International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 5274-5285

https://doi.org/10.70517/ijhsa464367

Research on Generative Artificial Intelligence Empowering Smart Classrooms and Thinking Development

Chenglin Yang[1], Guang Xing[1], Weiqian Ma[1] and Jia Tai[1,*]
[1] School of Art and Design, Anhui Broadcasting Movie and Television College, Hefei, Anhui, 230011, China
Corresponding authors: (e-mail: 15755181880@163.com).

Abstract The current process of education informatization is deepening, and the demand for intelligent technology in the field of education is becoming more and more urgent. This paper proposes a deep composite recommendation model VAE-GAN-DCR based on variational autoencoder and generative adversarial network, and explores the effect of generative artificial intelligence in smart classroom. Methodologically, the model combines the decoder of VAE with the generator of GAN, improves the traditional VAE model by introducing a priori distributions that depend on item features, and optimizes the reconstruction error by using the feature transfer of the GAN discriminator to achieve accurate recommendation of educational resources. At the same time, the Williams Creative Tendency Measurement Scale is used to evaluate the teaching effect of students in Zhanjiang Early Childhood Teacher Training College. The results show that the VAE-GAN-DCR model performs well on three datasets, in which the Recall@20 value is increased by 12.15% and the NDCG@100 value is increased by 12.94% on the Movielens-1M dataset. The educational application experiment shows that the experimental group is significantly better than the control group in creative thinking activities and creative tendency, and the score of the total creative tendency scale reaches 2.61.The conclusions show that generative artificial intelligence technology can effectively improve the precision of educational resources recommendation and the development of students' creativity, and provide a powerful support for the construction of smart classroom.

Index Terms Generative Artificial Intelligence, Smart Classroom, Variable Score Autoencoder, Generative Adversarial Network, Educational Resource Recommendation, Creative Thinking Activity

I. Introduction

Artificial Intelligence (AI), as an emerging technology, is developing rapidly and showing great potential in various fields. Among them, generative AI, as one of the important branches of AI, is not only widely used in the fields of art and literature, but also shows innovative application prospects in the field of education [1], [2]. Generative AI is a technology based on machine learning, which generates new content, such as images, music, text, etc., by learning a large amount of data [3], [4]. Generative AI can be categorized into rule-based approaches and neural network-based approaches. Rule-based approaches rely on manually formulated rules to generate content, while neural network-based approaches generate content by training neural network models, the most common of which uses generative adversarial networks [5]-[7].

The application of generative AI in the field of education is mainly reflected in personalized teaching and virtual experiment simulation [8]. Generative AI can generate personalized teaching materials based on students' interests, abilities and learning styles [9]. By analyzing students' learning data and feedback, generative AI can customize teaching content and learning paths to fit students' characteristics [10], [11]. This kind of personalized teaching method can improve students' learning effect, stimulate learning interest and reduce learning pressure [12]. In addition, in the case of limited laboratory conditions, generative AI can help students perform experimental operations and observations by simulating and generating virtual experimental scenarios [13], [14]. Generative AI can generate real, various types of conditions of experimental environments based on physical laws and experimental data, providing students with more practical opportunities and deepening their understanding of experimental knowledge and scientific principles [15]-[17].

This study proposes a smart classroom solution based on the VAE-GAN deep composite recommendation model, which constructs a deep learning model that can effectively handle the task of recommending educational resources by integrating the generative capability of the variational autoencoder and the optimization mechanism of the generative adversarial network. The research includes two levels: the technical level constructs the VAE-GAN-DCR recommendation model, introduces the item-dependent a priori distribution by improving the traditional VAE structure, and optimizes the reconstruction error by using the GAN discriminator to realize the accurate modeling and recommendation of educational resources; the application level carries out the validation of the educational practice, and evaluates the enhancement effect of generative AI technology on the creative thinking ability of the students through comparative experiments , providing empirical support for the popularization and application of smart classroom.

II. Method

II. A. Deep learning based recommendation algorithms

II. A. 1) Variational autoencoders

Self-encoders are algorithms that compress data and then downsize it, usually through a neural network, which can potentially replicate the input information and then output it. There are two main components of an auto-encoder, a coding component that processes the input information and a decoding component that hopes to recover the compressed information.

The self-encoder is able to reconstruct the input data, but when the reconstruction process is too perfect, the final data obtained will be the same as the initial data, which defeats the purpose of the self-encoder because the whole model does not gain any particular useful properties during the construction process. So in order to avoid the above problem, some constraints of the model need to be added so that the final value obtained is different from the very beginning.

The proposed variational autoencoder is an improvement on the more desirable generative model, the base model uses a variational Bayesian approach, which allows for approximate a posteriori algorithms and training to be done on standard probabilistic graphical models in a very efficient manner, and also allows for parsing the results obtained from the a posteriori algorithms and optimizing them. In existing research, the method of analyzing the approximation of a posteriori expectation is more common with mean-field, but this method is not well implemented in the variational autoencoder, and then researchers derived the method of maximizing the log-likelihood variational lower bounds of the data based on the structure of the autoencoder to deal with obtaining the a posteriori expectation. In addition, the variational lower bound evaluator SGVB was derived, which handles the a posteriori values by approximating them using the standard stochastic gradient descent method.

The $q_{\varphi}(z \mid x)$ of the encoder part is close to the real data value after training, but the distribution $p(z \mid x)$ after a posteriori manipulation has not been handled in a better way, X in the figure is the input value of the encoder part, and Z is the latent vector obtained after the distribution of the orthotropic distribution, and the mean processing, the N-dimensional latent vector is used as a new image with the same distribution in the decoder part $p_{\theta}(x \mid z)$ as input values for the decoder part $p(x \mid z)$, while a new image with the same distribution is obtained after training. After $z \mid x \sim N(\mu_{z|x}, \Sigma z \mid x)$ the potential vector Z is obtained from the sampled data, and the decoder partially after $x \mid z \sim N(\mu_{x|z}, \Sigma x \mid z)$ the sampled data is obtained from $x \mid z$.

II. A. 2) Adversarial Generative Networks

Generative Adversarial Networks [18] are proposed in the hope that the model can utilize the original dataset to obtain the latest differentiated data samples. Many generative models in deep learning are conceived by randomly sampling existing data samples to obtain new samples, but the generative adversarial network utilizes a one-to-one correspondence from random noise Z to training data X. The random noise Z in the process basically obeys a normal distribution, and Goodfellow utilizes a multilayer perceptual machine to construct a generative network $G\left(Z;\theta_g\right)$. The input is random noise and the output is an image.

In order to illustrate the training process of the GAN model in detail, the training mechanism of GAN will be analyzed in detail in this section, and it will be introduced in two parts next:

The first part is to train the generative network. The main purpose of this part is to pass the real data after training to get the fake data, and make this part of the fake data can not be discriminated by the discriminative network, in order to achieve this effect, the training function will be maximized D(G(Z)), which is also minimized 1-D(G(Z)).

The second part is to train the discriminative network. In this part, we want to strengthen the discriminative network's ability to discriminate the trained samples G(Z) by minimizing D(G(Z)), and compared to the real samples, we want to get the real samples with higher probability, so we maximize D(X).

However, both parts of the training process require fixed parameters, so for the discriminative network, the objective function needs to be maximized: $\log D(x) + \log(1 - D(G(Z)))$, after which the final result is obtained:

$$\min_G \max_D V(D,G) = E_{X-P_{data}(x)}[\log D(x)] + E_{X-P(z)}[\log(1 - D(G(z)))] \quad (1)$$

Also for a specific criterion in the discriminative network, it is necessary to fix the generative network to maximize $V(G,D)$:

$$\begin{aligned}\max V(G,D) &= \int_x P_{data}(x)\log(D(x))dx + \int_z P_z(z)\log(1-G(z))dz \\ &= \int_x [P_{data}(x)\log(D(x)) + p_g(x)\log(1-D(x))]dx\end{aligned} \tag{2}$$

Want to make the above equation reach the target maximum so that each of these x is:

$$P_{data}(x)\log(D(x)) + p_g(x)\log(1-D(x)) \tag{3}$$

Since $x, p_{data}(x), p_g(x)$ are fixed values, and also wanting to get the maximum value, there is, when $p_{data}(x)$, $p_g(x)$ are arbitrary non-zero values, and $D(x)\in[0,1]$, the above maximum value is at $p_{data}(x)/(p_{data}(x)+p_g(x))$ can be obtained, while making the whole model an optimal solution:

$$P_{data}(x)D_G(x)\frac{P_{data}(x)}{P_{data}(x)+p_g(x)} \tag{4}$$

When making improvements to the generative network, the best results can be obtained when the generative network is only $p_{data}(x)=p_g(x)$. Only then can the generative network be trained on the original data samples.

II. B.VAE-GAN based recommendation model for educational resources

II. B. 1) Model structure

This chapter proposes a deep composite recommendation model based on Variational Autoencoder and Generative Adversarial Network, referred to as VAE-GAN-DCR, which includes both user-side and item-side models. The basic idea is to combine the decoder of the VAE model with the generator of the GAN model to realize the combined recommendation of the two deep models.

Let the variable $u\in U=\{1.....m\}$ denote a set of users, and the variable $v\in V=\{1....n\}$ denote a set of items, and the partially observable user-item rating matrix is defined as $R\in R^{m\times n}$. Each user u is associated with a rating matrix row $r_u=(R_{u1}...R_{un})\in R^n$ is associated with r_u, which represents the ratings of all items by user u; similarly, each item v is associated with a rating matrix column $r_v=(R_{1v}...R_{mv})\in R^m$, and r_v denotes the ratings of all users for item v.

II. B. 2) VAE structure

The variational autoencoder includes two parts, the encoder and the decoder. Let the parameter in the encoder be λ and the parameter in the decoder be τ, then the parameter λ can complete the mapping of the input data x to the hidden variable z, and the parameter τ completes the mapping of the hidden variable z to the reconstruction x'. Thus the encoder can be denoted by $q_\lambda(z\mid x)$ and the decoder can be denoted by $p_\tau(x\mid z)$, respectively:

$$z\sim Enc(x)=q_\lambda(z\mid x),\quad x'\sim Dec(z)=p_\tau(x\mid z) \tag{5}$$

In the VAE model, the decoder, i.e., represents a generative model where, in the absence of auxiliary information, the hidden layer representation $z_v\in R^g$ of item v conforms to a Gaussian prior with mean zero $z_v\sim N(0,I_g)$. This chapter improves on the basic VAE model by adding item-dependent auxiliary information as a prior distribution for the hidden variables z_v:

$$z_v\sim N(y_v,S_v) \tag{6}$$

where y_v denotes the mean vector of item v; S_v denotes the covariance matrix of item v. The auxiliary information dependent on the item here includes the multimedia features and textual features of the item. After training this auxiliary information into a feature vector representation, it is directly embedded into the model during the training process.

Next, we use the mapping function $f_\tau(.)\in R^m$ to map the hidden layer representation z_v to the distribution $\pi(z_v)$ over m users:

$$\pi(z_v)\propto\exp\{f_\tau(z_v)\} \tag{7}$$

where the nonlinear function $f_\tau(\cdot)$ is a maximal value function constrained by the parameter τ. The mapped hidden layer representation can be mapped to the distribution $n(z_v)\in\Delta^{m-1}((m-l)-simplex)$ over m users. Then, assume that the observed user-item rating matrix columns $r_v \in R^m$ are sampled from a polynomial distribution obeying probability $\pi(z_v)$:

$$r_v \sim Mult(N_v, \pi(z_v)) \tag{8}$$

In the above equation, $N_v = \Sigma_j r_{jv}$. Then the decoder generation for item v can be expressed as:

$$\log p_\tau(r_v \mid z_v) = \sum_j r_{jv} \log \pi_j(z_v) \tag{9}$$

Since the distribution of the hidden variables is not directly available, it is not possible to use the EM algorithm directly for variational inference. To solve this problem, it is common to introduce an identification model in the encoder, replacing the uncertain true posterior distribution $p_\tau(r_v \mid r_v)$ with $q_\lambda(z_v \mid r_v)$:

$$q_\lambda(z_v \mid r_v) = N(\mu_\lambda(r_v), diag\{\sigma_\lambda^2(r_v)\}) \tag{10}$$

where λ is the parameter of the encoder, from which the encoder computes the parameters $\mu_\lambda(r_v), \sigma_\lambda(r_v) \in R^g$. That is, with r_v as input, the encoder outputs the variational parameters of the corresponding variational distribution $q_\lambda(z_v \mid r_v)$.

II. B. 3) GAN Architecture

The basic structure of generative adversarial network includes generative network and discriminative network, in which the former can randomly generate observation data according to the relevant information to fool the discriminative network; while the latter's role is to judge the source of input data, and distinguish real data from fake samples as far as possible [19]. In the GAN model, in order to realize the distinction between real data and generated data, the discriminative network is able to learn a very rich similarity measure of the elements. Therefore, in order to improve the metrics of the VAE model so that it can be better applied to recommender systems, this chapter introduces the GAN model to transfer the features learned by the discriminative network to the abstract reconstruction error of the VAE model.

Let r_v denote the input rating data, and γ be the discriminator parameters. The discriminator network $D_\gamma(r_v) \in [0,1]$ is used to predict the probability that the input scoring data r_v comes from the real scoring data, and $1 - D_y(p_\tau(r_v \mid z_v)$ denotes the probability that r_v is a sample of scoring data generated from the generative model. Its objective function is:

$$L(GAN) = \log(D_\gamma(r_v)) + \log(1 - D_\gamma(p_\tau(r_v \mid z_v))) \tag{11}$$

Since the cell-level reconstruction error method used in the VAE model does not have a high prediction accuracy for the scoring data, this chapter uses the discriminator representation of the reconstruction error of the GAN model to replace the reconstruction error in the VAE model. Specifically, we use $D_\gamma l(r_v)$ to denote the hidden representation of the l th layer of the discriminator to obtain the following $D_\gamma l(r_v)$ Gaussian observation model:

$$p(D_\gamma l(r_v) \mid z_v) \sim N(D_\gamma l(r_v) \mid D_\gamma l(r_v'), I_k) \tag{12}$$

where $\tau_v' \sim p(\tau_v \mid z_v)$ denotes the output of the decoder, and the reconstruction error of the VAE model after the above equation is $E_{q_\lambda(z_v|r_v)}[\log p_\tau(D_\gamma l(r_v) \mid z_v)]$, so the extended ELBO becomes:

$$\begin{aligned} & L(r_v; \tau, \lambda, \gamma, y_v, S_v) \\ & = E_{q_\lambda(z_v|r_v)}[logp_\tau(D_\gamma l(r_v) \mid z_v)] - KL(q_\lambda(z_v \mid r_v) \,\|\, p(z_v; y_v, S_v)) \end{aligned} \tag{13}$$

Here, the original Gaussian prior distribution is replaced by an item-dependent prior distribution, i.e., the variational posterior distribution $q_\lambda(z_v \mid r_v)$ is regularized to approximate the item-featured dependent prior distribution $p(z_v; y_v, S_v)$. After adding auxiliary information about the items, the KL scatter is computed as follows:

$$KL(q_\lambda(z_v \mid r_v) \,\|\, p(z_v; y_v, S_v)) = \log \frac{S_v}{\sigma_\lambda(r_v)} + \frac{\sigma_\lambda^2(r_v) + (\mu_\lambda(r_v) - y_v)^2}{2S_v^2} - \frac{1}{2} \tag{14}$$

where $q_\lambda(z_v \mid r_v)$ is the posterior distribution that obeys the parameter λ, and $p(z_v; y_v, S_v)$ is the mean γ_v and variance S_v of the dependent on the prior distribution of item characteristics. Ultimately, the optimization objective of the whole model is:

$$L = L(GAN) + L(\tau_v; \tau, \lambda, \gamma, y_v, S_v) \tag{15}$$

II. B. 4) Training process

In updating the τ parameter, we weight the reconstruction error of the VAE model and the network error of the GAN model. Specifically, the decoder receives error signals from both $L(GAN)$ and $L(r_v)$, where this chapter introduces the parameter θ to weigh the VAE model's ability to reconstruct the data against the ability to mislead the GAN discriminator. Again, since the weighting operation is performed only when the parameters of the VAE decoder are updated, the θ parameter is not used throughout the model:

$$\tau \leftarrow -\nabla_{\tau}(\theta L(r_v) + L(GAN)) \tag{16}$$

III. Results and Discussion

III. A. Model Performance Analysis

III. A. 1) Data set processing

When processing textual information, Doc2vec is used to train word vectors; data that is too sparse in the dataset is processed, and users with fewer than 5 ratings are removed from the process. The ratings are expressed as 0 or 1 to directly reflect the user's preference for the item. The rating information in the dataset is from 0 to 5. A rating higher than 3.5 is labeled as 1, which means that the user is interested in the item; a rating lower than 3.5 is labeled as 0, which means that the user is not interested in it or does not like it.

III. A. 2) Experimental evaluation indicators

D-VAE-GAN based recommendation algorithms are designed to address the performance of top-K recommendations, and the two evaluation metrics, Recall@K and NDCG@K, are commonly used in top-K based recommendation algorithms, which are used in most of the top-K based recommendation algorithm experiments. Therefore, Recall@K and NDCG@K are chosen as the evaluation metrics for the D-VAE-GAN model for the experiments in this chapter.

Let the set of items that user u has evaluated in the test data be Te(u), and the list of item recommendations given to the user by the recommender system on the training set Tr(u) be L(u). Then the recall is defined as shown in Equation (17):

$$Recall = \frac{\sum_{u \in U} |L(u) \cap Te(u)|}{\sum_{u \in U} |Te(u)|} \tag{17}$$

The Recall metric represents the number of items that a user actually wants to be recommended as a percentage of all items of interest to that user. The accuracy and recall of a recommendation algorithm are positively correlated. If the accuracy and recall metrics are larger, the accuracy and performance of the recommendation algorithm will be higher. Assuming that a recommendation algorithm has high accuracy and recall metrics, it means that the user's favorite items were accurately recommended. A good recommender system will be able to recommend items of interest to users while also allowing them to discover items they would normally want to know about but would have a hard time discovering or thinking about.

The normalized discounted cumulative gain (NDCG) is defined as shown in equation (18):

$$NDCG = \frac{1}{N} \sum_{u \in U} \frac{1}{\log_2(p_u + 1)} \tag{18}$$

Normalized discounted cumulative gain indicates that it is important to find those items that are more visible to the user, i.e., to emphasize "sequentiality". Where N is the total number of users, p_u is the position of the user's real visit value in the recommendation list, and if the value does not exist in the recommendation list, then $p_u \to \infty$.

III. A. 3) Experimental comparison

In order to verify the effectiveness of the algorithms in this chapter, the following models are selected for comparison experiments:

(1) Non-negative Matrix Factorization (NMF): trained using alternating least squares, NMF is a linear latent factorization model and usually has better performance than stochastic gradient descent.

(2) Neural Collaborative Filtering (NCF): uses a multi-layer neural network to learn information about user-item interactions, and is a matrix decomposition algorithm that allows access to nonlinear relationships.

(3) Recommendation Modeling with Dual Autoencoders (ReDa): uses autoencoders to generate potential user and item feature matrices through representation learning with dual autoencoders.

(4) Deep Collaborative Variational Autoencoder (CVAE): a deep Bayesian model that uses a Bayesian network to find its probability distribution and a variational autoencoder to extract feature information.

(5) Trust-based Singular Value Decomposition (TrustSVD): this model is a collaborative filtering that adds explicit and implicit effects of user trust and item ratings to the SVD model.

III. A. 4) Analysis of experimental results

The experiments in this chapter preprocess the two datasets.Firstly, the two datasets are divided into five groups respectively, and one group is randomly selected as the validation set and the other four groups are used as the training set. The dimension of the hidden layer is set to 500. Finally, the optimization is still performed using Adam's algorithm. Observe the experiments in this chapter using Tanh function and Relu function to increase the occurrence according to Epoch as shown in Fig. 1 and Fig. 2. The performance of the training algorithm using the two activation functions is compared using NDCG@100 as a metric.

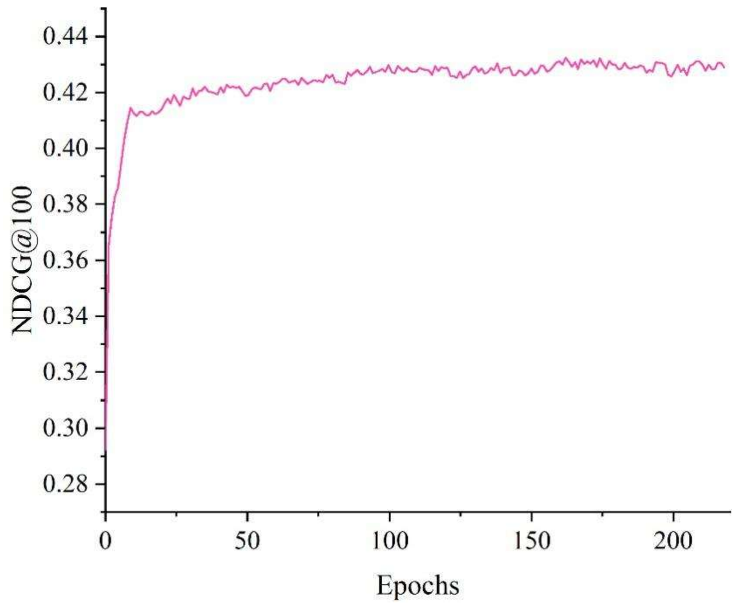

Figure 1: The number of training times based on the relu function

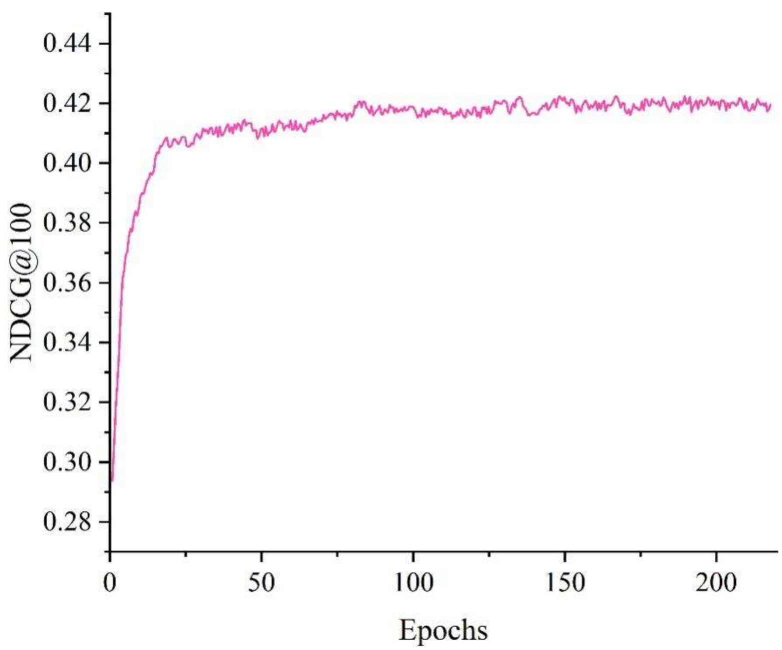

Figure 2: The number of training times based on the tanh function

In order to be able to train better models and improve efficiency, experiments are conducted to study the effect of the number of training times on the experimental results, as well as to explore the number of training times that have the best results, experiments are conducted on datasets with different sparsity, and the results of the VAE-GAN-DCR experiments are shown in Fig. 3. It can be seen that the RMSE value decreases as the number of times increases, while after reaching 100 iterations, the RMSE value starts to level off, and it is already close to smooth at 170-250 times. It can be concluded that during model training, between 170-250 times are used to achieve better results. Therefore, the number of epochs is set to 170 times in this chapter.

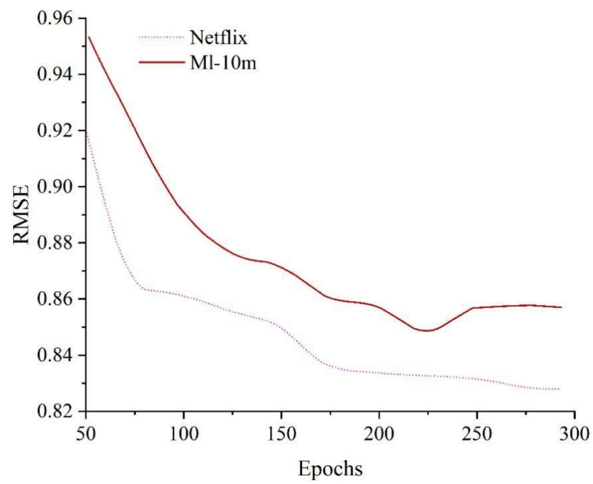

Figure 3: RMSE and the number of training times

In order to better compare the experimental results, this chapter uses ranking-based Top-N recommendations to evaluate the recommendation results. Recall is used to explore the effect of recommendation list length on recommendation accuracy. The comparison between the NMF model and the VAE-GAN-DCR model with different recommendation list lengths is explored at intervals of 10, as shown in Figure 4. From the figure, it can be concluded that the recall increases continuously with the list length and the recommendation accuracy is higher after 40 recommendation list length.

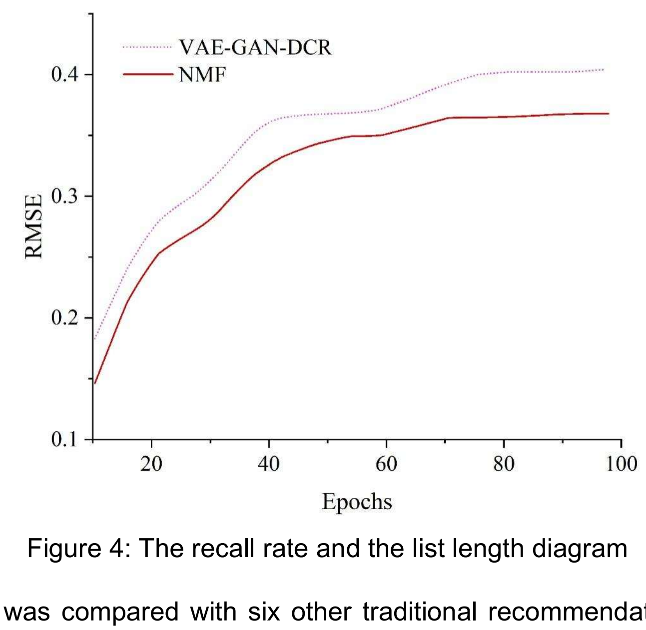

Figure 4: The recall rate and the list length diagram

Then, the VAE-GAN-DCR was compared with six other traditional recommendation algorithms on Recall and NDCG. Table 1 shows the experimental results of six recommendation algorithms on the Movielens-1M dataset. As can be seen from the table, compared with other recommendation algorithms, the performance of the recommendation algorithm of VAE-GAN-DCR based on the deep composite recommendation model based on variational autoencoder is improved, and compared with other models, the Recall@20 value is increased by 12.15%

on average, the Recall@50 value is increased by 6.11% on average, and the NDCG@100 value is increased by 12.94% on average.

Table 1: Performance comparison of six recommended algorithms

Model	Recall@20	Recall@50	Recall@100
NMF	0.24925	0.32914	0.30826
NCF	0.26751	0.39235	0.37291
ReD	0.33013	0.43473	0.44513
CVAE	0.34292	0.43794	0.41397
TrustSVD	0.35083	0.45916	0.45834
VAE-GAN-DCR	0.37835	0.47197	0.47151

Table 2 shows the experimental results of these three evaluation indicators on the above recommendation algorithm under the Movielens-10M dataset. Compared with other models, the algorithm in this chapter has a certain improvement in each evaluation index, with an average increase of 18.22% in Recall@20, 9.34% in Recall@50, and 17.96% in NDCG@100, which fully verifies the effectiveness and feasibility of the recommendation algorithm proposed in this chapter.

Table 2: Performance comparison of six recommended algorithms

Model	Recall@20	Recall@50	Recall@100
NMF	0.31646	0.35793	0.30571
NCF	0.35297	0.41943	0.38459
ReD	0.36344	0.43273	0.42521
CVAE	0.38202	0.40641	0.40254
TrustSVD	0.39424	0.43782	0.44941
VAE-GAN-DCR	0.41032	0.45971	0.44637

Based on the Netflix dataset, the experimental results of these three evaluation indicators on the above recommendation algorithm. Compared with other models, Table 3 shows that the Recall@20 value increased by 8.68%, the Recall@50 value increased by 7.94%, and the NDCG@100 value increased by 15.29% on average. Based on the above three tables, it is shown that the model in this chapter has been improved to a certain extent in different datasets and evaluation indicators, which fully verifies the recommendation accuracy of the deep composite recommendation model VAE-GAN-DCR based on variational autoencoder.

Table 3: Performance comparison of six recommended algorithms

Model	Recall@20	Recall@50	Recall@100
NMF	0.34021	0.42735	0.37292
NCF	0.36537	0.45506	0.42016
ReD	0.36929	0.46563	0.43683
CVAE	0.37423	0.47679	0.44874
TrustSVD	0.38137	0.47468	0.45686
VAE-GAN-DCR	0.38821	0.48893	0.47688

III. B. Analysis of the Educational Effectiveness of Applying Generative Artificial Intelligence

III. B. 1) Experimental design

Taking Zhanjiang Preschool Normal College as an example, students of different majors and different grades taught by the same teacher in the same semester, with an average age of 21 years old, were divided into two groups, A and B. In terms of the content of the teaching module, the teaching of "Robotics" course was mainly carried out for Group A, and the teaching of "Creative Programming" course was carried out for Group B. The teaching of "Robotics" course refers to the process of teaching with the help of the Misiqi development board combined with programming, 3D printing and other contents, mainly group collaboration; The "Creative Programming" course refers to the program teaching based on the Scratch platform, such as simple games, motion drawing, character modeling, variables, etc., which is mainly based on individual learning.

III. B. 2) Experimental methods

This study mainly used the questionnaire method to conduct pre- and post-teaching measurement experiments, the questionnaire is from Williams Creative Tendency Measurement Scale, which is mainly used to measure the level of creative tendency of students in groups A and B before and after the administration of teaching. In the process of filling out the questionnaire, students are easily affected by the different teaching methods of teachers, the different knowledge levels of students, the psychological state when filling out the questionnaire, the different specialties, the different grades, the teaching time and other factors. For this reason, this study tries to avoid the influence of some interfering factors, being able to achieve the same teacher teaching the control group and the experimental group; the two groups coming from the same class and the same specialty; the control group and the experimental group using the same measurement tools; and determining the main modules according to the teaching plan of the instructor, according to the percentage of the course content, and so on.

The Williams Creative Tendencies Measurement Scale has 50 questions and includes four items: adventurous, curious, imaginative, and challenging. Four scores can be obtained after the test, and the total score of the level of creative tendency is the sum of the four scores. The scale has three options for each question: “not at all”, “partially”, and “completely”, with scores of 1, 2, and 3 respectively, and the higher the total score, the higher the level of creativity. The higher the total score, the higher the level of creativity. The overall level of adventurousness, curiosity, imagination, challenge, and creativity is categorized into four levels: very weak, weak, average, and very strong in order to include the distribution of all scores.

Table 4: Test of independent sample t measured by the experimental group and the control group

Variable term	Group	Average	Standard deviation	T value	P value
Inventory of total	Experimental group	8.44	0.95	-1.23	0.267
	Control group	8.72	1.13		
Creative thinking	Experimental group	14.39	2.02	-1.04	0.319
	Control group	14.89	2.22		
Fluid force	Experimental group	12.43	0.91	0.61	0.568
	Control group	12.24	1.72		
Openness	Experimental group	24.64	4.17	-0.79	0.457
	Control group	25.44	4.46		
Workareness	Experimental group	8.21	1.67	0.49	0.651
	Control group	8.05	1.43		
Originality	Experimental group	13.52	4.26	-0.51	0.629
	Control group	14.02	4.02		
Precision	Experimental group	11	4.26	-1.73	0.094
	Control group	12.63	3.67		
Title	Experimental group	16.21	4.84	-0.71	0.505
	Control group	16.95	4.01		
Creative tendency	Experimental group	2.5	0.35	-1.15	0.294
	Control group	2.56	0.32		
Risk	Experimental group	2.57	0.35	-1.53	0.156
	Control group	2.66	0.34		
Curiosity	Experimental group	2.42	0.46	0.04	0.957
	Control group	2.42	0.37		
Imagination	Experimental group	2.59	0.47	-1.32	0.198
	Control group	2.71	0.45		
Challenge	Experimental group	2.42	0.37	-0.65	0.529
	Control group	2.46	0.41		

III. B. 3) Analysis of results

(1) Sample homogeneity test of the pre-test of the Creative Power Scale for two groups of college students

The summary of the independent samples t-test of the pre-test of the Creative Power Scale is shown in Table 4, which shows that the pre-tests of the experimental group and the control group did not reach the level of significance in the creative thinking activities (fluency, openness, variability, originality, precision and title), creative tendencies (adventurousness, curiosity, imagination, and challenge), and the pre-tests of the Total Creative Power Scale (t=-

1.23, -1.04, 0.61, -0.79, 0.49, -0.51, -1.73, -0.71, -1.15, -1.53, 0.04, -1.32, -0.65, p>0.05). The above results showed that the experimental and control groups possessed homogeneity in the creativity scale, which can be further analyzed in experimental teaching.

(2) The effect of applying generative artificial intelligence on college students' creative thinking activities

The mean and standard deviation of the pre and post-test raw scores of the experimental group and the control group on the Creative Thinking Activity Scale are presented as shown in Table 5 to serve as the basic information for further analysis. Table 5 shows that the experimental group showed an increase in the post-test scores of fluency, openness, adaptability, originality and title, and a decrease in the post-test scores of total creative thinking activities and precision, while the control group showed a decrease in the post-test scores of total creative thinking activities, fluency, openness, adaptability, originality, precision and title.

Table 5: The average and standard deviation of the experimental group and the control group

Variable term	group	Presurvey mean	Standard deviation	Backmeasured mean	Posterior standard deviation
Creative thinking	Experimental group	14.84	2.08	14.95	2.62
	Control group	15.34	2.28	13.34	3.04
Fluid force	Experimental group	12.88	1.1	13.17	1.18
	Control group	12.69	1.78	12.04	2.67
openness	Experimental group	25.09	4.23	25.9	6.83
	Control group	25.89	4.52	23.33	6.77
Workareness	Experimental group	8.66	1.73	9.35	2.45
	Control group	8.5	1.49	8.26	1.83
Originality	Experimental group	13.97	4.32	13.96	4.49
	Control group	14.47	4.08	11.88	3.8
Precision	Experimental group	11.45	4.32	10.44	5.48
	Control group	13.08	3.73	8	3.89
Title	Experimental group	16.66	4.9	16.9	4.37
	Control group	17.4	4.07	16.55	4.74

Table 6: Independent sample t test of the experimental group and the control group's creative thinking activity

Variable term	Group	Average	Standard deviation	T value	P value
Creative thinking	Experimental group	14.96	2.64	2.45*	0.021
	Control group	13.35	3.06		
Fluid force	Experimental group	13.18	1.2	2.37*	0.022
	Control group	12.05	2.69		
Openness	Experimental group	25.91	6.85	1.57	0.131
	Control group	23.34	6.79		
Workareness	Experimental group	9.36	2.47	2.18*	0.039
	Control group	8.27	1.85		
Originality	Experimental group	13.97	4.51	2.11*	0.045
	Control group	11.89	3.82		
Precision	Experimental group	10.45	5.5	2.15*	0.042
	Control group	8.01	3.91		
Title	Experimental group	16.91	4.39	0.33	0.761
	Control group	16.56	4.76		

The independent samples t-test of the post-test of creative thinking activities is shown in Table 6, in which the experimental group and the control group reached the level of significant difference in the total scale of creative thinking activities, fluency, openness, adaptability, originality, and sophistication ($t = 2.45, 2.37, 2.18, 2.11, 2.15$, $p < 0.05$). Further statistical results found that the experimental group performed significantly better than the control group in terms of total creative thinking activity scale, fluency, variability, originality and precision (mean of the experimental group 14.96, 13.18, 9.36, 13.97, 10.45, and the mean of the control group 13.35, 12.05, 8.27, 11.89, 8.01), which means that The phenomenon of difference between the experimental and control groups in terms of the total creative thinking activity scale, fluency, variability, originality and sophistication can be affected by the teaching strategies. In addition, the experimental and control groups did not reach the level of significant difference in openness and titles (t=1.57, 0.33, p>0.05), indicating that there is no significant difference between the experimental and control groups in terms of the impact of teaching strategy intervention on openness and titles.

(3) The effect of applying generative AI on the creative tendency of college students

Table 7 shows the mean and standard deviation of the raw scores of the experimental group and the control group in the pre- and post-test of the Creative Tendency Scale as the basic information for further analysis. As can be seen from the table, the experimental group showed an increase in the post-test scores of total creative tendency scale, adventurousness, curiosity, imagination and challenge, while the control group showed an increase in the post-test scores of total creative tendency scale, curiosity and challenge, and a decrease in the post-test scores of adventurousness and imagination.

Table 7: The average and standard deviation of the experimental group and the control group

Variable term	Group	Presurvey mean	Standard deviation	Backmeasured mean	Posterior standard deviation
Creative tendency	Experimental group	1.95	0.31	2.18	0.46
	Control group	2.01	0.28	1.91	0.42
Risk	Experimental group	2.02	0.31	2.35	0.57
	Control group	2.11	0.30	1.86	0.47
Curiosity	Experimental group	1.87	0.42	2.18	0.62
	Control group	1.87	0.33	2.02	0.48
Imagination	Experimental group	2.04	0.43	2.3	0.56
	Control group	2.16	0.41	1.83	0.42
Challenge	Experimental group	1.87	0.33	1.88	0.51
	Control group	1.91	0.36	1.92	0.39

Table 8 shows the summary table of the independent samples t-test of the post-test of creative tendencies, in which the experimental and control groups reached the level of significant difference in terms of the total scale of creative tendencies, adventurousness, and imagination (t=3.35, 4.73, and 4.85, p<0.05). In addition, the statistical results also pointed out that the experimental group performed significantly better than the control group in terms of the total scale of creative tendency, adventurousness, and imagination (the mean of the experimental group was 2.61, 2.77, and 2.72, while the mean of the control group was 2.33, 2.28, and 2.25), which indicated that the total scale of creative tendency, adventurousness, and imagination of the experimental group and the control group would be affected by the application of the generative Artificial Intelligence to produce the phenomenon of difference. In addition, there is no significant difference between the experimental and control groups in terms of curiosity and challenge (t=1.45, -0.45, p>0.05), which means that the effect of applying generative AI intervention on the curiosity and challenge of the experimental and control groups is not significant.

Table 8: Independent sample t test of the experimental group and the control group

Variable term	Group	Presurvey mean	Standard deviation	Backmeasured mean	Posterior standard deviation
Creative tendency	Experimental group	2.61	0.43	3.35*	0.003
	Control group	2.33	0.39		
Risk	Experimental group	2.77	0.54	4.73*	0.001
	Control group	2.28	0.44		
Curiosity	Experimental group	2.63	0.59	1.45	0.157
	Control group	2.44	0.45		
Imagination	Experimental group	2.72	0.53	4.85*	0.000
	Control group	2.25	0.39		
Challenge	Experimental group	2.35	0.48	-0.45	0.641
	Control group	2.34	0.36		

IV. Conclusion

The deep composite recommendation model based on variational autoencoder and generative adversarial network shows significant advantages in the field of educational resources recommendation. Experimental results show that the VAE-GAN-DCR model surpasses the traditional recommendation algorithm in multiple evaluation indicators, and the Recall@50 value on the Netflix dataset is increased by 7.94% on average, which fully verifies the effectiveness of the model. By introducing the prior distribution that depends on item features and the feature transfer mechanism of GAN discriminator, the model successfully solves the data sparsity and reconfiguration error optimization problems faced by traditional recommender systems. The educational application experiments further confirm the value of generative AI technology in smart classroom construction. Comparative analysis shows that students in the experimental group applying the technology excel in creative thinking activities, in which the key indicators of adaptability, originality and precision reach significant levels, and the score of the total creative tendency scale is 2.61, which is significantly higher than that of the control group, which is 2.33. This result shows that generative AI can not only improve the accuracy of educational resources recommendation, but also effectively stimulate students' innovative thinking and exploration spirit. The model shows good adaptability in dealing with the complexity and diversity of educational data, which provides a technical guarantee for personalized learning and accurate teaching.

Funding

This work is supported by Anhui Provincial Scientific Research Plan Preparation Project (2024AH052311). This work is supported by 2023 China University Industry-University-Research Innovation Fund (2023RY026).

References

[1] Vig, S. (2024). Intersection of generative artificial intelligence and copyright: an Indian perspective. Journal of Science and Technology Policy Management.

[2] Koh, E., & Doroudi, S. (2023). Learning, teaching, and assessment with generative artificial intelligence: towards a plateau of productivity. Learning: Research and Practice, 9(2), 109-116.

[3] Dimitrieska, S. (2024). Generative artificial intelligence and advertising. Trends in economics, finance and management journal, 6(1), 23-34.

[4] Shukla, S. (2023). Creative computing and harnessing the power of generative artificial intelligence. Available at SSRN 5173567.

[5] Dwivedi, R., & Elluri, L. (2024). Exploring Generative Artificial Intelligence Research: A Bibliometric Analysis Approach. IEEE Access.

[6] Christensen, J., Hansen, J. M., & Wilson, P. (2025). Understanding the role and impact of Generative Artificial Intelligence (AI) hallucination within consumers' tourism decision-making processes. Current Issues in Tourism, 28(4), 545-560.

[7] Leaver, T., & Srdarov, S. (2023). ChatGPT isn't magic: The hype and hypocrisy of Generative Artificial Intelligence (AI) rhetoric. M/c Journal, 26(5).

[8] Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. Journal of science education and technology, 32(3), 444-452.

[9] Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Journal of AI, 7(1), 52-62.

[10] Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. Sustainability, 15(17), 12983.

[11] Preiksaitis, C., & Rose, C. (2023). Opportunities, challenges, and future directions of generative artificial intelligence in medical education: scoping review. JMIR medical education, 9, e48785.

[12] Hwang, G. J., & Chen, N. S. (2023). Exploring the potential of generative artificial intelligence in education: applications, challenges, and future research directions. Journal of Educational Technology & Society, 26(2).

[13] Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. Education Sciences, 13(9), 856.

[14] Holmes, W., & Miao, F. (2023). Guidance for generative AI in education and research. UNESCO Publishing.

[15] Escotet, M. Á. (2024). The optimistic future of Artificial Intelligence in higher education. Prospects, 54(3), 531-540.

[16] Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. ECNU Review of Education, 6(3), 355-366.

[17] Ahmad, N., Murugesan, S., & Kshetri, N. (2023). Generative artificial intelligence and the education sector. Computer, 56(6), 72-76.

[18] Kaixuan Ju,Tingting Zhao,Yuting Luo,Xiaomin Ma,Jiaying Liu,Yuntian Feng & Zhihua Wang. (2025). Force chain generation and evaluation of granular materials based on discrete element simulation and generative adversarial network. Powder Technology,461,121100-121100.

[19] Wenbo Li,Yang Yang Dong & Chunxi Dong. (2025). Radar working mode amplitude waveform augmentation based on generative adversarial networks. Neurocomputing,639,130285-130285.