Matching Point of Interests and Travel Blog with Multi-view Information Fusion

Shuokai Li[†]
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences
lishuokai18z@ict.ac.cn

Jingbo Zhou[*]
Business Intelligence Lab
Baidu Research, China
zhoujingbo@baidu.com

Jizhou Huang
Hao chen
Baidu Inc., China
huangjizhou01@baidu.com
chenhao13@baidu.com

Fuzhen Zhuang
Institute of Artificial Intelligence, Beihang University
SKLSDE, School of Computer Science, Beihang University
zhuangfuzhen@buaa.edu.cn

Qing He[†,*]
Institute of Computing Technology, Chinese Academy of Sciences
University of Chinese Academy of Sciences
heqing@ict.ac.cn

Dejing Dou
BCG X, China
dejingdou@gmail.com

ABSTRACT

The past few years have witnessed an explosive growth of user-generated POI-centric travel blogs, which can provide a comprehensive understanding of a POI for people. However, evaluating the quality of the POI-centric travel blogs and ranking the blogs is not a simple task without domain knowledge or actual travel experience on the target POI. Nevertheless, our insight is that the user search behavior related to the target POI on the online map service can partly valid the rationality of the POIs appearing in the travel blogs, which helps for travel blogs ranking. To this end, in this paper, we propose a novel end-to-end framework for travel blogs ranking, coined **M**atching P**OI** and **T**ravel Blogs with Multi-view **InF**ormation (MOTIF). Concretely, we first construct two POI graphs as multi-view information: (1) the search-level POI graph which reflects the user behaviors on the online map service; and (2) the document-level POI graph which shows the POI co-occurrence frequency in travel blogs. Then, to better model the intrinsic correlation of the two graphs, we adopt Mutual Information Maximization to align the search-level and document-level semantic spaces. Moreover, we leverage a pair-wise ranking loss for POI-document relevance scoring. Extensive experiments on two real-world datasets demonstrate the superiority of our method.

CCS CONCEPTS

• **Information systems → Mobile information processing systems**; **Spatial-temporal systems**; **Information retrieval query processing**.

† The authors are at the Key Lab of Intelligent Information Processing of Chinese Academy of Sciences. This work was done when the first author was an intern in Baidu Research under the supervision of Jingbo Zhou.
∗ Corresponding authors.

SIGIR '23, July 23–27, 2023, Taipei, Taiwan, China

ACM ISBN 978-1-4503-9408-6/23/07.
https://doi.org/10.1145/3539618.3592016

KEYWORDS

Point of Interest; Document Ranking; Graph Neural Network; Mutual Information Maximization

ACM Reference Format:
Shuokai Li, Jingbo Zhou, Jizhou Huang, Hao Chen, Fuzhen Zhuang, Qing He, and Dejing Dou. 2023. Matching Point of Interests and Travel Blog with Multi-view Information Fusion. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23), July 23–27, 2023, Taipei, Taiwan, China.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3539618.3592016

1 INTRODUCTION

Nowadays, more and more people share travel literature on online websites, e.g., Tripadvisor, Booking.com, and Ctrip. Among various travel literature, POI-centric travel blogs, which focus on the travel experience of a target POI (e.g. presenting the POIs related to the Palace Museum), are especially useful for place assessment, to help people get the gist of the target POI quickly and offer convenience for other travelers. (Hereafter, we denote the POI-centric travel blogs as travel blogs for simplicity.)

For such travel blogs including the travel experience of a target POI, a high-quality blog should have: (1) the proper textual descriptions of the target POI; (2) the POIs related to the target POI with the corresponding introductions, which provide guidance for users' travel route planning. To be noticed, the related POIs do not mean the nearby POIs (see Section 5.2 in detail), and without domain knowledge and actual travel experience of the target POI, judging the related POIs is not a simple task. Though existing query-document ranking methods [2, 11] effectively model the relevance between the target POI and textual description of the target POI, they can not be directly applied to such POI-document ranking task. They fail to evaluate the quality of POIs included in travel blogs, which leads to poor blog ranking results.

Nevertheless, the user search behaviors on the online map service provide a practical source for validating the quality of travel blogs to some extent. For example, considering a travel blog containing

"Jingshan Park", and its target POI is "Beihai Park". From the search log, we find that "Beihai Park" and "Jingshan Park" are often co-searched in a search session, thus the two POIs are relevant and the quality of this travel blog is prone to be high. In light of this observation, the multi-view POI information (POI in documents and in search logs) jointly provides a new opportunity for travel blogs ranking.

To rank the high-quality travel blogs, in this paper, we propose a novel end-to-end framework, named **M**atching P**OI** and **T**ravel Blogs with Multi-view **In**Formation (MOTIF). We first construct two POI graphs as multi-view POI information for modeling the relevance between POIs: the search-level POI (co-occurrence) graph and the document-level POI (co-occurrence) graph. However, there is a natural semantic gap between the two kinds of graph signals. For the document-level graph, only the travelers who have visited the related POIs can write travel blogs, and their audiences are those who plan to have a trip to similar POIs. Therefore, the document-level semantic space includes domain-specific knowledge of POIs. In contrast, everyone can have a search record on an online map service, thus the search-level semantic space includes people's common sense. In light of this observation, we apply the Mutual Information Maximization [21, 26] method to bridge the semantic gap between these two graphs. The core idea is to force the representations of POIs from the two graphs to be close given high-quality documents. Based on the aligned semantic representations, the intrinsic correlations of the multi-view POI information are well explored, and we learn better POI representations. Then we leverage a pair-wise ranking loss for calculating the relevance between POI and document. The reason to choose the pair-wise loss instead of the absolute value of the relevance score is that we care about the ranking position of the documents. The contributions are as follows:

- To the best of our knowledge, we are the first to investigate the novel task, POI-centric travel blogs ranking.
- We incorporate multi-view POI information to the documents ranking task and adopt Mutual Information Maximization to align the multi-view POI representations.
- Our MOTIF achieves the best performances against the baselines on all metrics, shedding light on real-world applications.

2 RELATED WORK

POI Retrieval. It can be generally grouped into two categories: query POI matching (generating a relevance score for query and POI) [27–29] and POI auto-completion (suggesting a dynamic list of POIs as a user types each token) [4, 7, 17]. For query POI matching, Yuan et al. [28] integrate rich spatial-temporal factors of POIs and dynamic user preferences, and Yuan et al. [27] propose an incremental framework for online query-POI matching. For POI auto-completion, Huang et al. [7] model user's personal input habit and Fan et al. [4] adopt the meta-learning framework. Moreover, Huang et al. [8] utilize heterogeneous information for multilingual POI retrieval. Several studies have been conducted to explore the correlation between various modalities and Points of Interest (POIs), including but not limited to location-POI matching [9], voice-POI matching [6] and tag-POI matching [30]. However, all these works are different from our task, as we consider computing the relevance score between POIs and travel blogs, which considers both text data and multi-view user behaviors on POIs.

Text Retrieval. There are mainly three conventional categories for text retrieval: point-wise (logistic regression [12]), pair-wise (RankSvm [12] and BPR [19]), and list-wise (ListNet [2] and AdaRank [25]). With the development of deep learning, Huang et al. [11] propose the Deep Structured Semantic Models (DSSM), and Shen et al. [20] and Palangi et al. [18] improve the ability of the semantic extractor of DSSM. However, all these works do not consider the multi-view POI information when learning relevance scores.

Another related work is [3], which answers POI-based questions by tourism reviews. They recommend POIs to answer questions, rather than compute the relevance score between POIs and travel blogs.

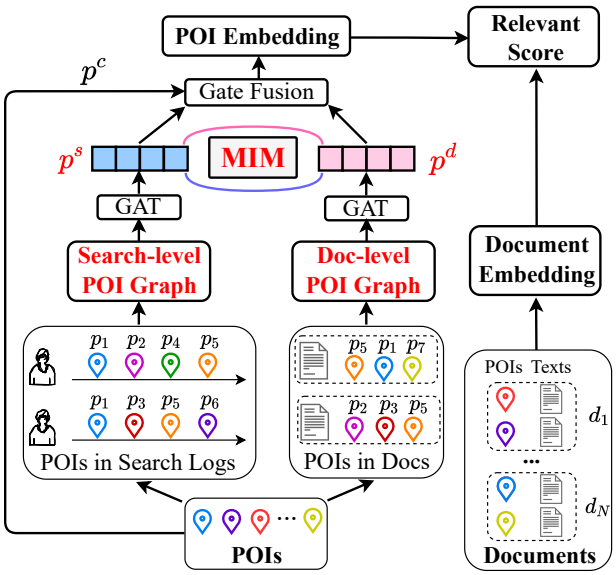

Figure 1: The overview of our model.

3 PROBLEM FORMULATION

A POI-centric travel blog is a document that consists of several POIs together with their text descriptions: $d_i = ((p_1, t_1), ..., (p_K, t_K))$. To be noticed, the text descriptions could be anything about the POIs (e.g., the history of the POI, the route planning to the POI, or personal visiting experience of the POI). Given a POI p_i with the corresponding documents $(d_1, d_2, ..., d_N)$, our goal is to rank the N documents by a relevance score for each POI-document pair (p_i, d_i).

We also consider two graphs as multi-view POI information. The node of them are POIs, and the weights are from multiple sources. For the search-level graph $\mathcal{G}_s$, the weights are the co-occurrence times of two POIs in search logs within a time session. For the document-level graph $\mathcal{G}_d$, the weights are the times that two POIs appear in the same document.

4 METHOD

The outline of our model is shown in Fig. 1. In the following, we will illustrate each component in our model: learning multi-view POI information with mutual information maximization, learning document representation, and assigning relevant scores for POI-document pairs.

4.1 POI Learning

Like previous methods [7, 28], we first extract some characteristic features for each POI. Then we incorporate two POI graph data and

fuse the semantic spaces of the two graphs. Finally, we combine the three parts to derive the POI representation $\boldsymbol{p}_i$.

4.1.1 POI Characteristic Features Learning. For each POI, the characteristic features representation $\boldsymbol{p}_i^c$ consists of four types of information which are POI address, primary category (e.g. historic sites, museums), second category (e.g. suitable for picnic, summer resort), and professional tag (i.e. a short sentence that describes the special features of the POIs, which is written by professional experts).

4.1.2 Search-level POI Graph Learning. From the search logs, users' search behaviors in a period of time reveal similar information need [10, 16, 31]. To model the POIs' relation in search-level semantic space, we use a 2-gram sliding window like [8] and apply the cosine similarity [24] to calculate the weight of two POIs by co-occurrence frequency:

$$\mathrm{S}_{i,j} = \frac{|W(p_i, p_j)|}{\sqrt{|W(p_i)| \cdot |W(p_j)|}}, \tag{1}$$

where $|W(p_i, p_j)|$ is the number of sliding windows that contain both p_i and p_j, and $|W(p_i)|$ is the number of windows that contains p_i.

Then we learn the representations of POIs in the search-level semantic space with the advantage of Graph Attention Networks (GAT [22]), which could aggregate information from the neighbors to the target node with a scalar attention coefficient α_{ij}:

$$\alpha_{ij} = \frac{\mathrm{S}_{i,j} \cdot \exp(\sigma_g(\boldsymbol{w}_a[\boldsymbol{W}_s\boldsymbol{p}_i^s || \boldsymbol{W}_s\boldsymbol{p}_j^s]))}{\sum_{k \in \mathcal{N}_{p_i}} \mathrm{S}_{i,j} \cdot \exp(\sigma_g(\boldsymbol{w}_a[\boldsymbol{W}_s\boldsymbol{p}_i^s || \boldsymbol{W}_s\boldsymbol{p}_k^s]))}, \tag{2}$$

where $\boldsymbol{p}_i^s$ is the representation of POI p_i in the serch-level POI graph, and is randomly initialized.

Then with the learned coefficients, a weighted sum between POI p_i and its neighbors $\mathcal{N}_{p_i}$ is the updated representation of p_i:

$$\boldsymbol{p}_i^s = \sum_{p_j \in \mathcal{N}_{p_i}} \alpha_{ij} \boldsymbol{W}_s \boldsymbol{p}_j^s. \tag{3}$$

4.1.3 Document-level POI Graph Learning. From the travel blogs (documents) which consist of several POIs, we could also model POIs' relation in document-level semantic space. Like Eq. 3, we could also learn the document-level POI representation $\boldsymbol{p}_i^d$.

4.1.4 Mutual Information Maximization. In order to bridge the semantic gap between search-level and document-level semantic spaces, we use Mutual Information Maximization (MIM) technique [21, 26]. Its core idea is to maximize the Mutual Information of two variables X and Y (i.e. $MI(X, Y)$). However, $MI(X, Y)$ is hard to compute, thus MIM tries to maximize the lower bound:

$$MI(X, Y) \geq \mathbb{E}_P[\log g(x, y)] + \mathbb{E}_N[\log(1 - g(x, y'))], \tag{4}$$

where $\mathbb{E}_P$ and $\mathbb{E}_N$ denote the expectation over positive and negative samples, respectively, and $g(\cdot)$ is the discriminator function that outputs the probability score modeled by a neural network.

In our setting, we view the POIs that co-appear in highly liked documents (i.e. the number of likes is large than 100) as the positive samples, and randomly choose some negative samples. For the POIs that co-occur in highly like documents, we pull their two graph representations close through a transformation matrix:

$$g(\boldsymbol{p}_i^s, \boldsymbol{p}_i^d) = \sigma(\boldsymbol{p}_i^s \cdot \mathbf{T} \cdot \boldsymbol{p}_i^d), \tag{5}$$

where $\sigma(\cdot)$ is the sigmoid function. By integrating Eq. 5 into Eq. 4, we can derive the MIM loss over all highly liked documents, and minimize the loss as a pre-training task.

Finally, the overall representation of POI p_i is the weighted sum of the two graph embeddings:

$$\begin{aligned} \boldsymbol{p}_i &= \lambda \cdot \boldsymbol{p}_i^s + (1 - \lambda) \cdot \boldsymbol{p}_i^d, \\ \lambda &= \sigma(\boldsymbol{W}_g[\boldsymbol{p}_i^s || \boldsymbol{p}_i^d]), \end{aligned} \tag{6}$$

where $\boldsymbol{W}_g$ is the gate transformation. Moreover, we also add the feature embedding $\boldsymbol{p}_i^c$ into $\boldsymbol{p}_i$.

4.2 Document Learning

Given the document which consists of several POIs together with their text descriptions, i.e. $d_i = ((p_1, t_1), ..., (p_K, t_K))$, we use the two parts (i.e. POI part and text part) to learn the document embedding $\boldsymbol{d}_i$.

For the first part, we average the K feature embeddings of POIs to learn the first part embedding: $\boldsymbol{d}_i^p = \mathrm{Mean}(\boldsymbol{p}_k^c)$.

For the k-th text description t_k, which includes several tokens, we first use BERT [13] to extract the token representations. Then these token representations are fed into an LSTM Network [5] followed by mean pooling. In this way, the text description t_k is transformed into a d-dimensional hidden representation $\boldsymbol{t}_k$. Then we aggregate K text descriptions by attention mechanism [1] to learn text embedding $\boldsymbol{d}_i^t$:

$$\begin{aligned} \boldsymbol{d}_i^t &= \sum \alpha_k \boldsymbol{t}_k, \\ \alpha_k &= \mathrm{Softmax}(\boldsymbol{b}^{\mathrm{T}} \tanh(\boldsymbol{W}_t \boldsymbol{t}_k)). \end{aligned} \tag{7}$$

Finally, we concatenate $\boldsymbol{d}_i^p$ and $\boldsymbol{d}_i^t$ and then followed by a fully connected layer to learn document representation $\boldsymbol{d}_i$.

4.3 Model Training

Given a POI p and its corresponding documents $(d_1, ..., d_N)$, we aim to rank the high-quality documents at the top position. Thus, we adopt the pair-wise loss Bayesian Personalized Ranking (BPR) [19] to train our model. First, we rank the N documents according to their like numbers. Specifically, d_i is more popular than d_j when $i < j$. Then the training loss is defined as:

$$\mathcal{L} = -\sum_{i,j} \log(\sigma(\boldsymbol{p}^{\mathrm{T}}\boldsymbol{d}_i - \boldsymbol{p}^{\mathrm{T}}\boldsymbol{d}_j)). \tag{8}$$

When testing, we directly set $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{d}_i$ as the relevance function between POI and document.

Our training procedure includes two stages. For pre-training stage, we first use MIM loss to learn better POI representations $\boldsymbol{p}^s$ and $\boldsymbol{p}^d$. Then during the next stage, we use BPR loss $\mathcal{L}$ to rank documents.

5 EXPERIMENT

5.1 Experimental Settings

Datasets. To extract the travel blogs, we crawl the documents on the traveling websites, which are Mafengwo[1] and Ctrip[2]. We only preserve the documents which include multiple POIs and take the two sentences near POIs as their text description. The search logs of POIs are obtained from Baidu map service in China. Here the ground-truth

[1] http://www.mafengwo.cn/
[2] https://www.ctrip.com/

Table 1: The overall performances. The marker * indicates that the improvement is statistically significant compared with the best baseline (pairwise t-test with p-value < 0.01).

Dataset	Beijing								Chengdu							
Method	SR@1	SR@3	SR@5	MRR@3	MRR@5	nDCG@1	nDCG@3	nDCG@5	SR@1	SR@3	SR@5	MRR@3	MRR@5	nDCG@1	nDCG@3	nDCG@5
Popularity	12.99	35.06	67.53	0.2294	0.3009	0.2576	0.3822	0.5302	12.39	41.16	77.78	0.2593	0.3404	0.2815	0.4059	0.5450
Distance	11.68	36.36	71.43	0.2229	0.3015	0.2762	0.4066	0.5473	13.33	42.22	80.03	0.2556	0.3389	0.3074	0.4355	0.5713
Random	15.58	42.85	72.72	0.2749	0.3379	0.2817	0.4185	0.5413	15.94	44.31	75.37	0.2876	0.3592	0.3187	0.4421	0.5873
BPR-CNN	22.07	54.54	75.32	0.3550	0.4017	0.3547	0.4776	0.5957	21.36	52.33	82.08	0.3566	0.4280	0.3586	0.5107	0.6298
ListRank	23.37	58.44	78.63	0.3679	0.4293	0.3722	0.5187	0.6320	22.82	53.79	82.36	0.3781	0.4476	0.3734	0.5243	0.6549
BPR-LSTM	25.37	58.74	81.82	0.3976	0.4485	0.4052	0.5381	0.6477	24.73	55.19	84.13	0.3963	0.4531	0.3963	0.5586	0.6592
BPR-BERT	27.60	62.33	83.40	0.4361	0.4811	0.4250	0.5604	0.6647	28.17	58.86	82.77	0.4155	0.4715	0.4172	0.5722	0.6785
HGAMN	28.23	61.03	83.68	0.4335	0.4919	0.4326	0.5669	0.6714	31.74	61.48	83.90	0.4763	0.5068	0.4583	0.6011	0.6932
MOTIF -Sea	29.35	62.44	83.77	0.4396	0.4883	0.4476	0.5893	0.6821	31.17	62.39	83.23	0.4573	0.5118	0.4633	0.5987	0.6961
MOTIF -MIM	31.64	64.70	85.12	0.4559	0.5022	0.4894	0.6137	0.7080	36.56	68.69	84.35	0.5107	0.5464	0.5203	0.6275	0.7142
MOTIF	**34.01***	**69.89***	**88.43***	**0.4953***	**0.5369***	**0.5097***	**0.6372***	**0.7273***	**43.64***	**71.30***	**87.88***	**0.5603***	**0.5971***	**0.5684***	**0.6634***	**0.7434***

value of the document quality is the user like numbers on websites. By doing this, we harvest two real-world datasets, namely **Beijing** and **Chengdu**. Both of them include 100 POIs. Beijing contains 5,001 documents and Chengdu contains 1,788 documents.

Baselines. We compare MOTIF with the following representative baselines: (1) ***Popularity*** ranks the documents according to the popularity of POIs. (2) ***Distance*** ranks the documents according to the average distance between the target POI and POIs in documents. (3) ***Random*** provides a random ranking list. (4) ***BPR-CNN*** [14] encodes the text by CNN and uses BPR loss [19] for ranking. Only POI characteristic feature $\boldsymbol{p}^c$ is used. (5) ***ListRank*** [2, 23] adopts list-wise loss to rank documents. Concretely, we use Cross Entropy to measure the distance between the label and the predicted score list. (6) ***BPR-LSTM*** [5] incorporates LSTM with BPR loss to capture the text representations for documents. (7) ***BPR-BERT*** [13] uses the pre-training model BERT as the token representation extractor. (8) ***HGAMN*** [8] adopts user search behaviors to retrieve relevant POIs. We reserve the heterogeneous graph attention module for POI search graph.

Implementation Details. To validate our model, each test sample includes one POI and 7 corresponding documents. In addition, we use the number of likes as the measurement of document quality. Then we dset the dimension of embedding as 128 and the batch size of 32, and employ the Adam [15] optimizer.

Evaluation Protocol. We adopt Success Rate (SR), MRR, and nDCG at Top-K ($K = 1, 3, 5$) for evaluation. SR denotes the average percentage of the highest-likes document (in 7 documents) ranked at or above the position K in the rank list. MRR and nDCG evaluates the ranking quality.

5.2 Results

Overall Performance. The ranking results on two datasets are shown in Table 1. The experimental results reveal several insightful observations. (1) Our MOTIF significantly outperforms all the baselines by a large margin on both two datasets, which verifies that the effectiveness of incorporating the multi-view graph data of POIs and MIM learns better POI representations by bridging the semantic gap between two semantic spaces. (2) Among all the baselines, HGAMN performs best. The reason is that HGANM incorporates the user search behaviors into POI representations. (3) Given the three text encoders CNN, LSTM, and BERT, we can find BPR-BERT achieves the best performance. It demonstrates that the pre-training model BERT could extract text information better. (4) Comparing BPR-LSTM with ListRank, we can see that the pair-wise loss is more suitable for the document ranking task. (5) Popularity and distance perform even worse than a random ranking list. We think the reasons are two perspectives. First, people care about not only the POIs included in documents, but also text descriptions. However, the two baselines omit the text descriptions of POIs. Second, for the POIs included documents, the hot POIs and the POIs near the target POI may not receive more likes. Thus, we need a neural model to learn high-quality POI representations.

Ablation Study. In our MOTIF, we incorporate two graph data and use MIM to align the two graph representations. Here we would like to examine the effectiveness of each part and show the results in Table 1. To be noticed, MOTIF -Sea trains MOTIF without POI search graph embedding and MOTIF -MIM denotes training our model without MIM pre-training. For the two graphs, MOTIF -Sea beats BPR-BERT and MOTIF -MIM beats MOTIF -Sea, which shows the usefulness of incorporating user search behaviors and POI co-occurrence frequency in travel blogs. For the MIM, MOTIF bridges the gap between two semantic spaces and learns better POI representations, thus beats MOTIF -MIM.

6 CONCLUSION

In this paper, we first investigated how to rank POI-centric travel blogs. As the blogs include related POIs of the target POI, we leveraged the user search behavior related to the target POI for validating the quality of documents, and proposed an end-to-end framework named MOTIF. It incorporates the search-level and document-level POI graphs, and then adopts MIM to bridge the gap between two semantic spaces. Extensive experiments conducted on two real-world datasets verified the superiority of our MOTIF.

ACKNOWLEDGEMENT

The work is partially supported by the National Key Research and Development Program of China under Grant No. 2022YFC3303302. This work is also supported in part by the National Natural Science Foundation of China under Grant No.61976204 and 62176014.

REFERENCES

[1] Dzmitry Bahdanau, Kyung Hyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. In *ICLR*.

[2] Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. 2007. Learning to rank: from pairwise approach to listwise approach. In *ICML*. 129–136.

[3] Danish Contractor, Krunal Shah, Aditi Partap, Parag Singla, and Mausam Mausam. 2021. Answering POI-recommendation Questions using Tourism Reviews. In *CIKM*. 281–291.

[4] Miao Fan, Yibo Sun, Jizhou Huang, Haifeng Wang, and Ying Li. 2021. Meta-Learned Spatial-Temporal POI Auto-Completion for the Search Engine at Baidu Maps. In *SIGKDD*. 2822–2830.

[5] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.

[6] Jizhou Huang, Haifeng Wang, Shiqiang Ding, and Shaolei Wang. 2022. DuIVA: An Intelligent Voice Assistant for Hands-free and Eyes-free Voice Interaction with the Baidu Maps App. In *SIGKDD*. 3040–3050.

[7] Jizhou Huang, Haifeng Wang, Miao Fan, An Zhuo, and Ying Li. 2020. Personalized prefix embedding for poi auto-completion in the search engine of baidu maps. In *SIGKDD*. 2677–2685.

[8] Jizhou Huang, Haifeng Wang, Yibo Sun, Miao Fan, Zhengjie Huang, Chunyuan Yuan, and Yawen Li. 2021. HGAMN: Heterogeneous Graph Attention Matching Network for Multilingual POI Retrieval at Baidu Maps. In *SIGKDD*. 3032–3040.

[9] Jizhou Huang, Haifeng Wang, Yibo Sun, Yunsheng Shi, Zhengjie Huang, An Zhuo, and Shikun Feng. 2022. ERNIE-GeoL: A Geography-and-Language Pre-trained Model and its Applications in Baidu Maps. In *SIGKDD*. 3029–3039.

[10] Jizhou Huang, Haifeng Wang, Wei Zhang, and Ting Liu. 2020. Multi-task learning for entity recommendation and document ranking in web search. *TIST* 11, 5 (2020), 1–24.

[11] Po-Sen Huang, Xiaodong He, Jianfeng Gao, Li Deng, Alex Acero, and Larry Heck. 2013. Learning deep structured semantic models for web search using clickthrough data. In *CIKM*. 2333–2338.

[12] Thorsten Joachims. 2002. Optimizing search engines using clickthrough data. In *SIGKDD*. 133–142.

[13] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *NAACL*. 4171–4186.

[14] Yoon Kim. 2014. Convolutional neural networks for sentence classification. In *EMNLP*. 1746–1751.

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

[16] Shuangli Li, Jingbo Zhou, Tong Xu, Hao Liu, Xinjiang Lu, and Hui Xiong. 2020. Competitive analysis for points of interest. In *SIGKDD*. 1265–1274.

[17] Ying Li, Jizhou Huang, Miao Fan, Jinyi Lei, Haifeng Wang, and Enhong Chen. 2020. Personalized query auto-completion for large-scale POI search at Baidu Maps. *TALLIP* 19, 5 (2020), 1–16.

[18] Hamid Palangi, Li Deng, Yelong Shen, Jianfeng Gao, Xiaodong He, Jianshu Chen, Xinying Song, and R Ward. 2014. Semantic modelling with long-short-term memory for information retrieval. *arXiv preprint arXiv:1412.6629* (2014).

[19] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *UAI*.

[20] Yelong Shen, Xiaodong He, Jianfeng Gao, Li Deng, and Grégoire Mesnil. 2014. Learning semantic representations using convolutional neural networks for web search. In *Web Conference*. 373–374.

[21] Fan-Yun Sun, Jordan Hoffman, Vikas Verma, and Jian Tang. 2019. InfoGraph: Unsupervised and Semi-supervised Graph-Level Representation Learning via Mutual Information Maximization. In *ICLR*.

[22] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph Attention Networks. In *ICLR*.

[23] Shiyao Wang, Qi Liu, Tiezheng Ge, Defu Lian, and Zhiqiang Zhang. 2021. A Hybrid Bandit Model with Visual Priors for Creative Ranking in Display Advertising. In *Web Conference*. 2324–2334.

[24] Tianyi Wu, Yuguo Chen, and Jiawei Han. 2007. Association mining in large databases: A re-examination of its measures. In *ECML-PKDD*. 621–628.

[25] Jun Xu and Hang Li. 2007. Adarank: a boosting algorithm for information retrieval. In *SIGIR*. 391–398.

[26] Yi-Ting Yeh and Yun-Nung Chen. 2019. QAInfomax: Learning Robust Question Answering System by Mutual Information Maximization. In *EMNLP*. 3370–3375.

[27] Zixuan Yuan, Hao Liu, Junming Liu, Yanchi Liu, Yang Yang, Renjun Hu, and Hui Xiong. 2021. Incremental Spatio-Temporal Graph Learning for Online Query-POI Matching. In *Web Conference*. 1586–1597.

[28] Zixuan Yuan, Hao Liu, Yanchi Liu, Denghui Zhang, Fei Yi, Nengjun Zhu, and Hui Xiong. 2020. Spatio-temporal dual graph attention network for query-poi matching. In *SIGIR*. 629–638.

[29] Ji Zhao, Dan Peng, Chuhan Wu, Huan Chen, Meiyu Yu, Wanji Zheng, Li Ma, Hua Chai, Jieping Ye, and Xiaohu Qie. 2019. Incorporating semantic similarity with geographic correlation for query-POI relevance learning. In *AAAI*, Vol. 33. 1270–1277.

[30] 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 *SIGKDD*. 1752–1761.

[31] Jingbo Zhou, Tao Huang, Shuangli Li, Renjun Hu, Yanchi Liu, Yanjie Fu, and Hui Xiong. 2021. Competitive relationship prediction for points of interest: A neural graphlet based approach. *TKDE* 34, 12 (2021), 5681–5692.