ACL.IJCAI.SIGIR(AIS)

2018论文预讲会

Constructing Narrative Event Evolutionary Graph for Script Event Prediction

Zhongyang Li, Xiao Ding and Ting Liu
Research Center for Social Computing and Information Retrieval
Harbin Institute of Technology
{zyli, xding, tliu}@ir.hit.edu.cn

1. Task

Characters

X=**Customer** Y=**Waiter**

Event Context

walk(X, restaurant), seat(X), read(X, menu), order(X, food), serve(Y, X, food), eat(X, food), make(X, payment), ____

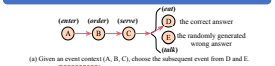
c_1: signed(X, scholarship)
c_2: drive(X, mile)
c_3: spent(X, time)
c_4: recalled(X, designer)
c_5: **leave(X, restaurant)**

?

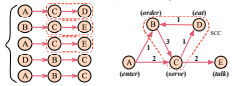
2. Previous Work

- Event pair-based methods
  - PMI: [Chambers and Jurafsky, 2008]
  - Bigram: [Jans et al., 2012]
  - EventComp: [Granroth-Wilding and Clark, 2016]
- Event chain-based methods
  - Bilinear and LSTM Language Model
  - PairLSTM: [Wang et al., 2017]

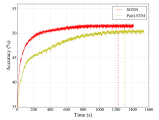
3. Motivation



(a) Given an event context (A, B, C), choose the subsequent event from D and E.



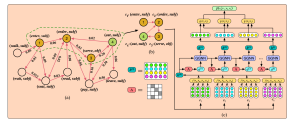
(b) Training event chains. (c) Event graph based on event chains in (b).



SGNN quickly reaches a stable high accuracy.

SGNN outperforms PairLSTM from start to the end.

4. Methods



- Construct an narrative event evolutionary graph (NEEG) based on extracted narrative event chains
- Network Representation Learning by a scaled graph neural network (SGNN) to solve the inference problem on the event graph

Step One: Construct NEEG

- Extract narrative event chains
  - The same extraction pipeline as Granroth-Wilding and Clark, 2016
- Exploit a suitable event representation
  - Predicate-GR: (eats, subj)
  - Verb and its grammatical dependency relation to the chain entity
  - Abstract and is a generalization of many concrete events
- Construct NEEG
  - Regard each predicate-GR bigram in the training event chains as a directed edge, and compute the probability as :

$$u(v_j|v_i) = \frac{\text{count}(v_i, v_j)}{\sum_k \text{count}(v_i, v_k)}$$

Step Two: Network Representation Learning

- Learn the initial event representations
- Update event representations using SGNN
  - Inputs to SGNN are two matrices $h^{(0)}$ and A
  - SGNN behaves like the widely used gated recurrent unit (GRU)

$$a^{(t)} = A^\top[h^{(t-1)} \; 1]$$
$$z^{(t)} = \sigma(W^z a^{(t)} + U^z h^{(t-1)})$$
$$r^{(t)} = \sigma(W^r a^{(t)} + U^r h^{(t-1)})$$
$$\tilde{s}^{(t)} = \tanh(Wa^{(t)} + U[r^{(t)} \odot h^{(t-1)}])$$
$$h^{(t)} = (1 - z^{(t)}) \odot h^{(t-1)} + z^{(t)} \odot \tilde{s}^{(t)}$$

- Choose the correct subsequent event

5. Experiments

	Training	Development	Test
#Documents	830,643	103,583	103,805
#Chains for NEEG	5,997,385	-	-
#Chains for SGNN	140,331	10,000	10,000

Methods	Accuracy
Random	20.00
PMI [Chambers and Jurafsky, 2008]	30.52
Bigram [Jans *et al.*, 2012]	29.67
Word2vec [Mikolov *et al.*, 2013]	42.23
DeepWalk [Perozzi *et al.*, 2014]	43.01
EventComp [Granroth-Wilding and Clark, 2016]	49.57
PairLSTM [Wang *et al.*, 2017]	50.83
SGNN-attention (without attention)	51.56
SGNN (ours)	**52.45**
SGNN+PairLSTM	52.71
SGNN+EventComp	54.15
SGNN+EventComp+PairLSTM	**54.93**

6. Conclusions

- We propose constructing narrative event evolutionary graph to solve the problem of script event prediction
- We present a Scaled Graph Neural Network (SGNN) to model the event interactions and learn better event representations
- Experimental results show that our model achieves the best performance