J Inf Process Syst, Vol.17, No.2, pp.352~368, April 2021
https://doi.org/10.3745/JIPS.01.0068

ISSN 1976-913X (Print)
ISSN 2092-805X (Electronic)

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

POI Recommendation Method Based on Multi-Source Information Fusion Using Deep Learning in Location-Based Social Networks

Liqiang Sun*

Abstract
Sign-in point of interest (POI) are extremely sparse in location-based social networks, hindering recommendation systems from capturing users' deep-level preferences. To solve this problem, we propose a content-aware POI recommendation algorithm based on a convolutional neural network. First, using convolutional neural networks to process comment text information, we model location POI and user latent factors. Subsequently, the objective function is constructed by fusing users' geographical information and obtaining the emotional category information. In addition, the objective function comprises matrix decomposition and maximisation of the probability objective function. Finally, we solve the objective function efficiently. The prediction rate and F1 value on the Instagram-NewYork dataset are 78.32% and 76.37%, respectively, and those on the Instagram-Chicago dataset are 85.16% and 83.29%, respectively. Comparative experiments show that the proposed method can obtain a higher precision rate than several other newer recommended methods.

Keywords
Convolutional Neural Network, Emotional Information, Geographical Information, Latent Factor Modelling, Location-Based Social Network, Objective Function

1. Introduction

Owing to the expansion of the Internet, information resources have grown exponentially, resulting in information overload. This prevents users from quickly obtaining useful knowledge from massive amounts of information. A recommendation system [1-5] can generate a list of recommended items for users or predict users' preference for a specific item without providing users with clear demand information.

Deep learning-based recommendation algorithms can take implicit, explicit, and auxiliary information data as input [6]. Deep network models and deep learning techniques are used to learn hidden feature representations of users and items. However, check-in point of interest (POI) are extremely sparse [7,8] in location-based social networks [9,10], hindering the learning of deep features and resulting in poor model effectiveness.

To address the problem of extremely sparse check-in points in social networks, a content-aware POI

Manuscript received April 23, 2020; first revision June 5, 2020; second revision July 3, 2020; accepted July 11, 2020.

Corresponding Author: Liqiang Sun (qingdaoqchm@163.com)
* Dept. of Information Technology, Qingdao Vocational and Technical College of Hotel Management, Qingdao, Shandong, China (qingdaochm@163.com)

recommendation algorithm (POIRA) is proposed based on a convolutional neural network (CNN). The main contributions of this study are summarised as follows:

1) A new representation learning method is proposed. This method integrates heterogeneous multi-source data into the recommendation model to learn the features with higher discriminative power and improve the precision of the model.
2) A two-layer stacked sparse autoencoder (SAE) network is proposed. The network connects two SAEs training activation vectors and softmax and fine-tunes weights to optimise the loss function, thereby improving the precision of the model.

The remainder of this paper is organised as follows: related work on social network recommendation systems and the motivation for writing this paper are described in Section 2. The architecture and detailed process of our proposed method are presented in Section 3. Experimental verification using recommended system datasets and comparative analysis with other existing methods are detailed in Section 4. Conclusions and future work are presented in Section 5.

2. Related Works

Researchers have proposed numerous methods to address the problem of extremely sparse check-in points in social networks. For example, in [11], an adaptive POI recommendation method was proposed by combining user activities and spatial features. This method performed adaptive operations based on user activities, improving model efficiency. In [12], a recommendation algorithm was proposed based on a deep neural network that uses basic data of users and projects. This method improved the recognition rate of the model by fusing the user feature matrix and item feature matrix. In [13], a model based on neural networks (preference and context embedding) was proposed. The model smoothed the sparse data by using neighbouring users and POI and constructed a context information graph to jointly learn the embedding of users and POI, improving the precision of the model. In [14], a POIRA was proposed based on an edge computing environment. The algorithm analysed users' personalised preference functions on the edge server and embedded geographical information into the framework to obtain POI, thereby improving the precision of the model. However, the latent features learned by these methods when the data are too sparse are not particularly effective.

In [15], a social recommendation algorithm was proposed based on stochastic gradient matrix decomposition. The algorithm improved the precision of the model by fusing social network information. In [16], a recommendation algorithm was proposed by combining probabilistic semantic clustering analysis and collaborative filtering. The algorithm improved the precision of the model by increasing the mining for the sequence of items with time dimension. In [17], a recommendation method was proposed based on group trust and user weight analysis. The algorithm improved the precision of the model by identifying trustworthy users and integrating rent-based trust and user-based trust. In [18], a collaborative filtering recommendation algorithm was proposed based on information theory and bi-clustering. The algorithm improved model efficiency by introducing information entropy and bi-clustering into collaborative work and extracting local dense scoring modules. However, these methods use a bag-of-words model to process comment information, ignoring the semantic context information of comment information.

In [19], a matrix factorisation recommendation algorithm (segmentation-based matrix factorisation [SPMF]) was proposed based on social trust and preference segmentation. The algorithm improved the precision of the model by distinguishing trust relationships and differences in preference domains. In [20], an autoencoder-based multi-criteria recommendation algorithm (AE-MCCF) was proposed. The algorithm used multi-criteria preferences to represent the relationship between users nonlinearly, improving the precision of the model. In [21], a location regularisation recommendation algorithm was proposed based on social networks. The algorithm improved the precision of the model by limiting the ranking loss and assuming that the nearest neighbour POI is more easily accessed by similar users. However, the comment information of these methods has not been fully utilised in related POI studies.

In [22], a context-specific sequence-aware POIRA with multi-gate cyclic units was proposed. The algorithm improved the precision of the model by separately processing the influence for the context of each category. In [23], a feature-space separated factorisation model (FSS-FM) was proposed. The algorithm improved the precision of the model by representing the POI feature space as separate slices, adding spatial and temporal information, and other contexts. However, when the data set is large, these methods are prone to overfitting.

In [24], a total time tensor decomposition model was proposed based on POI recommendations. The model used linear combination operators to aggregate time latent features on different time scales, improving the precision of the model. However, the model runs slowly.

In [25], an adaptive POIRA was proposed based on a multi-order Markov model. The algorithm improved the precision of the model by predicting users' next-favourite POI. However, this method is not suitable for long-term system prediction.

Most of the existing recommendation methods are based on word bags or document topic models to process comment text, which can only provide a shallow understanding of user preferences. This study focuses on the content-aware POIRA based on a deep CNN, to fuse the users' social relationships and geographical information factors of location POI and to optimise the objective function after fusing features.

3. Preliminaries

3.1 Text Expression

For each independent POI, we sort out the related text information. For POI_j, all comment texts related to POI_j and all description texts for POI_j are stored in the same document, d_j. We consider it as text information on POI_j. For the POI_j document, the s word, w_{js}, in d_j is applied with the embedded function of words to obtain a n-dimensional word vector expression, $\eta(w_{js})$, regarding word w_{js}. Therefore, we let the embedded vector matrix describing all words in document d_j related to POI_j be denoted by $\vec{d}_j$ to represent the initial feature vector for text semantics of POI_j, as shown in formula (1).

$$\vec{d}_j = \eta\left(w_{j1}\right) \oplus \eta\left(w_{j2}\right) \oplus \cdots \oplus \eta\left(w_{js}\right) \oplus \cdots \oplus \eta\left(w_{jN}\right) \tag{1}$$

Here, $\eta()$ is an embedded function of words and is called the embedding function. It matches each word to a fixed-length feature space. This is a vectorisation process. The embedded expression result of the s word, w_{js}, is located in the s column of matrix $\vec{d}_j$. $\oplus$ is the connection operator, and N is the number of words in document d_j. The result of the embedded function for words represents different features of words in different dimensions. Therefore, the values of projection for words in different dimensions represent different semantics. Hence, the initial feature vector expression, $\vec{d}_j$, for the semantics of texts can be obtained.

First, the convolutional layer is used to extract the semantic features of texts. The convolutional layer includes multiple neurons, which can use the convolution operator to extract new features from the initial feature vector, $\vec{d}_j$, of text semantics obtained previously. To extract features locally, we set an $a \times a$ sized convolution kernel (set as 3×3), where $a = m, n$. Because each convolution kernel has fixed parameters, the extracted features will be relatively single. Thus, the model sets multiple different convolution kernels to obtain different features. Subsequently, the output obtained after each convolution kernel can be regarded as each feature obtained after the feature extraction of texts. Convolution kernel $K_c, c = \{0,1 \ldots k\}$, is then used to extract feature z_c. Thus, the convolution feature generated by the d_j document under the action of convolution kernel K_c is z_c^j, as shown in formula (2).

$$z_c^j = f\left(K_c * \vec{d}_j + b_c\right) \tag{2}$$

Here, $f()$ is the activation function, $*$ is the convolution operator, and b_c is the deviation term of convolution. $K_c \in R^{h \times t}$; the window size is $h \times t$.

Subsequently, rectified linear units (ReLUs) are selected to further enhance the expressive power of CNNs. Compared with other activation functions, the ReLUs function is simple to operate. The function has suitable sparse activation, wide excitability boundaries, and unilateral inhibition. Furthermore, ReLUs are more computationally efficient for deep neural networks and are more suitable for deep network learning and training. The specific ReLU function is shown in formula (3).

$$f(x) = \max\{0, x\} \tag{3}$$

To reduce the dimension of features, reduce the number of parameters, and avoid overfitting, the next step is to undergo a process of pooling a layer. In the pooling layer, the maximum pooling is selected to obtain the required features, as shown in formula (4).

$$Maxpool(r_s) = \max_{i \in r_s} a_i \tag{4}$$

Here, r_s represents the size of the pooling window area, and a_i represents the i feature in window r_s. Thus, the new features obtained become

$$l_c = \max\left\{z_c^1, z_c^2, \cdots, z_c^{n-h+1}\right\}. \tag{5}$$

Next, the features generated by all convolution kernels after passing through the maximum pool are

$$L = \{l_1, l_2, \ldots l_c, \ldots, l_k\}, \tag{6}$$

where k is the number of convolution kernels.

Finally, the fully connected layers in the CNN are used to integrate the feature vectors obtained after multiple convolutional layers and pooling layers. In addition, the high-level semantic output, including the described object and its comment information in the text information, is obtained. That is, the expression learns the result of POI_j for the text semantics. This is also the output of the CNN, as shown in formula (7).

$$CNN_d\left(\vec{d}_j\right)=f\left(W_d\times L_d+g_d\right) \tag{7}$$

Here, $\vec{d}_j$ denotes the POI_j initial feature vector of text semantics, and W_d is the weight matrix above the fully connected layer related to the text. L_d is the feature vector generated before being fully connected, and g_d represents the skew associated with texts in the fully connected layer. The characteristic dimension of $CNN_d\left(\vec{d}_j\right)$ is d'.

3.2 Image Expression

Image p_j describing the size of POI_j as $m\times n$ comprises $m\times n$ pixels. A scale-invariant feature transform [26] is used to obtain several "key point" sets on image p_j. Each "key point" can be expressed as a 128-dimensional vector. For the entire image set, the k-means results of these "key points" can be used to generate several clusters, and each "cluster" serves as a "visual word." Finally, all the "clusters" are used to create a "visual word dictionary." In different images, each "key point" is assigned to the cluster closest to it. In this way, each image can be represented by a vector of size z in the end. The value in vectors indicates the number of times the 'visual word' in each image appears in the image. At this time, image p_j can be expressed as a vector expression of $\vec{p}_j=(a_1,a_2,\cdots a_z)$. The obtained image vector, p_j, is used as the initial feature vector of image features to perform deeper feature extraction as input in the deep learning process in the next step.

After processing the entire CNN, the output of image p_j of POI_j can be obtained, that is, the expression learns the result of POI_j for the image features, as shown in formula (8).

$$CNN_p\left(\vec{p}_j\right)=f\left(W_p\times L_p+g_p\right) \tag{8}$$

Here, $\vec{p}_j$ represents the initial feature vector of POI_j for image features. W_p is the weight matrix on the fully connected layer related to images, and L_p is the feature vector generated after passing through the pooling layer related to images. g_p is the image-related skew in the fully connected layer, and the feature dimension of $CNN_p\left(\vec{p}_j\right)$ is p'.

In summary, the initial feature vector, $\vec{d}_j$, of text semantics and the initial feature vector, $\vec{p}_j$, of image features are added to the CNN as inputs. This includes not only the feature extraction results of text and image information but also the cognition and understanding of the association between texts, images, and users' behaviour. These initial feature vector expressions of text semantics and image features contain richer features and information than text semantics.

3.3 Bayesian Personalized Ranking Model

Assumption: User *u* prefers item I_i over item I_j. While observing the (u, I_i) rating pair, it is impossible to observe rating pair (u, I_j); *U* and *I* represent the user set and item set, respectively. This definition assumes the following:

$$\hat{\delta}_{u,I_i}(\Lambda) \succ \hat{\delta}_{u,I_j}(\Lambda), I_i \in I_u^+, I_j \in I / I_u^+ \quad , \tag{9}$$

where Λ represents the parameter set of the ranking function, and $\hat{\delta}_{u,I_i}(\Lambda)$ and $\hat{\delta}_{u,I_j}(\Lambda)$ are the predicted scores of the ranking function.

Therefore, based on all partial order pairs f_u and the Bayesian theory, the posterior distribution of all partial order pairs f_u is obtained. The likelihood function is as follows:

$$p(\Lambda | \succ_u) \infty p(\succ_u | \Lambda) p(\Lambda) \quad , \tag{10}$$

where

$$p(\succ_u | \Lambda) = \prod_{(u, I_i, I_j \in D_x)} p(I_i \succ_u I_j | \Lambda) \tag{11}$$

$$p(I_i \succ_u I_j | \Lambda) := \eta\left(\hat{z}_{u,I_i,I_j}(\Lambda)\right) . \tag{12}$$

$\hat{z}_{u,I_i,I_j}(\Lambda)$ represents any real-valued function concerning the model parameter, Λ. It also reflects the preference relationship of user *u* for items I_i and I_j. η is defined as follows:

$$\eta(x) := \frac{1}{1+e^{-x}} . \tag{13}$$

Combining the above formulas, the final objective function is

$$\begin{aligned} BPR-OPT &= \ln p(\Lambda | \succ_u) \\ &= \ln p(\succ_u | \Lambda) p(\Lambda) \\ &= \ln \prod_{((u,I_i,I_j) \in D_x)} \eta\left(\hat{z}_{u,I_i,I_j}\right) p(\Lambda) \\ &= \sum_{((u,I_i,I_j) \in D_x)} \ln \eta\left(\hat{z}_{u,I_i,I_j}\right) + \ln p(\Lambda) \\ &= \sum_{((u,I_i,I_j) \in D_x)} \ln \eta\left(\hat{z}_{u,I_i,I_j}\right) - \lambda_\Lambda \|\Lambda\|^2 \quad , \end{aligned} \tag{14}$$

where λ_Λ is the model-specific regularisation parameter, and $D_x = \{(u, I_i, I_j) | I_i, I_j \in I; u \in U\}$ is the training dataset. The Bayesian personalized ranking model uses stochastic gradient descent (SGD) to optimise the objective function.

$$\frac{\partial BPR-OPT}{\partial \Lambda}=\sum_{(u,I_i,I_j)\in D_x}\frac{\partial}{\partial \Lambda}\ln \eta\left(\hat{z}_{u,I_i,I_j}\right)-\lambda_\Lambda \frac{\partial}{\partial \Lambda}\|\Lambda\|^2$$
$$\infty \sum_{(u,I_i,I_j)\in D_x}\frac{-e^{-\hat{z}_{u,I_i,I_j}}}{1+e^{-\hat{z}_{u,I_i,I_j}}}\cdot\frac{\partial}{\partial \Lambda}\hat{z}_{u,I_i,I_j}-\lambda_\Lambda \Lambda \tag{15}$$

4. Proposed Method

The framework of the proposed content-aware POIRA model based on a CNN is shown in Fig. 1. It first fuses geographical information, user social relations, and comment content information to construct the objective function. Subsequently, a two-layer stacked sparse autoencoder network is used for training to generate hidden features and item hidden features. Finally, a recommendation list is generated.

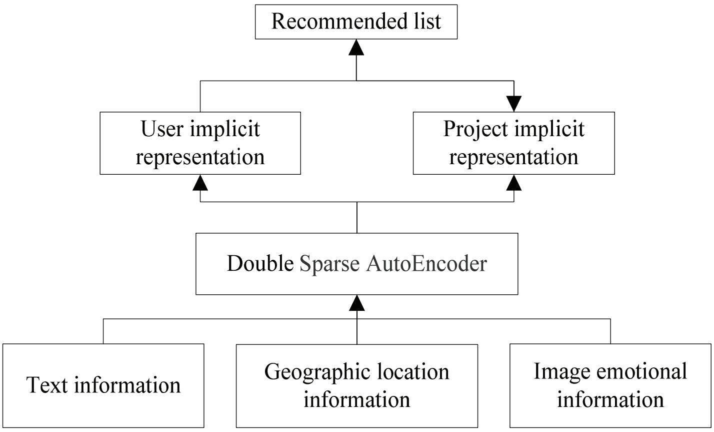

Recommended list
User implicit representation
Project implicit representation
Double Sparse AutoEncoder
Text information
Geographic location information
Image emotional information

Fig. 1. Overall framework of content-aware POI recommendation algorithm model based on the convolutional neural network.

4.1 Geographical Information Modelling

The preference of user u_i for several neighbouring positions of position l_j represents the preference of user u_i for position l_j. The geographic location weighting strategy [27] is used to complete the missing geographical information in the matrix decomposition model. The objective function minimisation problem can be expressed by the following formula.

$$\min_{U,L}\frac{1}{2}\left(\mathrm{I}\odot\left(\mathrm{R\text{-}ULH}^T\right)^2\right) \tag{16}$$

$$\mathrm{H}=\beta \mathrm{UL}^T+(1-\beta)\mathrm{G}^T,\mathrm{G}\in R^{n\times n},\mathrm{G}_{j,k}=\frac{sim\left(l_j,l_k\right)}{z\left(l_j\right)} \tag{17}$$

β is a weight parameter that controls the influence of adjacent positions. $sim(l_j,l_k)$ represents the geographical weight of adjacent position l_k at position l_j. $z(l_j)$ is a regularisation term defined as

$$Z\left(l_j\right)=\sum_{l_k \in E\left(l_j\right)} sim\left(l_j,l_k\right) \tag{18}$$

$sim(l_j, l_k)$ uses a Gaussian function, as shown in formula (19):

$$sim\left(l_j,l_k\right)=e^{\frac{\left\|l_j-l_k\right\|^2}{\rho^2}},\forall l_k \in E\left(l_j\right) \tag{19}$$

To distinguish the geographical range, a geographical area distance variable, *F*, is proposed because the user is less likely to check-in in the area farther from the current location. $E(l_j)$ represents the adjacent position of l_j. In the experiment, we set $F=10000$ according to experience. If the location to be recommended is not in the user's current location, $E(l_j)$, the location is not considered.

4.2 Image Emotional Classification

For the image-based emotional analysis method in social media, a double SAE, as shown in Fig. 2, is constructed. First, an SAE network is trained [28]. When a new sample, x, is input into the network, vector h_1 composed of the activation values of the units for the hidden layer can represent x, as shown in formula (20), where $i = 1$. The number of hidden layer nodes of the first SAE is set to 1000, and 1000 dimensional features are used as the new feature vector, h_1. This feature is called a first-order feature. This first-order feature is used as the input of the second SAE. The hidden layer sets the number of nodes to 256, and the 256-dimensional features trained are regarded as second-order features (h_2). Two of these SAE networks use the nonlinear activation function, ReLUs, as shown in formula (21).

$$h_i = f\left(W_i x + b_i\right) \tag{20}$$

$$\sigma_{RuLu}\left(h_i\right)=\max\left(0,h_i\right) \tag{21}$$

Next, we connect h_1 and h_2 trained in the two SAE networks with a Softmax classifier to fine-tune weights. In this way, the output of the last fully connected layer can be converted into the emotion classification probability of images:

$$p_i=\frac{\exp\left(h_i\right)}{\sum_i \exp\left(h_i\right)},i=1,\ldots,n, \tag{22}$$

where h_i is the output of the last fully connected layer. The loss function of the recognition probability is a multiclass cross-entropy loss function:

$$L=-\sum_i y_i \log\left(p_i\right), \tag{23}$$

where y_i is the true label of i images.

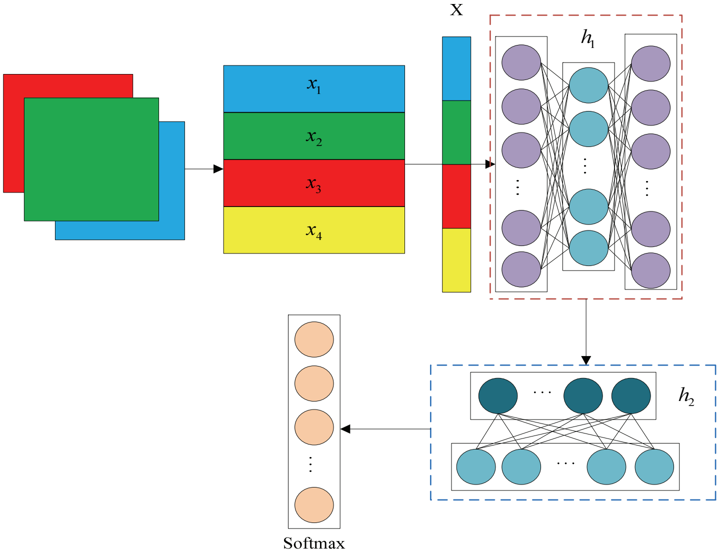


Fig. 2. Two-layer stacked sparse autoencoder network.

4.3 User Preference and Location POI Attribute Modelling

The CNN can use the pre-trained word embedding model to more deeply understand the content of the position interest review paper. The text features are expressed as $CNN(W, C_S)$, where C_S is the set of review texts and W is the internal weight of the CNN. u_i denotes users, and C_S indicates posted comments. $P = (f_{is} = 1|u_i, C_S)$, where f_{is} indicates whether C_S is issued by u_i. Then, $P = (f_{is} = 1|u_i, C_S)$ is specifically defined as follows:

$$P\left(f_{is}=1\left|u_i,c_s\right.\right)=\frac{\exp\left(u_i^T\cdot C\cdot CNN\left(W,C_s\right)\right)}{\sum_{C_k\in C}\exp\left(u_i^T\cdot C\cdot CNN\left(W,C_k\right)\right)}, \tag{24}$$

Here, $C \in R^{K\times d}$ is the interaction matrix between the content features of comment texts and the potential features of users. The matrix can be used to distinguish users' latent feature u_i, that is, whether comment C_S is issued by user u_i. The Softmax function is used to normalise weights [29], which has the meaning of probabilistic interpretation and can realise the output of the network and interpret it as the posterior probability of the classification target variable. To extract the users' latent feature vector, u_i, we convert formula (24) to the objective function.

$$\sum_{i=1}^{n}\sum_{c_k\in C_{u_i}}\log P\left(f_{is}=1\left|u_i,c_k\right.\right) \tag{25}$$

Similarly, probability $P\left(h_{jp}=1\left|l_j,c_{\mathrm{P}}\right.\right)$ that comment c_P is related to position l_j is defined as

$$P\left(h_{jp}=1\left|l_j,c_P\right.\right)=\frac{\exp\left(l_j^T\cdot P\cdot CNN\left(W,C_p\right)\right)}{\sum_{C_k\in C}\exp\left(l_i^T\cdot P\cdot CNN\left(W,C_k\right)\right)}, \tag{26}$$

Here, h_{jp} indicates whether comment c_P is associated with position l_j. $P \in R^{K\times d}$ represents the interaction matrix between the features of comment content and the potential features of location POI. The interaction matrix can be used to distinguish the potential features, l_j, of location POI, that is, whether comment c_P is associated with location l_j. Formula (26) is converted into the following objective function:

$$\sum_{j=1}^{m} \sum_{c_k \in C_{l_j}} \log P\left(h_{jk} = 1 \middle| l_j, c_k\right) \quad (27)$$

.

4.4 Constructing the Objective Loss Function

The comprehensive formula of the target loss function of the POIRA algorithm is obtained as follows:

$$L^* = \arg\min\nolimits_{\Omega} \left(\begin{array}{l} \sum\limits_{u_v \in U, l_v \in L_v^+, l_s \in L_{vl}^s} \ln\sigma\left(\hat{y}_{u_v,l_v} - \hat{y}_{u_v,l_s}\right) + \\ \sum\limits_{u_v \in U, l_s \in L_{vl}^s, l_g \in L_{vl}^g,} \ln\sigma\left(\hat{y}_{u_v,l_s} - \hat{y}_{u_v,l_g}\right) \\ + \sum\limits_{u_v \in U, l_g \in L_{vl}^g, l_w \in L_v^-} \ln\sigma\left(\hat{y}_{u_v,l_g} - \hat{y}_{u_v,l_w}\right) - \lambda_\Omega \left\|\Omega\right\|^2 \end{array} \right). \quad (28)$$

Here, the regularisation term, $\lambda_\Omega \left\|\Omega\right\|^2$, is used to avoid overfitting during the learning process, and λ_Ω is the regularisation parameter.

Finally, according to the matrix decomposition technique of predicting $\hat{y}$, the following formula is obtained:

$$\begin{aligned} \hat{y}_{u_v,l_v} &= Q_{u_v} \cdot H_{l_v}^T + b_{l_v} = \sum_f^{\varpi} q_{u_v,f} \times h_{l_v,f} + b_{l_v} \\ \hat{y}_{u_v,l_s} &= Q_{u_v} \cdot H_{l_s}^T + b_{l_s} = \sum_f^{\varpi} q_{u_v,f} \times h_{l_s,f} + b_{l_s} \\ \hat{y}_{u_v,l_g} &= Q_{u_v} \cdot H_{l_g}^T + b_{l_g} = \sum_f^{\varpi} q_{u_v,f} \times h_{l_g,f} + b_{l_g} \\ \hat{y}_{u_v,l_w} &= Q_{u_v} \cdot H_{l_w}^T + b_{l_w} = \sum_f^{\varpi} q_{u_v,f} \times h_{l_w,f} + b_{l_w} \end{aligned} , \quad (29)$$

Here, Q_{u_v} represents the latent feature matrix of user v. $H_{l_v}, H_{l_s}, H_{l_g}, H_{l_w}$ represent the latent feature matrix of POI l_v, l_s, l_g, l_w, respectively [30]. $b_{l_v}, b_{l_s}, b_{l_g}, b_{l_w}$ represent the deviation items of POI, l_v, l_s, l_g, l_w, respectively. ϖ denotes the dimension of the matrix decomposition. Row vector Q_{u_v} represented by each row of matrix q_{u_v} corresponds to the feature vector of each user. Row vectors $H_{l_v}, H_{l_s}, H_{l_g}, H_{l_w}$ represented respectively by each row of matrices $h_{l_v}, h_{l_s}, h_{l_g}, h_{l_w}$ correspond to the feature vector of each POI.

4.5 Loss Function Optimisation and Learning

Using SGD to optimise formula (28), we obtain the following formula:

$$\begin{aligned}\frac{\partial L^*}{\partial \Omega} = &\sum_{(u_v,l_v,l_s,l_g,l_w)} \frac{\partial}{\partial \Omega} \ln \sigma\left(\left(\hat{y}_{u_v,l_v} - \hat{y}_{u_v,l_s}\right)\right) \\ &+ \frac{\partial}{\partial \Omega} \ln \sigma\left(\left(\hat{y}_{u_v,l_s} - \hat{y}_{u_v,l_g}\right)\right) \\ &+ \frac{\partial}{\partial \Omega} \ln \sigma\left(\left(\hat{y}_{u_v,l_g} - \hat{y}_{u_v,l_w}\right)\right) - \lambda_\Omega \frac{\partial}{\partial \Omega}\|\Omega\|^2 \\ &\infty \sum_{(u_v,l_v,l_s,l_g,l_w)\in D_r} \frac{-e^{-\left(\hat{y}_{u_v,l_v}-\hat{y}_{u_v,l_s}\right)}}{1+e^{-\left(\hat{y}_{u_v,l_v}-\hat{y}_{u_v,l_s}\right)}} \cdot \frac{\partial}{\partial \Omega}\left(\hat{y}_{u_v,l_v} - \hat{y}_{u_v,l_s}\right) \\ &+ \frac{-e^{-\left(\hat{y}_{u_v,l_s}-\hat{y}_{u_v,l_g}\right)}}{1+e^{-\left(\hat{y}_{u_v,l_s}-\hat{y}_{u_v,l_g}\right)}} \cdot \frac{\partial}{\partial \Omega}\left(\hat{y}_{u_v,l_s} - \hat{y}_{u_v,l_g}\right) \\ &+ \frac{-e^{-\left(\hat{y}_{u_v,l_g}-\hat{y}_{u_v,l_w}\right)}}{1+e^{-\left(\hat{y}_{u_v,l_g}-\hat{y}_{u_v,l_w}\right)}} \cdot \frac{\partial}{\partial \Omega}\left(\hat{y}_{u_v,l_g} - \hat{y}_{u_v,l_w}\right) \\ &-\lambda_{r,q} Q_{u_v} - \lambda_{r,h}\left(H_{l_v} + H_{l_s} + H_{l_g} + H_{l_w}\right) \\ &-\lambda_{r,b}\left(b_{l_v} + b_{l_s} + b_{l_g} + b_{l_w}\right)\end{aligned} \quad (30)$$

,

Here, $\lambda_{r,q}$ represents the regularisation parameter of Q_{u_v}, $\lambda_{r,h}$ denotes the regularisation parameter of $H_{l_v}, H_{l_s}, H_{l_g}, H_{l_w}$, and $\lambda_{r,b}$ represents the regularisation parameter of $b_{l_v}, b_{l_s}, b_{l_g}, b_{l_w}$. The algorithm parameters are updated to obtain the following formula based on the aforementioned gradient:

$$\Omega \leftarrow \Omega + \gamma \cdot \frac{\partial L^*}{\partial \Omega}, \quad (31)$$

Here, γ is the learning rate. The pseudocode of the POIRA is shown in Algorithm 1.

Algorithm 1. POIRA algorithm

```
Input: $G_r(L, N_r), S_r(U, E_r), D_r$
Output: model parameters $\Omega$
1. Initialise $\Omega$ based on the normal distribution;
2. for $u_v \in U$;
3. Calculate $L_v^+, L_{vl}^g, L_{vl}^s, L_v^-$;
4. end;
5. for $u_v \in U$;
6. Traverse each 5-tuple $(u_v, l_v, l_s, l_g, l_w) \in D_r$ in $D_r$;
7. Calculate $\frac{\partial L^*}{\partial \Omega}$ according to formula (30);
8. Iteratively update parameter $\Omega$ according to formula (31);
9. end for; // If the iteration converges, exit the loop, otherwise jump to step 5
10. return $\Omega$;
```

Spatial coding is mainly used in the POIRA algorithm, and the storage space does not change with the size of the data; therefore, the spatial complexity of the POIRA algorithm is $O(1)$.

5. Experimental Results and Analysis

To verify the effectiveness of our proposed content-aware POIRA based on CNNs, sufficient experimental evaluations are performed on the Instagram-NewYork (NY) and Instagram-Chicago (CHI) datasets. The SPMF proposed in [19], AE-MCCF proposed in [20], FSS-FM proposed in [23], and POIRA are compared through experiments. These methods are implemented on Python 3.0.

5.1 Experimental Dataset

NY and CHI are datasets of two different regions on Instagram, which allows each user to associate geotags with pictures and related comments by posting pictures using their mobile phones. To improve the performance of this recommendation model, we first filter out the selfies and unclear virtual images. We then remove text and images that have no actual content or meaning. A maximum of 50 pictures and 100 text comments are retained. Users who have less than 10 check-in POI are also filtered out. We use 70% of the dataset for training, 10% for verification, and 20% for testing. The details of the two datasets are shown in Table 1.

Table 1. Details of Instagram datasets

Category	Instagram-NewYork (NY)	Instagram-Chicago (CHI)
Number of users	2,932	3,213
Number of POI	8,743	10,438
Number of pictures	171,281	192,065
Number of texts	126,534	157,153

5.2 Evaluation Criteria

The prediction rate and F1 value are used as evaluation criteria, and the detailed definitions are shown in formulas (32) and (34).

$$Precision = \frac{TP}{TP + FP} \tag{32}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{33}$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \tag{34}$$

Here, *TP* represents the number of samples predicted to be positive and actually positive. *FP* represents the number of samples predicted to be positive but actually negative. *FN* represents the number of samples predicted to be negative but actually positive.

5.3 Effect Verification of Geographical Information, Users' Social Relationship, and Comment Content Information

Precision is used in the experiment to measure the effectiveness of the model, abbreviated as $\mathrm{P@k}$. For a target user, u_i, $\mathrm{P@k}$ indicates the proportion of test access POI that will be included in the top k recommended POI. $Q(u_i)$ represents the POI that user u_i has checked in, and $E(u_i)$ represents the top k POI recommended. $\mathrm{P@k}$ is defined as follows:

$$\mathrm{P@k}=\frac{1}{|V|}\sum_{i=1}^{V}\frac{|Q(u_i)\cap E(u_i)|}{k} \quad (35)$$

,

where V represents the number of users in the test data. To verify the effectiveness of the content-aware POIRA based on CNNs, experiments are performed on the NY and CHI datasets using several existing methods. In the experiment, three components—geographical information, users' social relationship, and comment content information—are analysed and compared with those of the POIRA. The experimental results are listed in Table 2.

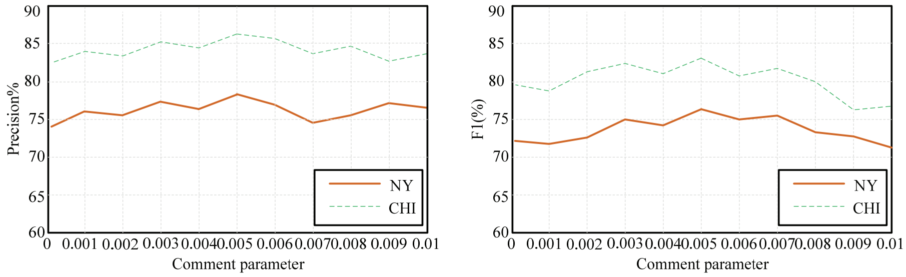
As can be seen in Table 2, the first three components are all critical to POI recommendations. The fusion of the third component is helpful in improving recommendation precision. Users are affected by various contextual information in real life. Therefore, the user's preference prediction cannot be modelled based on one aspect. The POIRA fully utilises the context information of various POI to solve the problem of sparse data in POI recommendation.

Table 2. POI recommendation algorithm based on NY and CHI datasets and the other three recommendation performance comparison (unit: %)

Method	Component	Instagram-NewYork (NY)	Instagram-Chicago (CHI)
SPMF	Geographical information	64.21	73.42
	Users' social relationship	65.78	75.16
	Comment content information	68.49	78.57
	POIRA	71.83	79.63
AE-MCCF	Geographical information	71.36	73.16
	Users' social relationship	70.47	75.49
	Comment content information	72.06	76.28
	POIRA	75.49	81.24
FSS-FM	Geographical information	72.28	73.61
	Users' social relationship	73.06	75.84
	Comment content information	75.31	78.26
	POIRA	76.15	83.47
POIRA	Geographical information	70.87	75.32
	Users' social relationship	70.54	77.56
	Comment content information	73.16	80.24
	POIRA	78.32	85.16

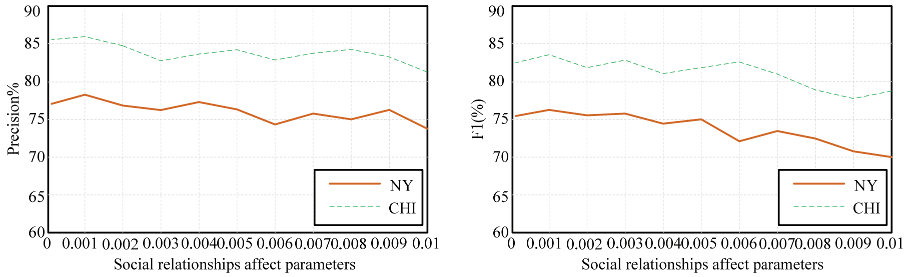
5.4 Impact Analysis of Parameter Changes

There are three important parameters in the POIRA: (1) control comment parameter, (2) control social relationships affect parameter, and (3) geographical neighbourhood relational weighting parameter. When studying and analysing these parameters, we change the value of one of the parameters while fixing the other parameters to analyse its impact on the final recommendation results and the sensitivity of POIRA to parameters. The experimental results are shown in Figs. 3–5.



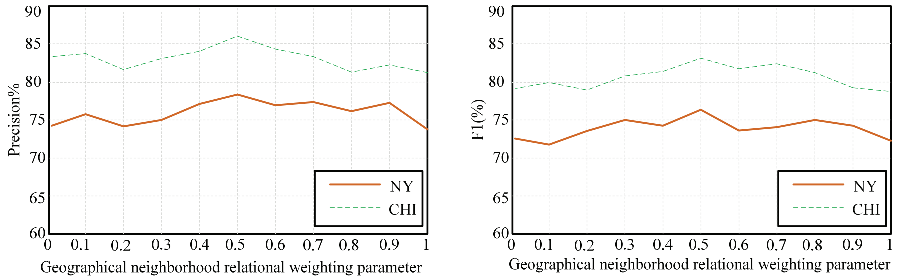
(a) (b)

Fig. 3. Performance analysis of comment parameter: (a) precision and (b) F1 value.



(a) (b)

Fig. 4. Performance analysis of social relationships affect parameters: (a) precision and (b) F1 value.



(a) (b)

Fig. 5. Performance analysis of geographical neighbourhood relational weighting parameter: (a) precision and (b) F1 value.

As can be seen in Figs. 3 and 4, when the comment parameter is 0.05 and the social relationships affect parameter is 0.001, this model achieves the best recommendation effect. This indicates that users may mention only some potential topics, but not all in one review. As can be seen in Fig. 5, a good result is obtained when the geographical neighbourhood relational weighting parameter is 0.5. This shows the importance of geographical neighbourhood relational weighting parameters in measuring users' preference for the recommended POI and the characteristics of the geographical neighbourhood.

5.5 Comparison with Other Methods

To verify the superiority of our algorithm, it is comprehensively compared with existing methods on the NY and CHI datasets. We ensure that the experiment on the training object is conducted under conditions unrelated to those of the test object. The experimental results are presented in Table 3.

Table 3. Comparison of prediction values on NY and CHI datasets with existing methods (unit: %)

		SPMF	AE-MCCF	FSS-FM	POIRA
Instagram-NewYork (NY)	k=1	72.54	73.67	75.48	78.32
	k=5	71.06	72.56	74.29	75.18
	k=10	68.13	69.29	70.12	72.44
Instagram-Chicago (CHI)	k=1	79.28	81.84	83.16	85.16
	k=5	77.51	79.23	81.33	82.67
	k=10	73.67	75.81	77.29	80.17

As can be seen in Tables 3 and 4, compared with several other existing POI recommendation algorithms, the POIRA algorithm shows the best recommendation performance in terms of the F1 value and precision rate. As the number of POI increases, the precision rate and F1 value continue to decline. This is because the POIRA algorithm recommends more POI to users, which helps users find more POI. This will encourage users to be more willing to check in POI, thereby improving the precision and F1 of the model.

Table 4. Comparison of F1 values on NY and CHI datasets with existing methods (unit: %)

		SPMF	AE-MCCF	FSS-FM	POIRA
Instagram-NewYork (NY)	k=1	71.83	73.49	74.15	76.37
	k=5	69.58	70.34	72.02	74.29
	k=10	65.29	67.28	68.27	71.08
Instagram-Chicago (CHI)	k=1	78.63	81.24	82.47	83.29
	k=5	76.21	77.48	78.22	81.48
	k=10	71.93	73.27	75.09	77.26

6 Conclusion

In this study, a new content-aware POIRA based on a CNN is proposed. Based on the matrix decomposition model, the algorithm uses CNNs to extract the content features of comment texts and model the comment information of POI. Furthermore, it integrates users' social relationships and geographical information to recommend location POI. Experimental results show that the content-aware POIRA based on a CNN improves the precision of the recommendation effectively.

In practical applications, the interest preferences of users often have an impact on POI owing to changes in time. However, this factor is not considered in the proposed method. Therefore, to the method of considering the time factor and modelling the dynamic changes of user interest preferences will be an important research direction in future work.

References

[1] P. Vilakone, K. Xinchang, and D. S. Park, "Personalized movie recommendation system combining data mining with the k-clique method," *Journal of Information Processing Systems*, vol. 15, no. 5, pp. 1141-1155, 2019. https://doi.org/10.3745/JIPS.04.0138

[2] J. Lin, Y. Li, and J. Lian, "A novel recommendation system via L0-regularized convex optimization," *Neural Computing and Applications*, vol. 32, no. 6, pp. 1649-1663, 2020.

[3] B. Kaya, "A hotel recommendation system based on customer location: a link prediction approach," *Multimedia Tools and Applications*, vol. 79, no. 3, pp. 1745-1758, 2020.

[4] N. Deepa and P. Pandiaraja, "Hybrid context aware recommendation system for e-health care by merkle hash tree from cloud using evolutionary algorithm," *Soft Computing*, vol. 24, pp. 7149-7161, 2020. https://doi.org/10.1007/s00500-019-04322-7

[5] A. Da'u and N. Salim, "Recommendation system based on deep learning methods: a systematic review and new directions," *Artificial Intelligence Review*, vol. 53, pp. 2709-2748, 2020. https://doi.org/10.1007/s10462-019-09744-1

[6] S. A. Thekdi and S. Chatterjee, "Toward adaptive decision support for assessing infrastructure system resilience using hidden performance measures," *Journal of Risk Research*, vol. 22, no. 8, pp. 1020-1043, 2019. https://doi.org/10.1080/13669877.2018.1440412

[7] L. Lejeune, J. Grossrieder, and R. Sznitman, "Iterative multi-path tracking for video and volume segmentation with sparse point supervision," *Medical Image Analysis*, vol. 50, pp. 65-81, 2018.

[8] R. V. Babu and P. Parate, "Robust tracking with interest points: a sparse representation approach," *Image and Vision Computing*, vol. 33, pp. 44-56, 2015.

[9] S. Brenner, S. Aksin Sivrikaya, and J. Schwalbach, "Who is on LinkedIn? Self-selection into professional online networks," *Applied Economics*, vol. 52, no. 1, pp. 52-67, 2020. https://doi.org/10.1080/00036846.2019.1638497

[10] E. del Castillo, A. Meyers, and P. Chen, "Exponential random graph modeling of a faculty hiring network: the IEOR case," *IISE Transactions*, vol. 52, no. 1, pp. 43-60, 2020. https://doi.org/10.1080/24725854.2018.1557354

[11] Y. Si, F. Zhang, and W. Liu, "An adaptive point-of-interest recommendation method for location-based social networks based on user activity and spatial features," *Knowledge-Based Systems*, vol. 163, pp. 267-282, 2019.

[12] J. W. Bi, Y. Liu, and Z. P. Fan, "A deep neural networks based recommendation algorithm using user and item basic data," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 4, pp. 763-777, 2020.

[13] C. Yang, L. Bai, C. Zhang, Q. Yuan, and J. Han, "Bridging collaborative filtering and semi-supervised learning: a neural approach for poi recommendation," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Halifax, Canada, 2017, pp. 1245-1254.

[14] K. Cao, J. Guo, G. Meng, H. Liu, Y. Liu, and G. Li, "Points-of-interest recommendation algorithm based on LBSN in edge computing environment," *IEEE Access*, vol. 8, pp. 47973-47983, 2020.

[15] T. W. Zhang, W. P. Li, L. Wang, and J. Yang, "Social recommendation algorithm based on stochastic gradient matrix decomposition in social network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 2, pp. 601-608, 2020.

[16] C. Xu, L. Xu, Y. Lu, H. Xu, and Z. Zhu, "E-government recommendation algorithm based on probabilistic semantic cluster analysis in combination of improved collaborative filtering in big-data environment of government affairs," *Personal and Ubiquitous Computing*, vol. 23, no. 3, pp. 475-485, 2019.

[17] C. H. Lai and Y. C. Chang, “Document recommendation based on the analysis of group trust and user weightings,” *Journal of Information Science*, vol. 45, no. 6, pp. 845-862, 2019. https://doi.org/10.1177%2F0165551518819973

[18] M. Jiang, Z. Zhang, J. Jiang, Q. Wang, and Z. Pei, “A collaborative filtering recommendation algorithm based on information theory and bi-clustering,” *Neural Computing and Applications*, vol. 31, no. 12, pp. 8279-8287, 2019.

[19] W. Peng and B. Xin, “A social trust and preference segmentation-based matrix factorization recommendation algorithm,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, article no. 272, 2019. https://doi.org/10.1186/s13638-019-1600-4

[20] Z. Batmaz and C. Kaleli, “AE-MCCF: an autoencoder-based multi-criteria recommendation algorithm,” *Arabian Journal for Science and Engineering*, vol. 44, no. 11, pp. 9235-9247, 2019.

[21] L. Guo, H. Jiang, and X. Wang, “Location regularization-based poi recommendation in location-based social networks,” *Information*, vol. 9, no. 4, article no. 85, 2018. https://doi.org/10.3390/info9040085

[22] K. U. Kala and M. Nandhini, “Context-category specific sequence aware point-of-interest recommender system with multi-gated recurrent unit,” *Journal of Ambient Intelligence and Humanized Computing*, 2019. https://doi.org/10.1007/s12652-019-01583-w

[23] L. Cai, J. Xu, J. Liu, and T. Pei, “Integrating spatial and temporal contexts into a factorization model for POI recommendation,” *International Journal of Geographical Information Science*, vol. 32, no. 3, pp. 524-546, 2018. https://doi.org/10.1080/13658816.2017.1400550

[24] S. Zhao, I. King, and M. R. Lyu, “Aggregated temporal tensor factorization model for point-of-interest recommendation,” *Neural Processing Letters*, vol. 47, no. 3, pp. 975-992, 2018.

[25] S. Liu and L. Wang, “A self-adaptive point-of-interest recommendation algorithm based on a multi-order Markov model,” *Future Generation Computer Systems*, vol. 89, pp. 506-514, 2018.

[26] T. N. Nizar, S. Supatmi, and E. P. Putro, “Scale invariant feature transform descriptor robustness analysis to brightness changes of robowaiter vision sensor system,” *IOP Conference Series: Materials Science and Engineering*, vol. 662, no. 5, article no. 052004, 2019. https://doi.org/10.1088/1757-899x/662/5/052004

[27] Y. H. Li, X. D. Wang, and Z. Wang, “Compressed sensing imaging system based on improved theoretical model and its weighted iterative strategy,” *Optics Communications*, vol. 439, pp. 76-84, 2019.

[28] S. Ozkan, B. Kaya, and G. B. Akar, “Endnet: Sparse autoencoder network for endmember extraction and hyperspectral unmixing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 1, pp. 482-496, 2018.

[29] L. Caturegli, M. Gaetani, M. Volterrani, S. Magni, A. Minelli, A. Baldi, et al., “Normalized Difference Vegetation Index versus Dark Green Colour Index to estimate nitrogen status on bermudagrass hybrid and tall fescue,” *International Journal of Remote Sensing*, vol. 41, no. 2, pp. 455-470, 2020. https://doi.org/10.1080/01431161.2019.1641762

[30] M. Dai, Y. Hou, C. Dai, T. Ju, Y. U. Sun, and W. Su, “Characteristic polynomial of adjacency or Laplacian matrix for weighted treelike networks,” *Fractals*, vol. 27, no. 5, article no. 1950074, 2019. https://doi.org/10.1142/S0218348X19500749

Liqiang Sun **https://orcid.org/0000-0001-8693-7256**

He received the master's degree in computer science and technology from Ocean University of China, Qingdao, China, in 2009. He is currently an associate professor of computer science with Qingdao Vocational and Technical College of Hotel Management, Qingdao, China. His current research interests include big data and software engineering.