You Are What and Where You Are: Graph Enhanced Attention Network for Explainable POI Recommendation

Zeyu Li[1], Wei Cheng[2], Haiqi Xiao[1], Wenchao Yu[2], Haifeng Chen[2], Wei Wang[1]
[1]University of California, Los Angeles, [2]NEC Labs America
{zyli,haiqixiao,weiwang}@cs.ucla.edu,{weicheng,wyu,haifeng}@nec-labs.com

ABSTRACT

Point-of-interest (POI) recommendation is an emerging area of research on location-based social networks to analyze user behaviors and contextual check-in information. For this problem, existing approaches, with shallow or deep architectures, have two major drawbacks. First, for these approaches, the attributes of individuals have been largely ignored. Therefore, it would be hard, if not impossible, to gather sufficient user attribute features to have complete coverage of possible motivation factors. Second, most existing models preserve the information of users or POIs by latent representations without explicitly highlighting salient factors or signals. Consequently, the trained models with unjustifiable parameters provide few persuasive rationales to explain why users favor or dislike certain POIs and what really causes a visit. To overcome these drawbacks, we propose GEAPR, a POI recommender that is able to interpret the POI prediction in an end-to-end fashion. Specifically, GEAPR learns user representations by aggregating different factors, such as structural context, neighbor impact, user attributes, and geolocation influence. GEAPR takes advantage of a triple attention mechanism to quantify the influences of different factors for each resulting recommendation and performs a thorough analysis of the model interpretability. Extensive experiments on real-world datasets demonstrate the effectiveness of the proposed model. GEAPR is deployed and under test on an internal web server. An example interface is presented to showcase its application on explainable POI recommendation.

ACM Reference Format:
Zeyu Li[1], Wei Cheng[2], Haiqi Xiao[1], Wenchao Yu[2], Haifeng Chen[2], Wei Wang[1]. 2020. You Are What and Where You Are: Graph Enhanced Attention Network for Explainable POI Recommendation. In *Proceedings of CIKM 2021: 30th ACM International Conference on Information and Knowledge Management (CIKM'21).* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Point of interest (POI) recommendation is a critical component in the recommender system family. *Point of interest* refers to locations that customers of online business directories or review forums are interested in. Such directories or forums are typically named as *location-based social network* (LBSN), e.g., Yelp and Foursquare, since users interact with each other in various ways such as co-reviewing, co-visiting, or direct connecting via friendship relations.

POI recommendation has a wide coverage of scenarios in which the advertised items have significant spatial attributes that strongly influence the user decisions. Properly recommending POI replies on precisely understanding user taste, POI's property, geolocation, and their correlations. Varying from simple to sophisticated, existing algorithms are painstakingly customized for more precise user preference modeling, POI profiling, and user-POI relevance estimation. In other words, the development of POI recommendation systems witnesses the utilization of multiple modalities of data to achieve more satisfactory POI recommendations.

That being said, we point out two prevalent shortcomings of existing models: (1) inadequate interpretable motivation analysis for POI visits, and (2) absent attribute study for users with a diverse background.

First, for motivation analysis, the ranking functions of existing approaches merely fuse the multi-modal information without explicitly quantifying or explaining which modalities are comparatively more important than the others and which are less relevant. However, quantitatively comprehending the key causes of the check-ins is valuable because it is able to measurably interpret the users' mind-sets on choosing the next POI to visit. For example, some users always check in places their friends have checked in or have suggested, while others tend to visit places that their peer group favors. Such numerical motivation importance measurements can also reasonably provide a clear answer to the following debate. Tobler's first law of geography [36], frequently cited by previous work [19, 44, 50], states that: "Everything is related to everything else, but near things are more related than distant things." But authors of GeoMF state the opposite: a user's visit to certain POI implies exactly her indifference to those nearby, otherwise she would have visited them instead in the first place [25]. With numerical motivation analysis, it becomes easy to capture and interpret the primary causes of user check-ins, i.e. the motivations, which also benefits LBSN on explaining their recommendations. In contrast, existing approaches are not adaptive enough to learn different motivations in a transparent way. They instead simply use unweighted additions [31, 52] or feature vector concatenations [43] to mingle the intermediate information and produce recommendations. Motivation importance is hardly revealed by these operations. Such discrepancy calls for an effective architecture that is elaborately developed for interpretable motivation analysis with explicit salience distribution on different motivation factors.

Second, existing POI recommendation methods largely ignore user attribute study which, however, is of great importance. The extensive literature of item-based recommender systems, e.g., movies

CIKM'21, Nov 1–5, 2021, Online

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/1122445.1122456

and books, have demonstrated the potential of user profile, demographics, and their complex joint effects to enhance recommendation accuracy [2, 11, 13, 23, 34, 41]. Such potential is also plausible in the context of POI recommendation. For example, the young population loves to try different restaurants in different locations while the seniors may have distance concerns. However, user attribute information has been underestimated even by the recent deep learning-based POI recommendation models [31, 52, 54, 55, 57] although deeper models have superior ability to fuse different information modalities and capture the corresponding importance. Therefore, it is necessary to incorporate the user attribute features to comprehensively cover possible motivation factors.

To address the two aforementioned concerns, we propose a Graph Enhanced Attention network for explainable POI Recommendation (short for *GEAPR*) in this paper that recommends POIs in an adaptive and interpretable way. GEAPR leverages not only geographical and social information but also user personal attributes and provides an end-to-end justification of the recommendation in the meantime. Specifically, we decompose the possible motivating causes into four factors:

Structural Context. A check-in can be motivated by neighboring users with high structural proximity in the social network since they have a similar social context. This type of stimulus is typically ignored as it is latent and implicit. We argue and experimentally demonstrate later that social structural context is a critical cause of visits.

Neighbor Impact. Impact from direct neighbors, i.e., friends, is another factor of interest since people are likely to trust their friends' suggestions and check-in POIs their friends did before. Previous works characterize neighbor impact by MF-based methods which fail to generate explanations simultaneously.

User Attributes. Check-in behaviors can also be spontaneous due to users' characteristics such as age, religion, income level, etc. For example, young users may choose to check-in the POIs that other young people love without external stimulus such as friends. GEAPR proposes to understand the underlying correlation between check-in behavior and attributes in a novel manner. Factorization machines-based models are dedicated to learning from attribute data. Therefore, GEAPR utilizes a factorization machines equipped with the attention mechanism to learn attribute features.

Geolocation Influence. Geolocation influence has a particularly strong impact on POI recommendations because it is intuitive that people are more aware of nearby restaurants, supermarkets, or museums, etc. than distant ones. In GEAPR, we fix the POI influence distribution parameterized by Manhattan distance and learn the user preference for each geographical unit.

Altogether, GEAPR takes advantage of the attention mechanism to quantify the influences of different factors for each resulting recommendation and performs a thorough analysis of the model interpretability. Some literature [15] states that attentions lack robustness to serve as an explanation. We acknowledge the statement but argue that interpretability reveals the *salience* of the factors the model captures from the complex statistics of training data. Also, to the best of our knowledge, generating a fully human-readable explanation as a by-product of the ranking score is yet technically infeasible since even users themselves are unable to articulate the exact reasons that motivate a visit to a POI.

Please note that the geolocation feature encoding is decoupled from the three other factors that only focus on the user's personal motivation. The main rationale is for the compatibility: although GEAPR is applied to the POI recommendation, it can be painlessly transplanted to geolocation-irrelevant recommendation scenarios by simply detaching the geolocation module. Examples include movie recommendation [53], question routing [24], and new friend recommendations, etc.

Here we summarize the major contributions of this work.

- We propose GEAPR, a POI recommender that is able to interpret the POI prediction in an end-to-end fashion. It specifically focuses on four factors, namely structural context, neighbor impact, user attributes, and geolocation influence, and quantifies their influences by numeric values as the feature salience indicators.
- User attributes are taken into consideration in GEAPR. To the best of our knowledge, this is the **first** work that incorporate attributes to POI recommendation.
- Attention mechanism is used to address the recommendation interpretability by means of finding significant factors which are more influential in POI recommendation compared with other features.
- Extensive experiments are conducted on two real-world datasets from Yelp. Experimental results demonstrate the effectiveness of the proposed model. GEAPR has been deployed and tested on the internal web server [1]. Testing results demonstrate the effectiveness of GEAPR.

2 RELATED WORK

2.1 POI Recommendation

POI recommendation is a popular task since it directly affects the revenue and reputation of POI platforms. Research on this topic has been fruitful [4–6, 12, 16, 18–22, 25, 27–31, 33, 35, 39, 40, 43–52, 54–57]. We categorize them into *traditional* POI models and *deep learning-based* ones and discuss their pros and cons by examples.

Traditional models. USG [44] is a collaborative filtering-based model for POI recommendation. It suggests that not only social connections but also geographical influences can help improve the accuracy of POI recommendation. Therefore, USG specifically looks at three complementary factors: user preference of POIs, social influences, and geographical influence. GeoSoCa [50] digs deeper into POIs' property that the *category* of POI is taken into consideration. Authors argue that category is critical information and it affects user preference since people have different biases towards different types of POIs. Therefore, GeoSoCa firstly employs the biases measurement to build personalized POI popularity. ASMF and ARMF [19] refer to augmented square error based MF and augmented ranking error based MF, respectively. Despite the minor difference in the selection of error function, they both focus on user relations from three dimensions which are generally defined as "friendships", namely social friends, location friends, and neighbor friends. The emphases on user friendships strongly indicate that users' preference can be greatly reshaped by and effectively learned from human-human connections.

[1] See Section 5.

Deep learning-based models. PACE [43] utilizes a multi-task learning architecture that models user context, POI context, and user-POI interaction simultaneously. Technically, it assigns a learnable embedding vector to each user and POI to capture their latent features and use a feed-forward layers-based deep network to predict "user context", "POI context", and check-ins. SAE-NAD [31] is composed of a self-attentive encoder (SAE) for user-POI interaction modeling and a neighbor-aware decoder (NAD) for geographical context modeling. SAE differentiates user preference degrees in multiple aspects by self-attention. NAD ensures only physically and preferentially nearby users' check-ins receive stronger weights in the POI recommendation. APOIR [55] signifies the first application of generative adversarial network (GAN) [10] on POI recommendation. The co-trained two sides of the mini-max game are the recommender aiming to suggest the most probable POI check-ins and the discriminator that separates the recommended POI from the true visits.

Comparison with GEAPR. All aforementioned previous approaches carefully attend to user preferences mining and POI profiling in terms of categories and geolocations. However, we notice two major disadvantages that deserve some improvements. First, attributes of individuals have long been ignored even though recommendation models such as factorization machine (FM) [34] has demonstrated the usefulness of user attributes to enhance accuracy. Existing algorithms have delicate design on user preference and geolocation modeling [31, 54] but lack latent attribute learning for users. Second, all previous models, either deep learning-based or MF-based, preserve the information of users or POIs by latent representations without explicitly highlighting salient factors or signals. Different information sources are integrated by simple operations such as addition, concatenation, or multilayer perceptrons (MLP). Consequently, the trained models with unjustifiable parameters fail to explains why users favor or dislike certain POIs and what really causes a visit. However, GEAPR is able to address both concerns as shown in Section 3 and Section 4.

Sequential POI recommendation. Sequential, or *successive*, POI recommendation (SPR) models [16, 35, 39, 48, 52, 57] is a separate branch of location-based recommendations from *general* POI recommendations such as GEAPR. They are essentially different use scenarios. General models emphasize the modeling of general user and POI characteristics whereas SPR models focus on time-sensitive check-in suggestions and temporal POI visit behavior mining.

2.2 Attention Mechanism

Attention mechanism learns a function that generates weights to intermediate features in the model pipeline and manipulates the information which will be fed into other internal modules. Researchers in neural machine translation first apply attention for better alignments between source and destination language [1]. Being capable to identify important features and feature interactions makes attention mechanism a reliable way to explain the "thinking" of machine learning models [9]. Therefore, attention is considered as the solution to "interpretablity" in various research scenarios including recommender systems [23, 41, 47], graph representation learning [38], computer vision [42], etc. Recent progress in natural language processing strengthens the point that attention is also beneficial to performance enhancement [7, 8, 37].

3 THE GEAPR MODEL

In this section, we demonstrate the technical details of GEAPR.

3.1 Overview

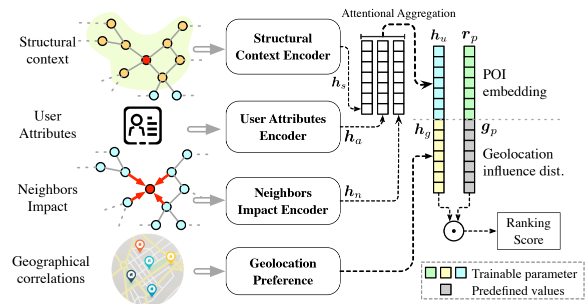


Figure 1: The overall pipeline of GEAPR.

Table 1: Partial notations for GEAPR that are essential.

Notation(s)	Definition
U, P, E	The sets of users, POIs, and friendship relations in an LBSN.
n_u, n_p	The total numbers of users and POIs in an LSBN.
G	The friendship graph for users. $G = \{U, E\}$.
$\mathcal{N}_G(u)$	The set of neighbors of user u in G and $\mathcal{N}_G(u) = \{v \mid (u, v) \in E\}$.
$\mathcal{F}, m$	The set of fields with m fields of user attributes. $m = \lvert\mathcal{F}\rvert$.
$\mathbf{M}_a$	The adjacency matrix of G, $\mathbf{M}_a \in \{0,1\}^{n_u \times n_u}$.
$\mathbf{M}_s$	The structural context matrix based on $\mathbf{M}_a$, $\mathbf{M}_s \in \mathbb{R}^{n_u \times n_u}$.
$\boldsymbol{h}_s, \boldsymbol{h}_n, \boldsymbol{h}_a$	Hidden vectors of structural context, neighbor impact, and attributes.
$\boldsymbol{h}_u$	The attentional aggregation of $\boldsymbol{h}_s$, $\boldsymbol{h}_n$, and $\boldsymbol{h}_a$.
$\boldsymbol{h}_g$	The geolocation preference of the user.
$\boldsymbol{g}_p$	The predefined geographical influence scores of POI to grids on a map.
$\boldsymbol{r}_p$	The hidden representation of POI semantics.
$s_{u,p}$	The score representing how likely user u will visit POI p.

The architecture of GEAPR is shown in Figure 1. Some important notations are summarized in Table 1. The inputs of GEAPR include the adjacency matrix $\mathbf{M}_a$ of the friendship graph of LBSN, structural context $\mathbf{M}_s$, the users' attributes $\mathcal{F}$, and the POI influence scores.

GEAPR uses three different architectures customized for the three factors on the user motivation side. Specifically, a dense neural network-based structural context encoder is utilized to learn the **structural context**, a graph neural network-based attentional friendship encoder is utilized to model the **neighbor impact**, and an attention-based latent factorization machine is utilized for preserving the **attribute interactions**.

These three sub-modules will generate three hidden feature representations individually as $\boldsymbol{h}_s$, $\boldsymbol{h}_n$, and $\boldsymbol{h}_a$. The information from three sources is then merged by an attentional aggregation [23] strategy which is able to reveal the relative salience among them. The merged motivation representation is then combined with geolocation features as constraints so that strongly relevant but distant POIs will be removed from the recommendation. GEAPR then takes the dot-product of the graph-enhanced user embedding and the POI embedding to generates a scalar score $s_{u,p}$ representing the likelihood of a user u visiting a POI p in the future.

In order to preserve reliable interpretability while making an accurate recommendation, the building blocks of GEAPR focus on attention-based algorithms. Although literature argues that attention lacks the potential to provide an "explanation" that agrees with human perception, it still reveals the *distribution of salience* which can be considered as a form of explanation. In addition, it is worth noting that unlike the tasks where ground truth is typically defined and easily accessible, formal explanations are unavailable from LBSN or public datasets for interpretable POI recommendation as ground truth. Therefore, measuring the "correctness" of the generated interpretation is impracticable.

3.2 Structural Context Factor

The structural context tries to model the commonality of the close neighbors of a certain user. Intuitively, the proximity of user characters and preferences can be propagated through a few hops of social connections to form cliques within the network. In order to capture social context from network structure, GEAPR utilizes Random Walk with Restart (RWR), a popular method widely used for learning community proximity [32]. The structural context features of a user are learned based upon his or her RWR representation.

Mathematically, given a network G of n_u nodes represented by its adjacency matrix $\mathbf{M}_a$ with $\mathbf{M}_{a,ij} = 1$ if nodes i and j are connected and otherwise 0, a starting user u_0 in U, the r-step RWR vector $\boldsymbol{p}^{(r)} \in \mathbb{R}^{n_u}$ is computed by

$$\boldsymbol{p}^{(r)} = \gamma \boldsymbol{p}^{(0)} + (1-\gamma)\boldsymbol{p}^{(r-1)}[\mathbf{D}^{-1}\mathbf{M}_a],$$

where γ denotes the probability that the random walk generator restarts from u_0, $\boldsymbol{p}^{(0)}$ denotes the corresponding row of u_0 in $\mathbf{M}_a$, and $\mathbf{D}$ denotes a diagonal matrix with $\mathbf{D}_{ii} = \sum_{j=1}^{n_u} \mathbf{M}_{a,ij}$.

Let R denote the maximum step of the RWR process, the summation of $\boldsymbol{p}^{(r)}$ is considered as the structural context. R is usually set as a small value such as 2 or 3 to make sure only *local* information is preserved in $\boldsymbol{h}'_s$ and $\boldsymbol{h}'_s = \sum_{r=1}^{R} \boldsymbol{p}^{(r)}$, $\boldsymbol{h}'_s \in \mathbb{R}^{n_u}$.

However, one problem of encoding the local context is the enormous dimension: the size of $\boldsymbol{h}'_s$ is the same scale as the user numbers. Therefore, GEAPR conducts dimension reduction to $\boldsymbol{h}'_s$ to generate $\boldsymbol{h}_s$, the latent features of structural context, by a multi-layer dense neural network with $\mathrm{ReLU}(x) = \max(0, x)$ as the activation function (using two layers as an example):

$$\boldsymbol{h}_s = \mathrm{ReLU}(\mathbf{W}_2^T(\mathrm{ReLU}(\mathbf{W}_1^T \boldsymbol{h}'_s + \boldsymbol{b}_1)) + \boldsymbol{b}_2), \quad \boldsymbol{h}_s \in \mathbb{R}^d,$$

where d is the dimension of hidden representations and $\{\mathbf{W}_i, \boldsymbol{b}_i\}$ are trainable parameters. $\mathrm{ReLU}(\cdot)$ introduces non-linearity and enhances the representation learning capacity for structural context.

3.3 Neighborhood Impact Factor

The second aspect of potential visit stimuli is the direct friends since one may naturally check in the POIs suggested by friends. We thereby focus on the understanding of impact from neighbors $\mathcal{N}_G(u)$ of a user u. Graph attention network (GAT) [38] provides an effective way to aggregate information from direct neighbors and compute the attention to pinpoint significant neighbors. Therefore, we encode neighborhood impact using an attention-based graph neural-network. Given a user u and the friends of u, $\mathcal{N}_G(u)$, the hidden neighbor feature $\boldsymbol{h}_n$ is

$$\boldsymbol{h}_n = \sigma\left(\sum_{j \in \mathcal{N}_G(u)} \alpha_{uj} \mathbf{W}_n \boldsymbol{v}_j\right).$$

$\sigma(\cdot)$ is typically a non-linear function such as $\mathrm{ReLU}(\cdot)$ or $\tanh(\cdot)$. $\mathbf{W}_n \in \mathbb{R}^{d_n \times d_p}$ is a learnable weight matrix for the attention network that maps all neighbor embeddings to a common space. $\boldsymbol{v}_j$ is derived by the average POI embeddings that user j visited before. The scalar α_{uj} is the weight from user j to u and GEAPR computes α_{uj} by Eq. (1) where *LeakyReLU* advances ReLU in that it allows shrunk negative signal to flow through, "||" denotes concatenation along an existing dimension, and $\boldsymbol{a} \in \mathbb{R}^{2d}$ is a trainable vector that helps compute the attention logits and $\mathbf{W} \in \mathbb{R}^{d \times d_p}$.

$$\alpha_{uj} = \frac{\exp\left(\mathrm{LeakyReLU}\left(\boldsymbol{a}^T[\mathbf{W}\boldsymbol{v}_u \| \mathbf{W}\boldsymbol{v}_j]\right)\right)}{\sum_{i \in \mathcal{N}_G(u)} \exp\left(\mathrm{LeakyReLU}\left(\boldsymbol{a}^T[\mathbf{W}\boldsymbol{v}_u \| \mathbf{W}\boldsymbol{v}_i]\right)\right)} \tag{1}$$

The set of attention weights α_{uj} demonstrates the influential neighbors. In addition to concatenation, other ways can also produce attention logits such as dot-product of $\boldsymbol{v}_j$ and $\boldsymbol{v}_u$, the matrix-dot-product $\boldsymbol{v}_j^T \mathbf{W} \boldsymbol{v}_u$, or the non-linear MLP of concatenation of $\boldsymbol{v}_j$ and $\boldsymbol{v}_u$. The original GAT [38] can handle multiple attention heads and multiple neighborhood hops. Increased head numbers can preserve information in more sub-spaces, and enlarged scopes of direct or nearby users bring in more local context, which both benefit the performance. However, in consideration of the interpretability, we simplify the settings to a single head and one-hop neighbors.

3.4 Attribute Interactive Factor

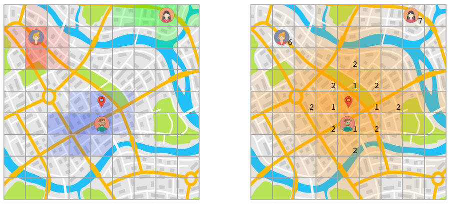
Apart from the effects of social structural context and direct neighbors, the personal attributes are also important factors to motivate the user to visit particular POIs. The combinatorial possibilities of feature interactions create diverse influences on the users' preference towards POIs, which has been thoroughly studied in feature-based recommender systems such as factorization machines (FM) [34], DeepFM [11], and xDeepFM [26], etc. In GEAPR, we combine feature-based FM methods with attention mechanism [41] to analyze feature interaction and maintain interpretability.

Embedding the categorical and numerical features into a lower-dimensional space is a prerequisite [11, 23, 26, 34, 41]. User attributes can be written as m fields $\{F_1, \ldots, F_m\}$ with different values, also known as *features*. We assign a trainable vector to each distinct feature $\boldsymbol{f}_c \in \mathbb{R}^{d_a}$ for categorical field and discretize the continuous value by bucketing and then treat the converted alternative as a categorical feature for the numerical field.

Given the feature embeddings, user attribute impact is model as

$$\boldsymbol{h}_a = \boldsymbol{w}_0 + \sum_{i=1}^{m} \beta_i \boldsymbol{f}_i + \sum_{i=1}^{m} \sum_{i=j+1}^{m} \lambda_{ij} \boldsymbol{f}_i \odot \boldsymbol{f}_j, \tag{2}$$

where $\boldsymbol{w}_0$ is the offset term, β_i, and λ_{ij} are the attention weights for first-order and second-order feature interactions. They are computed as follows, given feature matrix $\mathbf{F} = \{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m\}$, $\boldsymbol{\beta} = \mathrm{softmax}(\mathrm{ReLU}(\boldsymbol{q}_1^T \mathbf{F}))$, $\lambda_{ij} = \mathrm{softmax}(\boldsymbol{q}_2^T \mathrm{ReLU}(\mathbf{W}_a(\boldsymbol{f}_i \odot \boldsymbol{f}_j) + \boldsymbol{b}))$. Here $\boldsymbol{q}_1, \boldsymbol{q}_2 \in \mathbb{R}^{d_a}$, $\mathbf{W}_a \in \mathbb{R}^{d_a \times d_a}$ and $\boldsymbol{b} \in \mathbb{R}^{d_a}$ denote learnable tensors to build attention weights. $\odot$ is the element-wise multiplication. Once more, we use the attention weights as the information source for interpretability.



User geo-preference POI influence area

Figure 2: Geolocation encoding for users and POIs. Left: three trainable user preference geographical distributions. Right: an example of pre-defined POI influence score measured by Manhattan distance.

Eq. (2) contains an enumeration of the first-order features and second-order feature interactions which are incomprehensive compared with the exponential-sized feature interaction space. Studies have shown that the first two orders of features are already capable of contributing sufficient information for learning interactive features and adding higher-order features is only making a marginal information supplement. That being said, GEAPR still enjoys great compatibility with higher-order features interaction models [2, 23].

3.5 POI Geographical Influence

In GEAPR, we model the geographical influence features from two aspects: learnable user geolocational interest and predefined POI area influence as shown in Figure 2. Specifically, we first divide the city map of POIs into grids with n_{lat} units on the latitude axis and n_{long} units on the longitude axis. We model the geographical influence of a POI in grid p to a target grid t using the influential score $g_{p,t}$ [25] as

$$g_{p,t} = K\left(\frac{d_{\text{man}}(p,t)}{\sigma_g}\right)$$

where σ_g denotes the standard deviation of distances, $K(\cdot)$ denotes a standard normal distribution, and $d_{\text{man}}(a,b)$ measures the Manhattan distance from the grid a to grid b. Therefore, we can define the influential score vector $\boldsymbol{g}_p$ of POI p as $\boldsymbol{g}_p \in \mathbb{R}^{(n_{\text{long}} \cdot n_{\text{lat}})}$ which is essentially a flattened 2-dimensional influential score matrix. We are also curious about the geographical preference distribution of users which is defined as a learnable parameter representing user preference, $\boldsymbol{h}_g \in \mathbb{R}^{(n_{\text{long}} \cdot n_{\text{lat}})}$. Each user has one unique $\boldsymbol{h}_g$. We define the geographical influence correlation between users and POIs by taking the product $\boldsymbol{h}_g^T \boldsymbol{g}_p$. The overlap between user-preferred regions and POI influential regions can be selected and amplified by multiplication.

3.6 Objective and Optimization

After showing the derivations of the representation of the four causing factors, we show how to make predictions for future check-ins. We first aggregate $\boldsymbol{h}_s$, $\boldsymbol{h}_n$, and $\boldsymbol{h}_a$ by an attention mechanism as Eq. (3) and Eq. (4) since they all encode users motivation.

$$\boldsymbol{h}_u = \pi_s \cdot \text{ReLU}(\boldsymbol{h}_s) + \pi_n \cdot \text{ReLU}(\boldsymbol{h}_n) + \pi_a \cdot \text{ReLU}(\boldsymbol{h}_a) \quad (3)$$

$$\pi_{x \in \{s,n,a\}} = \frac{\exp(\boldsymbol{w}^T \text{ReLU}(\boldsymbol{h}_x))}{\sum_{x' \in \{s,n,a\}} \exp(\boldsymbol{w}^T \text{ReLU}(\boldsymbol{h}_{x'}))} \quad (4)$$

Then Eq. (5) computes the possibility of the potential check-in $s_{u,p}$ which is defined by the dot-product with motivation feature and geographical feature of users and POIs. If $\boldsymbol{r}_p$ represents the motivation-related POI semantics, then

$$s_{u,p} = [\boldsymbol{h}_u || \boldsymbol{h}_g] \cdot [\boldsymbol{r}_p || \boldsymbol{g}_p] = \boldsymbol{h}_u^T \boldsymbol{r}_p + \boldsymbol{h}_g^T \boldsymbol{g}_p. \quad (5)$$

The overall objective function is Eq. (6) which sums a ranking loss L_{rank} and a regularization loss L_{reg} weighted by a hyper-parameter. In GEAPR, we use L_2 norm as the regularization term.

$$L = L_{\text{rank}}(\mathcal{D}, \mathcal{D}') + cL_{\text{reg}} \quad (6)$$

We use negative sampling to implement the ranking term that specifically penalize on the negative samples $\mathcal{D}'$ while optimizing the positive samples $\mathcal{D}$. There are two standard ways to implement the ranking loss, L_{rank}, namely *pair-wise* or *point-wise* ranking loss.

Point-wise (PO) Loss. Point-wise loss forces the positive instances to approach an indicator 1 and pushes the negative instances to indicator 0 via a cross-entropy loss of binary classification. $y = 1$ if $(u,p) \in \mathcal{D}$, $y = 0$ if $(u,p) \in \mathcal{D}'$, and $\sigma(\cdot)$ is the sigmoid function.

$$L_{\text{rank-po}} = -\sum_{\mathcal{D},\mathcal{D}'} \left(y \log(\sigma(s_{u,p})) + (1-y)\log(1-\sigma(s_{u,p})) \right).$$

Pair-wise (PA) Loss. Pair-wise loss tries to capture the partial order relationships in the training data and maintain that order between the scores of positive instances and negative instances. We follow the method in RankNet [3] as the equation below with $(u,p) \in \mathcal{D}$, $(u,p') \in \mathcal{D}'$, and $\Delta_{u,p,p'} = s_{u,p} - s_{u,p'}$.

$$L_{\text{rank-pa}} = \sum_{\mathcal{D},\mathcal{D}'} -\Delta_{u,p,p'} + \log(1 + \exp(\Delta_{u,p,p'})).$$

We use Adam [17] to optimize the parameters since all modules in GEAPR are continuous and differentiable.

Complexity. We use $\delta(n)$ to denote the complexity of the multiplication of an n dimensional vector and an $n \times n$ dimensional matrix. We omit the detailed computation of the complexity and give the result as follows. Summing up complexity of the four modules and of the overall optimization, the forward pass complexity of GEAPR is $T = r\delta(n_u) + \delta(d) + n_u\delta(d_p) + m^2\delta(d_a) + O(n_{\text{long}} \times n_{\text{lat}})$. Further, considering that different dimensions are on the same scale, we rewrite T as

$$T = r\delta(n_u) + (n_u + m^2)\delta(d) + O(n_{\text{long}} \times n_{\text{lat}}).$$

If hyperparameters r, d, m, n_{long}, and n_{lat} are set, they can be viewed as constants. Then the complexity T of a training forward pass is proportional to $\delta(n_u)$, which equals to $O(n_u^2)$ and is common to graph neural networks [38].

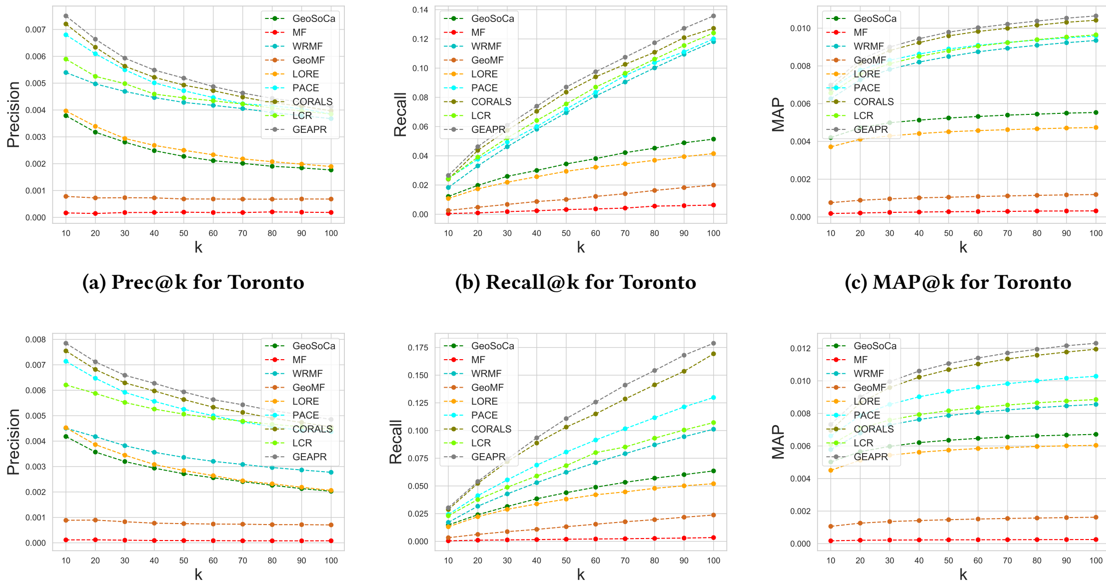
4 EXPERIMENTS

This section reports the evaluation of GEAPR on effectiveness and interpretability by its performance on real-world datasets and case studies of interpretation.

4.1 Experimental Settings

4.1.1 Dataset. We use Yelp Challenge[2] Round 13 dataset for effectiveness and interpretability tests. We divide the reviews by cities and subsets of "Toronto" and "Phoenix" are used due to larger sizes. Yelp dataset contains comprehensive details of the reviews, user attributes, user friendships, and POI locations. The friendship

[2]https://www.yelp.com/dataset



(a) Prec@k for Toronto (b) Recall@k for Toronto (c) MAP@k for Toronto

(d) Prec@k for Phoenix (e) Recall@k for Phoenix (f) MAP@k for Phoenix

Figure 3: Performance evaluation of GEAPR compared with baseline models.

network in the datasets serves as the information source of the neighbor impact encoder and the structural context encoder of GEAPR. The adjacency matrix is the input of the neighbor impact encoder and also helps to create the structural context. Yelp Round 13 dataset *does not* include the owners of the check-in records. We instead consider the abundant reviews to be equivalent to check-ins since intuitively each review corresponds with a past check-in. We sort all reviews grouped by users chronologically and filter out users with less than 10 reviews to avoid cold-start. We partition the review set of each individual into 9:1 where 90% data is used for training and the 10% rest for testing. Table 2 shows the statistics.

4.1.2 Metrics and Baseline Models. For effectiveness evaluation, we consider three metrics widely utilized in information retrieval and recommender systems, including Mean Average Precision at K (**MAP@k**), Precision at k (**Prec@k**), and Recall at k (**Recall@k**). For the i-th test sample, given the list of the top k POIs ranked according to the predicted scores $\boldsymbol{s}_i = (s_{i1}, s_{i2}, \ldots, s_{ik})$, and the m POIs that a particular user checked-in in the test set $T_i = \{t_{i1}, \ldots, t_{im}\}$, the three metrics are defined as follows where n is the number of test samples and $\boldsymbol{s}_{i,1-j}$ denotes the prefix sublist of $\boldsymbol{s}_i$ of length j. Prec@k $= \frac{1}{n}\sum_{i=1}^{n}\frac{|T_i \cap \boldsymbol{s}_i|}{k}$, Recall@$k$ $= \frac{1}{n}\sum_{i=1}^{n}\frac{|T_i \cap \boldsymbol{s}_i|}{|T_i|}$, MAP@$k$ $= \frac{1}{n}\sum_{i=1}^{n}\mathrm{AP}_i(T_i, \boldsymbol{s}_i)$, and $\mathrm{AP}_i(T_i, \boldsymbol{s}_i) = \frac{1}{k}\sum_{j=1}^{k}\frac{|T_i \cap \boldsymbol{s}_{i,1-j}|}{j}$. All three metrics take values in $[0, 1]$. Larger values represent better results. Prec@k and Recall@k measure how good the top-ranked POIs match with the ground truth and MAP@k signifies if the ground truth is ranked at higher positions. k is from 10 to 100.

Eight baselines are used for performance comparison including **MF**, **GeoMF**, **WRMF**, **LORE**, **GeoSoCa**, **PACE**, **CORALS**, and **LCR**. They include matrix factorization based models and deep-learning-based models. We provide brief descriptions for the baseline models for comparisons with GEAPR[4]. The parameter settings follow the default values in the source code or in their original papers. **MF**, the Matrix factorization model that decomposes the user-POI check-in matrix and make predictions by reconstruction. **GeoMF** [25], a geolocation-enhanced MF model that considers the geographical factors as constraints. **WRMF** [14], Weight regularized MF, incorporates both implicit and explicit check-ins for future check-in predictions. **LORE** [51] builds a location-location transition graph to specifically model the sequential influence of POIs to user preferences. In **GeoSoCa** [50], Geolocation, social connections, and POI categories are all considered to extract diverse information. **PACE** [43] is a deep learning-based multi-task learning algorithm that predicts POI, user context, and spot context simultaneously for better accuracy and robustness. **CORALS** [20] incorporates the modeling of reputation prediction of POIs in location recommendations. **LCR** [28], Local Collaborative Ranking, assumes that user-POI matrix is local low-rank rather than global low-rank to mitigate the data sparsity issue of POI recommendation.

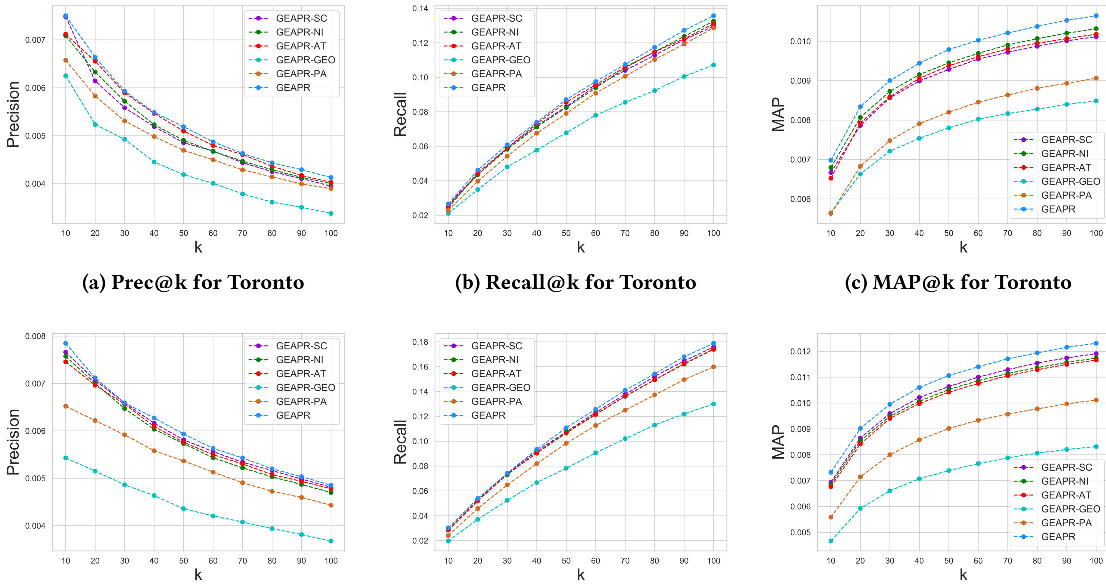
Some models mentioned in Section 2 are excluded from the comparison due to a lack of high-quality implementation or instructions for reproduction. Even so, our experiments are still able to

Table 2: Statistics of the datasets for evaluation.[3]

Dataset	#.User	#.POI	#.Reviews	#.U-Cxn	%.Reviews	%.U-Cxn
Toronto	9582	9102	234388	104402	2.687×10^{-3}	1.139×10^{-3}
Phoenix	11289	9633	249029	163900	2.290×10^{-3}	1.286×10^{-3}

[3] "#": *the count of*; "%": *the density of*; "U-Cxn": user friendship connections.

[4] The implementation of PACE: https://github.com/yangji9181/PACE2017; The implementation of the rest of baselines: http://spatialkeyword.sce.ntu.edu.sg/eval-vldb17.



(a) Prec@k for Toronto (b) Recall@k for Toronto (c) MAP@k for Toronto

(d) Prec@k for Phoenix (e) Recall@k for Phoenix (f) MAP@k for Phoenix

Figure 4: Ablation study of GEAPR compared with its variants.

provide coverage on both traditional methods and deep learning-based methods. In addition, sequential POI recommendation models [16, 35, 39, 48, 52, 57] are excluded as well due to different use cases as discussed earlier in Section 2.

4.1.3 Reproduction. The prototype of GEAPR was implemented by Python (3.6.8) and TensorFlow (1.14.0) and run with a 16 GB Nvidia Tesla V100 GPU embedded in a Nvidia DGX-1 server. The code is publicly available on GitHub[5]. and a comprehensive end-to-end instruction on how to run the code is also provided.

The hyper-parameter settings that generate the results in Section 4 are listed in Table 3. In order to prevent overfitting, dropout is employed in the graph attention network module for structural context modeling and the attentional factorization machines module for attribute impact modeling. The dropout rates are also shown in Table 3.

Table 3: Parameter settings for the experiments.

Parameters	Toronto	Phoenix
d_a, d_n, d_p, and d	{64,64,64,32}	
n_{long}, n_{lat}	{30, 30}	
R, γ, and c	{3, 0.05, 0.1}	
Loss function	Point-wise	
Regularization function	L_2, $c = 0.0001$	
Optimizer	Adam	
Learning rate	0.001	
Hidden layers of SC module	2 layers: {64,48}	
Negative sampling (P:N)	1:105	1:90
NI module dropout	0.3	0
AT module dropout	0.3	0.2

[5]The source code is available here: https://github.com/zyli93/GEAPR

4.2 Effectiveness

This section reports effectiveness by comparing GEAPR with the baselines for performance analysis and its own variants for ablation study. A parameter sensitivity study is also provided to cast light on the heuristics to parameter tuning. Please note that POI recommendation tries to identify all potential POIs from an enormous candidate base, which is essentially hard due to the unpredictability of users' minds. Users can receive multifaceted stimuli, a great proportion of which are implicit and difficult to capture based merely on the dataset. As a result, the numeric values of POI recommendation are **relatively small** which also applies to all state-of-the-art models.

4.2.1 Comparison with Baselines. The experimental results on effectiveness are shown in Figure 3. Point-wise (PO) loss is selected due to its superior performance and the results by pair-wise loss are shown in Section 4.2.2. It is demonstrated that GEAPR can achieve the state-of-the-art result by outperforming MF, GeoMF, WRMF, LORE, GeoSoCa, PACE, CORALS, and LCR on all three metrics. In other words, the ground truth of the test samples is effectively identified at high ranking positions. The advancement of deep learning has pushed the performance of this task to an extreme such that numerically small performance increases should also be considered **significant**. Therefore, GEAPR has made significant progress on accurate POI recommendation. As PACE is deep learning-based and requires larger input data density, its performance slightly drops under the settings of Table 2. CORALS requires dense review data for location reputation modeling and hence the density hurts its performance as well. The architecture that maintains interpretability requires GEAPR to *avoid unjustifiable* element-wise feature multiplication and aggregate information by weighted average pooling.

Hence some complex or subtle information is missed and the learning power of GEAPR is undermined. That being said, GEAPR still achieves great performances.

4.2.2 Ablation Study. Ablation study illustrates in Figure 4 the contributions of the four factors and the performance difference of the two ranking losses. "GEAPR-X" refers to the variant that differs from GEAPR by "X". X can be "SC" (no structural context), "NI" (no neighbor impact), "AT" (no user attributes), "GEO" (no geolocations), or "PA" (uses pair-wise loss). The curve of GEAPR is higher than other variants in all six subfigures. Four conclusions are drawn: (1) Attribute information provides performance gain indicating its usefulness. The reason of the relatively small contribution is the lack of the diversity of attribute information revealing user interests such as age and gender; (2) Removing the geolocation causes the largest performance deterioration. Therefore, the geographical information is the most influential factor as it is closely related to the POI task. It also shows that accurately mining geographical information is critical in accurate POI recommendation; (3) Taking away any factor will hurt the performance meaning that *all three non-geographical factors take effect and contribute uniquely* to the performance increase; (4) Pair-wise loss is not as useful as point-wise loss. Pair-wise loss relies on an advanced sampling strategy to sample data that is closer to the "decision hyperplane" with larger probabilities at the cost of extra computation. Instead, GEAPR draws negative samples uniformly with the goal of being generalizable to the diverse possible distributions of the data.

4.2.3 Parameter Sensitivity. We briefly introduce the parameter sensitivity study of GEAPR. None of the tuned parameters has a monotonic relationship with the evaluation results. For example, a greater negative sampling ratio will slow down the training and overly penalize on the unobserved check-ins so that the ground truth can also be mistakenly concealed, whereas a smaller ratio overfits the positive samples but underfits the negative, hurting the performance in another way. As for embedding size d, assigning d as 32, 64, or 128 only produces a little performance fluctuation but making it larger or smaller will deteriorate the recommendation accuracy. There is no theoretical guarantee on their optimality of the hyper-parameters used to derive the results in Figure 3. It is plausible to grid-search for other settings with better performances.

4.3 Interpretability

In this section, we demonstrate the interpretability of GEAPR by plotting the heat maps of $\boldsymbol{\pi} = \{\pi_s, \pi_n, \pi_a\}$, $\boldsymbol{\alpha}_u$ in the graph attention network module for neighbor impact, $\boldsymbol{\beta}$ and $\boldsymbol{\lambda}$ in the attribute influence module denoting the first- and second-order interaction importance, respectively, and $\boldsymbol{h}_g$ for geolocation feature. They are designed to probe the importance of features and generate interpretation accordingly.

4.3.1 User Motivation Study. We plot three examples on this topic shown in Figure 5, 6, and 7. The blue bars show the motivation breakdown $\boldsymbol{\pi}$; the green bars show the attributes importance $\boldsymbol{\beta}$; the variable-length dark orange bars show the friends count of a user and the weights learned from the neighbor impact module ($\boldsymbol{\alpha}_u$). The $\boldsymbol{\lambda}$ is not plotted since it is observed that the values in $\boldsymbol{\lambda}$ are almost identical, showing that the second-order features are orthogonal to

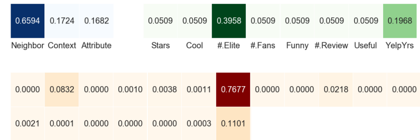


Figure 5: Example with significant neighbor impact. User neighbor ID are omitted for better visualization.

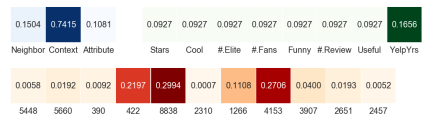


Figure 6: Example with significant structural context.

Figure 7: Example with significant user attribute.

the formation of user motivation. This fact is consistent with our perception since features provided by the dataset are mostly counts or scores (e.g., review counts, funny score, cool score, etc.) and are independent of each other by intuition.

Figure 5 shows an example with important *neighbor feature* since its weight is the largest in the blue heat map. From the 19 neighbors of the user, only one user has a huge impact on causing the user's motivation. The single peak of the user neighbor impact boosts of direct neighbor importance and decreases of the context importance. And the strong neighbor and context combined depress the weight of attribute since all weights add up to 1.

Figure 6 shows an example with important *structural context* features. The actual structural context is hard to depict since it is a dense vector with dimension size identical to the number of users. But we can still observe that the friends of this user are contributing more evenly impacts compared with Figure 5. That is, the user's visit preference is actually influenced by many neighbors and potentially the further neighbors, i.e., the users in the same structural context or community.

Figure 7 exemplify a case with both important *attribute* information and important *neighbor impact*. The lack of friendship connections (only 5 friends) and the single-peaked neighbor impact push the model to learn motivation from attributes. Therefore a heavyweight is put on `YelpYrs` that stands for the number of years the user had been on Yelp. It turns out that the user's "Yelping years" is 8 years, a time long enough to form a user's visiting habits. From the figures, we notice a common property of attribute weights (green bars) that `#.Elite` and `YelpYrs` are usually highlighted as opposed to other user attributes. `#.Elite` denotes how many times the user had been awarded as "Yelp elite". We suppose that these two features have tight bonds with user activity and are understood by the model as a signal of a generally stronger motivation.

In addition, we acknowledge the insufficiency of attribute information of the Yelp dataset even though other POI datasets do *not* provide attributes at all. Without informative features, the models

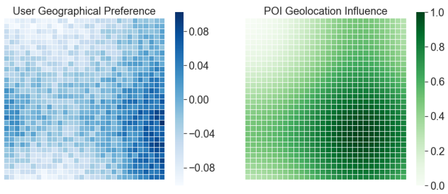


Figure 8: User geographical preference *v.s.* single POI influence distribution

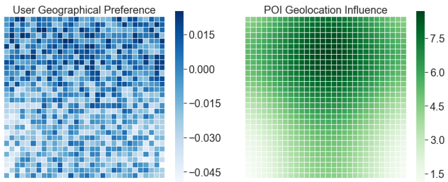


Figure 9: User geographical preference *v.s.* sum of POI influence distribution

can not produce convincing rationale but complex statistical signals from the data.

4.3.2 Geolocation Preference. As shown in Eq. (5), the geographical correlation with user and POI is modeled by a linear dot-product. That is, a greater preference value represents a stronger inclination to that zone. We exemplify the results by showing the learned user preference against (1) the predefined influence distribution of a single visited POI and (2) the normalized summation of the influence distributions of all visited POIs in Figure 8 and Figure 9. Each figure contains 30×30 grids each representing a grid in the real-world map of POIs.

In Figure 8, we observe that the user preference values concentrate at the bottom right corner which agrees with the influential center of the POI. In addition, the top left corner is not favored by the user and the POI influence figure also shows the same pattern of low influence. In Figure 9, the heat zones are generally aligned between the two figures. The upper halves of both user preference and POI influence are heated and the bottom parts are relatively inactive. This demonstrates that GEAPR is capable of capturing the geographical preference of a user and understand which parts the more favored than other parts.

5 INTERFACE OF DEPLOYED GEAPR

In this section, we showcase an example interface of a POI recommender system under test deployed on an internal web server in Figure 10. It is equipped with GEAPR as the recommendation engine. We take a screenshot and redact a number of sensitive fields such as the user name and the profile images of related friends.

We run a recommendation query and three results are returned with different rationales behind the recommendations. Specifically, the first recommendation result is mined from structural context since the user has a big group of similar users within his or her close social community that have visited this place. For the second recommendation, the user has three direct friends who have visited this POI and rated highly of it. Our model also ranks them based on their numerical contributions to the overall neighbor impact. For the third recommendation, both structural context and attribute information take effect. Specially, attribute factor encoder suggests that users with a big `YelpYrs` and a large number of `Elite` tend to visit this POI. Therefore, the user will be notified which attributes motivate the recommendations.

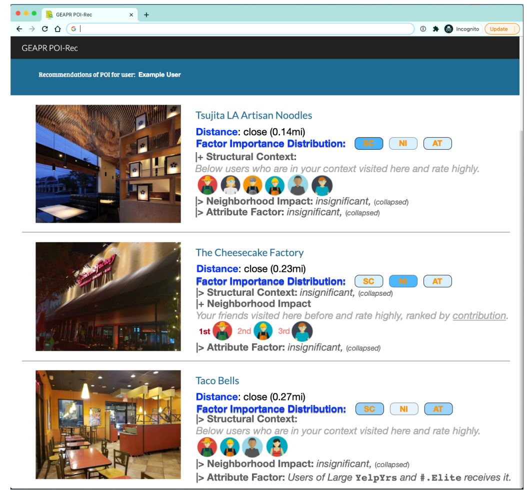


Figure 10: An interface of the explainable recommendations deployed on our internal web server. Due to the privacy leakage concern, this image has been edited: the user name is redacted and the user profile images are replaced with cartoon icons. The fields of *Factor Importance Distribution* show the significantly influential factors such as "SC" for the first recommended POI, "NI" for the second, and "SC"+"AT" for the third. Output of the four factors is presented underneath the POI name to explain the recommendations including distances (geographical factor), structural context factor, neighborhood impact, and attribute factor. The insignificant factors are collapsed in the presentation.

Although the system is still under test, we can draw the conclusion that GEAPR is effective for POI recommendation. Both users using the POI recommendation service and the service provide will benefit from the superior accuracy and good interpretability of GEAPR.

6 CONCLUSION

In this paper, we propose GEAPR, a graph-enhanced POI recommendation algorithm that incorporates user friendship network information in addition to user attributes and geolocation features. Specifically, GEAPR decomposes the motivation of user check-ins into four different aspects: social structural context, neighborhood impact, user attribute, and geolocation, and quantifies the importance of each feature. In addition, GEAPR employs the attention mechanism to generate interpretations that reveal the salient motivating factors, influential neighbors, informative attribute interactions, and heated geographical areas, etc. Experimental results demonstrate the effectiveness and interpretability of GEAPR.

We list the following potential improvements as future work: (1) It may be helpful to also build a POI graph by semantics and apply graph mining algorithms; (2) A better way to preserve non-linear geolocational features is needed to learn complex information.

REFERENCES

[1] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv e-prints* abs/1409.0473 (2014).

[2] Mathieu Blondel, Akinori Fujino, Naonori Ueda, and Masakazu Ishihata. 2016. Higher-order factorization machines. In *NIPS*.

[3] Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. 2005. Learning to rank using gradient descent. In *ICML*.

[4] Anirban Chakraborty, Debasis Ganguly, and Owen Conlan. 2020. Relevance Models for Multi-Contextual Appropriateness in Point-of-Interest Recommendation. In *SIGIR*.

[5] Chaochao Chen, Ziqi Liu, Peilin Zhao, Jun Zhou, and Xiaolong Li. 2020. Privacy preserving point-of-interest recommendation using decentralized matrix factorization. *arXiv preprint arXiv:2003.05610* (2020).

[6] Jingyuan Chen, Hanwang Zhang, Xiangnan He, Liqiang Nie, Wei Liu, and Tat-Seng Chua. 2017. Attentive collaborative filtering: Multimedia recommendation with item-and component-level attention. In *SIGIR*.

[7] Zihang Dai, Zhilin Yang, Yiming Yang, Jaime G. Carbonell, Quoc Viet Le, and Ruslan Salakhutdinov. 2019. Transformer-XL: Attentive Language Models beyond a Fixed-Length Context. In *ACL*.

[8] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[9] Leilani H Gilpin, David Bau, Ben Z Yuan, Ayesha Bajwa, Michael Specter, and Lalana Kagal. 2018. Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*.

[10] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*.

[11] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He. 2017. DeepFM: a factorization-machine based neural network for CTR prediction. *arXiv preprint arXiv:1703.04247* (2017).

[12] Mengyue Hang, Ian Pytlarz, and Jennifer Neville. 2018. Exploring student check-in behavior for improved point-of-interest prediction. In *KDD*.

[13] Xiangnan He and Tat-Seng Chua. 2017. Neural factorization machines for sparse predictive analytics. In *SIGIR*. 355–364.

[14] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *ICDM*.

[15] Sarthak Jain and Byron C. Wallace. 2019. Attention is not Explanation. In *NAACL*.

[16] Xu Jiao, Yingyuan Xiao, Wenguang Zheng, Hongya Wang, and Youzhi Jin. 2019. R2SIGTP: A novel real-time recommendation system with integration of geography and temporal preference for next point-of-interest. In *The World Wide Web Conference*.

[17] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[18] Guohui Li, Qi Chen, Bolong Zheng, Hongzhi Yin, Quoc Viet Hung Nguyen, and Xiaofang Zhou. 2020. Group-Based Recurrent Neural Networks for POI Recommendation. *ACM Transactions on Data Science* (2020).

[19] Huayu Li, Yong Ge, Richang Hong, and Hengshu Zhu. 2016. Point-of-interest recommendations: Learning potential check-ins from friends. In *KDD*.

[20] Ruirui Li, Jyun-Yu Jiang, Chelsea J-T Ju, and Wei Wang. 2019. CORALS: Who Are My Potential New Customers? Tapping into the Wisdom of Customers' Decisions. In *WSDM*.

[21] Shuangli Li, Jingbo Zhou, Tong Xu, Hao Liu, Xinjiang Lu, and Hui Xiong. 2020. Competitive Analysis for Points of Interest. In *KDD*.

[22] Xutao Li, Gao Cong, Xiao-Li Li, Tuan-Anh Nguyen Pham, and Shonali Krishnaswamy. 2015. Rank-geofm: A ranking based geographical factorization method for point of interest recommendation. In *SIGIR*.

[23] Zeyu Li, Wei Cheng, Yang Chen, Haifeng Chen, and Wei Wang. 2020. Interpretable Click-Through Rate Prediction through Hierarchical Attention. In *WSDM*.

[24] Zeyu Li, Jyun-Yu Jiang, Yizhou Sun, and Wei Wang. 2019. Personalized Question Routing via Heterogeneous Network Embedding. In *AAAI*.

[25] Defu Lian, Cong Zhao, Xing Xie, Guangzhong Sun, Enhong Chen, and Yong Rui. 2014. GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation. In *KDD*.

[26] Jianxun Lian, Xiaohuan Zhou, Fuzheng Zhang, Zhongxia Chen, Xing Xie, and Guangzhong Sun. 2018. xdeepfm: Combining explicit and implicit feature interactions for recommender systems. In *KDD*.

[27] Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. 2016. Predicting the next location: A recurrent model with spatial and temporal contexts. In *AAAI*.

[28] Wei Liu, Hanjiang Lai, Jing Wang, Geyang Ke, Weiwei Yang, and Jian Yin. 2020. Mix geographical information into local collaborative ranking for POI recommendation. *World Wide Web* (2020).

[29] Yiding Liu, Tuan-Anh Pham, Gao Cong, and Quan Yuan. 2017. An Experimental Evaluation of Point-of-interest Recommendation in Location-based Social Networks. *PVLDB* (2017), 1010–1021.

[30] Yong Liu, Wei Wei, Aixin Sun, and Chunyan Miao. 2014. Exploiting geographical neighborhood characteristics for location recommendation. In *CIKM*.

[31] Chen Ma, Yingxue Zhang, Qinglong Wang, and Xue Liu. 2018. Point-of-interest recommendation: Exploiting self-attentive autoencoders with neighbor-aware influence. In *CIKM*.

[32] Jingchao Ni, Shiyu Chang, Xiao Liu, Wei Cheng, Haifeng Chen, Dongkuan Xu, and Xiang Zhang. 2018. Co-regularized deep multi-network embedding. In *WWW*. 469–478.

[33] Hossein A Rahmani, Mohammad Aliannejadi, Mitra Baratchi, and Fabio Crestani. 2020. Joint Geographical and Temporal Modeling based on Matrix Factorization for Point-of-Interest Recommendation. In *European Conference on Information Retrieval*. Springer.

[34] Steffen Rendle. 2010. Factorization machines. In *ICDM*.

[35] Ke Sun, Tieyun Qian, Tong Chen, Yile Liang, Quoc Viet Hung Nguyen, and Hongzhi Yin. 2020. Where to Go Next: Modeling Long-and Short-Term User Preferences for Point-of-Interest Recommendation. In *AAAI*.

[36] Waldo R Tobler. 1970. A computer movie simulating urban growth in the Detroit region. *Economic geography* 46, sup1 (1970), 234–240.

[37] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *NIPS*.

[38] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).

[39] Qinyong Wang, Hongzhi Yin, Tong Chen, Zi Huang, Hao Wang, Yanchang Zhao, and Nguyen Quoc Viet Hung. 2020. Next Point-of-Interest Recommendation on Resource-Constrained Mobile Devices. In *Proceedings of The Web Conference*.

[40] Suhang Wang, Yilin Wang, Jiliang Tang, Kai Shu, Suhas Ranganath, and Huan Liu. 2017. What your images reveal: Exploiting visual contents for point-of-interest recommendation. In *Proceedings of The Web Conference*.

[41] Jun Xiao, Hao Ye, Xiangnan He, Hanwang Zhang, Fei Wu, and Tat-Seng Chua. 2017. Attentional factorization machines: Learning the weight of feature interactions via attention networks. *arXiv preprint arXiv:1708.04617* (2017).

[42] Tianjun Xiao, Yichong Xu, Kuiyuan Yang, Jiaxing Zhang, Yuxin Peng, and Zheng Zhang. 2015. The application of two-level attention models in deep convolutional neural network for fine-grained image classification. In *CVPR*.

[43] Carl Yang, Lanxiao Bai, Chao Zhang, Quan Yuan, and Jiawei Han. 2017. Bridging collaborative filtering and semi-supervised learning: a neural approach for poi recommendation. In *KDD*.

[44] Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *SIGIR*.

[45] Hongzhi Yin, Weiqing Wang, Hao Wang, Ling Chen, and Xiaofang Zhou. 2017. Spatial-aware hierarchical collaborative deep learning for POI recommendation. *IEEE Transactions on Knowledge and Data Engineering* (2017).

[46] Hongzhi Yin, Xiaofang Zhou, Bin Cui, Hao Wang, Kai Zheng, and Quoc Viet Hung Nguyen. 2016. Adapting to user interest drift for poi recommendation. *IEEE TKDE* (2016).

[47] Haochao Ying, Fuzhen Zhuang, Fuzheng Zhang, Yanchi Liu, Guandong Xu, Xing Xie, Hui Xiong, and Jian Wu. 2018. Sequential recommender system based on hierarchical attention networks. In *IJCAI*.

[48] Fuqiang Yu, Lizhen Cui, Wei Guo, Xudong Lu, Qingzhong Li, and Hua Lu. 2020. A Category-Aware Deep Model for Successive POI Recommendation on Sparse Check-in Data. In *Proceedings of The Web Conference*.

[49] Jia-Dong Zhang and Chi-Yin Chow. 2013. iGSLR: personalized geo-social location recommendation: a kernel density estimation approach. In *SIGSPATIAL*.

[50] Jia-Dong Zhang and Chi-Yin Chow. 2015. GeoSoCa: Exploiting geographical, social and categorical correlations for point-of-interest recommendations. In *SIGIR*. 443–452.

[51] Jia-Dong Zhang, Chi-Yin Chow, and Yanhua Li. 2014. Lore: Exploiting sequential influence for location recommendations. In *SIGSPATIAL*.

[52] Shenglin Zhao, Tong Zhao, Irwin King, and Michael R Lyu. 2017. Geo-teaser: Geo-temporal sequential embedding rank for point-of-interest recommendation. In *Proceedings of The Web Conference*.

[53] Zhou Zhao, Qifan Yang, Hanqing Lu, Tim Weninger, Deng Cai, Xiaofei He, and Yueting Zhuang. 2018. Social-Aware Movie Recommendation via Multimodal Network Learning. *IEEE Trans. Multimedia* (2018).

[54] Zhou Zhao, Qifan Yang, Hanqing Lu, Min Yang, Jun Xiao, Fei Wu, and Yueting Zhuang. 2017. Learning max-margin geoSocial multimedia network representations for point-of-interest suggestion. In *SIGIR*.

[55] Fan Zhou, Ruiyang Yin, Kunpeng Zhang, Goce Trajcevski, Ting Zhong, and Jin Wu. 2019. Adversarial point-of-interest recommendation. In *The World Wide Web Conference*.

[56] Jingbo Zhou, Shan Gou, Renjun Hu, Dongxiang Zhang, Jin Xu, Airong Jiang, Ying Li, and Hui Xiong. 2019. A collaborative learning framework to tag refinement for points of interest. In *KDD*.

[57] Xiao Zhou, Cecilia Mascolo, and Zhongxiang Zhao. 2019. Topic-enhanced memory networks for personalised point-of-interest recommendation. In *KDD*.