McGill

Point-of-Interest Recommendation: Exploiting Self-Attentive Autoencoders with Neighbor-Aware Influence

Chen Ma, Yingxue Zhang, Qinglong Wang and Xue Liu

McGill University, Montreal, Canada

CIKM 2018, Turin, Italy

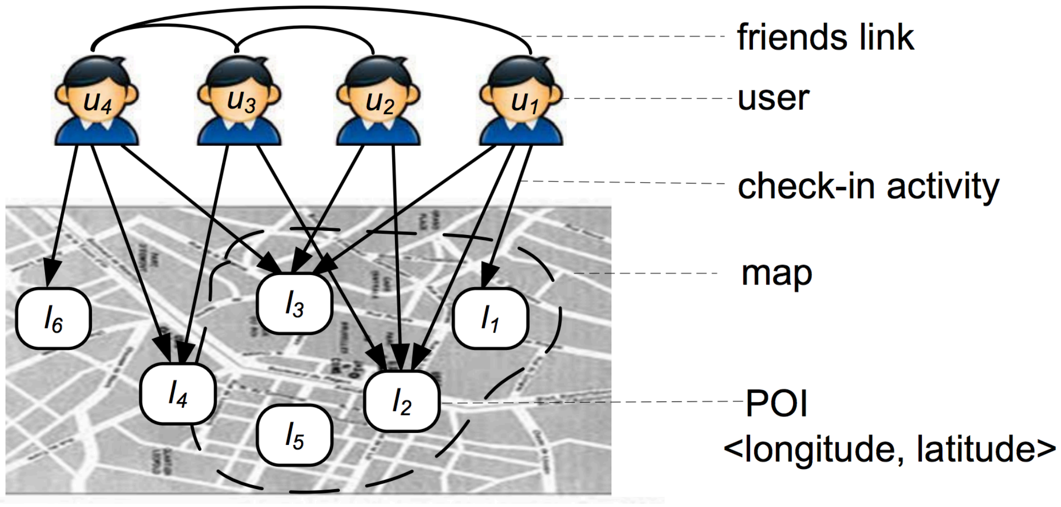
Background

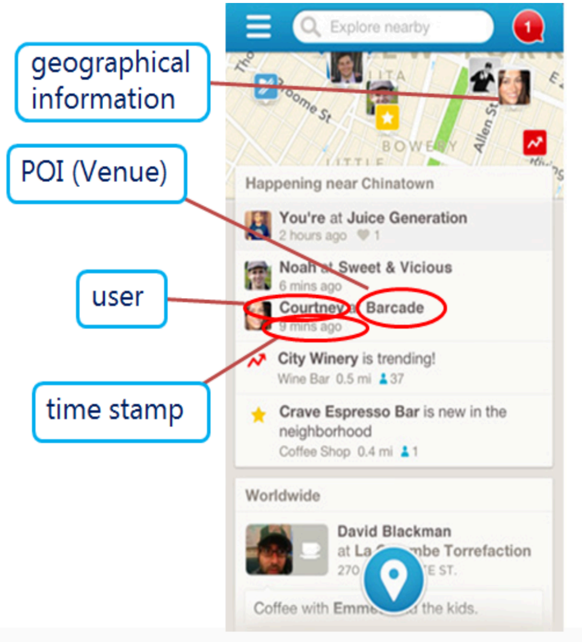
Many **location-based social networks** (LBSNs) have emerged in recent years, such as Yelp, Foursquare, Facebook Place.

- Yelp had a monthly average of **32 million** unique visitors Via the App
- More than **50 million** people use Foursquare every month

Background

In LBSNs, users can check-in and share their experience when they visit a location, namely, **Point-of-Interest** (POI)





Ye et al., Exploiting Geographical Influence for Collaborative Point-of-Interest Recommendation, SIGIR 2011
Bao et al., Recommendations in Location-based Social Networks: A Survey, Geoinformatica 2015

Background

The large amount of user-POI interactions facilitates a promising service – **personalized POI recommendation**

- Help users easily find the places they are **interested** in
- Improve the customer **satisfaction**
- Attract **potential visitors** for POI owners
- Increase **revenue** for POI owners and service providers
- ……

Challenges

Data Sparsity: the check-in data is extremely sparse

Dataset	Movielens10M	Netflix Prize	Check-in Data
Density	1.3%	1.2%	~0.1%

Implicit Feedback Property: check-ins are implicit feedback

Explicit Feedback: movie rating data	**Implicit Feedback:** check-in data
Users explicitly denote "like" or "dislike" with different scores	Only check-in frequency is available

Context Information: how to incorporate different context information?

- Geographical coordinates of POIs (key distinction: geographical influence)
- Timestamps of check-ins
- Text description of POIs

Related Work

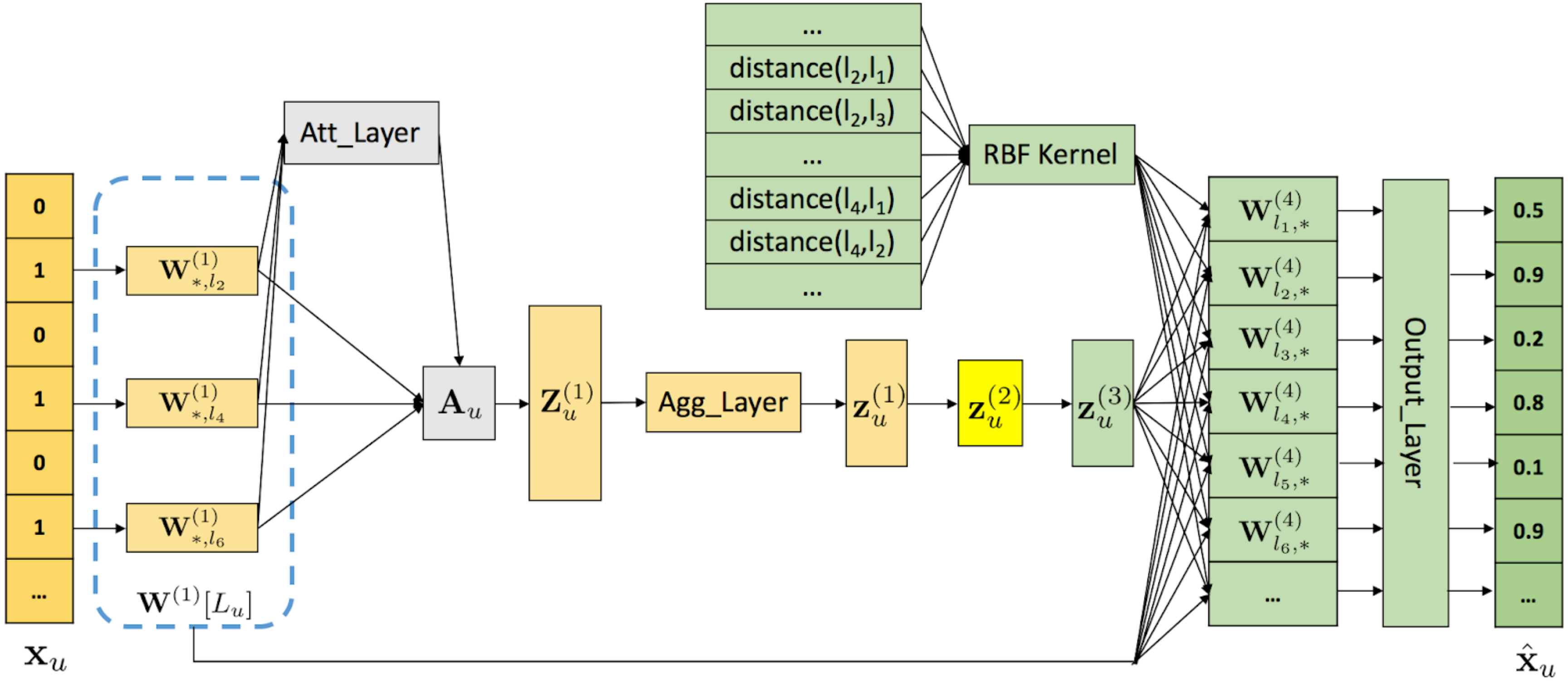
Methods	Major algorithm
USG (*Ye et al, SIGIR' 2011*)	Memory-based CF
MGMMF (*Cheng et al, AAAI' 2012*)	Poisson MF
GeoMF (*Lian et al, SIGKDD' 2014*)	Weighted MF
IRENMF (*Liu et al, CIKM' 2014*)	Weighted MF
RankGeoFM (*Li et al, SIGIR' 2015*)	BPR MF
ARMF (*Li et al, SIGKDD' 2016*)	Weighted MF

- Combine latent factors linearly
- Not distinguish user preference levels on visited POIs
- Not explicitly model the POI-POI relations

CF: Collaborative Filtering
MF: Matrix Factorization
BPR: Bayesian Personalized Ranking

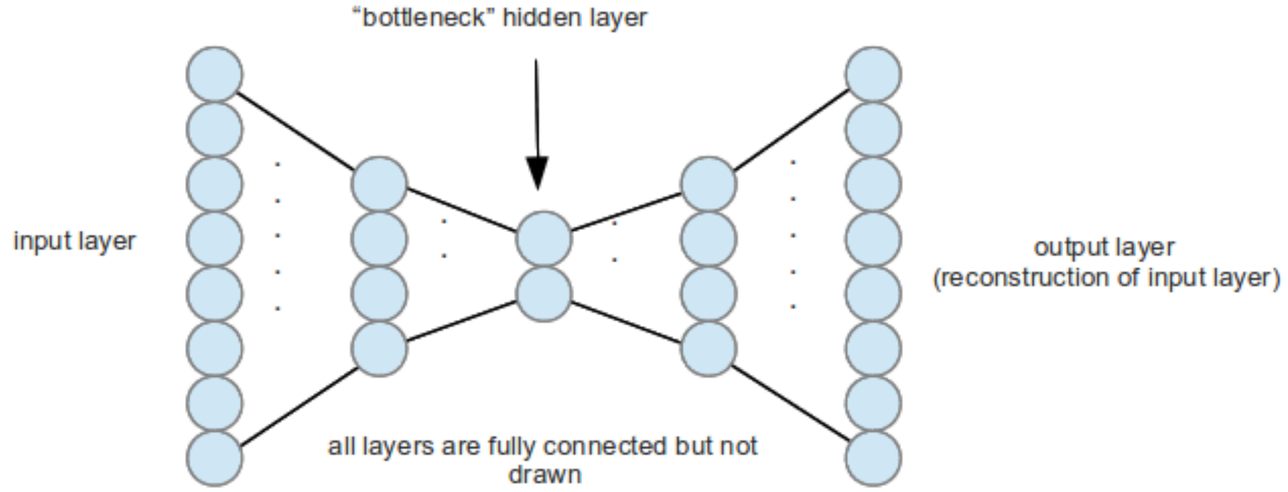
Model Overview

An **autoencoder**-based model, consisting of a **self-attentive encoder (SAE)** and a **neighbor-aware decoder (NAD)**



Preliminary

Autoencoder: an unsupervised neural network with an encoder and a decoder



$$enc: \begin{cases} \mathbf{z}_u^{(1)} = a_1(\mathbf{W}^{(1)}\mathbf{x}_u + \mathbf{b}^{(1)}) \\ \mathbf{z}_u^{(2)} = a_2(\mathbf{W}^{(2)}\mathbf{z}_u^{(1)} + \mathbf{b}^{(2)}) \end{cases} \quad dec: \begin{cases} \mathbf{z}_u^{(3)} = a_3(\mathbf{W}^{(3)}\mathbf{z}_u^{(2)} + \mathbf{b}^{(3)}) \\ \hat{\mathbf{x}}_u = a_4(\mathbf{W}^{(4)}\mathbf{z}_u^{(3)} + \mathbf{b}^{(4)}) \end{cases} \quad loss: \mathcal{L} = \sum_{i=1}^{M} ||\mathbf{x}_i - \hat{\mathbf{x}}_i||_2^2$$

http://nghiaho.com/?p=1765

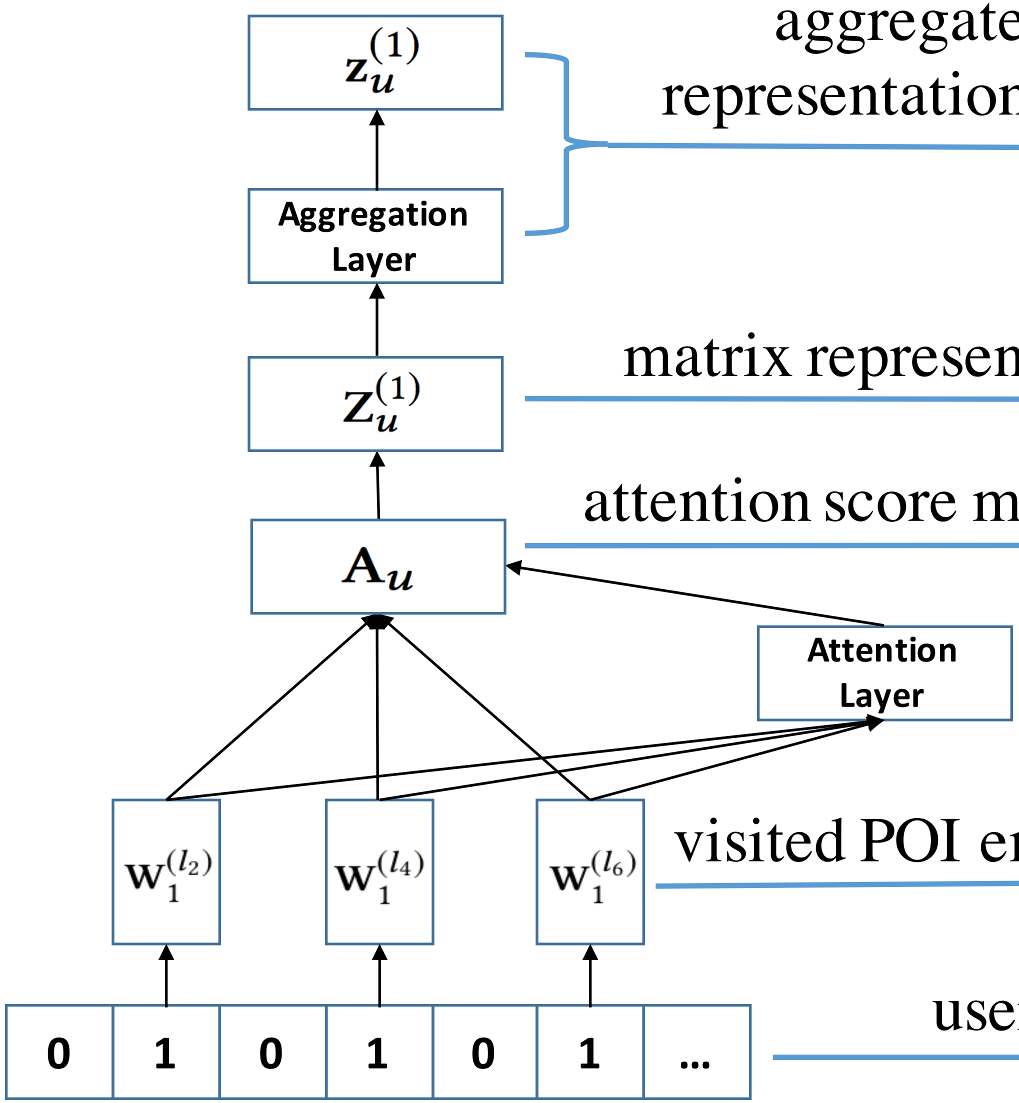
Self-attentive Encoder

- Previous works do not further discriminate user preference levels on visited POIs
- User preference is a complex sentiment

Flavor	Price	Environment
		

Some visited POIs are more representative than others and should contribute more to characterize users' preferences

Self-attentive Encoder



aggregate user hidden representations into one aspect

$$\mathbf{z}_u^{(1)} = a_t(\mathbf{Z}_u^{(1)\top}\mathbf{w}_t + \mathbf{b}_t)$$

matrix representation of users

$$\mathbf{Z}_u^{(1)} = \mathbf{A}_u\mathbf{W}_1^\top[L_u]$$

attention score matrix

$$\mathbf{A}_u = softmax(tanh(\mathbf{W}_a\mathbf{W}_1[L_u]))$$

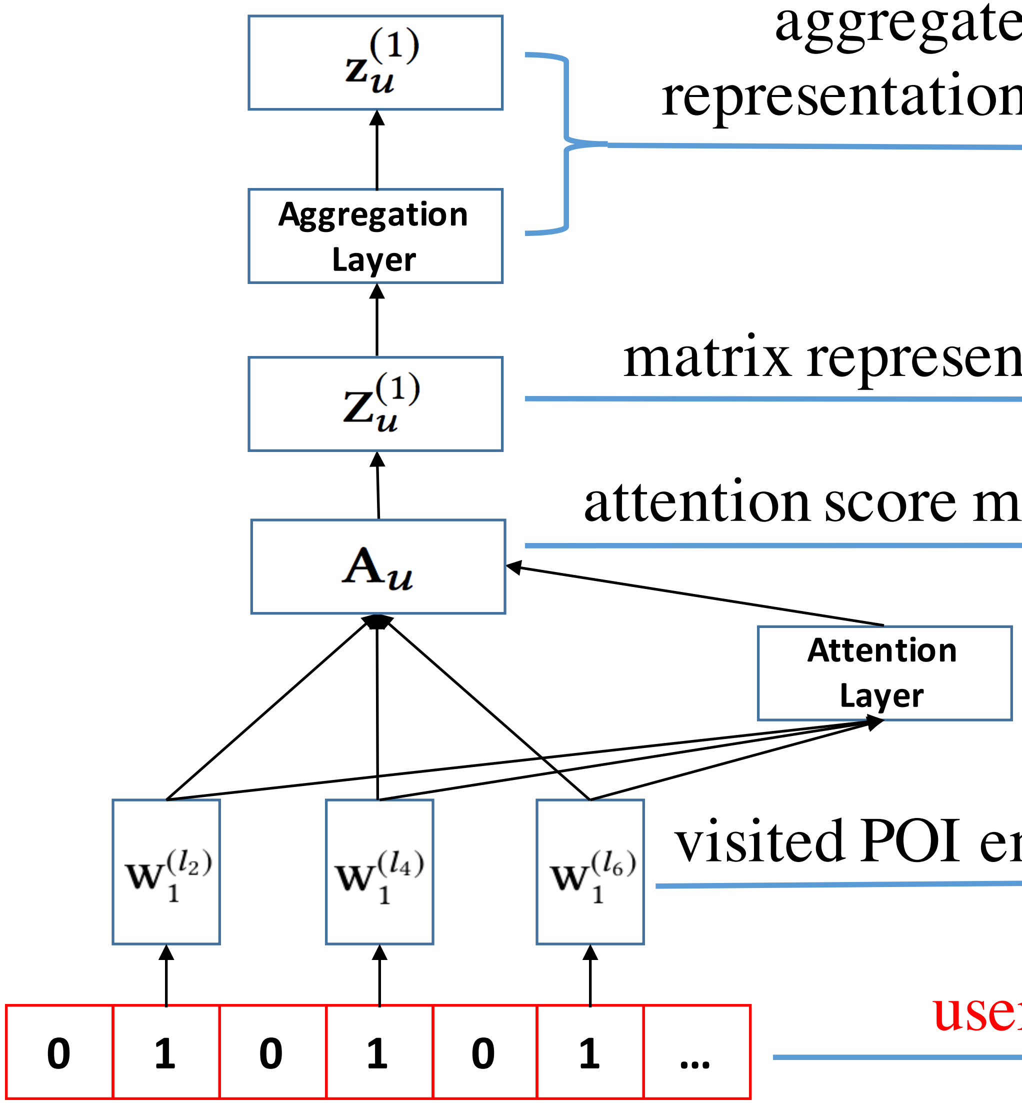
visited POI embeddings

$$\mathbf{W}_1[L_u] = (\mathbf{W}_1^{(l_1)}, \mathbf{W}_1^{(l_2)}, ..., \mathbf{W}_1^{(l_n)})$$

user visited POIs

$$L_u = \{l_1, ..., l_n\}$$

Self-attentive Encoder



aggregate user hidden representations into one aspect

$$\mathbf{z}_u^{(1)} = a_t(\mathbf{Z}_u^{(1)\top}\mathbf{w}_t + \mathbf{b}_t)$$

matrix representation of users

$$\mathbf{Z}_u^{(1)} = \mathbf{A}_u \mathbf{W}_1^\top [L_u]$$

attention score matrix

$$\mathbf{A}_u = softmax(tanh(\mathbf{W}_a \mathbf{W}_1[L_u]))$$

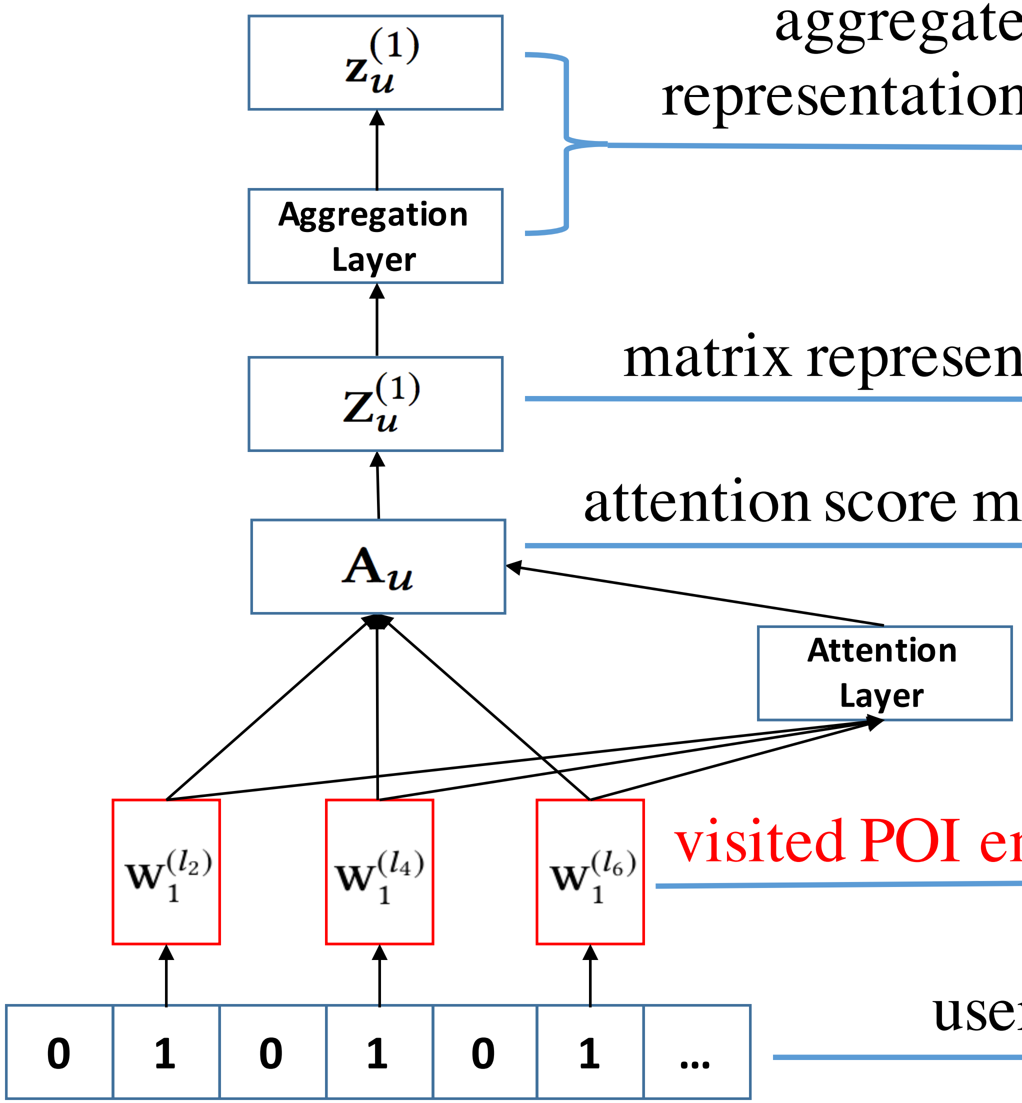
visited POI embeddings

$$\mathbf{W}_1[L_u] = (\mathbf{W}_1^{(l_1)}, \mathbf{W}_1^{(l_2)}, ..., \mathbf{W}_1^{(l_n)})$$

user visited POIs

$$L_u = \{l_1, ..., l_n\}$$

Self-attentive Encoder



aggregate user hidden representations into one aspect

$$\mathbf{z}_u^{(1)} = a_t(\mathbf{Z}_u^{(1)\top}\mathbf{w}_t + \mathbf{b}_t)$$

matrix representation of users

$$\mathbf{Z}_u^{(1)} = \mathbf{A}_u\mathbf{W}_1^{\top}[L_u]$$

attention score matrix

$$\mathbf{A}_u = softmax(tanh(\mathbf{W}_a\mathbf{W}_1[L_u]))$$

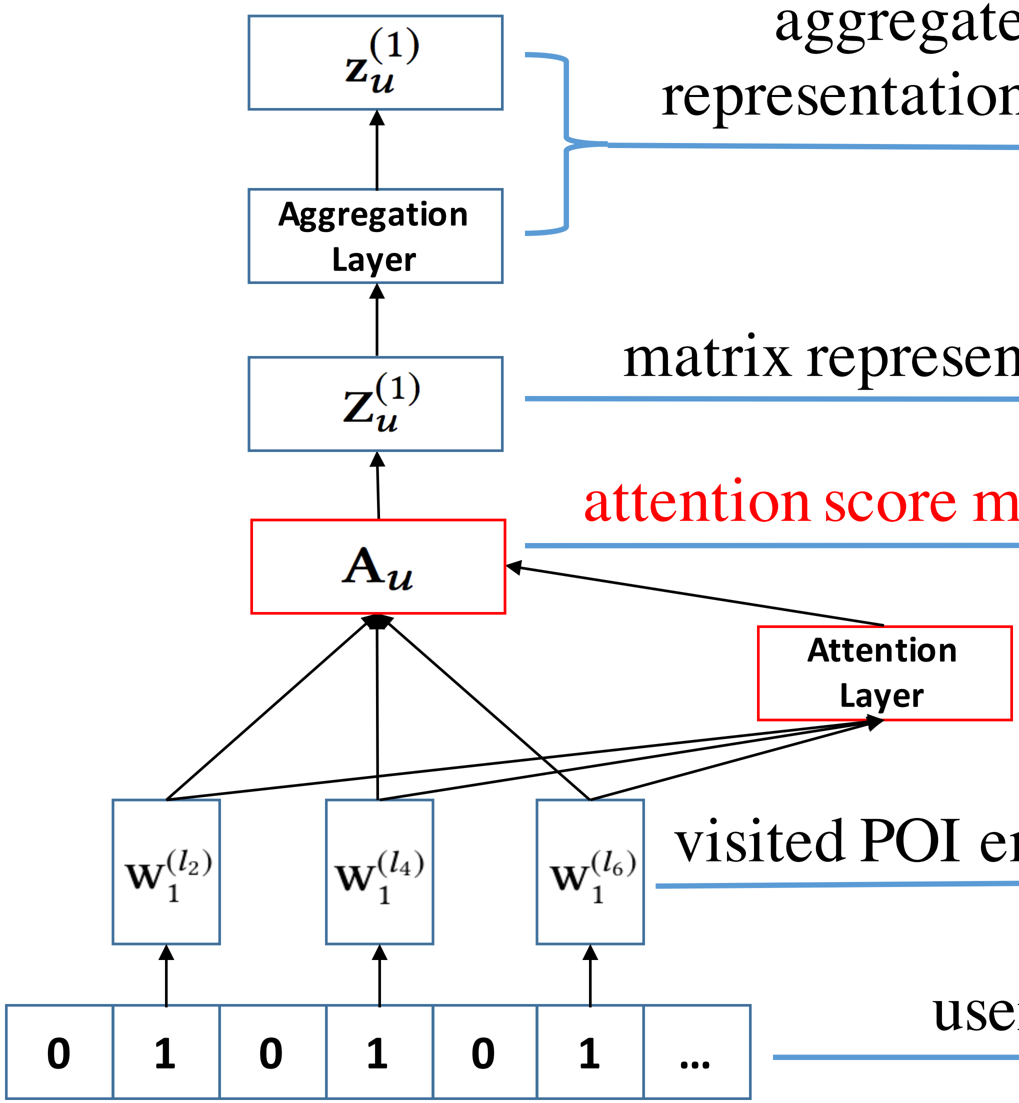
visited POI embeddings

$$\mathbf{W}_1[L_u] = (\mathbf{W}_1^{(l_1)}, \mathbf{W}_1^{(l_2)}, ..., \mathbf{W}_1^{(l_n)})$$

user visited POIs

$$L_u = \{l_1, ..., l_n\}$$

Self-attentive Encoder



aggregate user hidden representations into one aspect

$$\mathbf{z}_u^{(1)} = a_t(\mathbf{Z}_u^{(1)\top}\mathbf{w}_t + \mathbf{b}_t)$$

matrix representation of users

$$\mathbf{Z}_u^{(1)} = \mathbf{A}_u \mathbf{W}_1^\top [L_u]$$

attention score matrix

$$\mathbf{A}_u = softmax(tanh(\mathbf{W}_a \mathbf{W}_1[L_u]))$$

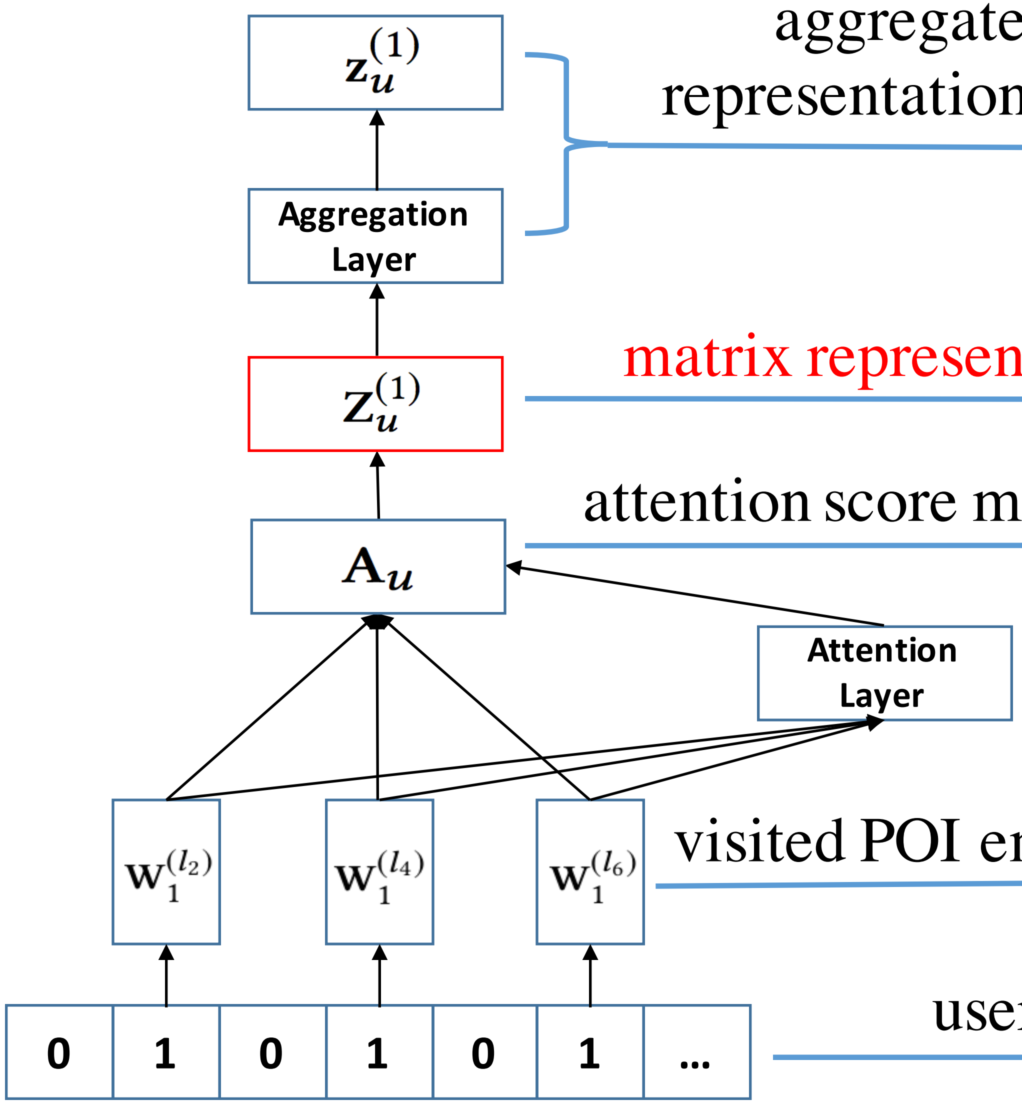
visited POI embeddings

$$\mathbf{W}_1[L_u] = (\mathbf{W}_1^{(l_1)}, \mathbf{W}_1^{(l_2)}, ..., \mathbf{W}_1^{(l_n)})$$

user visited POIs

$$L_u = \{l_1, ..., l_n\}$$

Self-attentive Encoder



aggregate user hidden representations into one aspect

$$\mathbf{z}_u^{(1)} = a_t(\mathbf{Z}_u^{(1)\top}\mathbf{w}_t + \mathbf{b}_t)$$

matrix representation of users

$$\mathbf{Z}_u^{(1)} = \mathbf{A}_u\mathbf{W}_1^{\top}[L_u]$$

attention score matrix

$$\mathbf{A}_u = softmax(tanh(\mathbf{W}_a\mathbf{W}_1[L_u]))$$

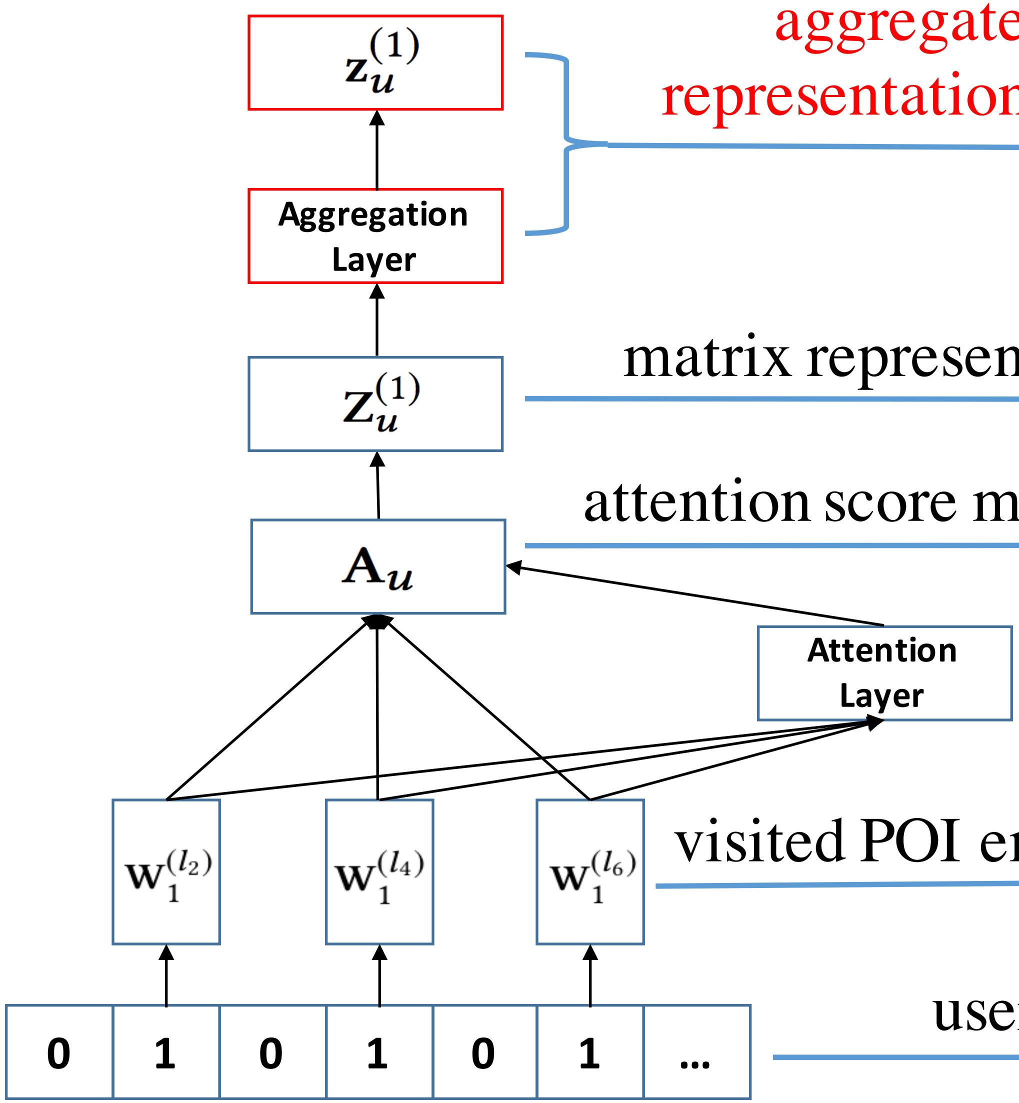
visited POI embeddings

$$\mathbf{W}_1[L_u] = (\mathbf{W}_1^{(l_1)}, \mathbf{W}_1^{(l_2)}, ..., \mathbf{W}_1^{(l_n)})$$

user visited POIs

$$L_u = \{l_1, ..., l_n\}$$

Self-attentive Encoder



aggregate user hidden representations into one aspect

$$\mathbf{z}_u^{(1)} = a_t(\mathbf{Z}_u^{(1)\top}\mathbf{w}_t + \mathbf{b}_t)$$

matrix representation of users

$$\mathbf{Z}_u^{(1)} = \mathbf{A}_u\mathbf{W}_1^\top[L_u]$$

attention score matrix

$$\mathbf{A}_u = softmax(tanh(\mathbf{W}_a\mathbf{W}_1[L_u]))$$

visited POI embeddings

$$\mathbf{W}_1[L_u] = (\mathbf{W}_1^{(l_1)}, \mathbf{W}_1^{(l_2)}, \ldots, \mathbf{W}_1^{(l_n)})$$

user visited POIs

$$L_u = \{l_1, \ldots, l_n\}$$

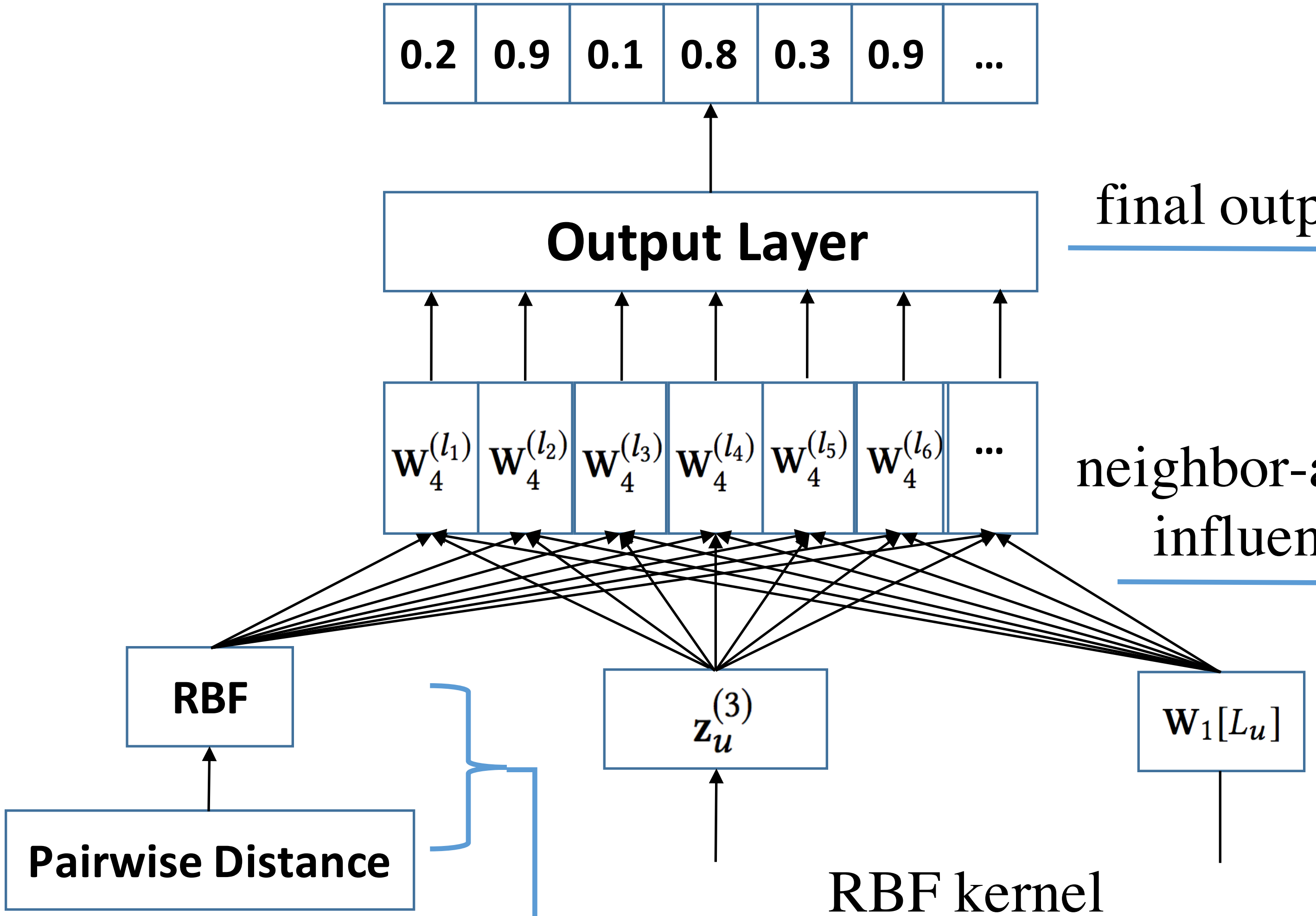
Neighbor-aware Decoder

- Explicitly capture the POI-POI relations, e.g., properties, similarity
- Incorporate the geographical influence by the RBF kernel
- Similar to *FISM* (*SIGKDD' 2013*) that applies the inner product to capture the similarity between POIs
- Similar to *word2vec*: given a set of POIs, how likely other POIs will be visited

Model the pairwise relations: the unvisited POIs that close to visited POIs are more likely to be checked-in

Kabbur et al., FISM: Factored Item Similarity Models for Top-N Recommender Systems, SIGKDD 2013

Neighbor-aware Decoder



final output

$$\hat{\mathbf{x}}_u = a_4(\mathbf{W}_4\mathbf{z}_u^{(3)} + \mathbf{b}_4 + \mathbf{p}_u)$$

$$\mathbf{p}_u = \sum_{j=1}^{n} \mathbf{P}_u^{(i,j)}, i = 1, 2, ..., N$$

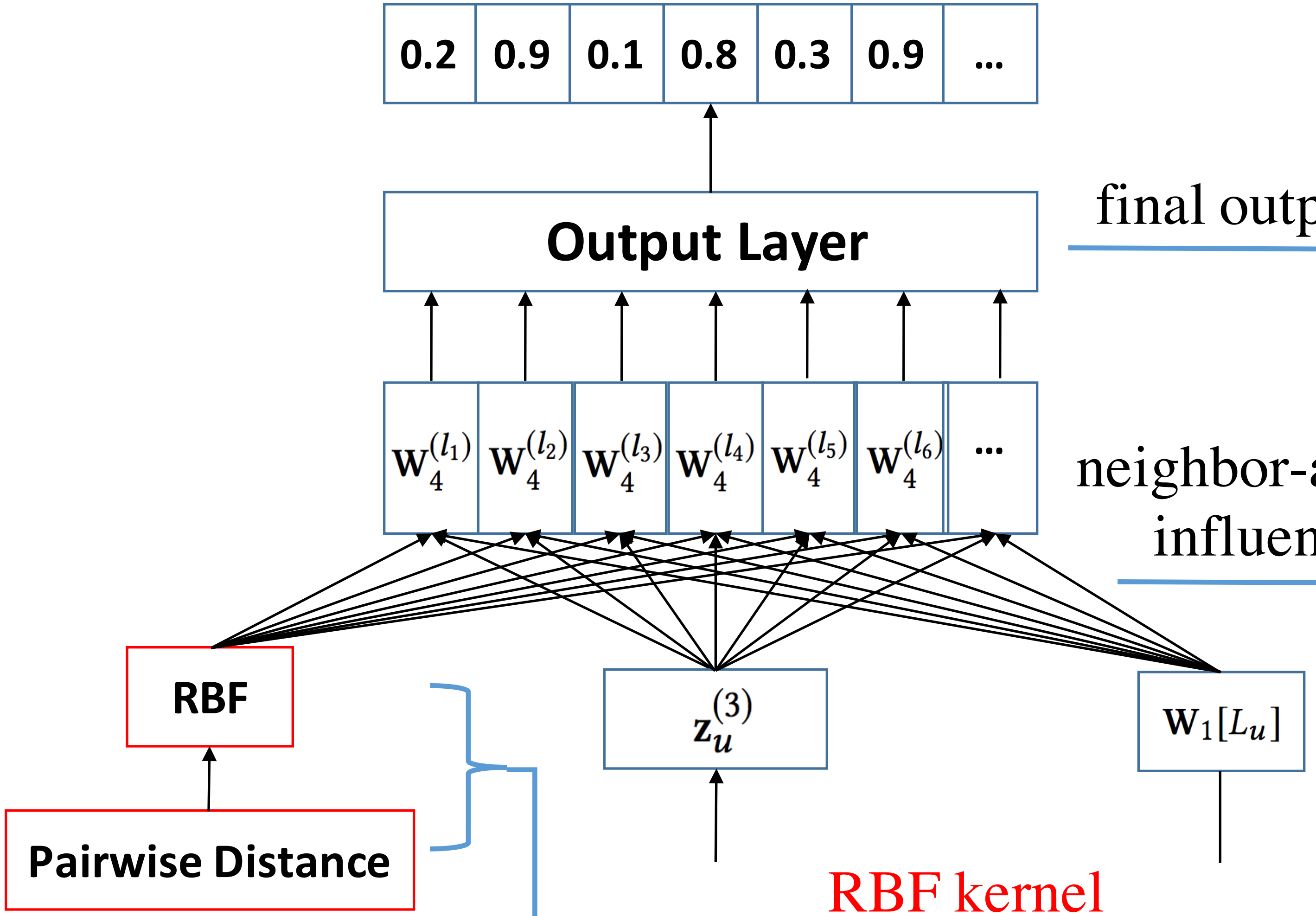
neighbor-aware influence

$$\mathbf{P}_u = (\mathbf{W}_4 \cdot \mathbf{W}_1[L_u]) \odot \mathbf{K}[L_u]$$

RBF kernel

$$K(l_i, l_j) = exp(-\gamma||\mathbf{l}_i - \mathbf{l}_j||^2)$$

Neighbor-aware Decoder



final output

$$\hat{\mathbf{x}}_u = a_4(\mathbf{W}_4\mathbf{z}_u^{(3)} + \mathbf{b}_4 + \mathbf{p}_u)$$

$$\mathbf{p}_u = \sum_{j=1}^{n} \mathbf{P}_u^{(i,j)}, i = 1, 2, ..., N$$

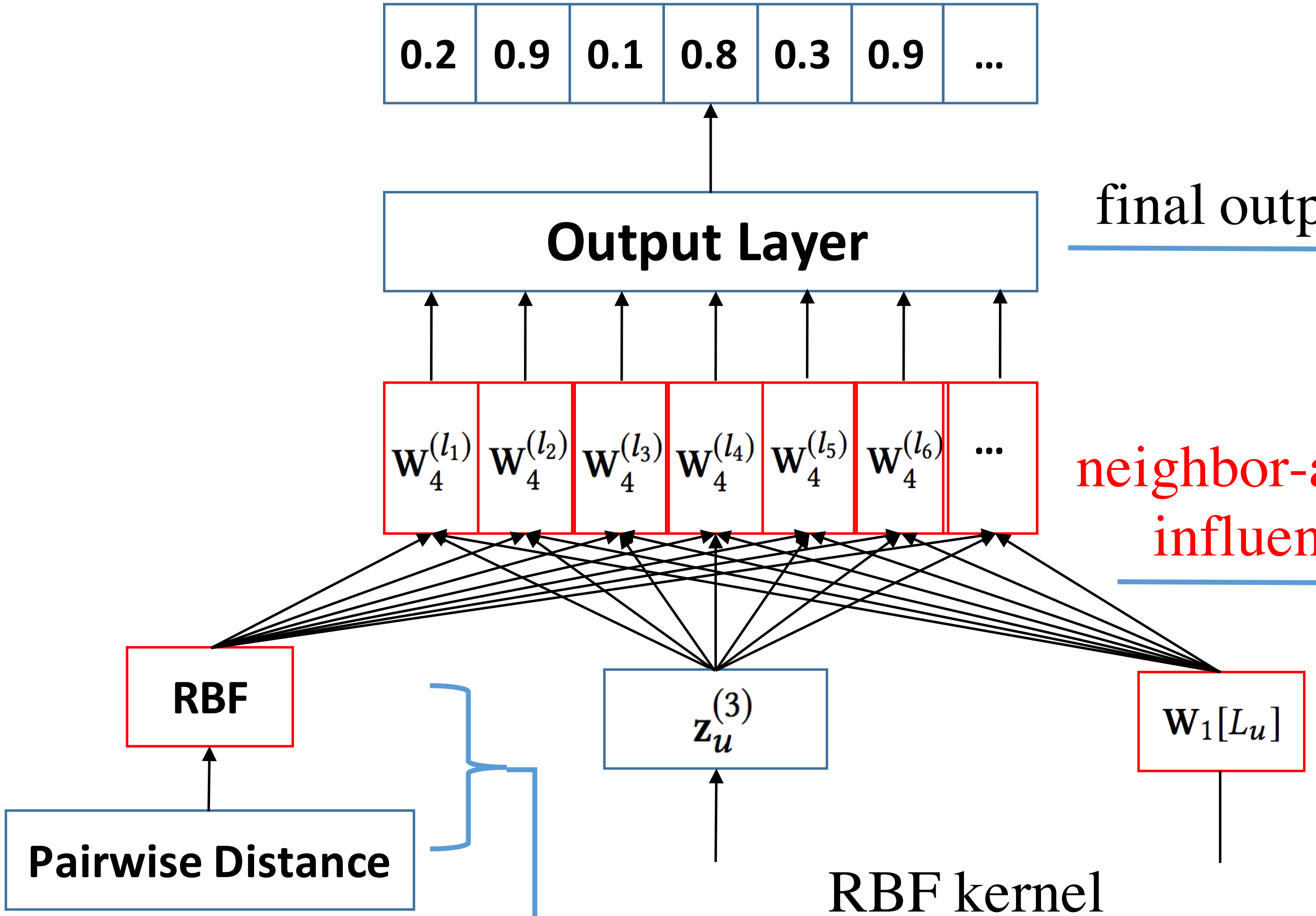
neighbor-aware influence

$$\mathbf{P}_u = (\mathbf{W}_4 \cdot \mathbf{W}_1[L_u]) \odot \mathbf{K}[L_u]$$

RBF kernel

$$K(l_i, l_j) = exp(-\gamma||\mathbf{l}_i - \mathbf{l}_j||^2)$$

Neighbor-aware Decoder



final output

$$\hat{\mathbf{x}}_u = a_4(\mathbf{W}_4\mathbf{z}_u^{(3)} + \mathbf{b}_4 + \mathbf{p}_u)$$

$$\mathbf{p}_u = \sum_{j=1}^{n} \mathbf{P}_u^{(i,j)}, i = 1, 2, ..., N$$

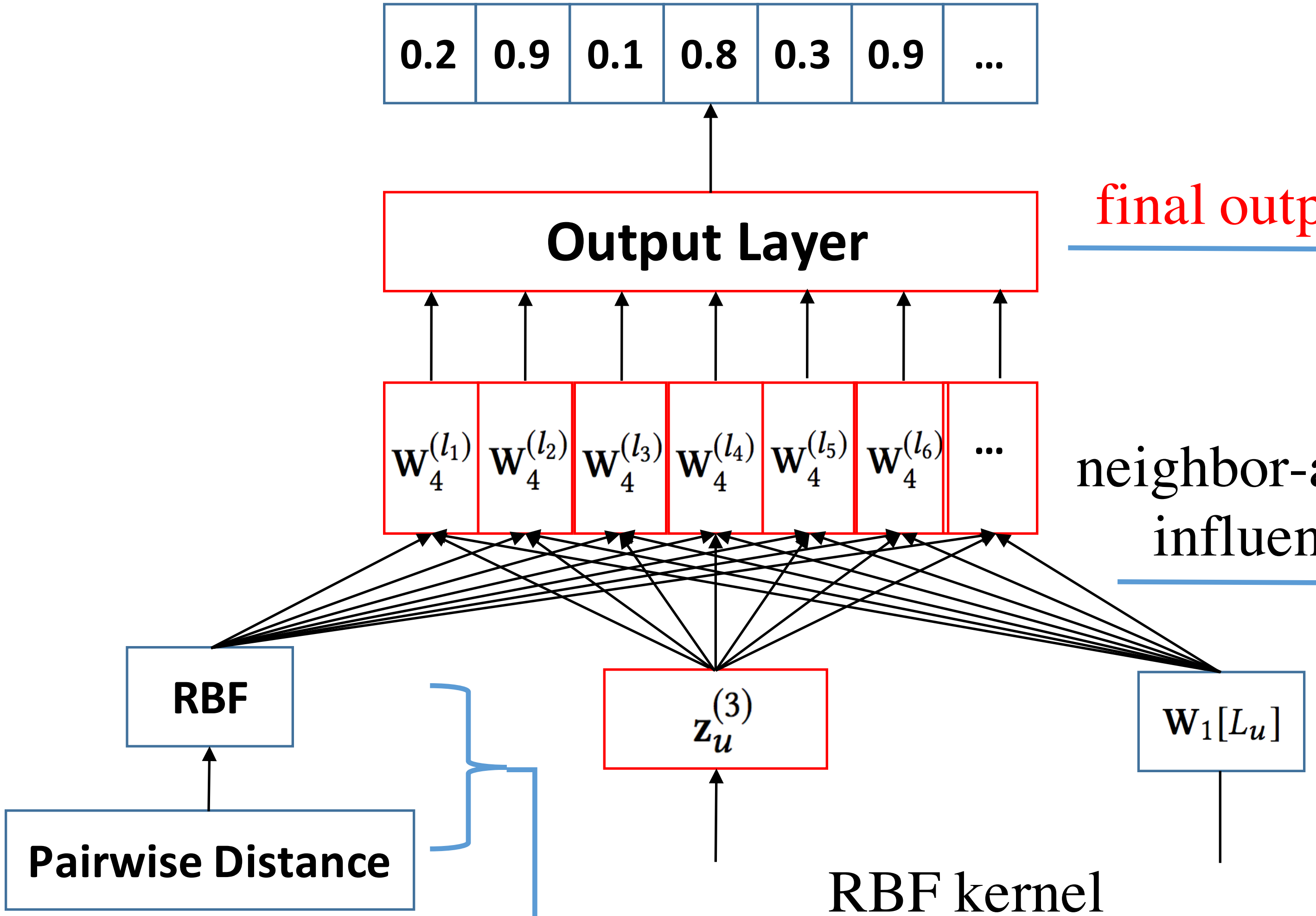
neighbor-aware influence

$$\mathbf{P}_u = (\mathbf{W}_4 \cdot \mathbf{W}_1[L_u]) \odot \mathbf{K}[L_u]$$

RBF kernel

$$K(l_i, l_j) = exp(-\gamma||\mathbf{l}_i - \mathbf{l}_j||^2)$$

Neighbor-aware Decoder



user preference

neighbor-aware influence

final output

$$\hat{\mathbf{x}}_u = a_4(\mathbf{W}_4\mathbf{z}_u^{(3)} + \mathbf{b}_4 + \mathbf{p}_u)$$

$$\boxed{\mathbf{p}_u = \sum_{j=1}^{n} \mathbf{P}_u^{(i,j)}, i = 1, 2, ..., N}$$

neighbor-aware influence

$$\mathbf{P}_u = (\mathbf{W}_4 \cdot \mathbf{W}_1[L_u]) \odot \mathbf{K}[L_u]$$

RBF kernel

$$K(l_i, l_j) = exp(-\gamma||\mathbf{l}_i - \mathbf{l}_j||^2)$$

Loss Function

The **weighted loss for implicit feedback**: the check-in frequency should reflect the user preference levels on POIs

$$\mathcal{L}_{WAE} = \sum_{u=1}^{M} \sum_{i=1}^{N} ||c_{u,i}(x_{u,i} - \hat{x}_{u,i})||_2^2 = ||\mathbf{C} \odot (\mathbf{X} - \hat{\mathbf{X}})||_F^2$$

$$c_{u,i} = \begin{cases} \varphi(r_{u,i}) + 1 & \text{if } r_{u,i} > 0 \\ 1 & otherwise \end{cases}$$

Evaluation

- Three datasets

Dataset	#Users	#POIs	#Check-ins	Density
Gowalla	43,074	46,234	1,720,082	0.0391%
Foursquare	24,941	28,593	939,317	0.0791%
Yelp	30,887	18,995	675,887	0.1099%

For each user, **20%** of her visiting locations are selected as testing.

- Evaluation Metrics
  - Precision@5, 10, 15, 20
  - Recall@5, 10, 15, 20
  - Mean Average Precision (MAP) @5, 10, 15, 20

Evaluation Baselines

WRMF: weighted regularized matrix factorization, ICDM' 2008

BPRMF: bayesian personalized ranking, UAI' 2009

Classical CF methods

MGMMF: multi-center Gaussian model fused with MF, AAAI' 2012

IRENMF: instance-region neighborhood MF, CIKM' 2014

RankGeoFM: ranking-based geographical factorization, SIGIR' 2015

POI recommendation methods

PACE: preference and context embedding, SIGKDD' 2017

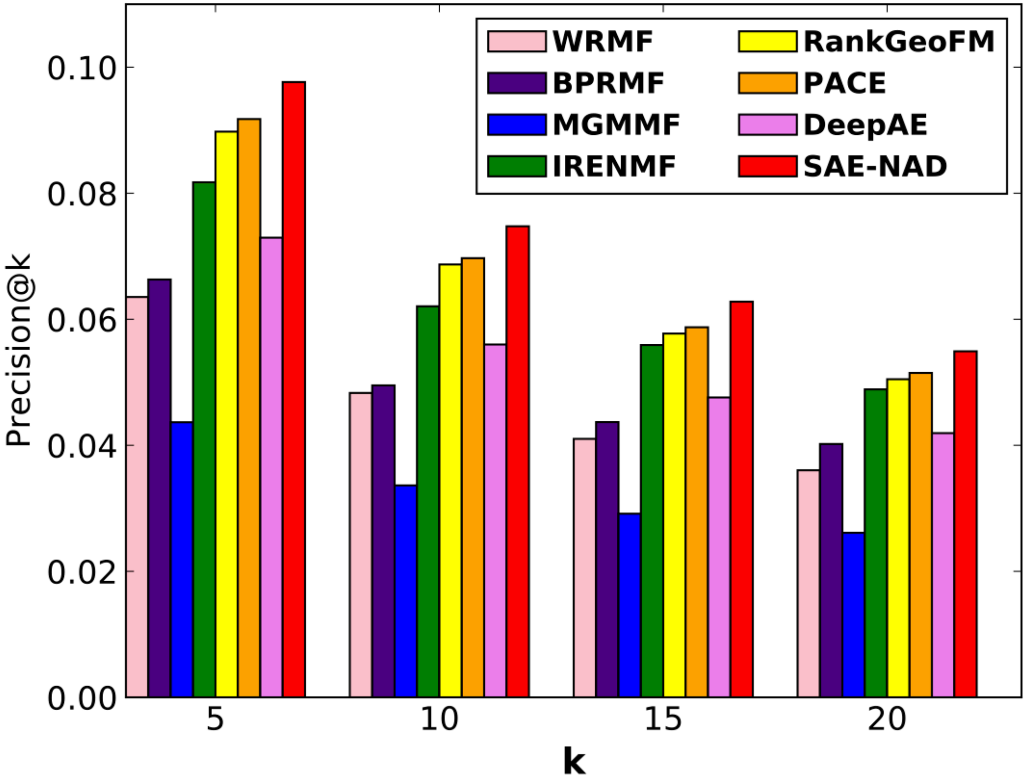
DeepAE: three-hidden-layer autoencoder with a weighted loss

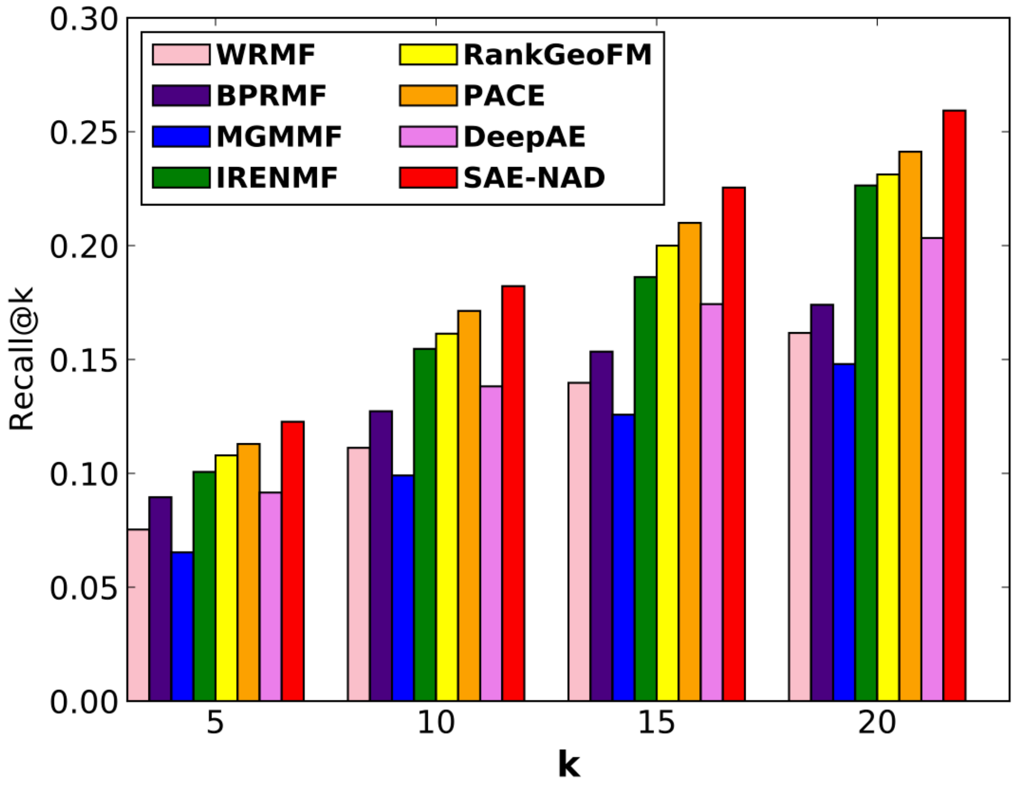
Deep learning based methods

Liu et al., An Experimental Evaluation of Point-of-interest Recommendation in Location-based Social Networks, PVLDB 2017

Evaluation Results

- On Gowalla dataset





1. The proposed method outperforms all other baseline methods on three datasets
2. By incorporating SAE and NAD, the proposed method largely increases the performance of DeepAE
3. Implicit feedback and geographical influence are important to model in POI recommendation

Evaluation Results

- Ablation study

Gowalla	P@10	R@10	MAP@10
WAE	0.05599	0.13819	0.06728
SAE-WAE	0.06039	0.14808	0.07257
NAD-WAE	0.07029	0.17915	0.08699

Foursquare	P@10	R@10	MAP@10
WAE	0.05961	0.11134	0.05632
SAE-WAE	0.06346	0.11813	0.06054
NAD-WAE	0.06598	0.12546	0.06333

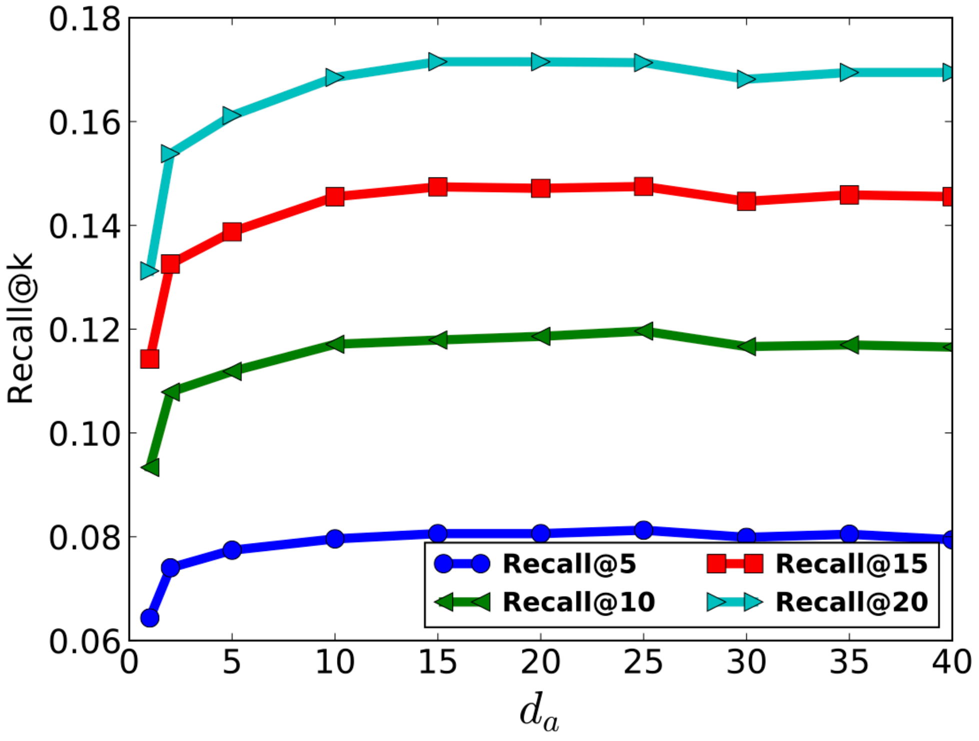
Yelp	P@10	R@10	MAP@10
WAE	0.03764	0.07386	0.03198
SAE-WAE	0.03951	0.07586	0.03307
NAD-WAE	0.04115	0.08016	0.03402

WAE: deep autoencoders with the weighted loss
SAE-WAE: the self-attentive encoder + WAE
NAD-WAE: the neighbor-aware decoder + WAE

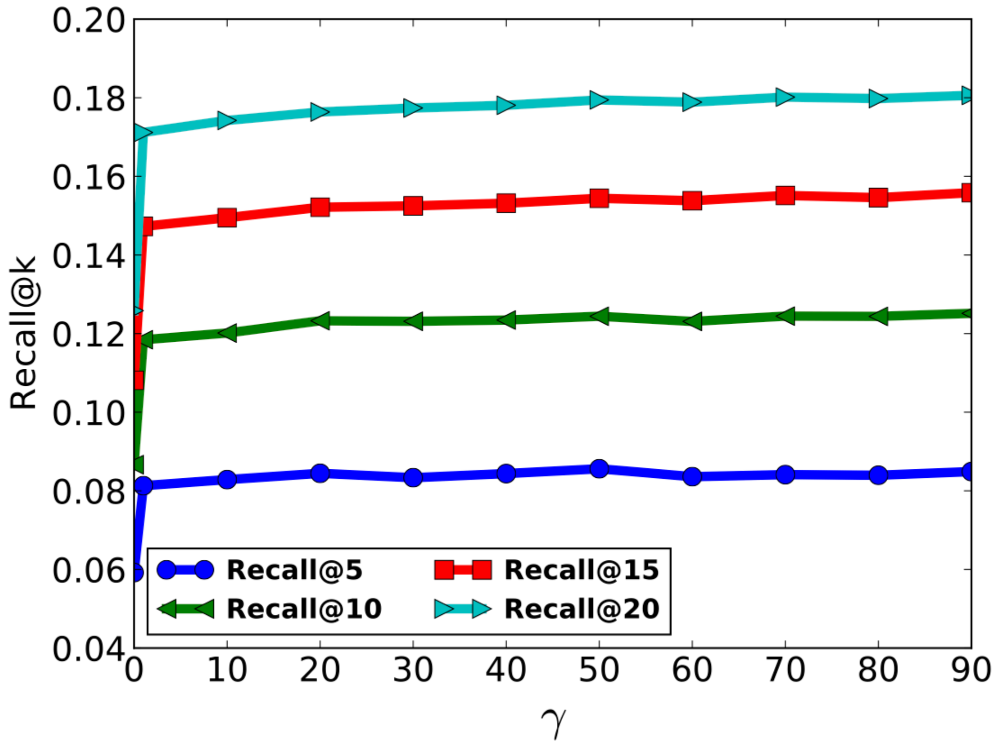
- SAE and NAD all improve the performance of WAE
- Our **NAD** plays a more important role for performance improvement

Evaluation Results

- Hyper-parameters on the Foursquare dataset



The dimension of attention vectors



The geographical correlation level

Conclusion

We propose an **encoder-decoder** based method, which consists of a **self-attentive encoder** and a **neighbor-aware decoder**, to model the complex interactions between users and POIs.

Experimental results show that the proposed method outperforms the state-of-the-art methods significantly for POI recommendation.

McGill

Thank you!

Q & A

Email: chen.ma2@mail.mcgill.ca
Code: https://github.com/allenjack/SAE-NAD
LibRec: https://www.librec.net/

SIGIR
Special Interest Group on Information Retrieval
SIGIR Student Travel Grants