Co-training Embeddings of Knowledge Graphs and Entity Descriptions for Cross-lingual Entity Alignment

Muhao Chen, Yingtao Tian, Kai-WeiChang, Steven Skiena, Carlo Zaniolo

Xiaofan Yan

Outline

- Abstract
- Introduction
- Modeling
- Experiments
- Conclusion and future work

Abstract

- **Multilingual** KG embeddings provide latent semantic representations of entities and structured knowledge with **cross-lingual inferences**.
- Challenge: low coverage of entity alignment
- Introduce an embedding-based approach which leverages a weakly aligned multilingual KG for semi-supervised cross-lingual learning using entity descriptions

Introduction

- KG for NLP-related task
  - Relation extraction
  - Ontology population
  - Question answering
  - Dialogue agents
  - Visual semantic labeling
  - Connect KG structures
    - Mutilingual
      - Knowledge alignment
      - Cross-lingual QA
      - Machine-translation

Introduction

- Challenge

  cross lingual knowledge

  formed as

  inter-lingual links(ILLs) → match cross lingual counterparts of entities

- Method
  - Propose a novel co-training-based approach KDCoE, which trains two component embedding models on multilingual KG structures and entity descriptions.

Introduction

- KG embedding model
  - Jointly train
    - Translational knowledge model
    - Linear-transformation-based alignment
  - To encode KG structure
- Description embedding
  - Employ
    - Attentive gated recurrent unit encoder
    - Multilingual word embeddings
  - To characterize multilingual entities embedding

Related work

- Monolingual KG embedding
  - TransE, TransH, TransR, TransD, TransA
    - Differences: forms of relation-specific projrctions
  - DistMult, HolE, ConvE
    - Non-translational models

Related work

- Multilingual KG embedding:
  - MTransE: connects monolingual model with a jointly trained alignment model
  - JAPE: strengthen MTransE based on the similarity of entity attributions
  - ITransE: self-training
  - LM, CCA: off-line training
  - OT: orthogonal-transformation based
- **Co-training**: combine multiple models to learn different views of the data to enhance the supervision by taking turn in suggesting more labels

Modeling

- KDCoE
  - KGEM -- capture embeddings with cross-lingual inferences for structured knowledge
  - DEM -- entity descriptions
  - Co-training -- propose new ILLs with high confidence to populate

Modeling

- Multilingual KG Embeddings
  - Preserve entities and relations in a separated embedding space
- Multilingual Entity Description Embeddings
  - An attentive gated recurrent unit encoder (AGRU) is used to encode the multilingual entity descriptions. On top of that, DEM is trained to collocate the description embeddings of cross-lingual counterparts.
- Iterative Co-training
  - The co-training of the two model components is conducted iteratively on the KG, where a small amount of ILLs is provided for training.

Experiments

- Cross-lingual entity alignment
  - match the same entities from different languages in KB

Language	En-Fr			En-De		
Metric	*Hit@1*	*Hit@10*	*MRR*	*Hit@1*	*Hit@10*	*MRR*
LM	1.02	2.21	0.014	1.37	2.14	0.015
CCA	1.80	3.54	0.021	2.19	3.42	0.025
OT	20.15	25.37	0.212	11.04	19.74	0.122
ITransE	10.14	11.59	0.106	6.55	11.44	0.076
MTransE-AC	4.49	8.67	0.051	5.56	8.50	0.060
MTransE-TV	5.12	7.55	0.055	3.62	8.12	0.053
MTransE-LT	27.40	33.98	0.309	17.90	31.59	0.225
KDCoE ($i2$)	37.70	45.01	0.405	29.80	41.66	0.322
KDCoE ($i3$)	43.77	53.07	0.463	30.99	43.02	0.334
KDCoE ($i4$)	46.17	54.85	0.487	32.20	44.58	0.346
KDCoE (term)	**48.32**	**56.95**	**0.496**	**33.52**	**45.47**	**0.349**

Experiments

- Zero-shot alignment
  - align entities that do not exist in the structure of KG.

Language	En-Fr			En-De		
Metric	$Hit@1$	$Hit@10$	MRR	$Hit@1$	$Hit@10$	MRR
Single-layer	0.97	1.80	0.013	0.36	2.10	0.010
CNN	1.19	6.91	0.036	1.28	4.63	0.019
GRU	18.45	27.65	0.204	11.23	24.48	0.165
AGRU-mono	5.08	18.27	0.096	5.03	14.90	0.085
AGRU-multi	26.92	44.69	0.337	19.34	45.69	0.269
KDCoE ($i1$)	27.69	48.69	0.346	19.52	45.84	0.274
KDCoE ($i2$)	28.82	52.58	0.350	20.37	46.35	0.279
KDCoE ($i3$)	30.83	55.91	**0.384**	21.28	48.49	0.283
KDCoE (term)	**30.96**	**56.93**	0.382	**21.97**	**50.02**	**0.285**

Experiments

- Cross-lingual KG Completion
  - Compare the KGEM of KDCoE against its monolingual counterpart TransE for KG completion, based on the sparser French and German versions.
  - Monolingual prediction (KDCoE-mono)
  - Cross-lingual prediction

Language	Fr				De			
Predict	Tail		Head		Tail		Head	
Metric	*Hit@10*	*MRR*	*Hit@10*	*MRR*	*Hit@10*	*MRR*	*Hit@10*	*MRR*
TransE	29.21	0.077	18.19	0.046	29.58	0.099	23.57	0.059
KDCoE-mono	31.05	0.092	16.88	0.053	29.13	0.124	27.63	0.106
KDCoE-cross	**37.21**	**0.139**	**22.23**	**0.093**	**34.17**	**0.134**	**31.05**	**0.143**

Conclusion and future work

- Propose a semi-supervised learning approach to co-train multilingual KG embeddings and the embeddings of entity descriptions for cross-lingual knowledge alignment.
- And able to enhance the traditional methods of KG completion