武漢大學
Wuhan University

Understanding User Generated Data

Tieyun Qian
School of Computer Science, Wuhan University

Email: qty@whu.edu.cn

武漢大學
WUHAN UNIVERSITY

Outline

- User Generated Data
- An overview on Sentiment Analysis
- An overview on Recommender Systems
- Our Work

Outline

- User Generated Data
- An overview on Sentiment Analysis
- An overview on Recommender Systems
- Our Work

User Generated Data

User generated content: explicit

- Texts
- Pictures
- Videos
- ……

Web2.0:
User receiving data

User generating data.

User generated data: both explicit and implicit

- Behaviors
- Interactions
- Users’ relations
- ……

User Generated Data

In this talk, I will focus on:

- Explicit texts
- Sentiment analysis

I will also focus on:

- Implicit interactions
- Recommender systems

Understanding User Generated Data

It is important to understand user generated data!

- For sellers: sales promotion, and quality improvement
- For buyers: comparison and analysis before making decision
- For government: public sentiment monitoring
- For marketing: financial analysis
- …………

Outline

- User Generated Data
- An overview on Sentiment Analysis
- An overview on Recommender Systems
- Our Work

An overview on Sentiment Analysis

SA： automatically detect the polarity of the texts

Granularity
- document
- sentence
- aspect
  - Application: more detailed and through analysis
  - Technique: model the semantic relation between the aspect and sentence

武漢大學
WUHAN UNIVERSITY

An overview on Sentiment Analysis

Two types of scenarios:

- Aspect term sentiment analysis (ATSA)

  to predict the polarity of an entity in the sentence

- Aspect category sentiment analysis (ACSA)

  to predict the polarity of an predefined aspect category

A mix of students and area residents crowd into this narrow, barely there space for its quick, tasty treats at dirt-cheap prices.

- ATSA

  space: negative, *treats*： positive, *prices*: positive

- ACSA

  food: positive, *price*: positive, *ambience*: negative

Main Problems in ABSA

- Data labeling is extremely expensive in ABSA
- Hard to train high-quality embeddings for the terms with low frequency
- Hard to locate the aspect category in the sentence

Outline

- User Generated Data
- An overview on Sentiment Analysis
- An overview on Recommender Systems
- Our Work

An overview on Recommender Systems

- Traditional recommender systems
- Sequential recommender systems (SRS) √
- Session based sequential recommender systems √

Main Problems in SRS

- Modeling long term interests
- Modeling short term interests
- Dealing with cold start issues

Outline

- User Generated Data
- An overview on Sentiment Analysis
- An overview on Recommender Systems
- Our Work

Our Work

- Sentiment analysis
- Recommender systems
- User profiling
- Representation learning

Transfer Capsule Network for Aspect Level Sentiment Classification

Zhuang Chen, Tieyun Qian*
School of Computer Science, WuHan University, China
{zhchen18, qty}@whu.edu.cn

Published at ACL 2019
(Research Long Paper)

TransCap - Motivations(1/17)

- The lack of labeled data is a major obstacle in ASC. Publicly available datasets for ASC often contain limited number of training samples.
- Document-level labeled data like reviews are easily accessible from online websites such as Yelp and Amazon. The accompanying rating scores can naturally serve as the sentiment labels.
- The document-level data contain useful sentiment knowledge for analysis on aspect level data since they may share many linguistic and semantic patterns.

TransCap - Architecture(2/17)

Definitions

- ► T_A : Aspect-level sentiment classification (ASC).
- ► T_D : Document-level sentiment classification (DSC).
- ► TransCap: Improve T_A with transferred knowledge from T_D.

TransCap - Architecture(3/17)

Preliminary

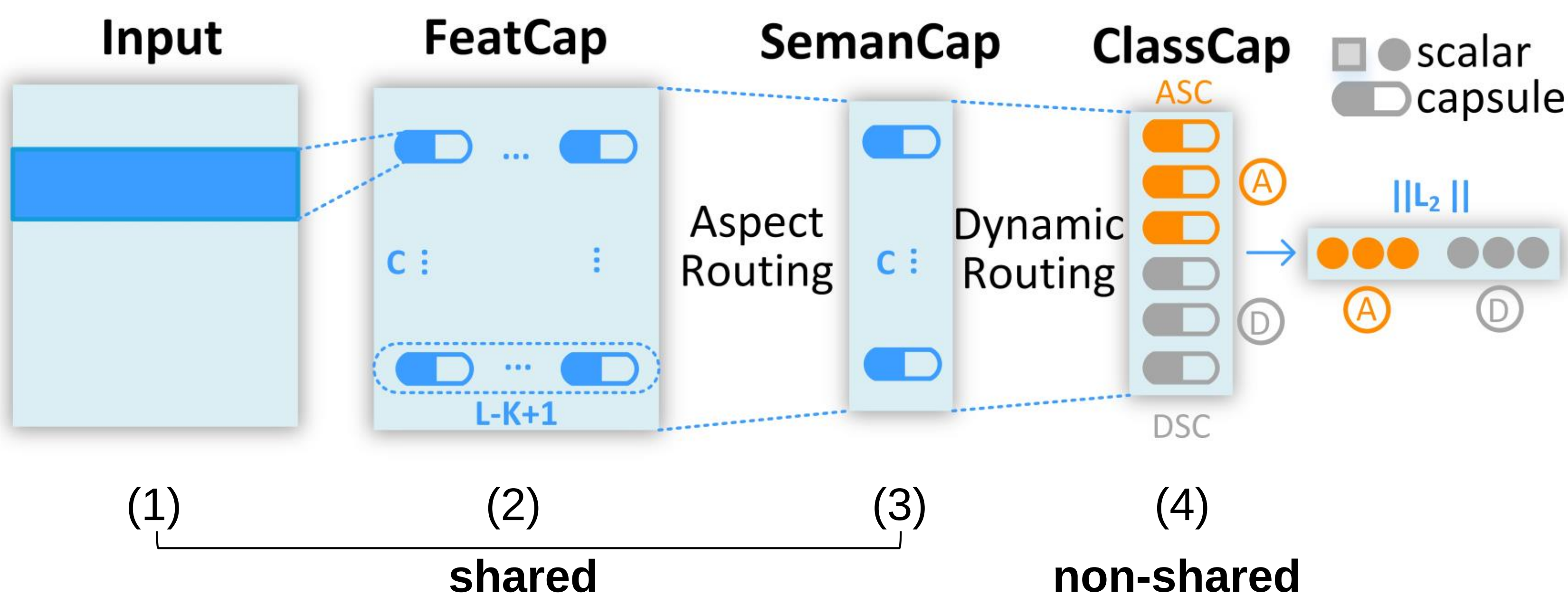
► CapsNet : CapsNet is first proposed for image classification in computer vision.

► Applied to NLP tasks : text classification, relation extraction...

► Why we use capsules : encapsulated features, separate classes, dynamic routing.

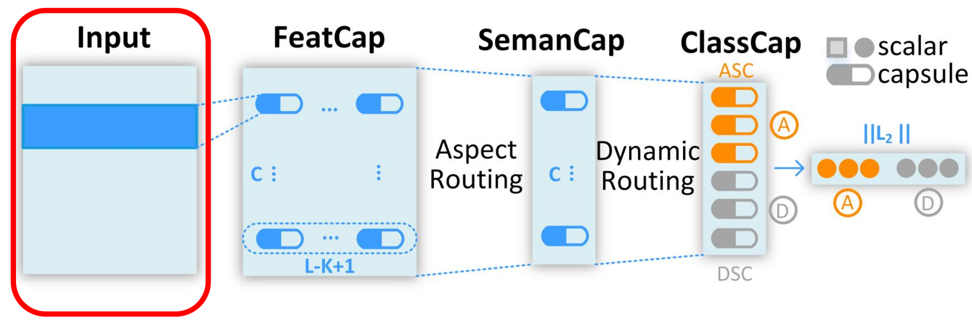
TransCap - Architecture$_{(4/17)}$

TransCap Overview



Transfer Capsule Network for Aspect Level Sentiment Classification (ACL2019)

TransCap - Architecture(5/17)



(1) Input Layer (shared)

- Look-up layer

  $\{e_1, ..., e_a, ..., e_L\} \in \mathbb{R}^{d_w \times L}$

- Position Information

  $\{l_1, ..., l_a, ..., l_L\} \in \mathbb{R}^{d_l \times L}$ $x_i = (e_i \oplus l_i)$

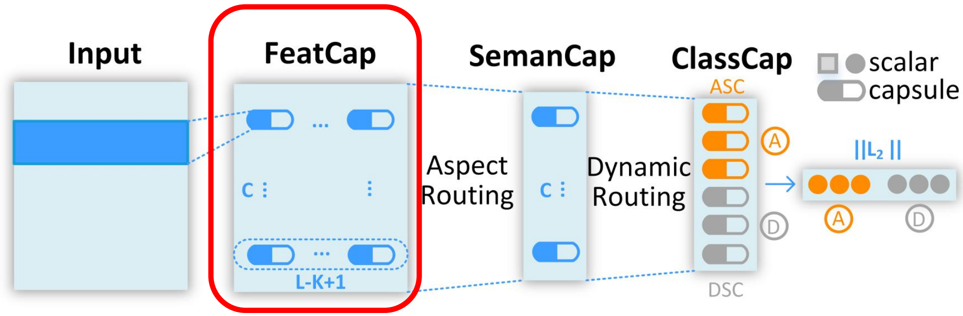
- Sentence Embedding

  $\mathbf{X} = \{x_1, ..., x_L\} \in \mathbb{R}^{d_h \times L}$

Transfer Capsule Network for Aspect Level Sentiment Classification (ACL2019)

WUHAN UNIVERSITY

TransCap - Architecture$_{(6/17)}$



(2) Feature Capsule Layer (shared)

- Filter Group
  $\mathbf{F} \in \mathbb{R}^{d_p \times (d_h \times K)}$

- Multiple Convolution Operations
  $\boldsymbol{r_i} = \mathbf{X}_{i:i+K} * \mathbf{F} + b$

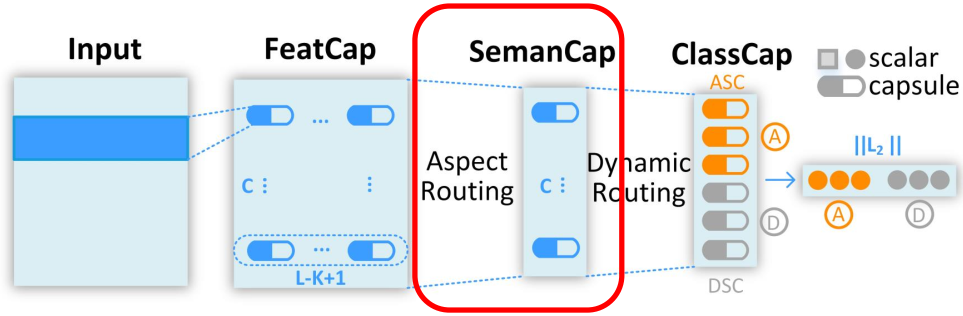
- Generated Feature Capsules (1 category of semantic meaning)
  $\boldsymbol{r} \in \mathbb{R}^{d_p \times (L-K+1)}$

- Repeat for Multiple Channels (C categories of semantic meaning)
  $\mathbf{R} = [\boldsymbol{r_1}, \boldsymbol{r_2}, ..., \boldsymbol{r_C}] \in \mathbb{R}^{C \times d_p \times (L-K+1)}$

Transfer Capsule Network for Aspect Level Sentiment Classification (ACL2019)

TransCap - Architecture$_{(7/17)}$

(3) Semantic Capsule Layer (shared)



- An Additional Convolution

$$a_i = sigmoid(\mathbf{X}_{i:i+K} * \mathbf{F}_a + \mathbf{T}_a \boldsymbol{e_a} + b_a)$$
$$a_i \in [0,1]$$

- <u>Aspect Routing</u>

weight $g_i = \begin{cases} a_i & \mathrm{X} \in \mathcal{C}_A \\ 1.0 & \mathrm{X} \in \mathcal{C}_D \end{cases} \in \mathbb{R}^{1\times(L-K+1)}$ $\quad \mathbf{G} = [\boldsymbol{g_1}, \boldsymbol{g_2}, ..., \boldsymbol{g_C}] \in \mathbb{R}^{C\times 1\times(L-K+1)}$

routing $\mathbf{P} = \mathbf{R} \odot \mathbf{G} \in \mathbb{R}^{C\times d_p \times (L-K+1)}$

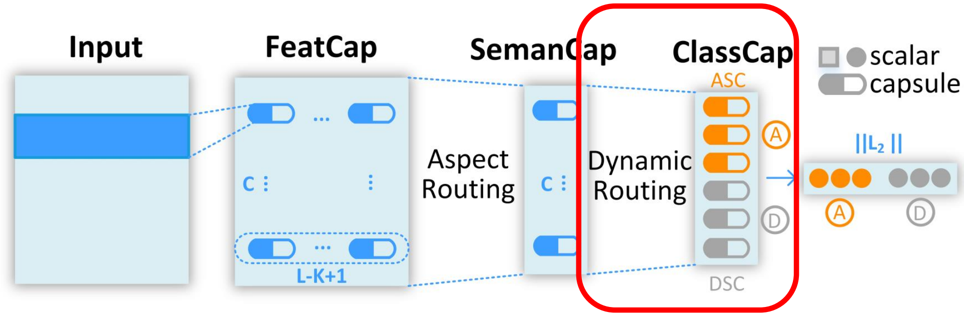
- Generated Semantic Capsules (<u>C</u> categories of semantic meaning)

condensation $\mathbf{U} = \overset{C\times d_p}{\underset{t=1}{max}} \mathbf{P}_t \in \mathbb{R}^{C\times d_p}$ $\quad \boldsymbol{u_i} \leftarrow \frac{\|\boldsymbol{u_i}\|^2}{1+\|\boldsymbol{u_i}\|^2}\frac{\boldsymbol{u_i}}{\|\boldsymbol{u_i}\|}$

Transfer Capsule Network for Aspect Level Sentiment Classification (ACL2019)

TransCap - Architecture$_{(8/17)}$

(4) Class Capsule Layer (non-shared)



- Separate Class Capsules
  6 capsules for 2 tasks & 3 polarities

- Generated Class Capsules
  prediction vector $\hat{\boldsymbol{u}}_{j|i} = \mathbf{W}_{ij}\, \boldsymbol{u}_i$ $\quad \mathbf{W}_{ij} \in \mathbb{R}^{d_c \times d_p}$
  weighted sum $\boldsymbol{s}_j = \sum_i c_{ij} \hat{\boldsymbol{u}}_{j|i}$ $\quad c_{ij} = \frac{exp(\boxed{b_{ij}})}{\sum_k exp(b_{ik})}$
  output vector $\boldsymbol{v}_j = squash(\boldsymbol{s}_j)$

- Dynamic Routing
  $a_{ij} = \hat{\boldsymbol{u}}_{j|i} \cdot \boldsymbol{v}_j \qquad \boxed{b_{ij}} \leftarrow b_{ij} + a_{ij}$

TransCap - Architecture$_{(9/17)}$

Training Procedure

- Data from T_D and T_A take turns to train TransCap model.
- Layer(1)(2)(3) transfer knowledge from T_D to T_A, layer(4) avoids mutual disturbance.
- Loss Function

each polartity $\mathcal{L}_j = Y_j max(0, m^+ - \|\boldsymbol{v_j}\|)^2 + \lambda(1 - Y_j) max(0, \|\boldsymbol{v_j}\| - m^-)^2$

single task $\mathcal{L}_S = \sum_{j=1}^{J} \mathcal{L}_j$

final loss $\mathcal{L} = \mathcal{L}_A + \gamma \mathcal{L}_D$

TransCap - Experiments (10/17)

Task	Dataset	Type	Pos.	Neu.	Neg.
$\mathcal{T}_A$	Restaurant	train	2164	633	805
		test	728	196	196
	Laptop	train	987	460	866
		test	341	169	128
$\mathcal{T}_D$	Yelp	train	10k	10k	10k
	Amazon	train	10k	10k	10k
	Twitter	train	10k	10k	10k

Datasets for T_A

- SemEval 2014 Task 4 : Restaurant & Laptop
- 80% for training, 20% for development

Datasets for T_D

- Yelp, Amazon and Twitter
- 30,000 balanced samples, all for training
- <3 : negative, =3 : neutral, >3 : positive (Yelp, Amazon)

Dataset Combinations

- {Y,A} : {Restaurant+Yelp, Laptop+Amazon} Relevant but Imprecise
- {T,T} : {Restaurant+Twitter, Laptop+Twitter} Precise but Irrelevant

武漢大學
WUHAN UNIVERSITY

TransCap - Experiments (11/17)

Results

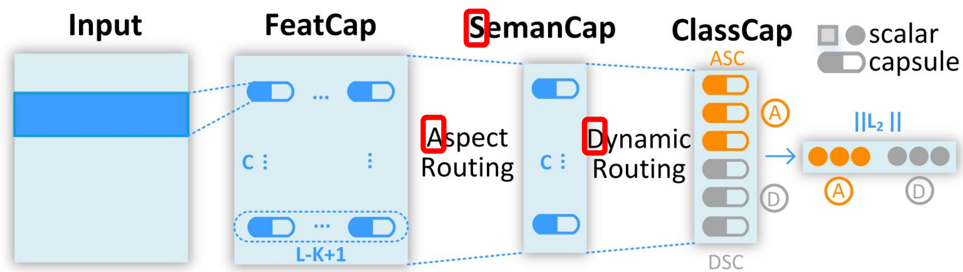
- Averaged results over 5 runs with random initialization.
- TransCap outperforms all baselines.
- {Y,A} and {T,T} achieve similar results.

	Model	Restaurant		Laptop	
		Acc.	F1	Acc.	F1
M1	ATAE-LSTM	78.38	66.36	69.12	63.24
M2	IAN	78.71	67.71	69.56	63.72
M3	AF-LSTM(CONV)	76.46	65.54	69.97	63.70
M4	AF-LSTM(CORR)	75.96	64.41	69.78	63.38
M5	PBAN	78.62	67.45	71.98	66.91
M6	MemNN	77.69	67.53	68.86	62.60
M7	RAM	78.41	68.52	<u>72.16</u>	66.97
M8	CEA	78.44	66.78	70.52	64.52
M9	DAuM	77.91	66.47	70.36	65.86
M10	IARM	77.73	66.66	68.63	63.30
M11	PRET+MULT	<u>78.73</u>	<u>68.63</u>	71.91	<u>68.79</u>
M12	GCAE	76.09	63.29	68.72	63.32
M13	TransCap{S}	78.84	69.70	72.65	68.77
M14	TransCap{Y,A}	**79.55**	**71.41**	73.51	69.81
M15	TransCap{T,T}	79.29	70.85	**73.87**	**70.10**

TransCap - Experiments (12/17)

Ablation Study

- “- A”: remove Aspect routing
- “- S”: remove Semantic capsules
- “- D”: remove Dynamic routing

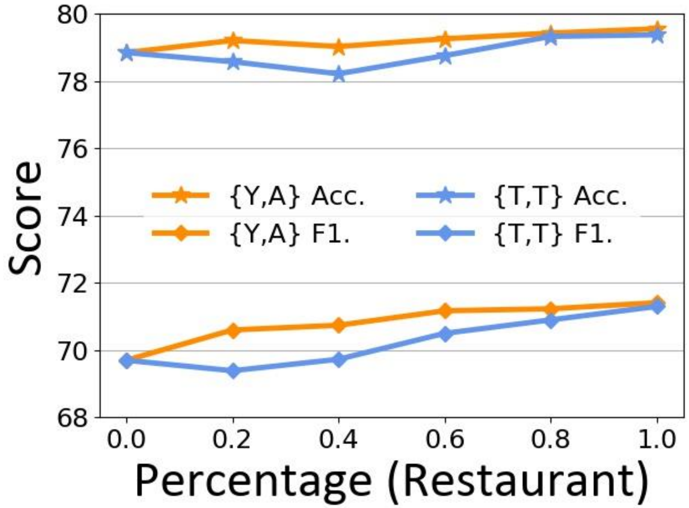


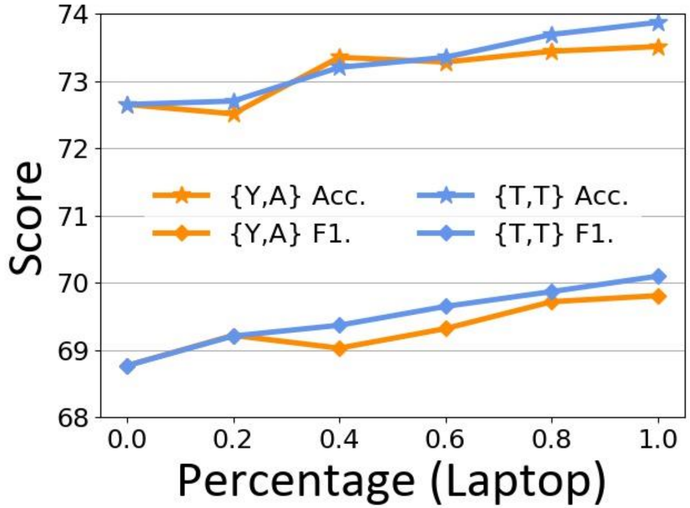
	Restaurant				Laptop			
	{Y,A}		{T,T}		{Y,A}		{T,T}	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
Ori.	79.55	71.41	79.29	70.85	73.51	69.81	73.87	70.10
- A.	3.75↓	**6.49**↓	**2.63**↓	**3.95**↓	2.98↓	5.34↓	**3.34**↓	**3.80**↓
- S.	**4.01**↓	5.14↓	1.45↓	2.08↓	2.35↓	3.64↓	2.40↓	2.15↓
- D.	2.80↓	4.06↓	0.54↓	1.01↓	**3.29**↓	**6.03**↓	1.14↓	1.75↓

TransCap - Experiments (13/17)

Parameter Analysis

► Influence of Auxiliary Corpus Size

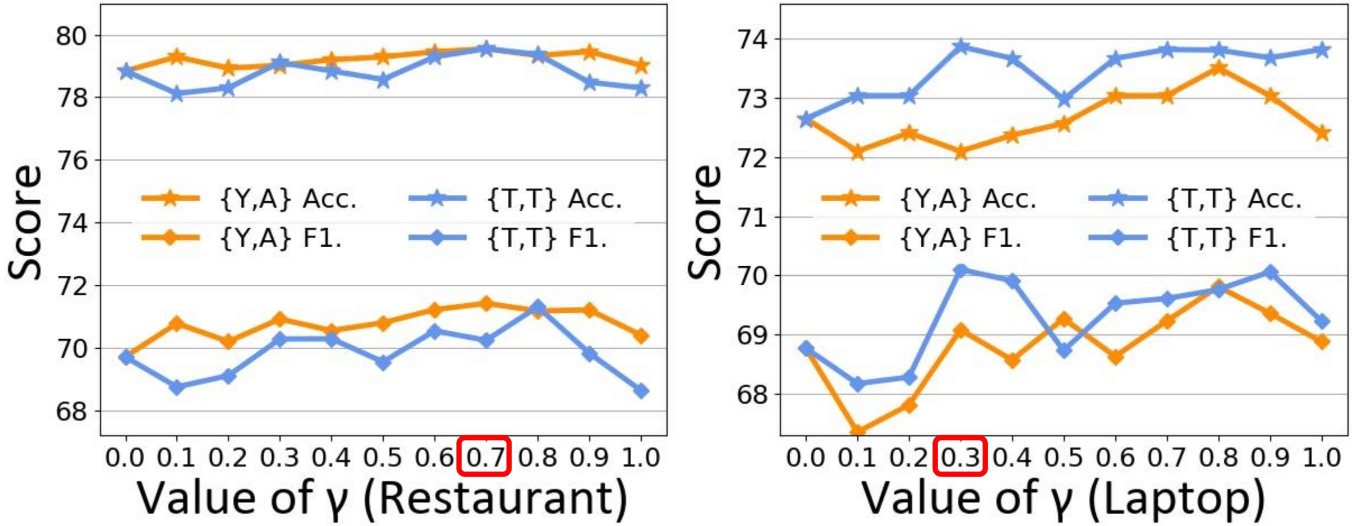




TransCap - Experiments (14/17)

Parameter Analysis

► Effects of Balance Factor γ



武漢大學 WUHAN UNIVERSITY

TransCap - Experiments (15/17)

Case Study

► Part 1 : What does TransCap transfer?

test sample (TransCap ✓ Others ✗)

1. *" It has so much more speed and the [screen]$_{pos}$ is very sharp. "*

"sharp" is a multi-polarity word

2. *" Once open, the [leading edge]$_{neg}$ is razor sharp. "*
3. *" [Graphics]$_{pos}$ are clean and sharp, internet interfaces are seamless. "*

- The training set in Laptop contains only 8 samples including "sharp" with 5 of them are labeled as negative.
- Amazon dataset contains 294 samples where "sharp" co-occurs with lots of different contexts.

TransCap - Experiments (16/17)

Case Study

► Part 2 : How does TransCap make decisions?

test sample

1. " Great [food]$_{pos}$ but the [service]$_{neg}$ is dreadful ! "

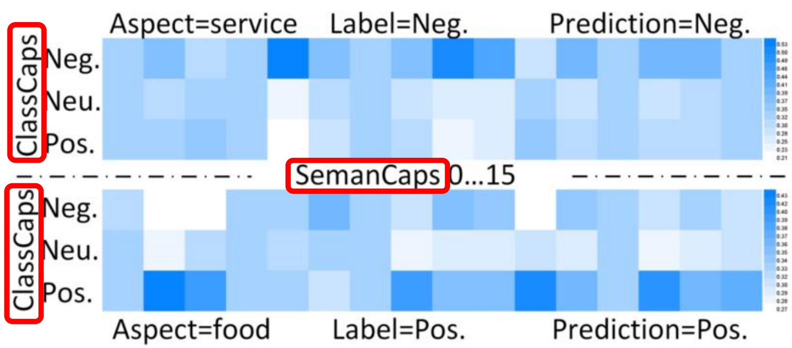


Figure. Visualization of coupling coefficients c_{ij} after dynamic routing.

TransCap - Experiments (17/17)

Case Study

► Part 3 : Can TransCap handle complicated patterns?

test sample (TransCap ✓ Others ✗)

1. " The [staff]$_{neg}$ should be a bit more friendly. "

- An euphemistic negative review towards the aspect [staff].
- TransCap generates and transfers sentence-level knowledge.

auxiliary sample from Yelp

2. " The pro-shop staff should be more polite when answering the phone...]$_{neg}$ "

Aspect Aware Learning for Aspect Category Sentiment Analysis

Peisong Zhu, Zhuang Chen, Haojie Zheng, Tieyun Qian*
School of Computer Science, WuHan University, China
{zhups24, zhchen18, zhenghaojie, qty}@whu.edu.cn

Published at TKDD 2019

AAL - Motivations$_{(1/11)}$

- For category-based ASC, 1) the first challenge is to locate the exact position of the aspect category, 2) and the second is to correlate opinion words with different categories in one sentence.

great	0. 04
food	0. 03
but	0. 01
the	0. 01
service	0. 01
was	0. 01
dreadful	0. 80
!	0. 09

Category： service
Label： Negative
Prediction：Negative

great	0. 07
food	0. 04
but	0. 02
the	0. 01
service	0. 01
was	0.02
dreadful	0.72
!	0.11

Category： food
Label： Positive
Prediction：Negative

- Categories have specific opinion words.

delicious → food
expensive → price
quiet → ambience

The fish is delicious.

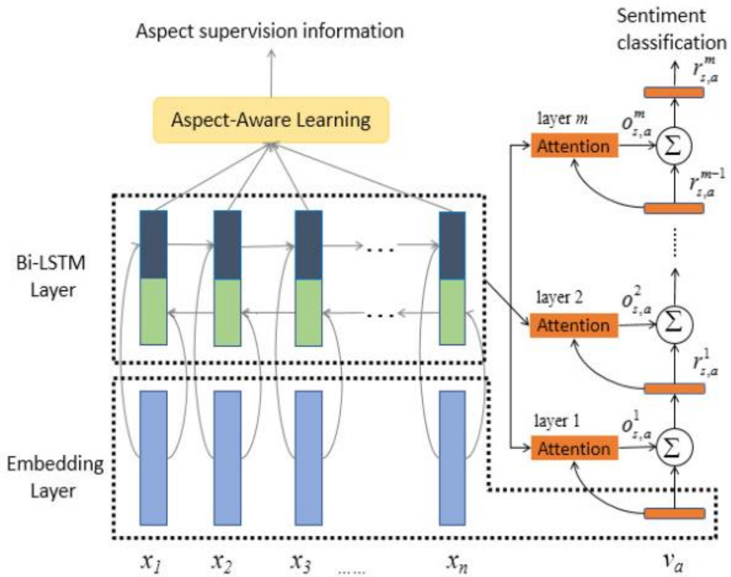
Great food but the service is dreadful.

AAL - Architecture$_{(2/11)}$

AAL Overview

We introduce additional supervision (Aspect-aware learning, AAL) to help correlate contexts with given aspect categories.

- ▶ AAL-LEX at word level
- ▶ AAL-SS at sentence level



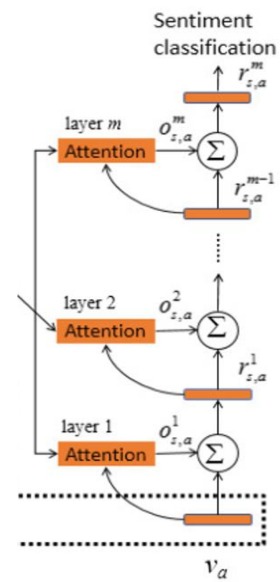
Aspect Aware Learning for Aspect Category Sentiment Analysis (TKDD19)

AAL - Architecture (3/11)

(1) Input Layer

- Look-up layer initialized by Glove
- Bi-LSTM

(2) Task 1 - Sentiment Prediction



- Attention $$o^m_{s,a} = \sum_{i=1}^{n} h_i \beta^m_{i,a},$$ $$\beta^m_{i,a} = softmax(tanh(W_s[h_i, r^{m-1}_{s,a}] + b_s)),$$
- Aggregation $$r^m_{s,a} = o^m_{s,a} + r^{m-1}_{s,a}$$
- Prediction $$\hat{p}(c|S,a) = softmax(W_c r^m_{s,a} + b_c)$$
- Task1 loss $$L_{senti} = -\sum_{(S,a)\in T} \sum_{c \in C} p(c|S,a) \cdot \hat{p}(c|S,a)$$

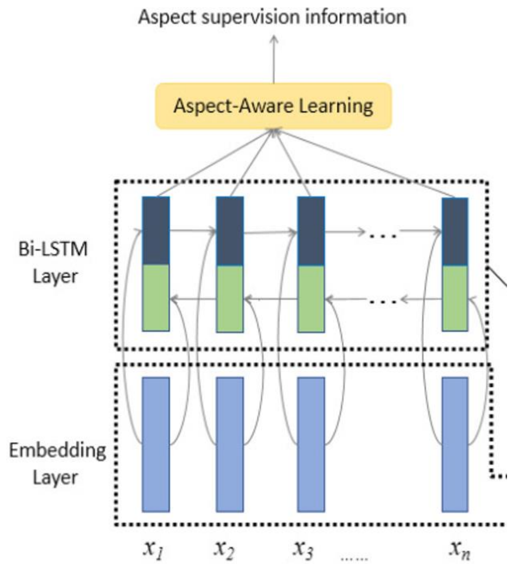
stack to multiple layers...

AAL - Architecture$_{(4/11)}$

(3) Task 2 - Category Prediction

AAL-LEX at word level

- Word-cate PMI $$PMI(w,a) = log\frac{N(w,a) \times N(S)}{N(w) \times N(a)}$$
- Co-occurrence prob $$p(a|w) = \frac{exp(PMI(w,a))}{\sum_{j=1}^{|A|} exp(PMI(w,a_j))}$$
- Normalization $$\hat{p}(a|w) = softmax(W_a h + b_a)$$
- Task2 loss $$L_{asplex} = -\sum_{S \in T}\sum_{w \in S}\sum_{a \in A} p(a|w) \cdot log\hat{p}(a|w)$$
- Final loss $$L_{aallex} = (1-\gamma_1)L_{senti} + \gamma_1 L_{asplex}$$



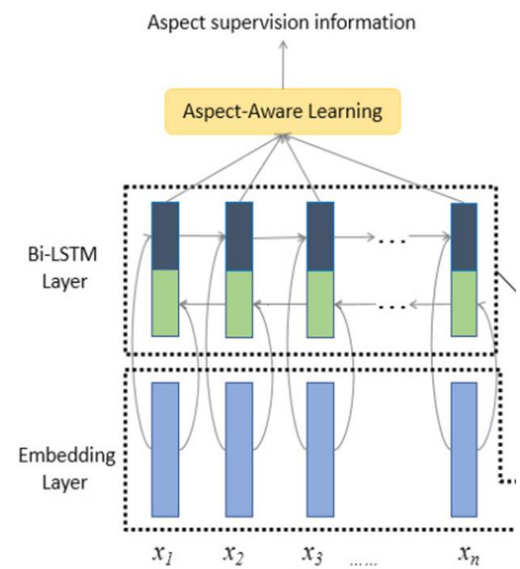
Aspect Aware Learning for Aspect Category Sentiment Analysis (TKDD19)

AAL - Architecture(5/11)

(3) Task 2 - Category Prediction

AAL-SS at sentence level

- Attention&aggregation $M = tanh(W_h H + b_h),$ $\alpha = softmax(u^T M), r_a = H\alpha^T$
- Category prob $\hat{p}(a|S) = \sigma(W_a r_a + b_a)$
- Task2 loss $L_{aspss} = -\sum_{S \in T} \sum_{a \in A} y(S,a) \cdot \hat{p}(a|S)$
- Final loss $L_{aalss} = (1 - \gamma_2) L_{senti} + \gamma_2 L_{aspss}$



Aspect Aware Learning for Aspect Category Sentiment Analysis (TKDD19)

AAL - Experiments (6/11)

Datasets

SemEval 2014, SemEval 2015, SemEval 2016

Data	Aspect	Positive		Negative		Netural	
		Train	Test	Train	Test	Train	Test
Restaurant -2014	food	867	302	209	69	90	31
	price	179	51	115	28	10	1
	service	324	101	218	63	20	3
	ambience	263	76	98	21	23	8
	anecdotes/miscellaneous	546	127	199	41	357	51
	Total	2179	657	839	222	500	94
Laptop -2015	general	366	187	154	71	6	15
	price	40	35	25	5	22	17
	quality	110	55	266	60	10	4
	operation_performance	154	82	111	76	9	5
	usability	106	32	42	26	10	11
	design_features	142	65	63	39	32	16
	portability	36	5	8	2	0	1
	connectivity	17	6	15	15	0	3
	miscellaneous	70	43	33	21	12	5
	Total	1041	513	717	315	101	77
Restaurant -2015	service#general	153	40	95	110	7	5
	food#quality	328	153	95	59	13	10
	restaurant#general	217	93	47	50	5	2
	drinks#style_options	23	4	1	2	0	0
	drinks#prices	11	2	4	3	0	0
	restaurant#prices	29	6	14	28	5	1
	ambience#general	127	45	22	17	8	6
	food#style_options	56	19	23	15	5	4
	restaurant#miscellaneous	41	19	17	12	3	7
	food#prices	25	8	22	19	1	2
	drinks#quality	31	7	1	4	1	1
	location#general	17	4	2	0	1	4
	food#general	0	0	1	0	0	0
	Total	1058	400	344	319	49	42
Restaurant -2016	restaurant#general	312	107	100	34	8	1
	service#general	194	66	206	70	12	7
	food#quality	480	186	153	24	23	12
	food#style_options	76	25	41	14	9	8
	drinks#style_options	27	11	3	1	0	0
	drinks#prices	13	0	7	3	0	0
	restaurant#prices	34	6	40	13	6	2
	restaurant#miscellaneous	57	16	27	13	13	4
	ambience#general	171	52	34	1	15	3
	food#prices	36	6	44	13	1	3
	location#general	21	11	1	0	6	2
	drinks#quality	39	20	5	1	2	0
	Total	1460	506	661	187	95	42

AAL - Experiments (7/11)

Results

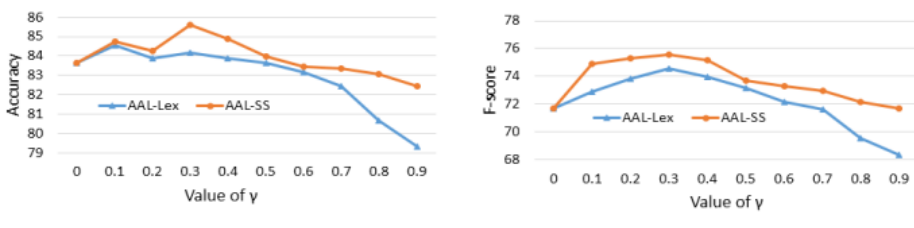
Method	Restaurant-2014				Laptop-2015			
	Acc.	Pre.	Rec.	F-score	Acc.	Pre.	Rec.	F-score
AE-LSTM	81.40	71.78	65.22	67.63	74.39	56.94	55.66	54.73
ATAE-LSTM	82.32	73.95	70.45	71.96	74.50	56.42	55.77	55.60
Tensor DyMemNN	80.99	73.22	65.04	68.10	75.66	60.88	55.12	53.47
Holo DyMemNN	80.37	72.00	67.76	69.62	76.08	66.57	55.37	53.05
CEA	82.94	73.23	69.01	70.81	74.50	59.35	55.82	56.51
GCAE	81.09	69.93	65.88	67.61	75.03	60.79	59.80	59.96
DAuM	81.50	75.51	64.66	67.92	76.19	50.00	54.63	52.21
AF-LSTM (CORR)	82.01	74.83	71.24	72.01	76.19	49.93	54.87	52.26
AF-LSTM (CONV)	82.22	72.41	**74.63**	73.32	76.29	50.04	54.77	52.29
AAL-No	83.63	71.25	72.22	71.67	75.17	60.34	56.26	57.61
AAL-LEX	84.17	75.25	73.96	74.57	75.87	62.18	58.14	59.25
AAL-SS	**85.61**	**78.03**	73.71	**75.54**	**78.29**	**67.75**	**59.82**	**60.00**

Method	Restaurant-2015				Restaurant-2016			
	Acc.	Pre.	Rec.	F-score	Acc.	Pre.	Rec.	F-score
AE-LSTM	76.08	52.39	52.52	51.42	80.95	62.62	58.74	58.25
ATAE-LSTM	76.87	51.13	54.34	52.61	81.09	71.09	55.65	58.86
Tensor DyMemNN	78.58	59.59	55.58	55.08	82.99	76.01	62.14	63.45
Holo DyMemNN	77.92	52.16	54.39	53.08	82.59	69.53	58.23	59.93
CEA	78.19	66.73	58.00	59.36	82.72	77.66	60.26	62.04
GCAE	77.53	60.93	57.42	58.22	82.59	72.26	61.32	61.86
DAuM	78.98	67.87	57.98	59.01	81.36	66.58	64.09	63.95
AF-LSTM (CORR)	77.40	51.55	54.25	52.84	83.27	71.17	61.31	62.20
AF-LSTM (CONV)	76.61	51.84	53.18	52.01	81.90	69.07	59.14	61.29
AAL-No	76.74	63.55	56.64	57.81	82.31	65.88	64.27	64.62
AAL-LEX	77.92	63.59	57.32	58.14	84.08	**80.00**	64.06	65.56
AAL-SS	**79.11**	**70.44**	**60.61**	**62.80**	**84.35**	77.68	**64.64**	**67.14**

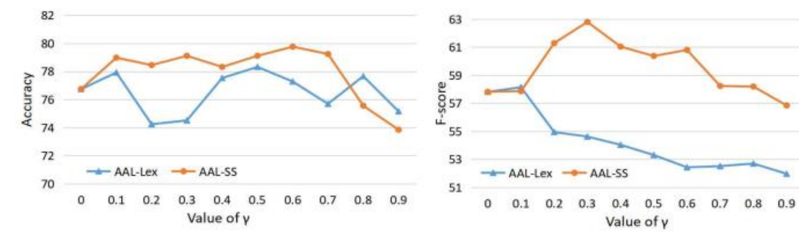
AAL - Experiments (8/11)

Parameter Analysis

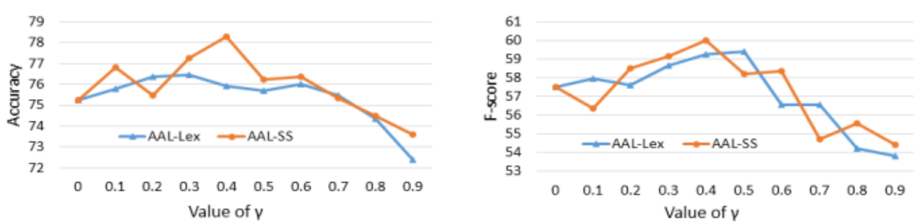
► Effects of Balance Factor



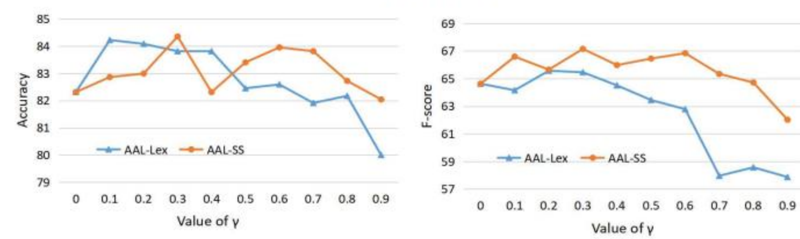
(a) Restaurant-2014



(c) Restaurant-2015



(b) Laptop-2015

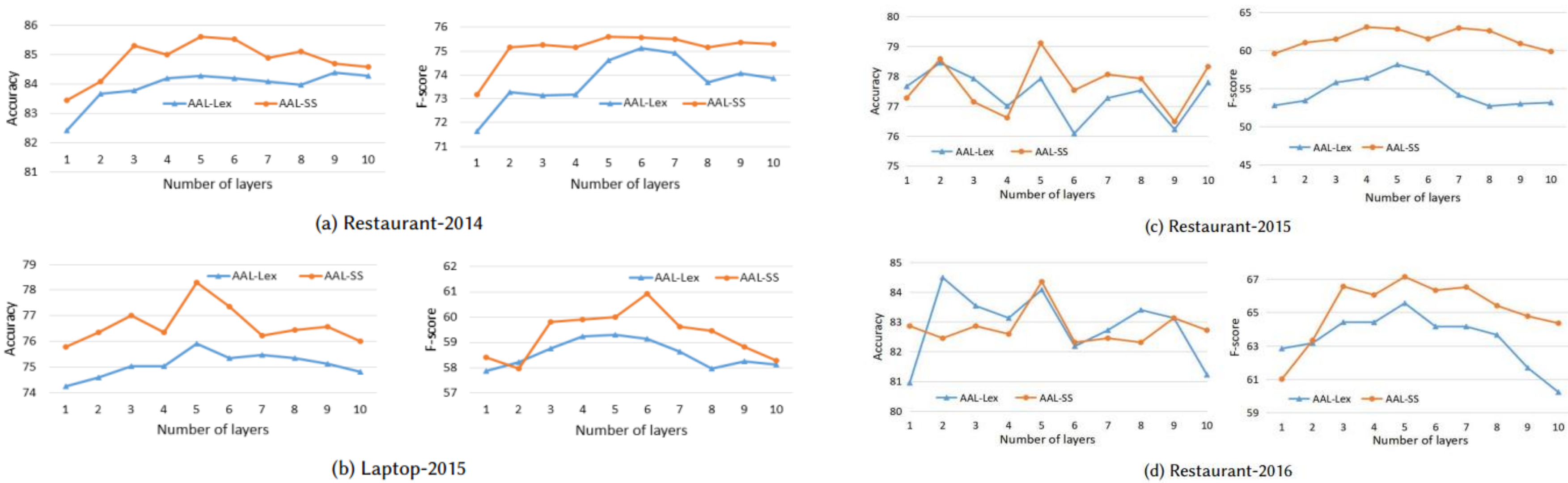


(d) Restaurant-2016

AAL - Experiments (9/11)

Parameter Analysis

► Effects of Layer Number



(a) Restaurant-2014

(b) Laptop-2015

(c) Restaurant-2015

(d) Restaurant-2016

武汉大学 WUHAN UNIVERSITY

AAL - Experiments (10/11)

Case Study

- Aspect Lexicon

Table 4. Aspect Lexicon on Restaurant-2014

Aspect Category	Top-10 aspect terms with the highest PMI
food	tasted, melted, burnt, roast, grilled, oily, shredded, martini, crispy, egg
price	inexpensive, reasonably, $, free, pricey, priced, cost, prices, price, reasonable
service	server, servers, smile, courteous, ignored, greeted, helpful, manager, phone, asked,
ambience	outdoor, paris, sleek, scene, music, cramped, laid-back, romantic, cozy, comfortable,
anecdote/miscellaneous	anniversary, stumbled, reading, based, month, somewhere, celebrate, opened, katz, located

Table 5. Aspect Lexicon on Laptop-2015

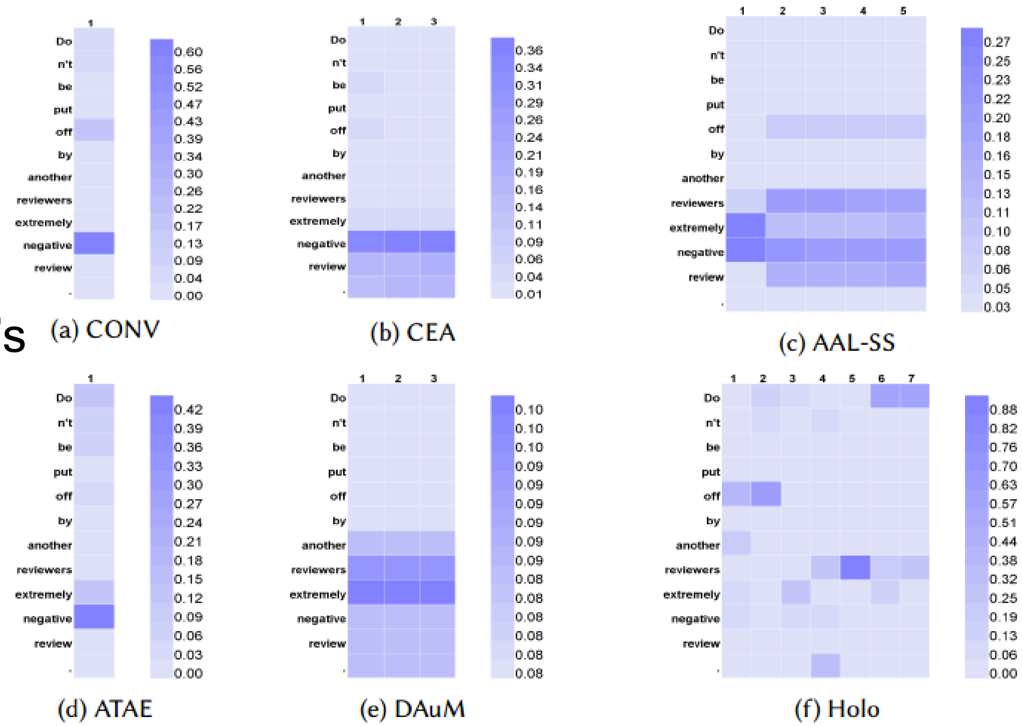
Aspect Category	Top-10 aspect terms with the highest PMI
general	gift, worst, loved, products, pleased, recommended, friends, hate, happy, stars
price	expensive, shipping, paid, price, cost, spent, $, worth, deal, money, fixed
quality	crap, defective, year, customer, told, loud, waited, piece, hot, quiet
operation performance	freezes, hrs, seconds, loads, flawlesslym, freaking, runs, applications, stopped, blue
usability	navigate, curve, learning, friendly, learn, switch, ease, window, user, os
design features	offers, sleek, feature, ram, sized, place, design, features, allow, ports
portability	travel, portability, carry, portable, meets, durable, student, fit, business, sit
connectivity	wifi, network, wireless, ethernet, port, connection, connect, plugged, stay, camera
miscellaneous	media, word, handle, facebook, microsoft, basic, games, gaming, stuff, trial

AAL - Experiments (11/11)

Case Study

- ▶ Attention Visualization

Example:
Do n't be put off by another reviewer's extremely negative reviews!



(a) CONV (b) CEA (c) AAL-SS

(d) ATAE (e) DAuM (f) Holo

Enhanced Aspect Level Sentiment Classification with Auxiliary Memory

Peisong Zhu, Tieyun Qian*
School of Computer Science, WuHan University, China
{zhups24, qty}@whu.edu.cn

Published at COLING 2018
(Research Long Paper)

DAuM - Motivations$_{(1/9)}$

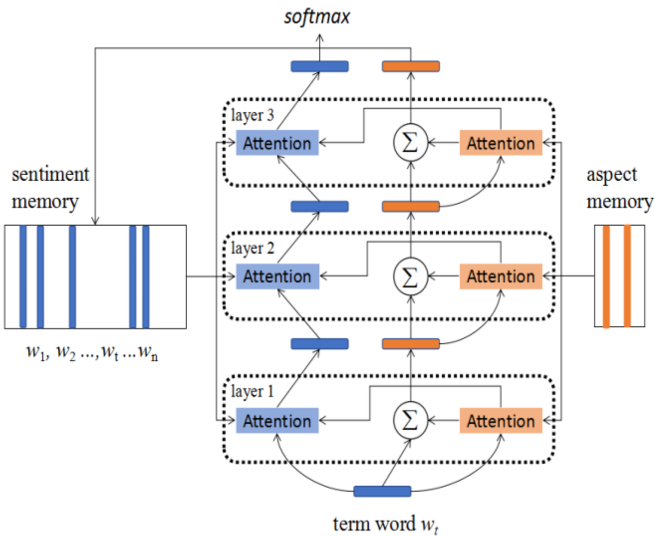
- For term-based ASC, it's hard to train high-quality embeddings for the terms with low frequency.
- For category-based ASC, categories do not explicitly occur in sentences, thus it is hard for a model to capture aspect related context words from the sentence.
- The terms and categories in ASC are closely related to each other. Existing methods fail to utilize the relevance between them because only very few datasets are annotated with both labels.

DAuM - Architecture$_{(2/9)}$

DAuM Overview

Deep memory network with an **Au**xiliary **M**emory

- Auxiliary Memory : Generate corresponding category embeddings for given terms, vice versa.
- Sentiment Memory : Both original and generated embeddings are fed into sentiment memory to collect relative information from context.



Enhanced Aspect Level Sentiment Classification with Auxiliary Memory (COLING18)

DAuM - Architecture$_{(3/9)}$

(1) Auxiliary Memory (e.g. term => category)

- Initialized memory: m_i
- Generate category

$$u_i = m_i \cdot W_a \cdot v_t$$

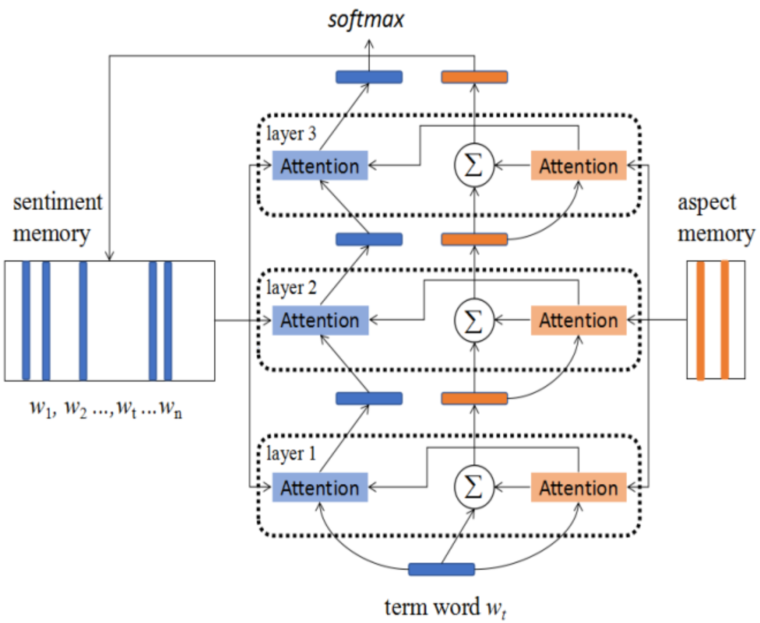
$$\alpha_i = \frac{exp(u_i)}{\sum_{j=1}^{k} exp(u_j)} \qquad o_a = \sum_{i=1}^{k} m_i \alpha_i$$

(2) Sentiment Memory

- Term attention $\beta_i^t = softmax(\tanh(W_s^t[x_i; v_t] + b_s^t))$
- Category attention $\beta_i^a = softmax(\tanh(W_s^a[x_i; o_a] + b_s^a))$
- Aggregation $\beta = (1 - \lambda)\beta^t + \lambda\beta^a$

$$o_s = \sum_{i=1}^{n} x_i \beta_i,$$

stack to multiple layers...



Enhanced Aspect Level Sentiment Classification with Auxiliary Memory (COLING18)

武漢大學 WUHAN UNIVERSITY

DAuM - Architecture$_{(4/9)}$

Loss Function

- Sentiment Prediction

$$P_c(s,w_t) = softmax(W_c o_s^f + b_c) \qquad L_{cla} = -\sum_{(s,w_t)\in T}\sum_{c\in C} y_c(s,w_t)\cdot logP_c(s,w_t)$$

- Semantic Relatedness (term & category) Regularization

$$L_{pre} = \sum_{(s,w_t)\in T}\ \sum_{i\in\{1\ldots n\}/\{t\}} max(0, 1 - o_a^f\cdot W_p\cdot v_t + o_a^f\cdot W_p\cdot x_i)$$

- Category Embedding Regularization

$$L_{reg} = \|M_a M_a^T - I\|$$

$$L_{final} = L_{cla} + \gamma_1 L_{pre} + \gamma_2 L_{reg}$$

DAuM - Experiments (5/9)

Datasets

SemEval 2014, SemEval 2016

Category-based

Term-based

Dataset	Set	Total	Pos.	Neg.	Neu.
Restaurants(ASC)	Train	2927	1806	697	424
	Test	973	657	222	94
	Dev	591	372	142	77
Restaurants(TSC)	Train	3017	1806	669	542
	Test	1120	728	196	196
	Dev	591	358	138	95
Laptops	Train	1934	823	730	381
	Test	638	341	128	169
	Dev	394	171	140	63
TweetNews	Train	2416	627	1163	626
	Test	1249	303	716	230
	Dev	498	125	233	140

武漢大學
WUHAN UNIVERSITY

DAuM - Experiments (6/9)

Results

Term-based

Method	Restaurants(TSC) Accuracy	Precision	Recall	F-score	Laptops Accuracy	Precision	Recall	F-score
ContextAVG	75.09	68.93	62.91	63.96	67.24	65.56	60.96	61.86
LSTM	74.28	68.72	61.89	62.21	66.46	65.04	60.54	61.72
TD-LSTM	75.63	69.18	63.05	64.16	68.18	66.86	61.15	62.28
AE-LSTM	76.25	69.76	63.21	64.32	68.97	67.12	61.32	62.50
ATAE-LSTM	77.23	70.83	63.95	64.95	68.65	66.98	61.18	62.45
MemNN	80.09	72.10	65.68	67.82	72.21	70.82	65.03	66.75
DAuM	**82.32**	**74.68**	**70.18**	**71.45**	**74.45**	**72.96**	**69.21**	**70.16**

Category-based

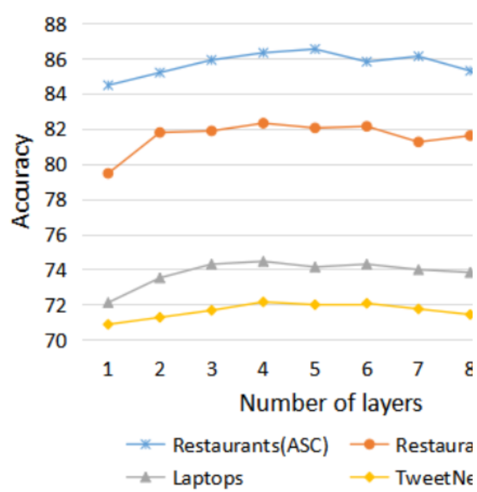
Method	Restaurants(ASC) Accuracy	Precision	Recall	F-score	TweetNews Accuracy	Precision	Recall	F-score
ContextAVG	80.37	72.86	67.58	68.72	65.25	57.52	51.96	52.88
LSTM	82.01	74.20	69.16	70.20	66.86	58.76	53.12	54.32
AE-LSTM	82.53	74.56	69.36	70.48	69.42	60.96	55.18	56.28
ATAE-LSTM	83.98	75.92	70.88	71.76	69.58	61.10	55.45	56.72
TAN	82.53	74.48	69.50	70.55	68.78	60.42	55.06	56.25
MemNN	84.28	76.06	70.24	72.38	70.14	62.21	57.05	58.62
DAuM	**86.33**	**77.54**	**73.82**	**75.16**	**72.14**	**64.52**	**58.96**	**60.24**

Enhanced Aspect Level Sentiment Classification with Auxiliary Memory (COLING18)

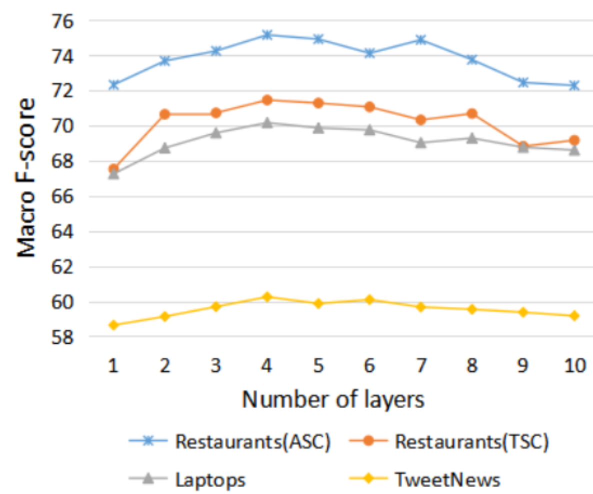
DAuM - Experiments (7/9)

Parameter Analysis

- Effects of Multiple Layers



(a) Accuracy

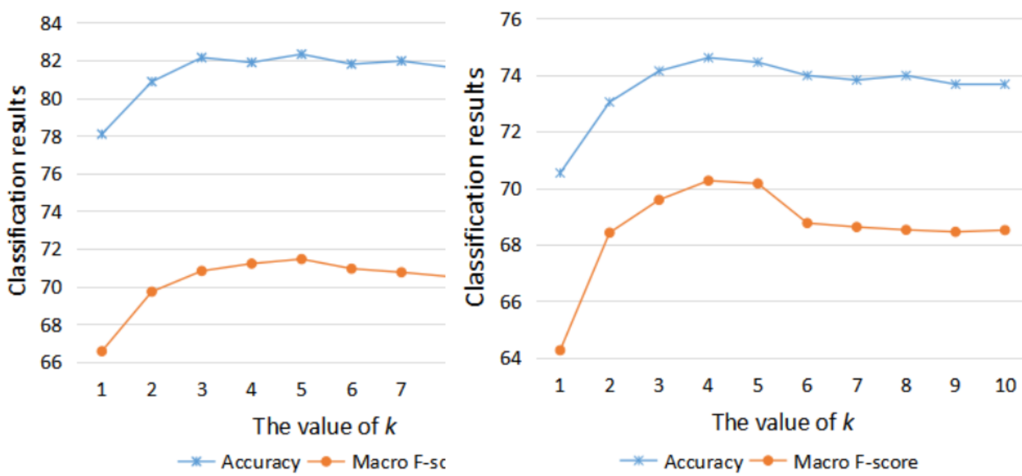


(b) Macro F-score

DAuM - Experiments (8/9)

Parameter Analysis

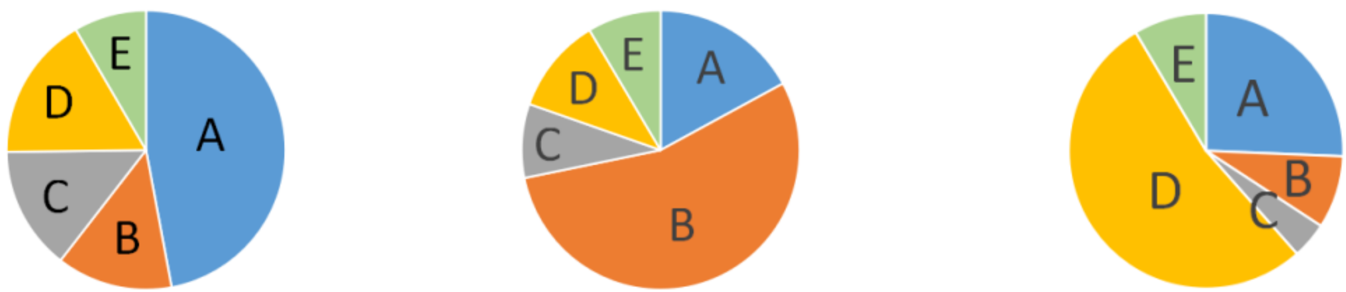
▶ Effects of Aspect Number k



(a) Results on Restaurants(TSC) (b) Results on Laptops

DAuM - Experiments (9/9)

Case Study



(a) Given "*pasta*" from *"food"*

(b) Given "*waitress*" from *"service"*

(c) Given "*price*" from *"price"*

Enhanced Aspect Level Sentiment Classification with Auxiliary Memory (COLING18)

武漢大學
WUHAN UNIVERSITY

Outline

- User Generated Data
- An overview on Sentiment Analysis
- An overview on Recommender Systems
- Our Work

Our Work

- Sentiment analysis
- Recommender systems
- User profiling
- Representation learning

武漢大學
WUHAN UNIVERSITY

Spatiotemporal Representation Learning for Translation-Based POI Recommendation

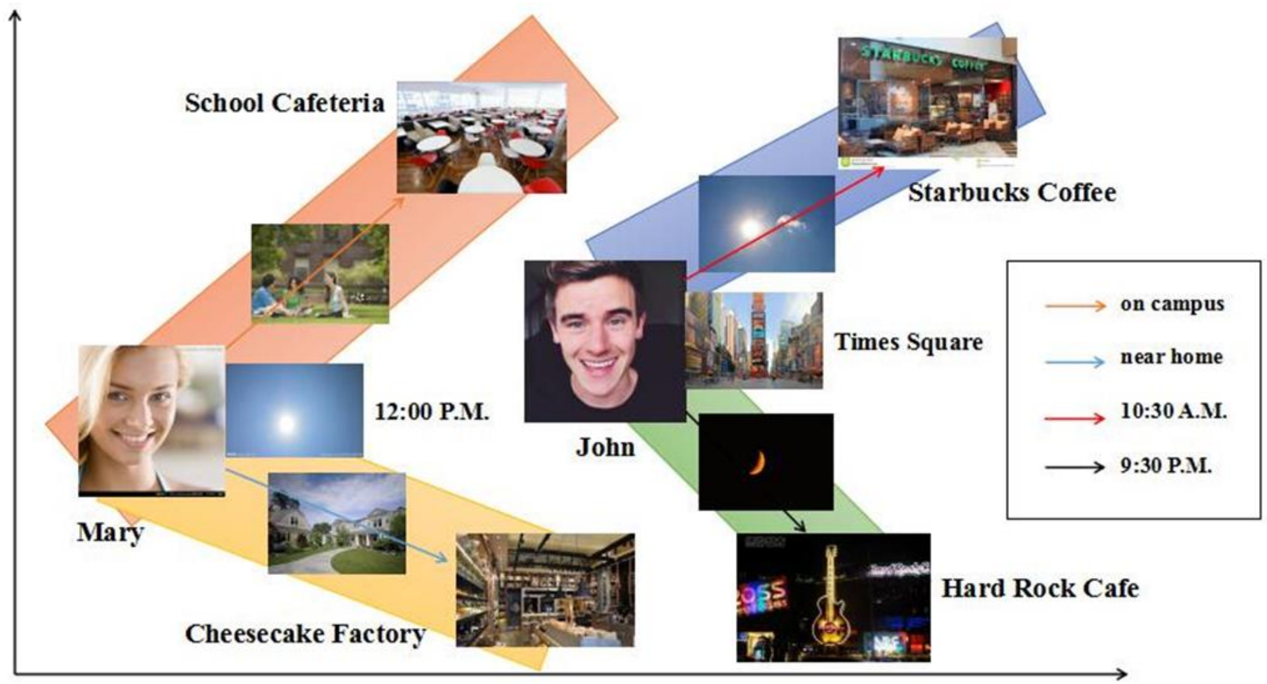
Tieyun Qian[1], Bei Liu[1],
Quoc Viet Hung Nguyen[2], Hongzhi Yin[3],

[1]Wuhan University, [2]Griffith University,
[3]The University of Queensland

Published at TOIS 2019
(Regular Research Paper)

Problems in POI recommendation (1/12)

- Highly spatio- and temporal- sensitive
- Cold start



STL – Main Contributions (2/12)

- We adopt the translation-based knowledge graph embedding techniques to model the spatiotemporal effects in POI recommendation. The joint modeling of spatiotemporal information also distinguishes our work from existing studies which consider these information in a separate way.

- We propose a new type of cold-start problem, cold-start spatiotemporal contexts, and develop effective methods to exploit and integrate various correlation information into the representation learning of cold-start users, items and spatiotemporal contexts to address cold start problems

STL – Proposed Model (3/12)

Basic idea:

We represent the spatio-temporal context < *t* , *l* > as a type of relation such that it captures the user's check-in behaviors at the specific time and location

"Mary + 12:00 P.M. on campus ⇒ School Cafeteria"
"Mary + 12:00 P.M. near home ⇒ Cheesecake Factory"
and
"John + 10:30 A.M. Times Square ⇒ Starbucks Coffee"
"John + 9:30 P.M. Times Square ⇒ Hard Rock Cafe"

A user *u* will reach an interested POI *vq* via a translation edge *tl*

$$\vec{u} + \vec{tl} \approx \vec{v_q}$$

Spatiotemporal Representation Learning for Translation-Based POI Recommendation (TOIS 2019)

STL – Proposed Model (4/12)

Dealing with cold start check-ins:

The cold-start spatial-temporal contexts refer to those new time-location pairs that have never appeared in the training dataset. Almost all individual spatial and temporal contexts are not new although their combinations are cold start.

We leverage the contexts for finding its nearest and farthest neighbors based on their spatial and temporal similarity.

$$sim_t(t_i, t_j) = \frac{\sum_{u \in U} \frac{\vec{a}_i \cdot \vec{a}_j}{||\vec{a}_i||_2 ||\vec{a}_j||_2}}{|U|}, \quad sim_l(l_i, l_j) = a * (d(l_i, l_j))^b,$$

$$sim_{tl}(\tau_i, \tau_j) = \alpha_1 \cdot sim_t(t_i, t_j) + \alpha_2 \cdot sim_l(l_i, l_j),$$

STL – Proposed Model (5/12)

Dealing with cold start users

We propose two geo-social correlation measures to incorporate a user's social and geographical contexts.

$$sim_u(u_q, u_k) = \lambda_1 \cdot sim_{soc}(u_q, u_k) + \lambda_2 \cdot sim_{geo}(u_q, u_k)$$

We employ the normalized ratio of common friends in two users' social circles as the similarity metric of social influence.

$$sim_{soc}(u_q, u_k) = \frac{|N(u_q) \bigcap N(u_k)| + 1}{\sum_{k=1}^{k=|N(u_q)|}(|N(u_q) \bigcap N(u_k)| + 1)},$$

The geographical similarity simgeo is similar to the definition in *siml.*

STL – Proposed Model (6/12)

Dealing with cold start POIs:

We propose two geo-semantic correlation measures to incorporate the semantic and geographical contexts of a POI.

$$sim_p(p_q, p_k) = \gamma_1 \cdot sim_{sem}(p_q, p_k) + \gamma_2 \cdot sim_{geo}(p_q, p_k)$$

The semantic similarity *simsem* is defined as the Jaccard coefficient between the tag set *T* (*pq*) and *T* (*pk*) for POI *pq* and *pk*

$$sim_{sem}(p_q, p_k) = \frac{|T(p_q) \bigcap T(p_k)|}{|T(p_q) \bigcup T(p_k)|}$$

STL – Evaluation (7/12)

We evaluate our methods on two real-life LBSN datasets: Foursquare and Gowalla

Table 2. Statistics of Two Datasets

	Foursquare	Gowalla
# of users	114,508	107,092
# of POIs	62,462	1,280,969
# of check-ins	1,434,668	6,442,892
#std time slots	24	24
# of locations	5,846	200
# of $< t, l >$ contexts	28,868	3,636

STL – Evaluation (8/12)

We compare our STA with 10 POI recommendation. They represent the state-of-the-art methods:

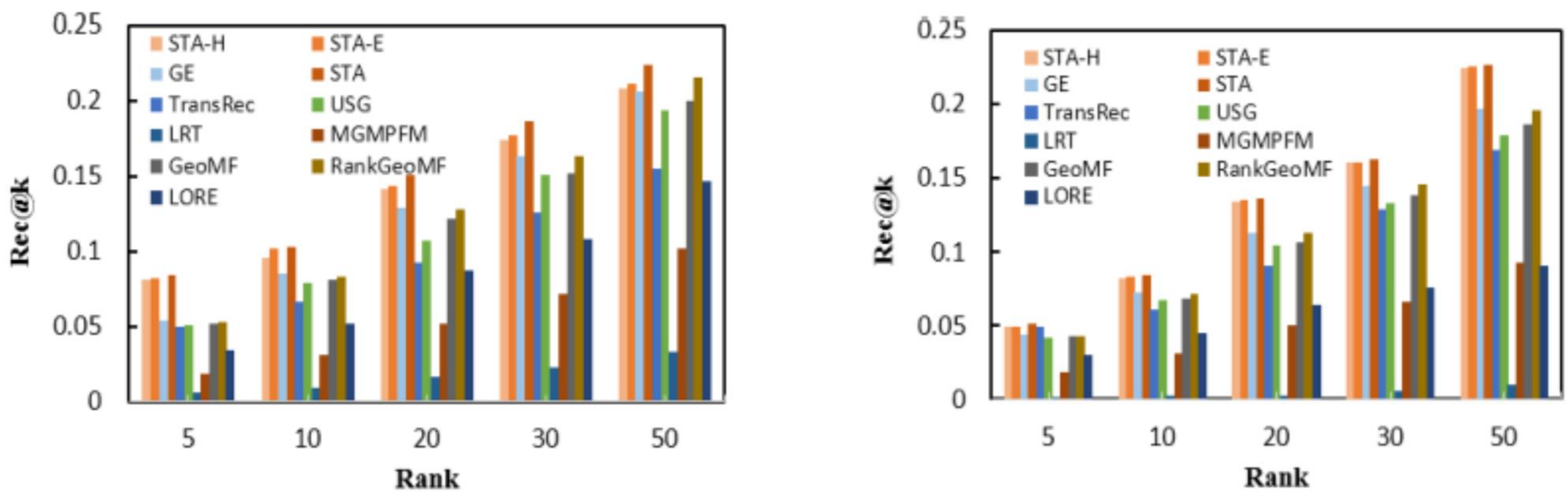
firstly, they cover four types of popular recommendation techniques, i.e., collaborative filtering, matrix factorization, distributed representation, and hybrid model;

secondly, they consider six important factors that influence user decision-making for choosing POIs, including user preference, temporal, geographical, social, content, and sequential influence.

WUHAN UNIVERSITY

STL – Evaluation (9/12)

Comparison Results



(a) Comparision on Foursquare Dataset

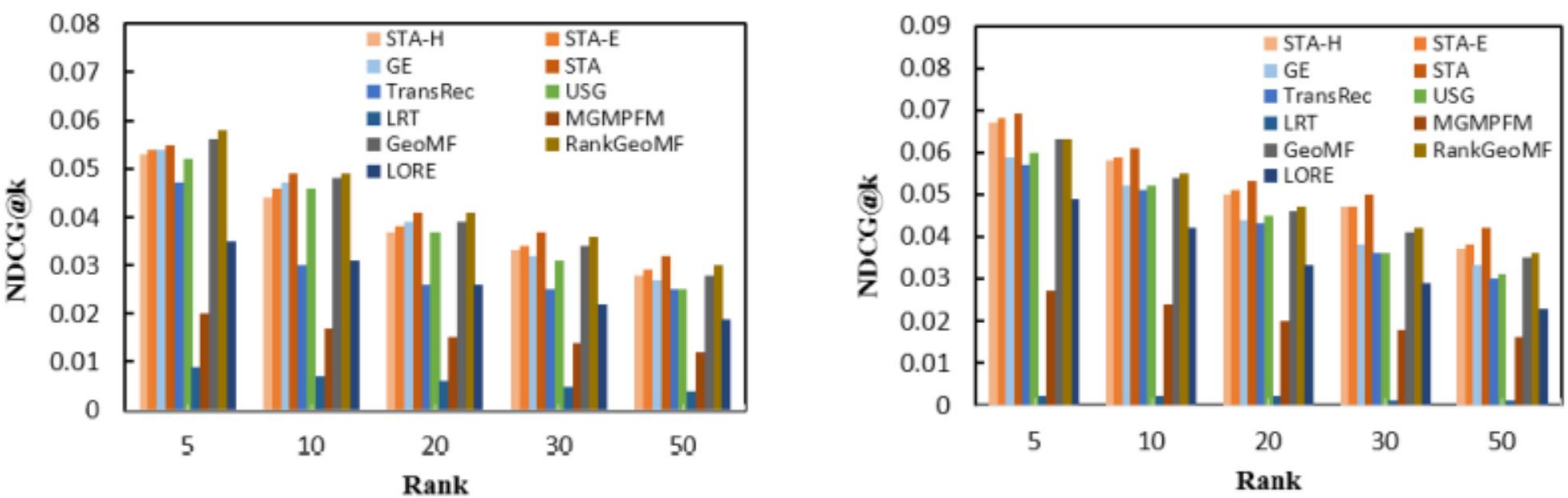
(b) Comparision on Gowalla Dataset

Fig. 2. Comparisons with the baselines on the two datasets in terms of Rec@K.

Spatiotemporal Representation Learning for Translation-Based POI Recommendation (TOIS 2019)

STL – Evaluation (10/12)

Comparison Results



(a) Comparision on Foursquare Dataset

(b) Comparision on Gowalla Dataset

Fig. 3. Comparisons with the baselines on the two datasets in terms of NDCG@K.

Spatiotemporal Representation Learning for Translation-Based POI Recommendation (TOIS 2019)

武漢大學 WUHAN UNIVERSITY

STL – Evaluation (11/12)

Sensitivity to data sparsity

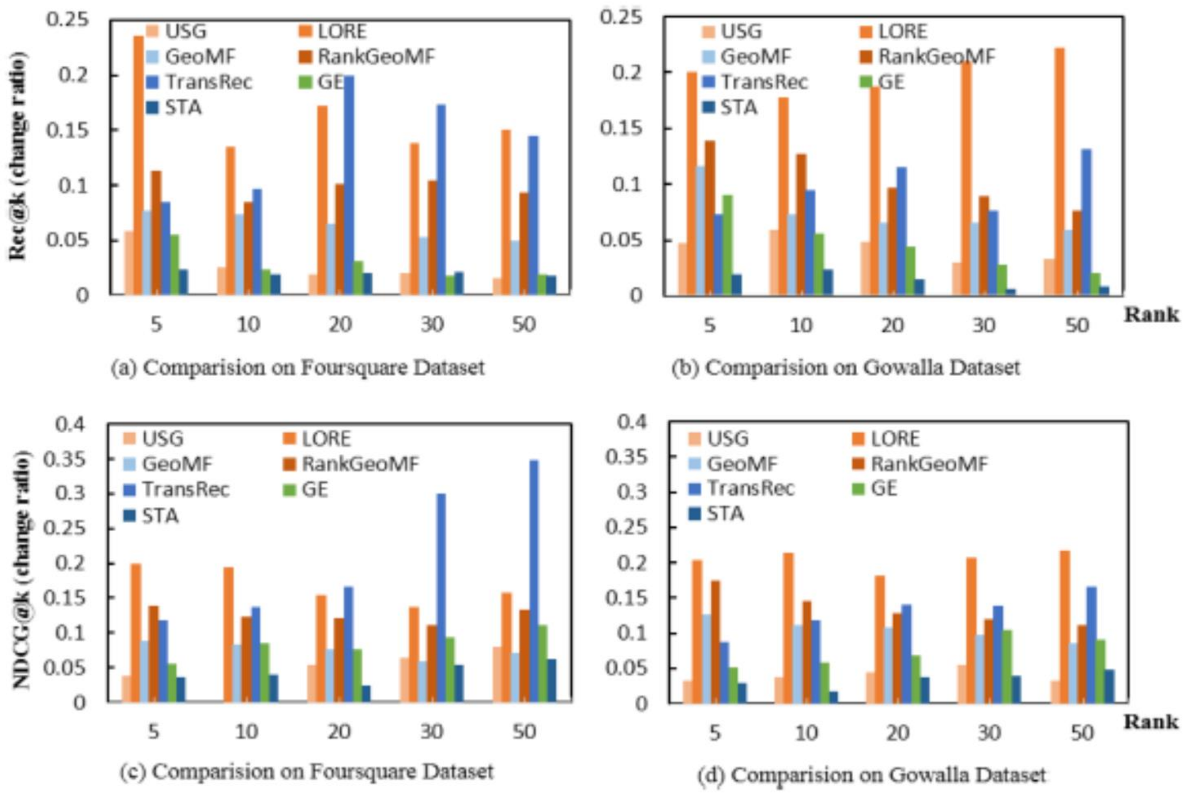


Fig. 4. Changes with the reduction of training data (smaller change ratios are better).

STL – Evaluation (12/12)

Results in cold-start scenario

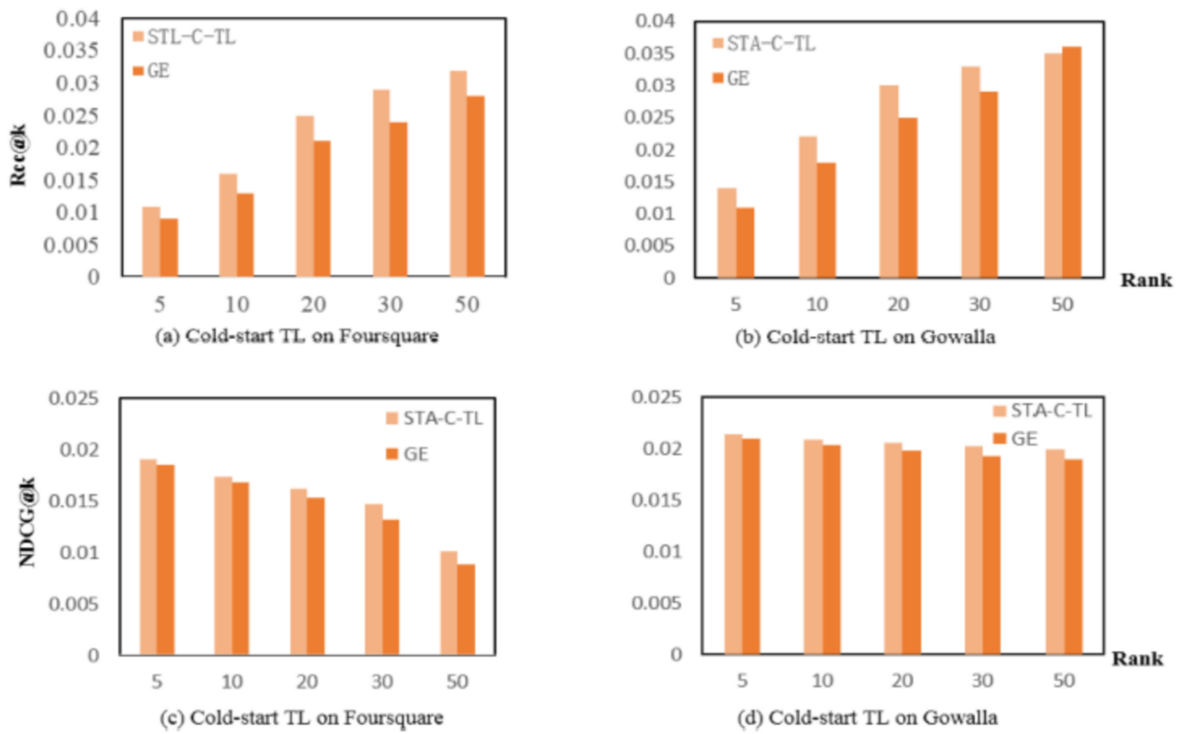


Fig. 5. Test for cold-start spatiotemporal contexts.

Spatiotemporal Representation Learning for Translation-Based POI Recommendation (TOIS 2019)

What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation

Ke Sun[1], Tieyun Qian[1], Hongzhi Yin[2], Tong Chen[2], Yiqi Chen[1], Ling Chen[3]

[1]Wuhan University , [2]University of Queensland, [3]University of Technology, Sydney

Published at CIKM 2019
(Research Long Paper)

Background

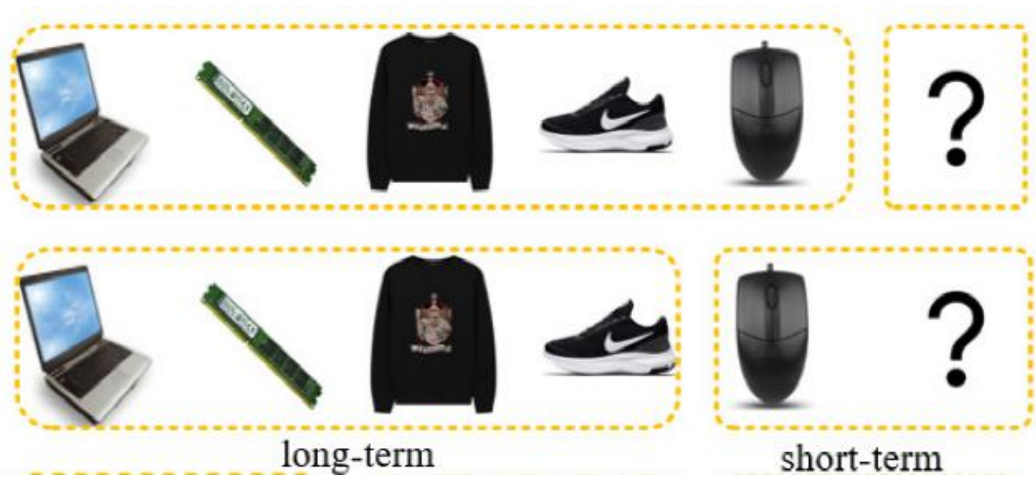
- Sequential recommendation
  - ➢ Modelling sequential dependencies of user-item interactions.
- Session-based recommendation
  - ➢ A subtask of sequential recommendation.
  - ➢ User transaction sequence is partitioned into sessions.

What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation (CIKM 2019)

Motivation

- Limitations of existing sequential methods:
  - ➢ Ignore long-term user preferences. (Fig. a)
  - ➢ Consider all historical sessions without any distinction. (Fig. b)

a. b.



What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation (CIKM 2019)

Our Contributions

- We propose a novel deep learning based sequential recommender framework for session-based recommendation by integrating both long-term and short-term user preferences in a unified way.

What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation (CIKM 2019)

Problem and Preliminary

- Problem Definition: Given a historical session sequence $\{S_1, \ldots, S_{n-1}\}$ and current session $S_n = \{i_1, \ldots, i_{t-1}\}$ of u, predict i_t.
- Preliminary: nonlocal structure

$$\mathbf{y}_i = \frac{1}{\boxed{C(X)}} \sum_{\forall j} \boxed{f(\mathbf{x}_i, \mathbf{x}_j)} g(\mathbf{x}_j),$$

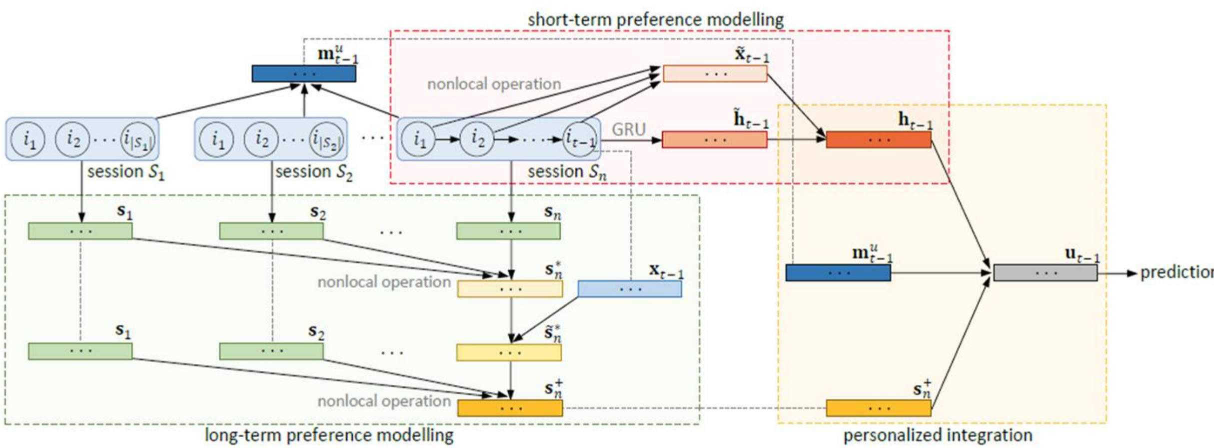
normalization factor

similarity scalar

What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation (CIKM 2019)

Architecture

- We design a two-layer nonlocal architecture to learn long-term user preferences from relevant historical sessions
- We integrate the nonlocal structure with a gated recurrent unit (GRU) to learn short-term user preferences.

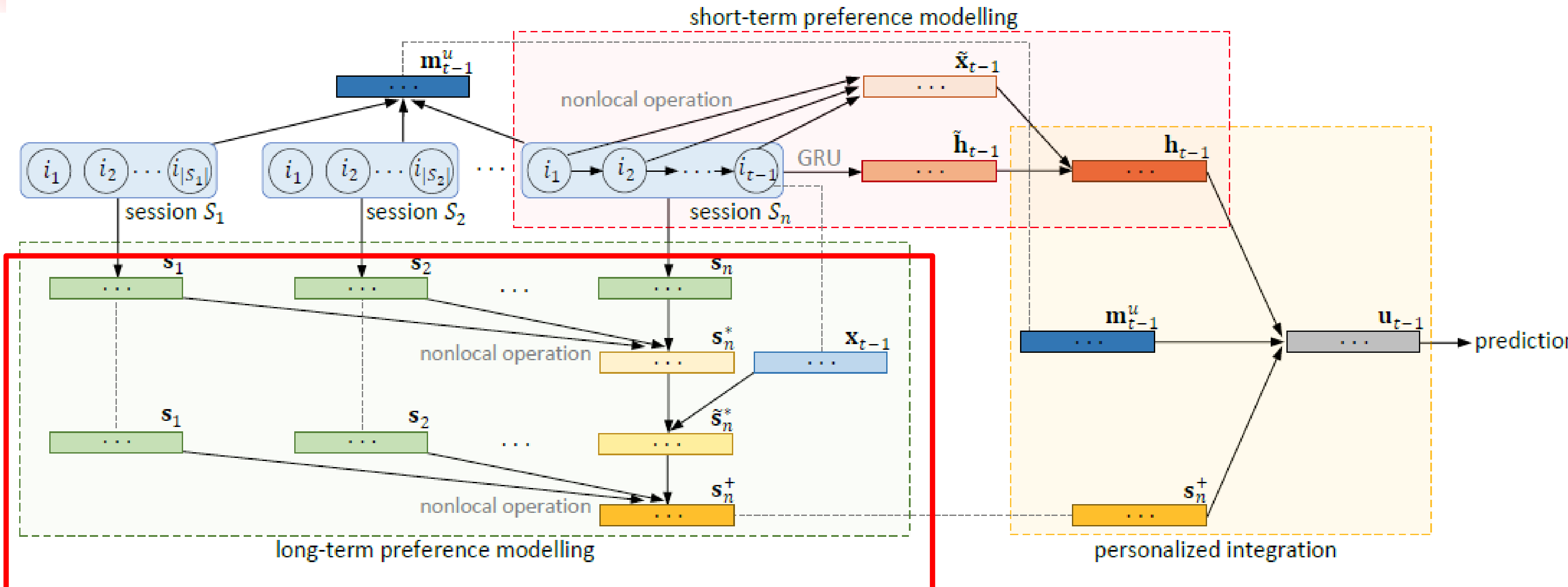


What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation (CIKM 2019)

Proposed Method

武漢大學 WUHAN UNIVERSITY

- Long-term preference modeling
  - ➢ two-layer nonlocal neural network
    - ✓ session-level: $\mathbf{s}_n^* = \frac{1}{\sum_{j=1}^{n-1} f(\mathbf{s}_n, \mathbf{s}_j)} \sum_{j=1}^{n-1} f(\mathbf{s}_n, \mathbf{s}_j) g(\mathbf{s}_j), \quad \mathbf{s}_m = \frac{1}{|S_m|} \sum_{t=1}^{|S_m|} \mathbf{x}_t,$



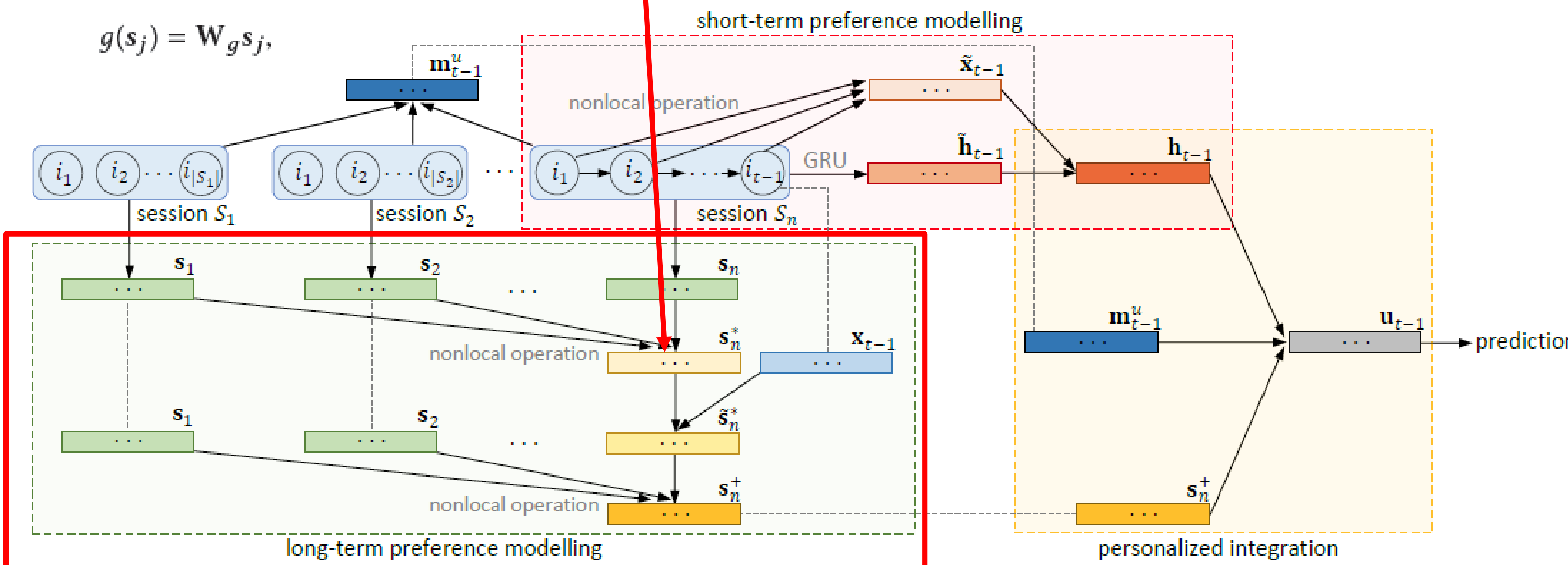
Proposed Method

Long-term preference modeling

- two-layer nonlocal neural network
  - ✓ session-level: $\boxed{\mathbf{s}_n^*} = \frac{1}{\sum_{j=1}^{n-1} f(\mathbf{s}_n, \mathbf{s}_j)} \sum_{j=1}^{n-1} f(\mathbf{s}_n, \mathbf{s}_j) g(\mathbf{s}_j), \qquad \mathbf{s}_m = \frac{1}{|S_m|} \sum_{t=1}^{|S_m|} \mathbf{x}_t,$

$$f(\mathbf{s}_n, \mathbf{s}_j) = e^{(\mathbf{W}_\theta \mathbf{s}_j)^T \mathbf{W}_\phi \mathbf{s}_n},$$

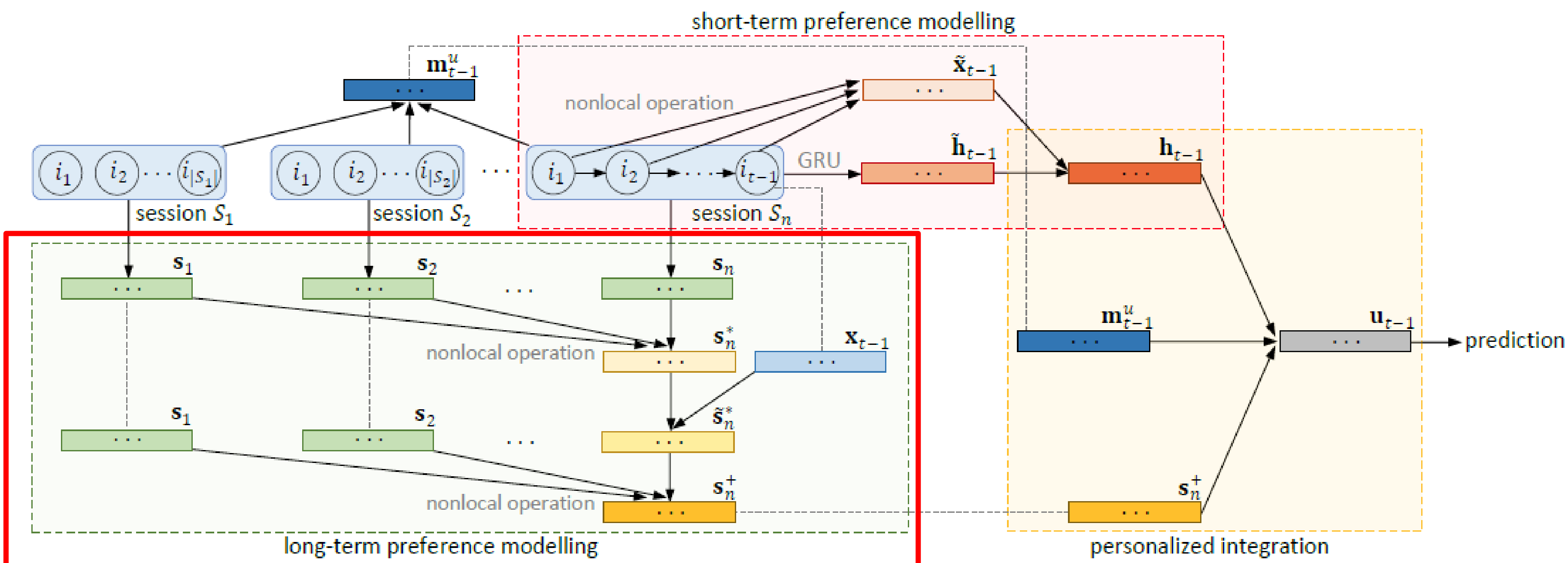
$$g(\mathbf{s}_j) = \mathbf{W}_g \mathbf{s}_j,$$



Proposed Method

Long-term preference modeling

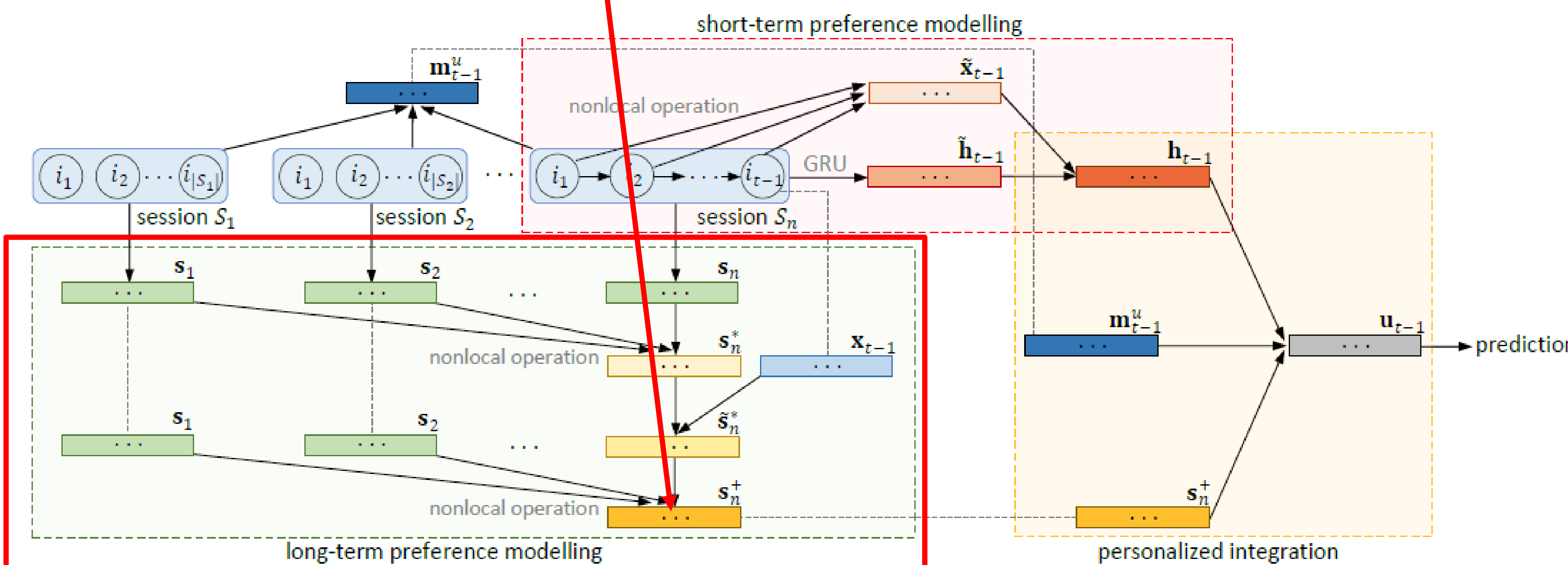
- two-layer nonlocal neural network
  - item-level: $\mathbf{s}_n^+ = \frac{1}{\sum_{j=1}^{n-1} f'(\tilde{\mathbf{s}}_n^*, \mathbf{s}_j)} \sum_{j=1}^{n-1} f'(\tilde{\mathbf{s}}_n^*, \mathbf{s}_j) g'(\mathbf{s}_j), \quad \tilde{\mathbf{s}}_n^* = \mathbf{s}_n^* + \mathbf{x}_{t-1},$



Proposed Method

Long-term preference modeling

- two-layer nonlocal neural network
  - ✓ item-level: $\boxed{\mathbf{s}_n^+} = \frac{1}{\sum_{j=1}^{n-1} f'(\tilde{\mathbf{s}}_n^*, \mathbf{s}_j)} \sum_{j=1}^{n-1} f'(\tilde{\mathbf{s}}_n^*, \mathbf{s}_j) g'(\mathbf{s}_j), \quad \tilde{\mathbf{s}}_n^* = \mathbf{s}_n^* + \mathbf{x}_{t-1},$

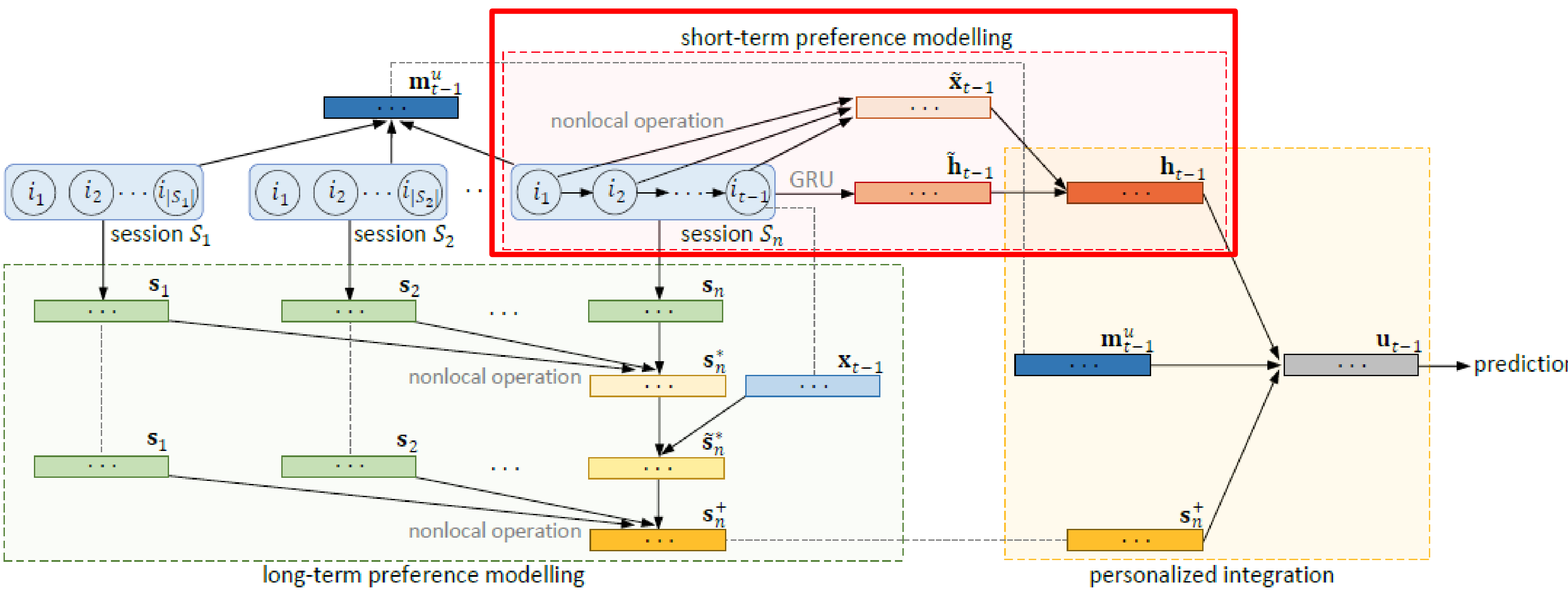


Proposed Method

Short-term preference modeling

- GRU and nonlocal neural network

$$\tilde{\mathbf{h}}_{t-1} = GRU(\mathbf{x}_{t-1}, \tilde{\mathbf{h}}_{t-2}), \qquad \tilde{\mathbf{x}}_{t-1} = \frac{1}{\sum_{j=1}^{t-2} f_x(\mathbf{x}_{t-1}, \mathbf{x}_j)} \sum_{j=1}^{t-2} f_x(\mathbf{x}_{t-1}, \mathbf{x}_j) g_x(\mathbf{x}_j),$$

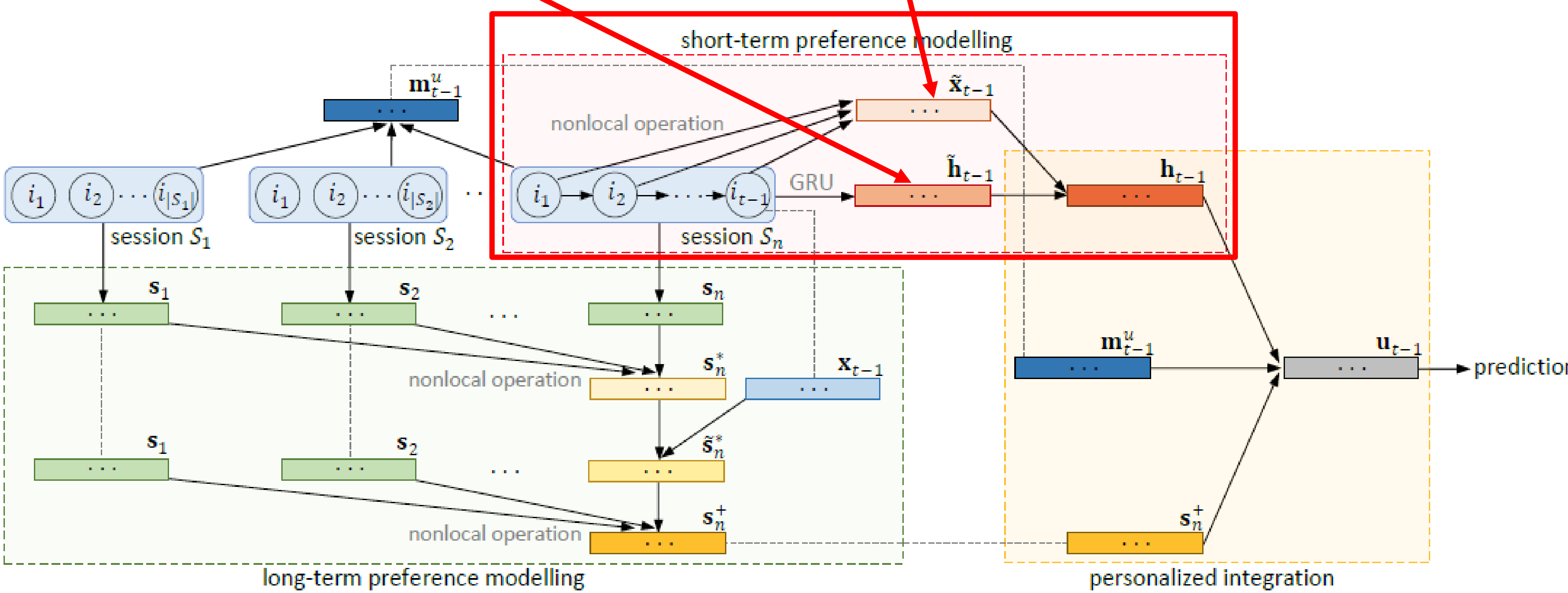


Proposed Method

Short-term preference modeling

- GRU and nonlocal neural network

$$\boxed{\tilde{\mathbf{h}}_{t-1}} = GRU(\mathbf{x}_{t-1}, \tilde{\mathbf{h}}_{t-2}), \qquad \boxed{\tilde{\mathbf{x}}_{t-1}} = \frac{1}{\sum_{j=1}^{t-2} f_x(\mathbf{x}_{t-1}, \mathbf{x}_j)} \sum_{j=1}^{t-2} f_x(\mathbf{x}_{t-1}, \mathbf{x}_j) g_x(\mathbf{x}_j),$$



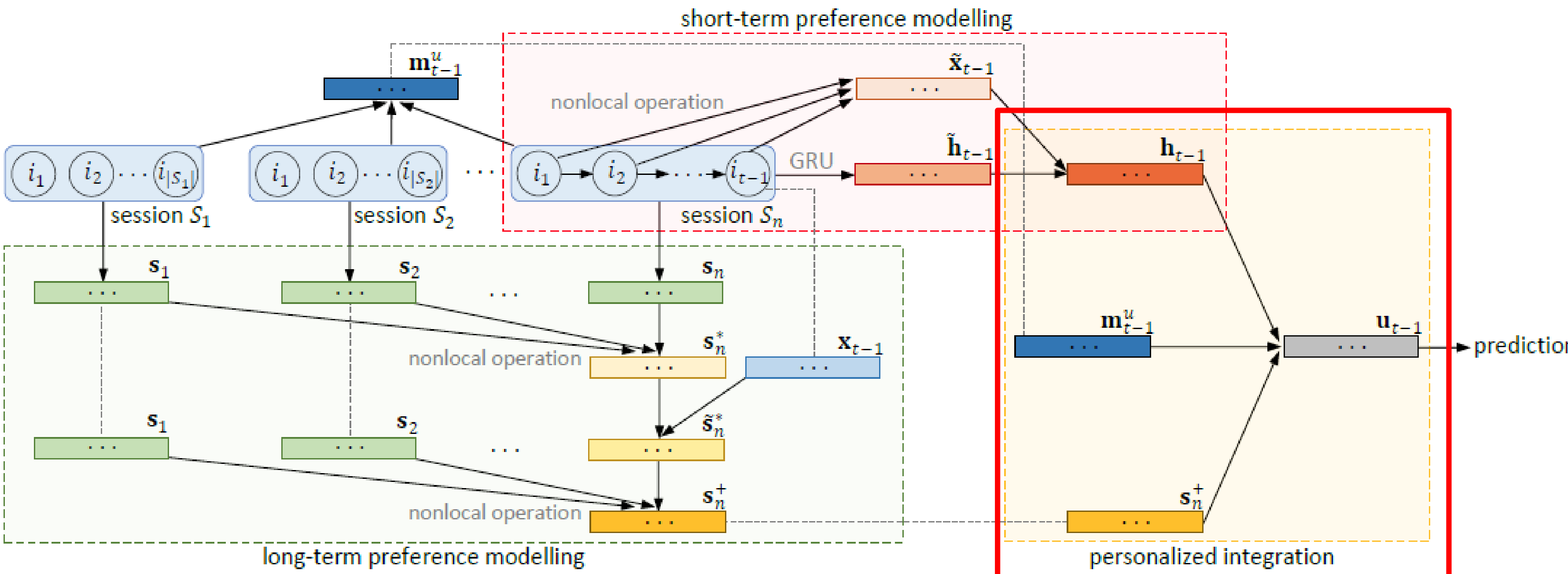
Proposed Method

Personalized Integration

- historical representation
- current representation

$$\mathbf{m}_{t-1}^{h} = \frac{1}{|I^h|}\sum_{j=1}^{|I^h|}\mathbf{x}_j$$

$$\mathbf{m}_{t-1}^{c} = \frac{1}{|I^c|}\sum_{j=1}^{|I^c|}\mathbf{x}_j$$

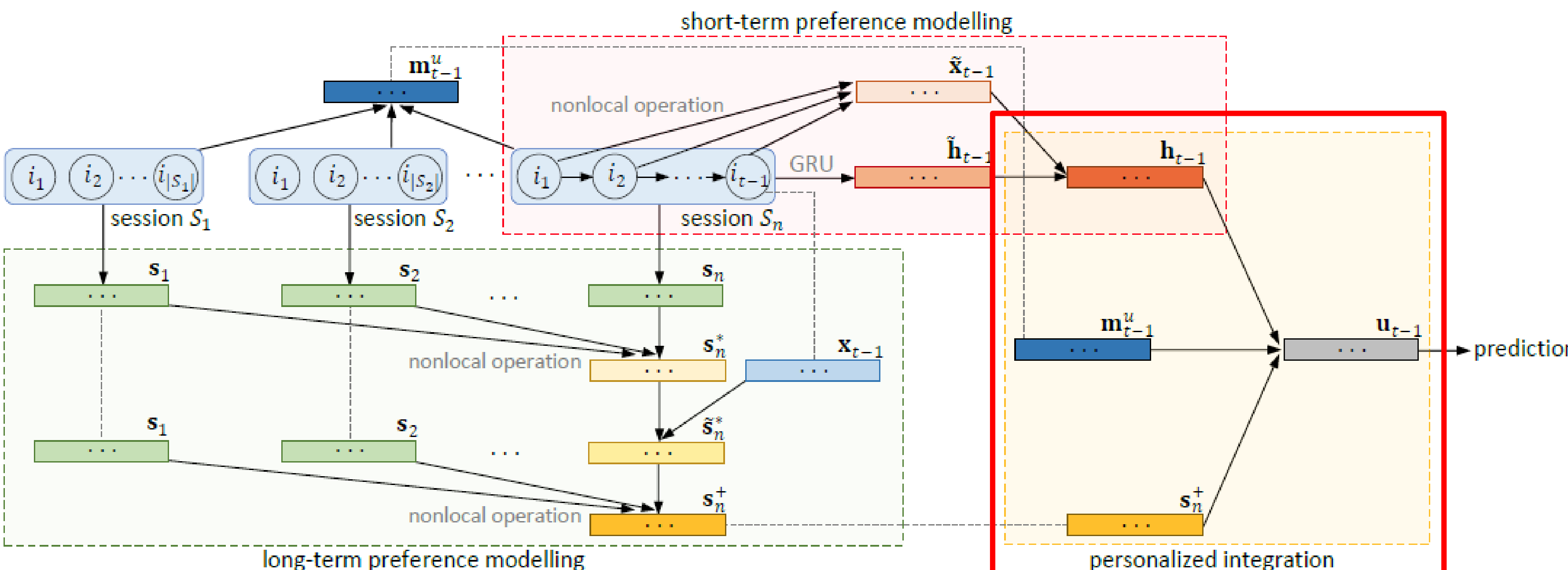


Proposed Method

Personalized Integration

- overall representation

$$\mathbf{m}^u_{t-1} = \lambda \mathbf{m}^h_{t-1} + (1-\lambda)\mathbf{m}^c_{t-1},$$

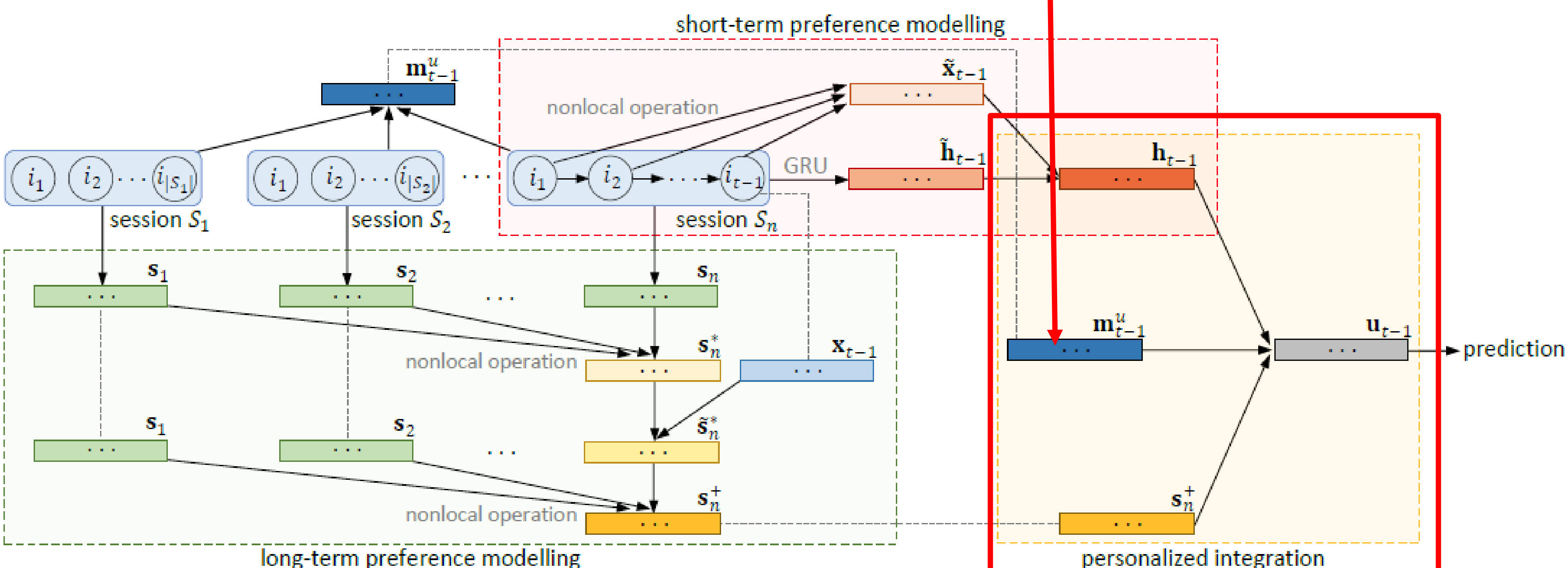


Proposed Method

Personalized Integration

- overall representation

$$\boxed{\mathbf{m}_{t-1}^{u}} = \lambda \mathbf{m}_{t-1}^{h} + (1-\lambda)\mathbf{m}_{t-1}^{c},$$

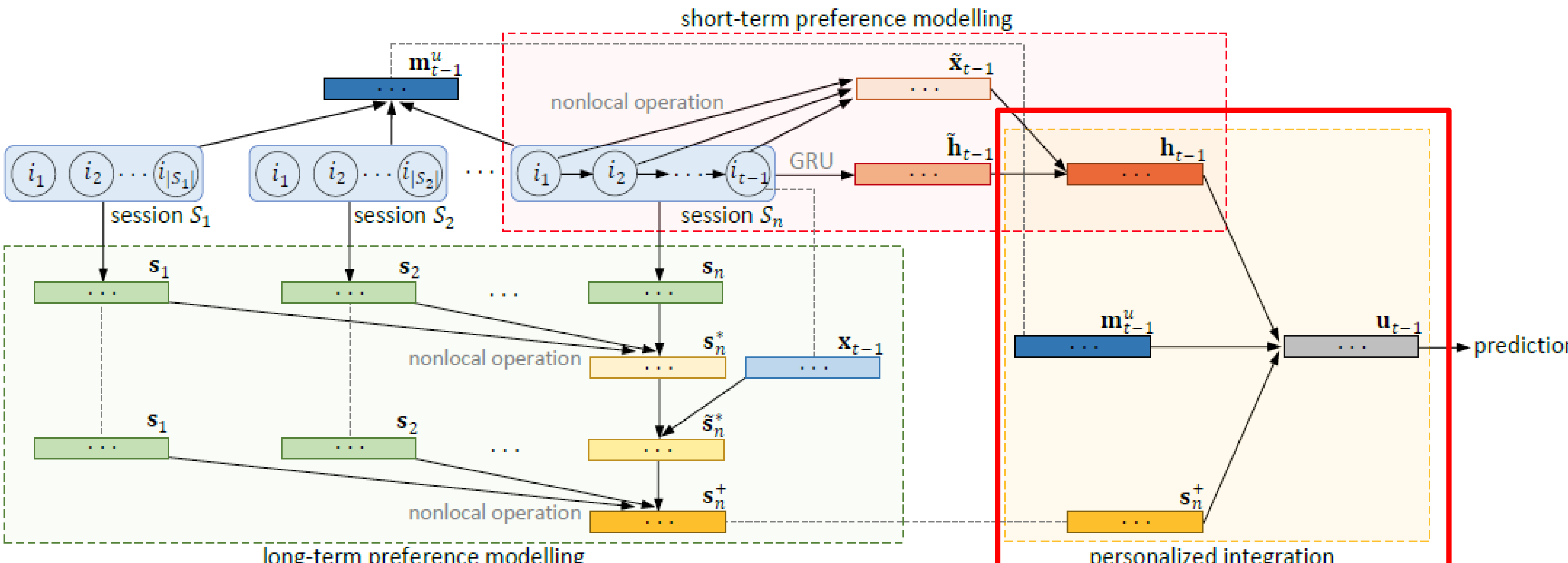


Proposed Method

Prediction

- personalized user preference

$$\mathbf{u}_{t-1} = \frac{1}{2}\left(\mathbf{W}_h(\mathbf{m}^u_{t-1} \oplus \mathbf{s}^+_n) + \mathbf{W}_c(\mathbf{m}^u_{t-1} \oplus \mathbf{h}_{t-1})\right),$$

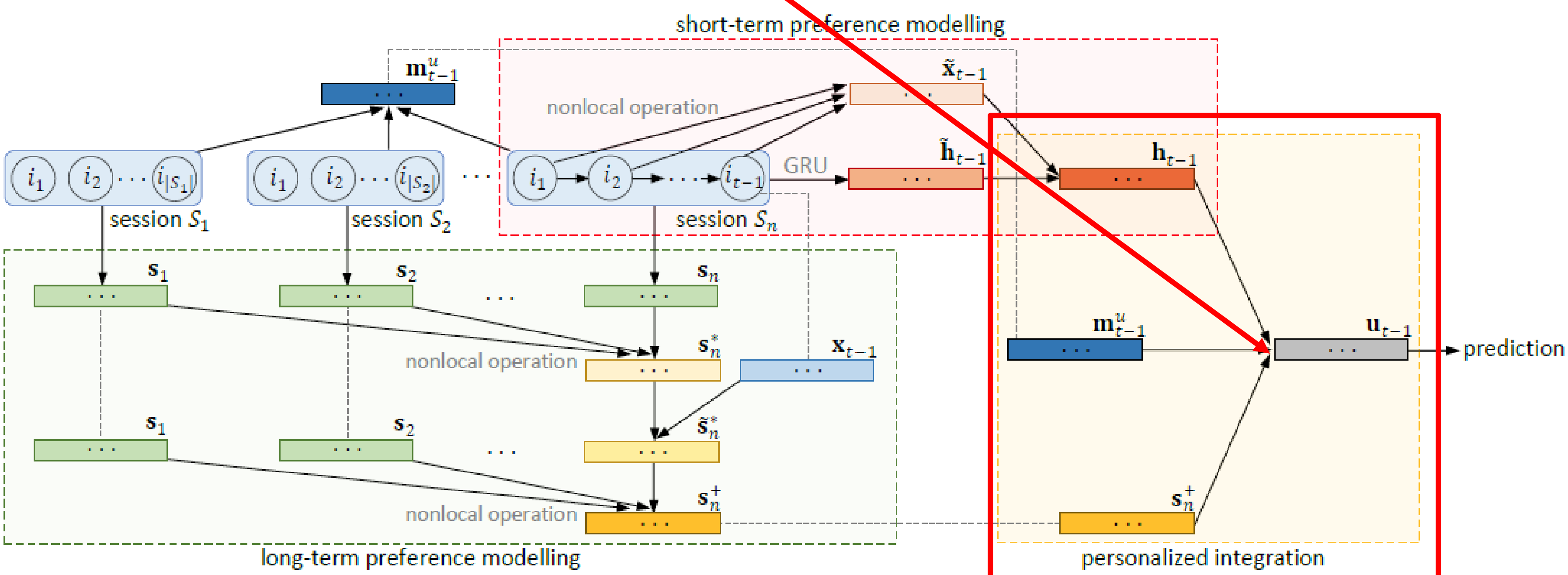


Proposed Method

Prediction

- personalized user preference

$$\boxed{\mathbf{u}_{t-1}} = \frac{1}{2}\Big(\mathbf{W}_h(\mathbf{m}^u_{t-1} \oplus \mathbf{s}^+_n) + \mathbf{W}_c(\mathbf{m}^u_{t-1} \oplus \mathbf{h}_{t-1})\Big).$$



Evaluation

- Dataset
  - Tmall (an E-commerce dataset)
  - Gowalla (a Point-Of-Interest recommendation dataset)
- Evaluation
  - Recall
  - MRR

Statistic	Tmall	Gowalla
#user	36,595	14,898
#item	26,576	15,291
avg. session length	2.903	3.267
#train session	105,560	166,683
#test session	35,613	14,892

What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation (CIKM 2019)

Evaluation

武漢大學
WUHAN UNIVERSITY

- Baseline methods
  - POP
  - Fossil [He et al, ICDM2016]
  - GRU4Rec [Hidasi et al, ICLR2015]
  - NARM [Li et al, CIKM2017]
  - STAMP [Liu et al, KDD2018]
  - HRNN [Quadrana et al, RecSys2017]
  - SHAN [Ying et al, IJCAI2018]
  - BINN [Li et al, KDD2018]

Consider short-term preference only

Consider both long and short-term preferences

Evaluation

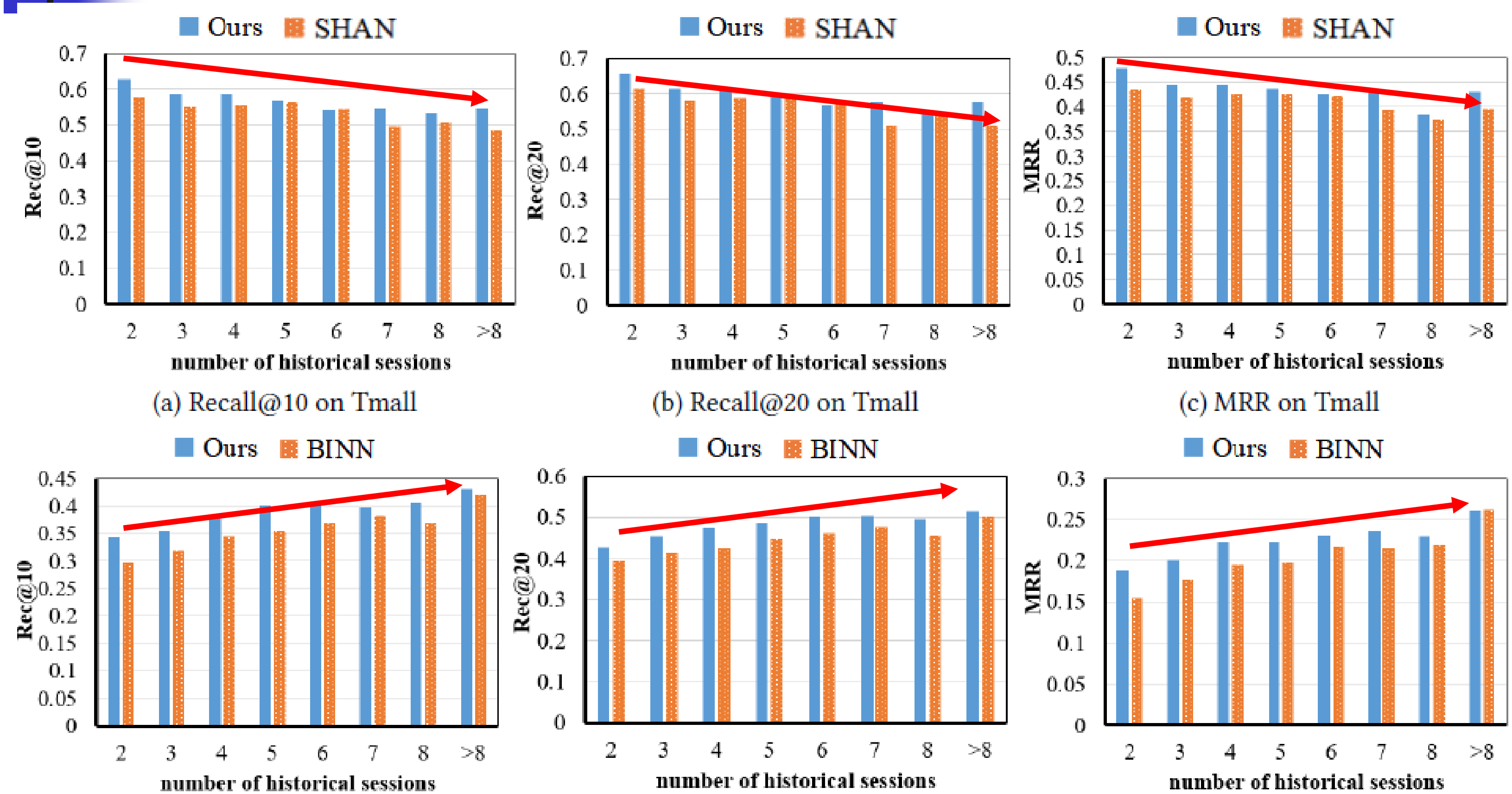
- Overall Performance
  - ➢ Short-term preference is more important in Tmall dataset.
  - ➢ Long-term preference is more important in Gowalla dataset.

Method	Tmall			Gowalla		
	Rec@10	Rec@20	MRR	Rec@10	Rec@20	MRR
POP	0.0213	0.0339	0.0085	0.0424	0.0638	0.0155
Fossil	0.1251	0.1523	0.0647	0.0894	0.1189	0.0342
HRNN	0.5200	0.5520	0.3973	0.1615	0.2015	0.0853
GRU4Rec	0.5736	0.6014	0.4601	0.3183	0.3820	0.1889
NARM	0.6117	0.6450	0.4683	0.3607	0.4441	0.2003
STAMP	0.6112	0.6442	0.4643	0.3395	0.4189	0.1805
SHAN	0.5834	0.6123	0.4409	0.3421	0.4214	0.1942
BINN	0.5312	0.5775	0.3896	0.3679	0.4549	0.2146
Ours	**0.6228**	**0.6530**	**0.4757**	**0.3954**	**0.4844**	**0.2301**

What Can History Tell Us? Identifying Relevant Sessions
for Next-Item Recommendation (CIKM 2019)

Evaluation

Impact of Different User History Lengths



(a) Recall@10 on Tmall (b) Recall@20 on Tmall (c) MRR on Tmall

(d) Recall@10 on Gowalla (e) Recall@20 on Gowalla (f) MRR on Gowalla

Evaluation

Analysis on Different Model Components

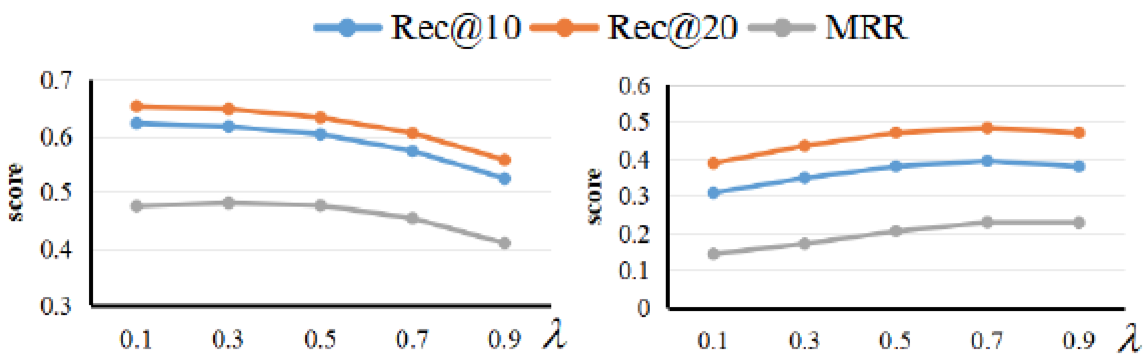
- Remove-LT: No long-term preference
- Remove-ST: No short-term preference
- Remove-PI: No personalized integration

Dataset	Method	Rec@10	Rec@20	MRR
Tmall	Default	0.6228	0.6530	0.4757
	Remove-LT	0.5827↓	0.6099↓	0.4704↓
	Remove-ST	0.5858↓	0.6135↓	0.4711↓
	Remove-PI	0.5768↓	0.6116↓	0.4362↓
Gowalla	Default	0.3954	0.4844	0.2301
	Remove-LT	0.3618↓	0.4470↓	0.2178↓
	Remove-ST	0.3813↓	0.4666↓	0.2252↓
	Remove-PI	0.3404↓	0.4255↓	0.1802↓

What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation (CIKM 2019)

Evaluation

- Analysis on Hyperparameter λ
  - Tmall: performance decreases when the value of λ grows.
  - Gowalla: performance increases when the value of λ grows.



(a) Tmall (b) Gowalla

What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation (CIKM 2019)

Conclusion

- We design a two-layer nonlocal neural network to precisely capture user's long-term preferences.
- We deploy the GRU network coupled with a nonlocal structure to model short-term preferences.
- We present a personalized strategy to adaptively combine the learned long and short-term preferences.

Our work on sentiment analysis

Zhuang Chen, Tieyun Qian. Transfer Capsule Network for Aspect Level Sentiment Classification. ACL, pp 547-556, 2019.

Peisong Zhu, Zhuang Chen, Haojie Zheng, Tieyun Qian. Aspect Aware Learning for Aspect Category Sentiment Analysis. ACM Transactions on Knowledge Discovery from Data (TKDD), 13 (6) , 2019.

Peisong Zhu,Tieyun Qian. Enhanced Aspect Level Sentiment Classification with Auxiliary Memory. COLING. 2018, pp 1077-1088.

Yunkai Yang, Tieyun Qian, Zhuang Chen. Aspect-Level Sentiment Classification with Dependency Rules and Dual Attention. ICONIP (2) 2019: 643-655.

Our work on recommender systems

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

Tieyun Qian, Bei Liu, Quoc Viet Hung Nguyen, and Hongzhi Yin. 2019. Spatiotemporal Representation Learning for Translation-Based POI Recommendation. ACM Trans. Inf. Syst. (TOIS) 37, 2, Article 18, 2019

Ke Sun, Tieyun Qian, Hongzhi Yin, Tong Chen, Yiqi Chen, Ling Chen. What Can History Tell Us? Identifying Relevant Sessions for Next-Item Recommendation. CIKM, 2019: 1593-1602.

Yile Liang, Tieyun Qian, Huilin Yu. Align Reviews with Topics for Rating Prediction. DASFAA 2019, DASFAA (3) 2019: 249-253

Sourcecode available at:

https://github.com/NLPWM-WHU

Thank you!

Any questions?