# Research of Art Point of Interest Recommendation Algorithm Based on Modified VGG-16 Network

Yi Liu*

Department of Arts and Sports, Henan Technical College of Construction, Zhengzhou 450000, China
byoungholee@qq.com

**Abstract**. Traditional point of interest (POI) recommendation algorithms ignore the semantic context of comment information. Integrating convolutional neural networks into recommendation systems has become one of the hotspots in art POI recommendation research area. To solve the above problems, this paper proposes a new art POI recommendation model based on improved VGG-16. Based on the original VGG-16, the improved VGG-16 method optimizes the fully connection layer and uses transfer learning to share the weight parameters of each layer in VGG-16 pre-training model for subsequent training. The new model fuses the review information and user check-in information to improve the performance of POI recommendation. Experiments on real check-in data sets show that the proposed model has better recommendation performance than other advanced points of interest recommendation methods.

**Keywords:** art point of interest, VGG-16, transfer learning, recommendation algorithm

## 1 Introduction

One obvious advance in social networking applications in recent years has been the introduction of spatial technology [1], which has facilitated the flourishing of location-based social networks (LBSN) such as Foursquare, Twinkle and Geolife. In LBSN, locations based on spatial technology are also called point-of-interest (POI), such as restaurants, shops and museums [2, 3]. However, the interest point recommendation faces many problems, such as sparse matrix, low recommendation accuracy and low model performance. Selecting the appropriate context information and recommendation method is an important to improve the recommendation accuracy and better mining user preferences [4]. In LBSN, point of interest recommendation integrates multiple context information. Liu et al. [5] proposed a SocialMF algorithm to decompose the fusion matrix of social relations and integrate the trust transmission mechanism, thus alleviating the problem of cold start. Zhao et al. [6] incorporated L1 regularization into the loss function of ListMLE algorithm, alleviating the sparse problem of sorted lists. Zhang et al. [7] explored the list-level sorting learning algorithm based on Group, divided labels with high and low relevance into groups and combined with unsupervised learning to generate more accurate labels. Zhang et al. [8] explored user check-in distribution in Gaussian modeling and optimized Poisson matrix decomposition process by using BPR. Yin et al. [9] used Gaussian kernel density estimation to calculate the influence of users' geographical factors on the recommendation of interest points.

In recent years, some researchers have integrated ranking learning into the recommendation process. Sequencing learning is to use machine learning methods to sort items or points of interest and output a list of points of interest that conform to user preferences [10-12]. According to the different forms of training data, it can be divided into point-level sorting learning, pair level sorting learning and list-level sorting learning [13]. Yin et al. [14] used point-ranking learning method to mine implicit feature matrix based on user implicit feedback. Wang et al. [15] defined the loss function on the set of item pairs and learned the ordering model considering the partial order relationship between items that users tend to prefer. Fang et al. [16] used the list-level method (ListNet) to process the feature matrix and compare the differences in sorting performance. Yin et al. [17] applied list-level ranking to restaurant recommendation and used the ranking model to learn the results generated by the base model [18] to provide a query and recommendation scheme for vertical search. To solve the above problems, this paper proposes a new art POI recommendation model based on improved VGG-16.

This paper is organized as follows. Section 2 detailed introduces the proposed art POI recommendation method. Section 3 gives the experiments and analysis. There is a conclusion in section 4.

---

* Corresponding Author

## 2 Proposed Art POI Recommendation Algorithm

The presented model in this section is shown in Fig. 1. The dotted border on the left represents the check-in preprocessing component. The dotted border on the right represents the comment information feature learning component. Input $(\tilde{u}, \tilde{i}, x)$ triples, in here, $\tilde{u} = [u_1, u_2, \cdots, u_m]$ denotes user set. $\tilde{i} = [i_1, i_2, \cdots, i_n]$ denotes a collection point of interest. $x = [x_1, x_2, \cdots, x_n]$ denotes the comments information collection. In particular, initialization parameters of user $\tilde{u}$ and POI $\tilde{i}$ are learned by SDAE. It gets the optimized user check-in feature $u_{L/2}$ and interest point check-in feature $i_{L/2}$. Potential feature vectors are obtained by improving CNN network to learn the comment information. Then, the feature of POI $v$ is obtained by integrating the check-in feature $i_{L/2}$ and the potential feature $z$ of POI, so the score of POI is predicted. The learning process for each component of the model is described in detail below.
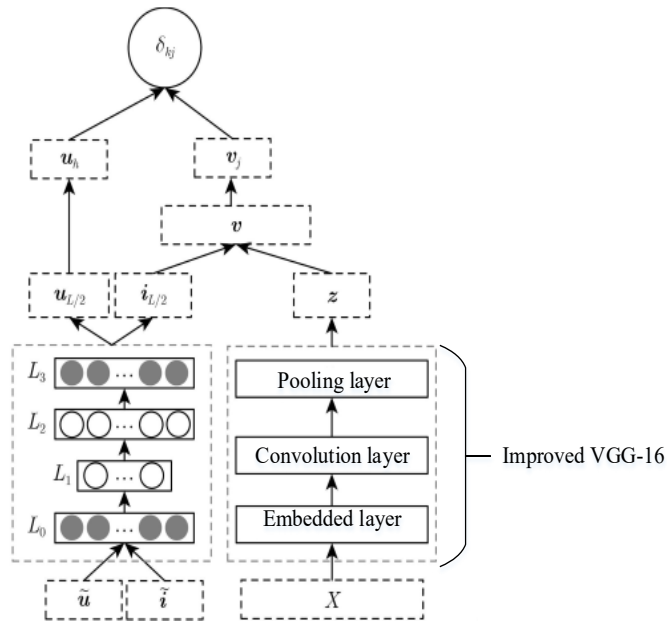


**Fig. 1.** Proposed POI system

### 2.1 Improved VGG-16 Network

VGGNet explores the relationship between the recognition performance and the depth of the network model on the basis of AlexNet [19]. In the process of exploration, it is found that smaller convolution kernel and deeper network can significantly improve the classification accuracy of the model, resulting in a significant decrease in the error recognition rate. VGGNet uses a 3×3 convolution kernel and deepens the layers of the model. The field of two 3×3 convolution kernels stacked together is equivalent to that of a 5×5 convolution kernel. The field of three 3×3 convolution kernels stacked together is equivalent to a 7×7 convolution kernel. Using a 3×3 convolution kernel in the same field can reduce the number of parameters. At the same time, a deeper network means that there will be more nonlinear transformations, and the learning ability of the network will be stronger and more features can be learned. VGGNet also adopts the multi-scale training method to increase the amount of data during training, preventing over-fitting and improving accuracy [13]. Based on the uniqueness of VGGNet network structure and its excellent performance in public data sets, VGG-16 network is selected as the feature extraction network to solve the recommendation of POI. The network structure is shown in Fig. 2.
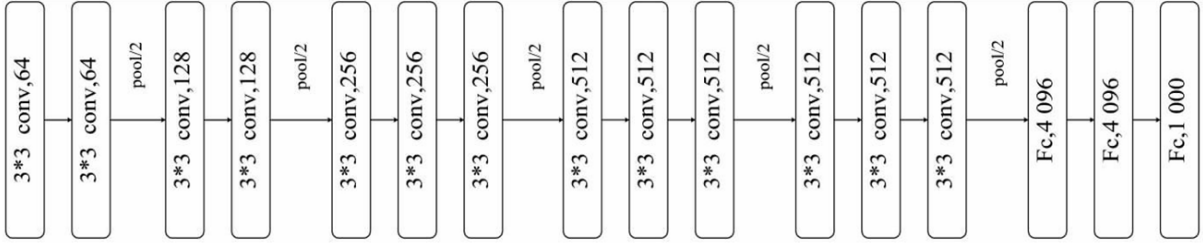
**Fig. 2.** Network of VGG-16

An improved VGG-16 model is designed to meet the requirements of accuracy and computational efficiency of POI recommendation. At the same time, the accuracy of the model is guaranteed and the computational cost of the model is reduced. The recommendation model of POI constructed in this paper still uses the convolution layer of the standard VGG-16 network structure. Sixteen convolution layers are separated by five pooling layers, and the input data can be extracted from the calculation of sixteen convolution layers and five pooling layers. The last full connection layer of the model simulates the logical cognitive process of the brain, and classifies the features extracted from the convolutional pooling layer to obtain the recognition tags [20]. Because the layers and parameters of VGG-16 model are designed for 1000 classification categories, the number of weight parameters is $65 \times 10^6$, and most of the parameters are distributed on the fully connected layer. The problem to be solved in this paper is the classification and recognition of the four states, which does not need so many weight parameters to participate in the calculation. Therefore, an improved method is proposed to replace the original three fully connection layers with two fully connection layers to design 4096 neurons in the first fully connected layer. The number of neurons in the second fully connection layer, namely the output layer, is four, which reduces the number and complexity of model parameters to improve model recognition accuracy and computational efficiency.

The designed model is shown in Fig. 3. All convolution computations use a convolution kernel with a size of $3 \times 3$, and the convolution step size is fixed as 1. All pooling operations use a maximum pooling of $2 \times 2$, and the pooling step size is fixed as 2. A fully connection layer is removed so that convolutional neural network can save a lot of computation and memory. Relu function is used as the activation function of learning features, which can make the acquired features more obvious and achieve better classification and recognition effect.
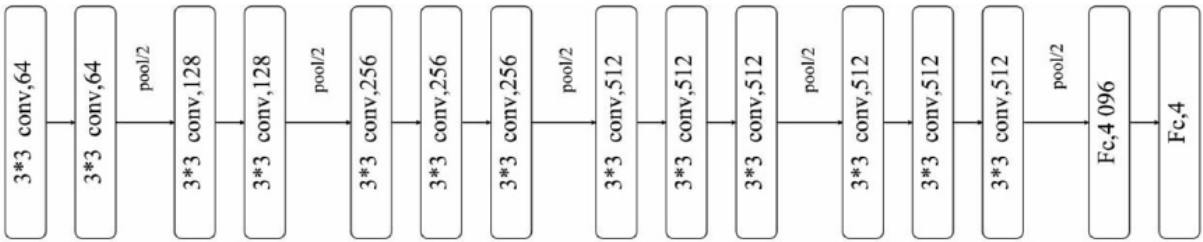


**Fig. 3.** Improved VGG-16

## 2.2 Comment Information Feature Learning Component

Convolutional layer: the set of comment information of any interest point is expressed as $x_i = \{w_1, w_2, \cdots, w_n\}$. The word vector model is used to map each word $w_i$ in $x_i$ to the corresponding word vector $w'_i \in R^{p \times 1}$ according to the order of occurrence. $x_i$ is converted to the word vector matrix $D \in R^{p \times l}$, which is invariant in word order.

$$D = [w'_0 \ w'_1 \cdots w'_{n-1}]. \qquad (1)$$

Where, D and $p$ represent the embedded dimension of the word $w_i$. $w'_i$ represents the word vector of the

words at point $i$ in $x_i$. $l$ is the length of $x_i$. Assuming that the size of the convolution window is $wd$, the context feature $D_{(:,i:(i+wd-1))}$ is obtained by convolving the j-th shared weight $W_c^j \in R^{p \times wd}$ with the content D of the $x_i$ through the convolution window. Namely,

$$c_i^j = f(W_c^j * D_{(:,i:(i+wd-1))} + b_c^j) \ . \tag{2}$$

Where, * represents the convolution operation. $b_c^j$ represents the bias term, and $f$ represents the activation function. The feature vector $c^j \in R^{l-wd+1}$ is obtained after multiple convolution, which is expressed as:

$$c^j = [c_1^j \ c_2^j \ \cdots \ c_{l-wd+1}^j] \ . \tag{3}$$

Pooling layer: Extracting the largest feature vector from the context feature vector represents the potential feature of the interest point. The contextual feature of the comment information $x_i$ of a point of interest is expressed as $d_j$, i.e.,

$$d_j = [\max(c_j^1), \max(c_j^2), \cdots, \max(c_j^{n_c})] \ . \tag{4}$$

Output layer: Feature vectors extracted from the pooling layer are mapped to vector spaces of different dimensions for different tasks. Here, feature vector $d_j$ is mapped to $k_d$ dimensional vector space for the recommendation task in this paper.

$$z = \tanh(W_{f_2}\{\tanh(W_{f_1}d_j + b_{f_1})\} + b_{f_2}) \ . \tag{5}$$

Where $z \in R^{k_d}$, $W_{f_1} \in R^{f \times n_c}$, $W_{f_2} \in R^{k_d \times f}$ are mapping matrices. $b_{f_1} \in R^f$, $b_{f_2} \in R^{k_d}$ are offset vectors. Finally, the context features of comment information $x_i$ of POI are extracted. Therefore, in the improved CNN for context features of learning POI, the objective function is expressed as:

$$\Re_{cnn} = \sum_i^n \| v_i - cnn(W, x_i)) \|^2 + \lambda_W^{cnn} \sum_i^{|W|} \| W_i \|^2 \ . \tag{6}$$

## 2.3 Sign-in Information Preprocessing Component

Given input $S^u$ and $S^v$ represent user preference feature and interest point feature respectively. The preference feature of each user is expressed as $s_i^u = [f_{i,1}, f_{i,2}, \cdots, f_{i,n}]^T$, and the feature of each interest point is expressed as $s_i^v = [f_{1,i}, f_{2,i}, \cdots, f_{m,i}]^T$. The denoising auto-encoder randomly destroys $S^u$ and $S^v$ to obtain $\widetilde{S}^u$ and $\widetilde{S}^v$. After the destruction, the preference feature of each user and the feature of each interest point are expressed as respectively: $\widetilde{s}_i^u = [\widetilde{f}_{i,1}, \widetilde{f}_{i,2}, \cdots, \widetilde{f}_{i,n}]^T$, $\widetilde{s}_i^v = [\widetilde{f}_{1,i}, \widetilde{f}_{2,i}, \cdots, \widetilde{f}_{m,i}]^T$. The specific process can be expressed as:

$$\left.\begin{aligned} h_u &= g(W_1\tilde{s}^u + b_u) \\ h_v &= g(V_1\tilde{s}^v + b_v) \\ \hat{s}^u &= f(W_2 h_u + b_{\hat{u}}) \\ \hat{s}^v &= f(V_2 h_v + b_{\hat{v}}) \end{aligned}\right\}.$$
(7)

Where $\tilde{s}^u$ and $\tilde{s}^v$ denote the destroyed $s^u$ and $s^v$. $\hat{s}^u$ and $\hat{s}^v$ denote the constructed $s^u$ and $s^v$. $h_u$ and $h_v$ represent potential representations of the input. W and V represent the weight matrix. $b$ is the offset vector. $g(\cdot)$ and $f(\cdot)$ represent activation functions. As shown in Fig. 1, multiple hidden layers of the stack denoising auto-encoder can be represented as:

$$h_l = g(W_l h_{l-1} + b_l).$$
(8)

Where $h_0$ is the destruction input, $l \in \{1,2,\cdots,L-1\}$. For inputs $\tilde{s}^u$ and $\tilde{s}^v$, the output of L-th layer is:

$$\left.\begin{aligned} \hat{s}^u &= f(W_L h_L + b_{\hat{u}}) \\ \hat{s}^v &= f(V_L h_L + b_{\hat{v}}) \end{aligned}\right\}.$$
(9)

The front L/2 layer acts as the encoder and the back L/2 layer acts as the decoder. Therefore, the loss function of reconstructed data is:

$$\begin{aligned} \arg\min_{W_l,V_l,b_l} & \parallel S^u - \hat{S}^u \parallel_F^2 + \parallel S^v - \hat{S}^v \parallel_F^2 \\ &+ \lambda_U (\sum_l \parallel W_l \parallel_F^2 + \parallel b_l \parallel_F^2) \\ &+ \lambda_V (\sum_l \parallel V_l \parallel_F^2 + \parallel b_l \parallel_F^2) \end{aligned}.$$
(10)

Where, $W_l$ and $V_l$ represent the weight matrix. $b_l$ is the offset vector. $\lambda_U$ and $\lambda_V$ denote the regularization parameters. Therefore, the final objective function of the proposed model is:

$$\begin{aligned} \Re = &-\sum_{hj}(R_{hj} - u_h^T v_j)^2 - \lambda_m^h \sum_{u=1}^{m}(s_h^u - \hat{s}_h^u)^2 \\ &-\lambda_n^j \sum_{j=1}^{n}(s_j^v - \hat{s}_j^u)^2 - \lambda_w^h \sum_h (\parallel W_l \parallel_F^2 + \parallel b_l \parallel_F^2) \\ &-\lambda_w^j \sum_h (\parallel V_l \parallel_F^2 + \parallel b_l \parallel_F^2) \\ &-\lambda_{cnn} \sum_{j=1}^{n} \parallel (\tilde{v}_j - cnn(W,x_j)) \parallel_F^2 \\ &-\lambda_w^{cnn} \sum_{j=1}^{|W_{cnn}|} \parallel W_{cnn} \parallel_F^2 - \lambda_h \sum_{h=1}^{m} \parallel u_h \parallel_F^2 \\ &-\lambda_j \sum_{j=1}^{n} \parallel v_j \parallel_F^2 \end{aligned}.$$
(11)

## 2.4 Parameter Learning

For the update of parameter $u_h$ and $v_j$, this paper uses the stochastic gradient descent method to learn the parameter. $\Re(U,V)$ represents the objective function. Assume that variables independent of $u_h$ and $v_j$ are constant. The update rule is:

$$\left. \begin{aligned} u_h &= u_h - \eta \frac{\partial}{\partial u_h} \Re(U,V) \\ v_j &= v_j - \eta \frac{\partial}{\partial v_j} \Re(U,V) \end{aligned} \right\} . \tag{12}$$

Where $\eta$ is the learning rate, and the gradient update result is:

$$\begin{aligned} \frac{\partial}{\partial u_h} \Re(U,V) &= -\sum_{hj} (R_{hj} - u_h^T v_j)(-v_j) \\ &\quad - \lambda_m^h \sum_{u=1}^{m} (s_h^u - \hat{s}_h^u) \frac{\partial \hat{s}_h^u}{\partial u_h} - \lambda_h u_h . \end{aligned} \tag{13}$$

$$\begin{aligned} \frac{\partial}{\partial v_j} \Re(U,V) &= -\sum_{hj} (R_{hj} - u_h^T v_j)(-u_h^T) \\ &= -\lambda_n^j \sum_{j=1}^{n} (s_j^i - \hat{s}_j^i) \frac{\partial \hat{s}_j^i}{\partial v_j} \\ &= -\lambda_{cnn} \sum_{j=1}^{n} (\tilde{v}_j - cnn(W, x_j)) \frac{\partial \tilde{v}_j}{\partial v_j} - \lambda_j v_j . \end{aligned} \tag{14}$$

By learning the potential feature representation of users and interest points, the potential feature representation and preference prediction of interest points are:

$$\hat{R}_{u,i} \approx (u_u)^T v_i = (u_u)^T (cnn(W, x_i) + s_{L/2}^i) . \tag{15}$$

The learning process of the proposed model is shown in Algorithm 1.

---

**Algorithm 1.** Proposed POI recommendation

---

**Input**: $x_i$, $S^u$, $S^v$, $\widetilde{S}^u$, $\widetilde{S}^v$, T, B

**Output**: $\mathfrak{R}$

(1) For $t < T$ do

(2) A comment matrix $x_i$ of interest points is randomly selected from the comments of POI for training, and the training batch size is $\beta_0$. The size of each batch is B. Calculating the $\mathfrak{R}_{cnn}$ loss during training.

(3) If $t > T$ or $\mathfrak{R}_{cnn}$ is enough small

(4) End

(5) For $t < T$ do

(6) The user-interest point pair $(\widetilde{s}_i^u, \widetilde{s}_i^v)$ of an interest point is randomly selected from the interest point score for training, and the training batch size is $\beta_1$. The size of each batch is B. Calculating the $\mathfrak{R}_{ui}$ loss during training.

(7) If $t > T$ or $\mathfrak{R}_{ui}$ is enough small.

(8) End

(9) Calculating the final loss value $\mathfrak{R} = \mathfrak{R}_{cnn} = \mathfrak{R}_{ui}$.

(10) Return $\mathfrak{R}$.

---

### 2.5 Complexity Analysis

When updating the potential feature vector of the convolutional weight learning text, the complexity of the convolutional neural network is $O(n_c plN)$. Where N represents the number of comment information of POI. $p$ represents the embedded dimension. $l$ indicates the length of a comment message for a POI. $n_c$ represents the number of learned context features. The complexity of the encoder updating users and points of interest is $O(2\eta k_1)$. Where, $\eta$ is the size of vocabulary, $k_1$ is the potential dimension of output. The complexity of update weights and deviations is $O(m\eta k_1 + n\eta k_1)$. Thus, the total complexity is $O(m\eta k_1 + n\eta k_1 + n_c plN + 2\eta k_1)$.

## 3 Experiments and Analysis

This paper uses two public Foursquare datasets to verify the validity of the proposed model. The Foursquare dataset is a tally of check-ins in Los Angeles (LA) and New York (NYC), United States, as shown in Table 1. For sign-in (comment), the method in reference [21] is adopted to preprocess text information.

**Table 1.** Data set collection

| Data set | LA | NYC |
|---|---|---|
| User number | 30208 | 47240 |
| POI number | 142798 | 203765 |
| Number of check-ins (comments) | 244861 | 388954 |
| User-location matrix density | $5.68 \times 10^{-5}$ | $4.04 \times 10^{-5}$ |

### 3.1 Evaluation Index

In this paper, accuracy rate and recall rate are used as evaluation indexes for location recommendation to evaluate the performance of the recommendation algorithm, represented by P@k and RC@k, respectively. For a user $u$, $dl$ represents the number of observed points of interest. $vl$ represents the number of POI that have been accessed, P@k and RC@k are defined as:

$$P@k = dl/k \ . \tag{16}$$

$$RC@k = dl/vl \ . \tag{17}$$

Where, the performance of the model under different parameters is evaluated for several times. In the experiment, P@1, P@5 and P010, as well as RC@1, RC@5, and RC@10, are selected as evaluation indicators.

### 3.2 Experimental Design

In this paper, three different strategies are used to verify the effectiveness of the proposed algorithm.

(1) Comparison with mainstream advanced algorithms. This paper compares with four advanced mainstream point of interest recommendation algorithms to verify the efficiency of the algorithm.

(2) Pre-processing test for potential features. The effectiveness of the proposed initialization method is verified by preprocessing the potential features of matrix interaction during matrix decomposition.

(3) Test the number of SDAE layers. Similar to reference [21], the influence of SDAE layers on model performance is verified.

Firstly, four points of interest recommendation algorithms are selected for comparison.

(1) IRenMF: Reference [22] proposed to use the geographical neighborhood relationship of two levels to conduct modeling of geographic location information. It assumed that the geographical neighborhood interest point of check-in was more suitable for recommending to users.

(2) ASMF: Reference [23] proposed a two-stage framework to recommend points of interest based on users' social information according to the category of friends (social friends, location friends and neighbor friends).

(3) LCARS: Reference [24] built a point of interest recommendation system based on the LDA topic model by utilizing users' check-in data and comment text information.

(4) CDL: A hierarchical Bayesian model was proposed in reference [25], which extracted the effective depth feature representation of content and captured the similarity and implicit relationship between objects and users. In this paper, the user check-in behavior is modeled based on this method.

Secondly, four classical initialization methods are compared with the proposed initialization method.

(1) Random initialization: Reference [18] proposed that the settings should be randomly based on users.

(2) Zero initialization: Reference [18] suggested that the initial value was set to zero.

(3) K-means initialization: Reference [22] proposed an initialization method based on k-means.

(4) Normalized-CUT initialization (NCUT): Reference [26] proposed a method based on NCUT for initialization.

Finally, the influence of different layers on the model recommendation performance in this paper is verified.

### 3.3 Parameter Setting

For different model parameters, they are set as:

In IRenMF: K=100, $\alpha$ =0.4, $\lambda_1 = \lambda_2 = 0.015$, $\lambda_3 = 1$, NN=10, clusters=50;

In ASMF: K=100, $\alpha$ =0.4, $\lambda_u = \lambda_v = 0.01$, $\lambda_q = 0.1$, $\varsigma = 0.4$, $\gamma = \varepsilon = 0.4$;

In LCARS: K=100, $\alpha = \alpha' = 50/K =$, $\beta = \beta' = 0.01$, $\gamma = \gamma' = 0.5$;

In the CDL with two layers: $\lambda_h = 0.1$, $\lambda_j = 1$, $\lambda_w^j = 0.0001$, $\lambda_n^j = 100$;

The setting of SDAE is same to the proposed method.

For the proposed model in this paper, SDAE takes a 30% noise wipe. The discard rate is 0.1 to achieve adaptive regularization and prevent overfitting. The number of hidden units K is 1000, and the number of intermediate layers is 200. The potential factors number is $|u_h| = |v_j| = |v_k| = 200$. The learning rate is 0.2. $\lambda_h = 0.01$, $\lambda_j = 0.1$, $\lambda_w^j = 0.0001$, $\lambda_n^j = 5$ can achieve good performance.

### 3.4 Result Analysis

The performance comparison between the proposed model in this paper and the existing mainstream advanced algorithm is shown in Fig. 4.
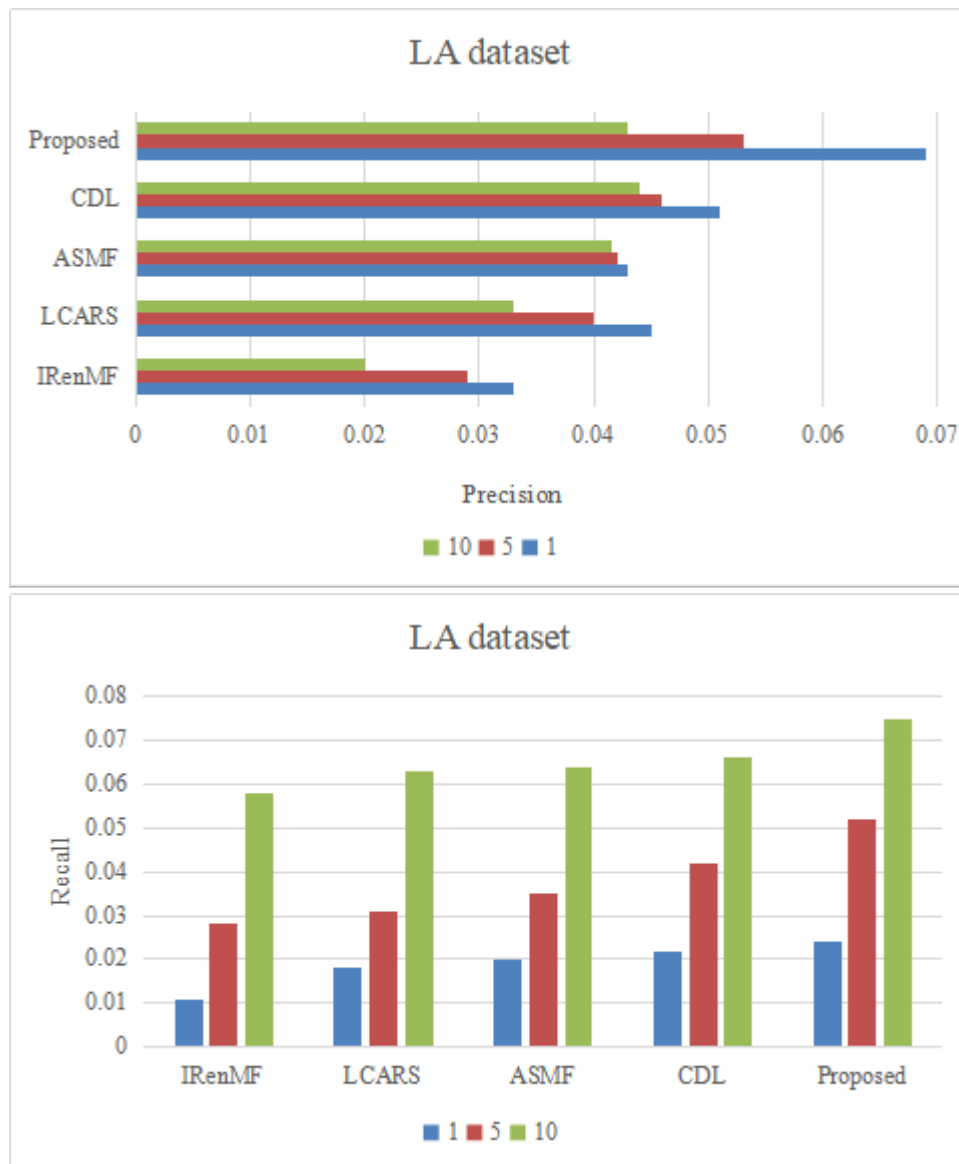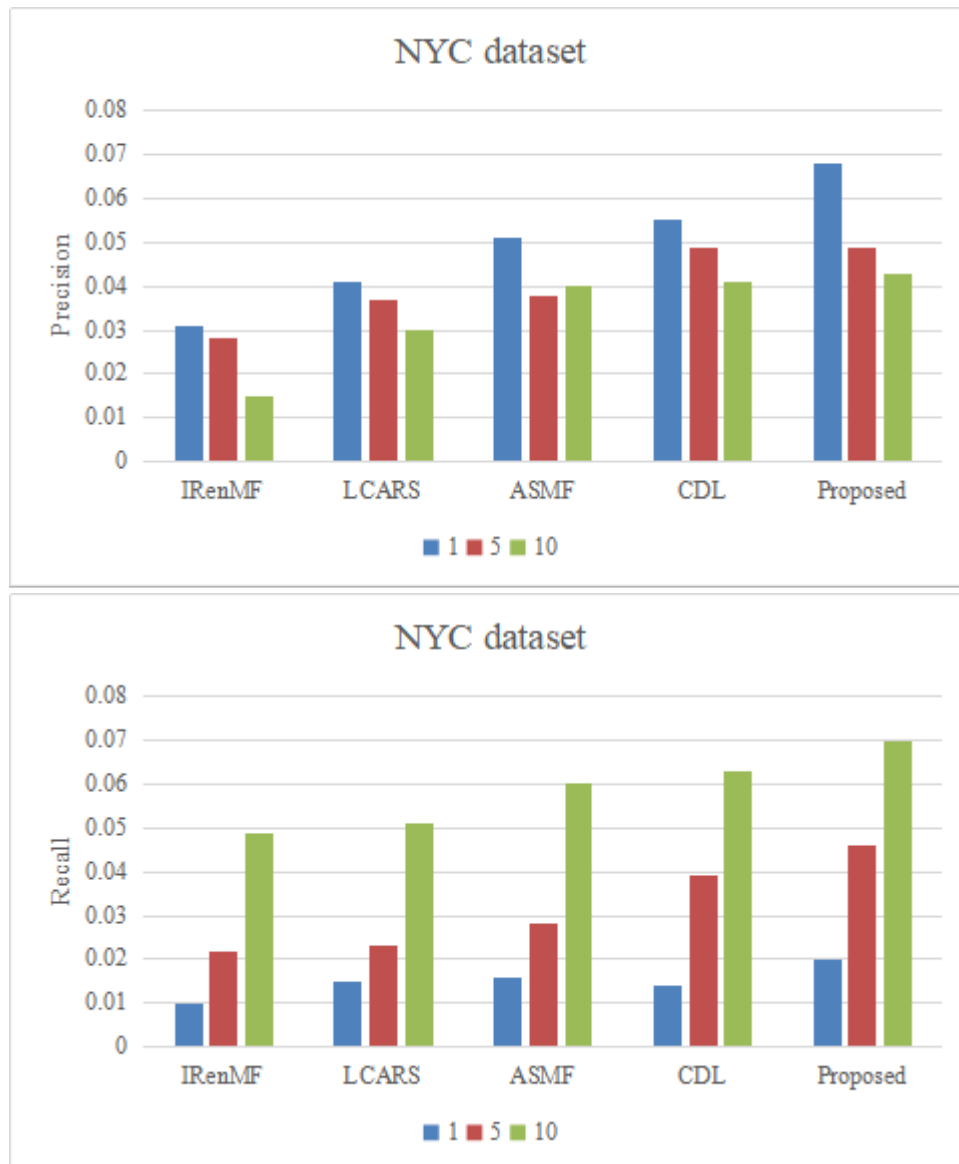
**Fig. 4.** The comparison with different methods on the LA and NYC data sets

(1) IRenMF: This method is greatly affected by missed data as shown in algorithm 1. The matrix density of user interest points in the two data sets in this paper is relatively low. Therefore, IRenMF shows the worst recommended performance relative to the other four methods as shown in Fig. 4.

(2) LCARS: This method is still affected by data sparsity. At the same time, when the data is too sparse, the LDA model cannot effectively learn the potential features of interest points. Compared with social information and geographic information, comment text information has the least impact on the performance of POI recommendation. As shown in Fig. 4, the LCARS method has the 3rd highest recommended performance compared to the other four methods.

(3) ASME: ASMF is not as good as IRenMF model as shown in Fig. 4. Compared with the other four methods, the ASMF method shows the fourth highest recommended performance. Because the ASMF focuses on using social information.

(4) CDL: CDL can produce better recommendation results than the above models. However, the influence of word order on the potential features of the comment text is not considered when learning the potential features of the comment text. As a result, the CDL model finally shows the second best recommendation result.

(5) Proposed method: As shown in Fig. 4, the proposed model performs the best in the recommendation performance based on two data sets. Proposed model fully considers the influence of word order and context infor-

mation on the representation of potential features when learning potential features of comment information, and overcomes the problems that LDA model is difficult to define prior distribution before modeling, and topic proportion cannot effectively represent the potential features of projects when comment information is very scarce. At the same time, when the proposed model conducts matrix decomposition to realize the user-point of interest feature interaction, the initial value of the feature is selected based on the improved neural network method, which avoids falling into the local optimal solution during optimization. Therefore, the proposed model ultimately shows the best recommendation effect.

In order to verify the validity of the proposed auto-encoder initialization, the proposed initialization method is compared with several classical initialization methods, including random initialization, zero initialization, K-means initialization and normalized cutting initialization. The comparison methods is shown in Fig. 5.
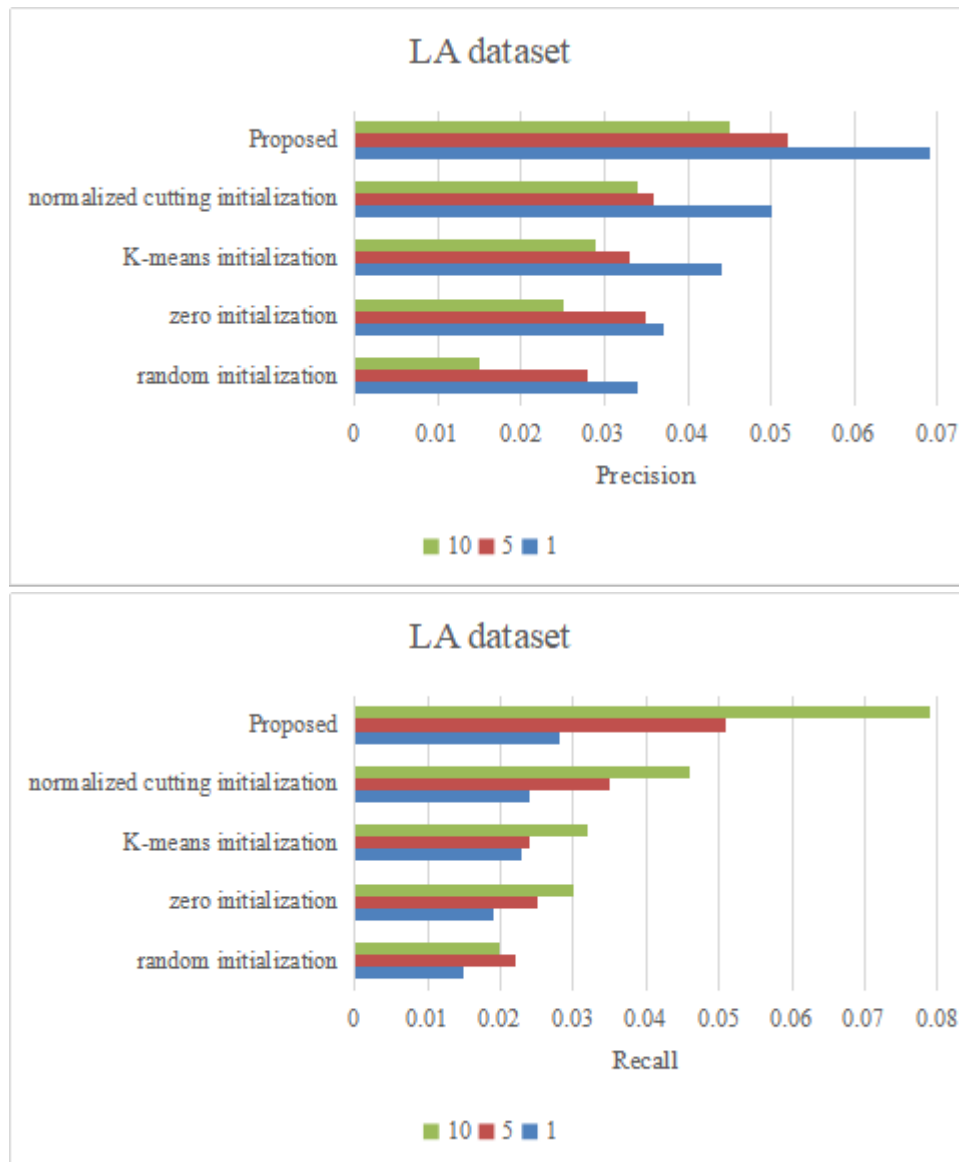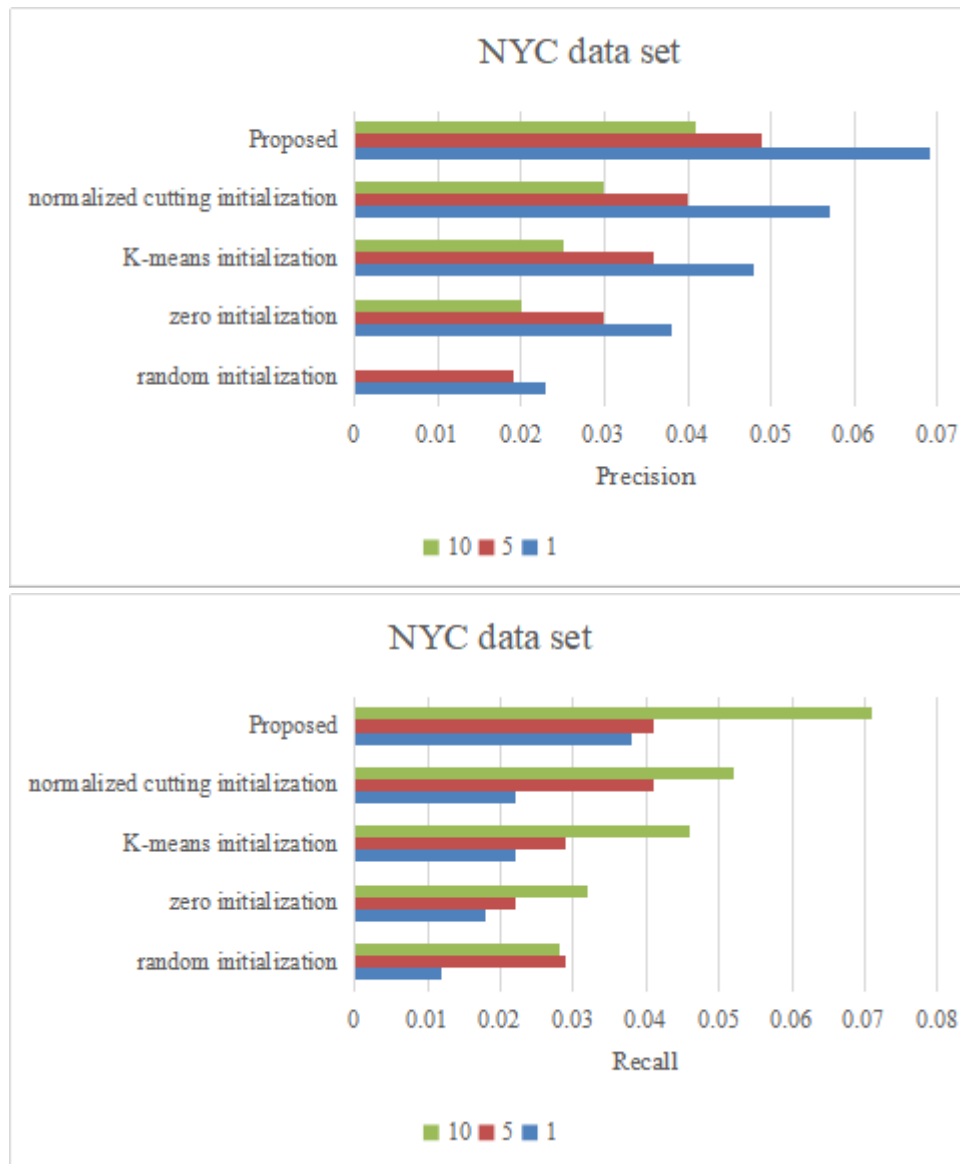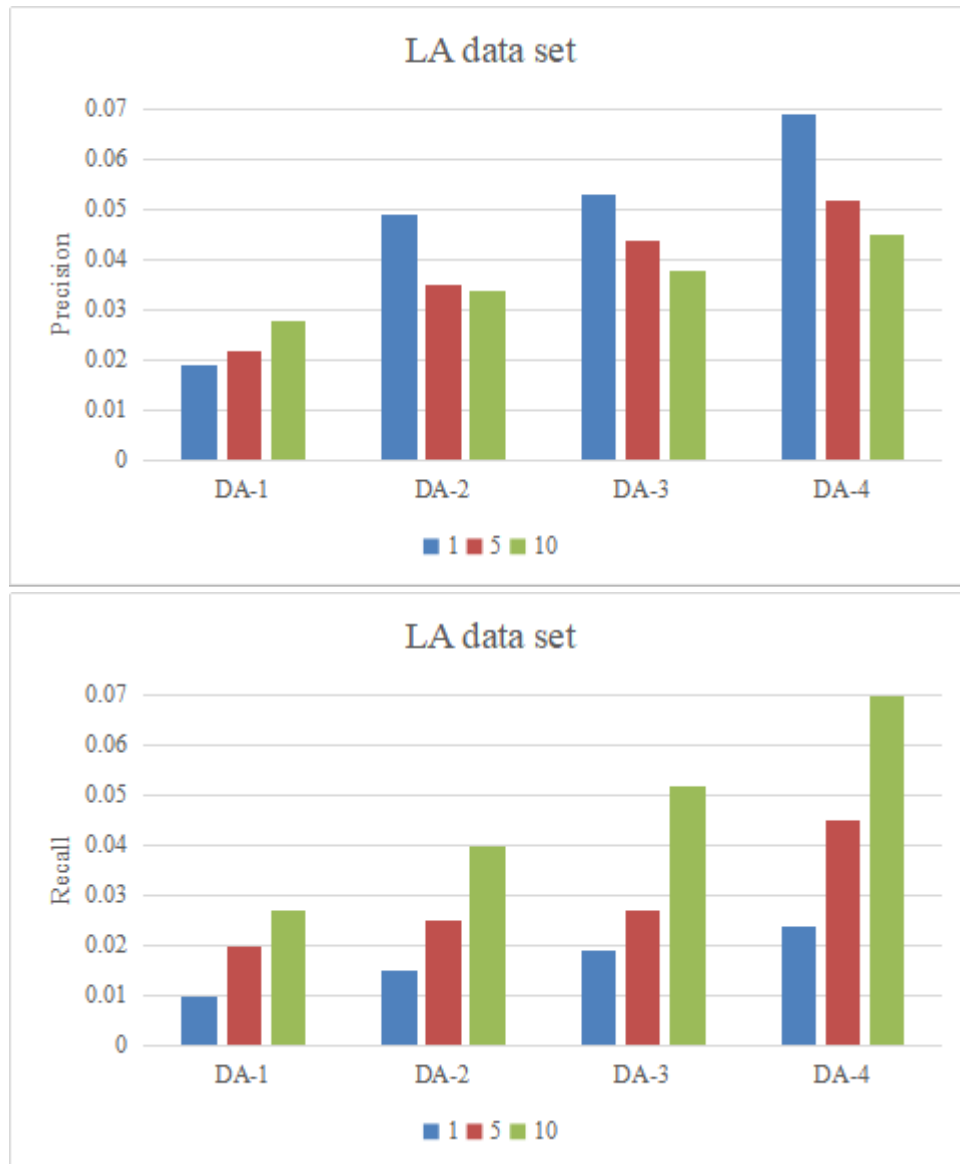
**Fig. 5.** Performance comparison of five initialization methods

As can be seen from Fig. 5, the pre-processing method of denoising encoder for the initialization model proposed in this paper is superior to other initialization methods and effectively avoids the local optimal problem caused by initial values in the process of matrix decomposition. The main reasons are as follows: 1) The model proposed in this paper uses the original users and feature vectors of interest points to learn the features of users and interest points, which is equivalent to the global minimization of the learning model; 2) The initialization vector of learning can better reflect the original similarity relationship between users and interest points.

In order to study the change influence of the encoder layer number on the proposed model in this paper, experiments with different levels are designed to verify the recommendation performance. Where, DA-1, DA-2, etc., represent the encoder layers, and the experimental results are shown in Fig. 6. Fig. 6 shows the change of recall rate and accuracy rate with the change of the encoder layer number based on LA and NYC. As can be seen from Fig. 6, when the layer number is 1 and 2, both recall rate and accuracy rate are quite low. However, as the number of layers increases, recall rate and accuracy rate are improved to some extent. That is, when the feedback matrix is very sparse, the performance of POI recommendation depends on the quality of potential feature representation. When the layer number is 3, the recall rate and accuracy reach the best values, indicating that the proposed model can effectively improve the accuracy of recommendation by using SDAE as a preprocessing component. When the number of layers is 4, it appears fitting phenomenon, which reduces the recommendation performance of the

proposed model. Therefore, it is reasonable to set the number of layers to three in this article.
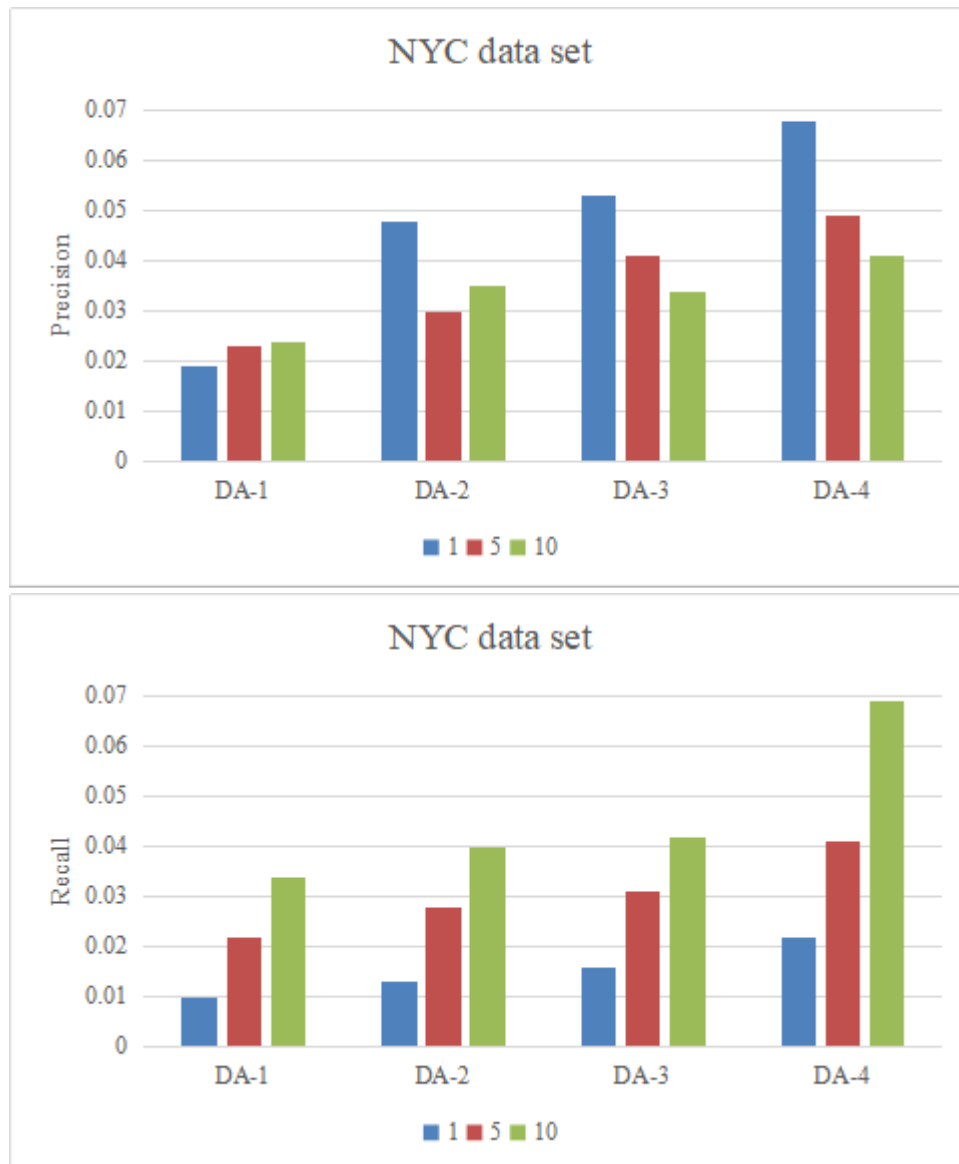
**Fig. 6.** Performance comparison of proposed model at different layers

## 4 Conclusions

This paper proposes a deep learning-based hybrid art POI recommendation algorithm. Deep autoencoders are used to learn the initial values of potential eigenvectors of users and points of interest during matrix decomposition. The proposed model adopts the improved convolutional neural network model (VGG-16) to learn the contextual features of the comment information, so as to extract more accurate feature representation and realize the modeling of the comment information. For the modeling of interaction between users and art POI in matrix decomposition model, the denoising autoencoder is used to learn the best initial value of potential feature vectors of users and points of interest to effectively avoid falling into the local optimal solution in the process of matrix decomposition. Finally, the matrix decomposition technology is used to fuse the above two models to provide users with interest point recommendation service. In the future, the fusion of multiple context information based on deep learning framework will be a direction for practical engineering application.

## 5 Acknowledgement

# References

[1] Y. Sun, S. Yin, H. Li, L. Teng, S. Karim, GPOGC: Gaussian Pigeon-Oriented Graph Clustering Algorithm for Social Networks Cluster, IEEE Access 7(2019) 99254-99262.

[2] Y.-F. Huo, B.-L. Chen, J. Tang, Y.-F. Zeng, Privacy-preserving point-of-interest recommendation based on geographical and social influence, Information Sciences 543(2021) 202-218.

[3] D.-S. Liu, L.-N. Shan, L. Wang, S.-L. Yin, H. Wang, C.-Y. Wang, P$^3$OI-MELSH: Privacy Protection Point of Interest Recommendation Algorithm Based on Multi-exploring Locality Sensitive Hashing, Frontiers in Neurorobotics, 2021. doi: 10.3389/fnbot.2021.660304.

[4] Y. Liu, An experimental evaluation of point-of-interest recommendation in location-based social networks, Proceedings of the Vldb Endowment 10(10)(2017) 1010-1021.

[5] T. Liu, J. Liao, Z. Wu, Y. Wang, J. Wang, Exploiting Geographical-Temporal Awareness Attention for Next Point-of-Interest Recommendation, Neurocomputing 400(2020) 227-237.

[6] X. Zhao, Z. Zhang, X. Bi, Y. Sun, A new point-of-interest group recommendation method in location-based social networks, Neural Computing and Applications 3(2020).

[7] Z. Zhang, C. Zou, R. Ding, Z. Chen, VCG: Exploiting visual contents and geographical influence for Point-of-Interest recommendation, Neurocomputing 357(2019) 53-65.

[8] J. Zhang, P.-P. Sun, F. Zhao, Q.-R. Guo, Y. Zou, Network Pseudohealth Information Recognition Model: An Integrated Architecture of Latent Dirichlet Allocation and Data Block Update, Complexity 2020(2020).

[9] M.-H. Yin, Y.-H. Liu, X. Zhou, G. Sun, A tensor decomposition based collaborative filtering algorithm for time-aware POI recommendation in LBSN, Multimedia Tools and Applications (2021) 1-21. https://doi.org/10.1007/s11042-021-11407-9

[10] J.-R. Chen, Y.-Q. Lu, F.-H. Shang, T.-T. Zhu, A novel recommendation scheme with multifactorial weighted matrix decomposition strategies via forgetting rule, Engineering Applications of Artificial Intelligence 101(10)(2021) 104191.

[11] J.-R. Chen, Y.-Q. Lu, F.-H. Shang, Y.-Y. Wang, A fuzzy matrix factor recommendation method with forgetting function and user features, Applied Soft Computing 100(1)(2021) 106910.

[12] D. Li, Z. Gong, D. Zhang, A Common Topic Transfer Learning Model for Crossing City POI Recommendations, IEEE Transactions on Cybernetics 49(12)(2019) 4282-4295.

[13] X. Pei, T. Wu, Convex-Nonnegative Matrix Factorization with structure constraints, in: 2013 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2013.

[14] S.-L. Yin, H. Li, L. Teng, Airport Detection Based on Improved Faster RCNN in Large Scale Remote Sensing Images, Sensing and Imaging 21 (2020).

[15] X.-W. Wang, S.-L Yin, H. Li, A Network Intrusion Detection Method Based on Deep Multi-scale Convolutional Neural Network, International Journal of Wireless Information Networks 27(4)(2020) 503-517.

[16] Z. Fang, A High-Efficient Hybrid Physics-Informed Neural Networks Based on Convolutional Neural Network, IEEE Transactions on Neural Networks and Learning Systems, (2021) doi: 10.1109/TNNLS.2021.3070878.

[17] S. Yin, Y. Zhang and S. Karim, Region search based on hybrid convolutional neural network in optical remote sensing images, International Journal of Distributed Sensor Networks 15(5) 2019.

M. Moravík, M. Schmid, N. Burch, V. Lis, M. Bowling, DeepStack: Expert-Level Artificial Intelligence in No-Limit Poker, Science 356(6337)(2017) 508.

[18] B. Rao, An Accurate Leukocoria Predictor Based On Deep VGG-Net CNN Technique, IET Image Processing (5)(2020).

[19] J. Yu, H. Li, S.-L. Yin, Dynamic Gesture Recognition Based on Deep Learning in Human-to-Computer Interfaces, Journal of Applied Science and Engineering 23(1)(2020) 31-38.

[20] S.-C. Li, X. Cheng, S. Su, H.-N. Sun, Exploiting organizer influence and geographical preference for new event recommendation, Expert Systems 34(2)(2017) e12190.

[21] L. Guo, H. Jiang, X. Wang, Location Regularization-Based POI Recommendation in Location-Based Social Networks, Information (Switzerland) 9(4)(2018) 85.

[22] H. Li, Y. Ge, R. Hong, H. Zhu, Point-of-Interest Recommendations: Learning Potential Check-ins from Friends, In: Proc. the 22nd ACM SIGKDD International Conference. ACM, 2016.

[23] Y.-D. Seo, Y.-S. Cho, Point of interest recommendations based on the anchoring effect in location-based social network services, Expert Systems with Applications 164(2020).

[24] J. He, X. Li, L.-J. Liao, Next point-of-interest recommendation via a category-aware Listwise Bayesian Personalized Ranking, Journal of computational science 28(2018) 206-216.

[25] M.-M. Kalayeh, M. Shah, Training Faster by Separating Modes of Variation in Batch-Normalized Models, IEEE Transactions on Pattern Analysis and Machine Intelligence 42(6)(2020) 1483-1500.