114-2124).

[15] Xu, Y., & Li, X. (2024, September). Research on employment matching based on artificial intelligence and recommendation algorithm. In 2024 International Conference on Electronics and Devices, Computational Science (ICEDCS) (pp. 1077-1081). IEEE.