Modeling Relational Data With Graph Convolutional Networks

School of Industrial and Management Engineering, Korea University

Jinsoo Bae

Contents

Research Purpose

❖ Modeling Relational Data With Graph Convolutional Networks (2018 European semantic web conference)

- 암스트레담 대학교, 캐나다 고등연구재단에서 연구하였고, 2022년 4월 3일 기준 2012회 인용
- Knowledge Graph Embedding에 Graph Convolutional Networks를 처음으로 적용한 논문
  - R-GCN(Relational GCN) 모델 구현을 통해 entity classification, link prediction task를 수행
  - Question and Answering, Information Retrieval 분야로 활용 가능

Modeling Relational Data with Graph Convolutional Networks

Michael Schlichtkrull*
University of Amsterdam
m.s.schlichtkrull@uva.nl

Thomas N. Kipf*
University of Amsterdam
t.n.kipf@uva.nl

Peter Bloem
VU Amsterdam
p.bloem@vu.nl

Rianne van den Berg
University of Amsterdam
r.vandenberg@uva.nl

Ivan Titov
University of Amsterdam
titov@uva.nl

Max Welling
University of Amsterdam, CIFAR†
m.welling@uva.nl

Research Purpose

❖ Knowledge Base는 지적 활동과 경험을 통해 축적된 전문 지식, 사실, 규칙들이 저장된 데이터베이스로, Triple (*subject, predicate, object*) 형태로 표현됨

❖ 그래프 데이터로 표현될 경우 *subject*와 *object*를 노드로, *relation*을 엣지로 표현

"Mikhai Baryshnikov was educated at Vaganova Academy"

Triple

(Mikhai Baryshnikov, educated_at, Vaganova Academy)

Subject	Predicate	Object
Head	Relation	Tail
(Node)	**(Edge)**	**(Node)**

Research Purpose

❖ Knowledge Base는 지적 활동과 경험을 통해 축적된 전문 지식, 사실, 규칙들이 저장된 데이터베이스로, Triple (*subject, predicate, object*) 형태로 표현됨

❖ 그래프 데이터로 표현될 경우 *subject*와 *object*를 노드로, *relation*을 엣지로 나타냄

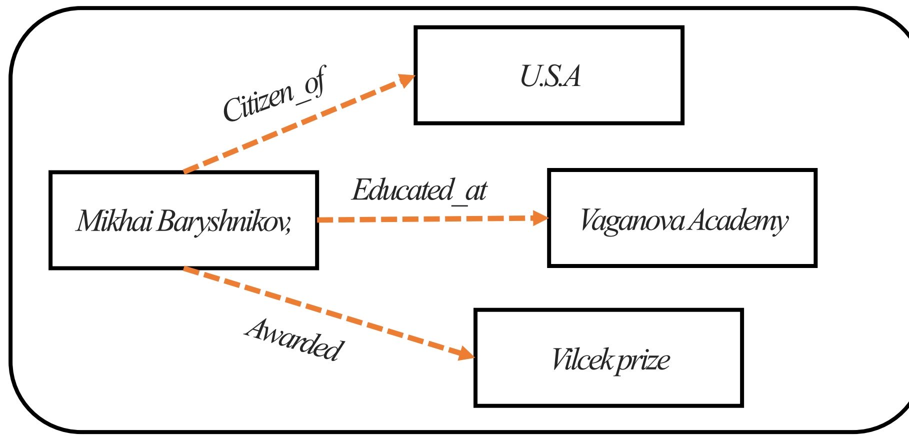
[Knowledge Base]

(Mikhai Baryshnikov, educated_at, Vaganova Academy)

(Mikhai Baryshnikov, citizen_of, U.S.A)

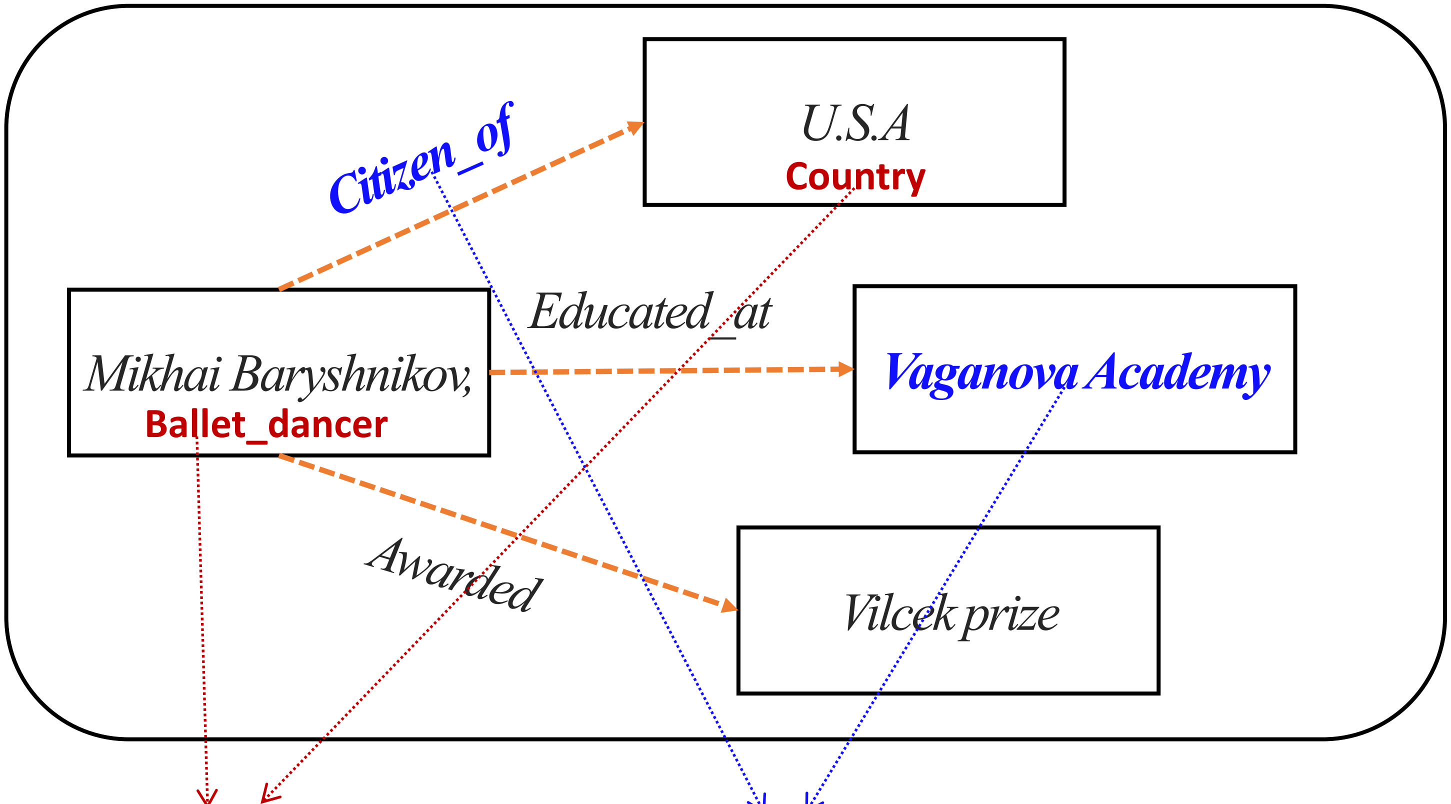
(Mikhai Baryshnikov, awarded, Vilcek prize)

[Knowledge Graph]



Research Purpose

❖ 본 논문에서는 R-GCN(Relational GCN) 모델 구현을 통해 Statistical Relational Learning(SRL) 분야의 Entity Classification, Link Prediction을 수행함



[Entity Classification and Link Prediction in Knowledge Graph]

Proposed Method

Relational GCNs

$W^{(l)} \rightarrow W^{l}_{r \in \mathcal{R}}$

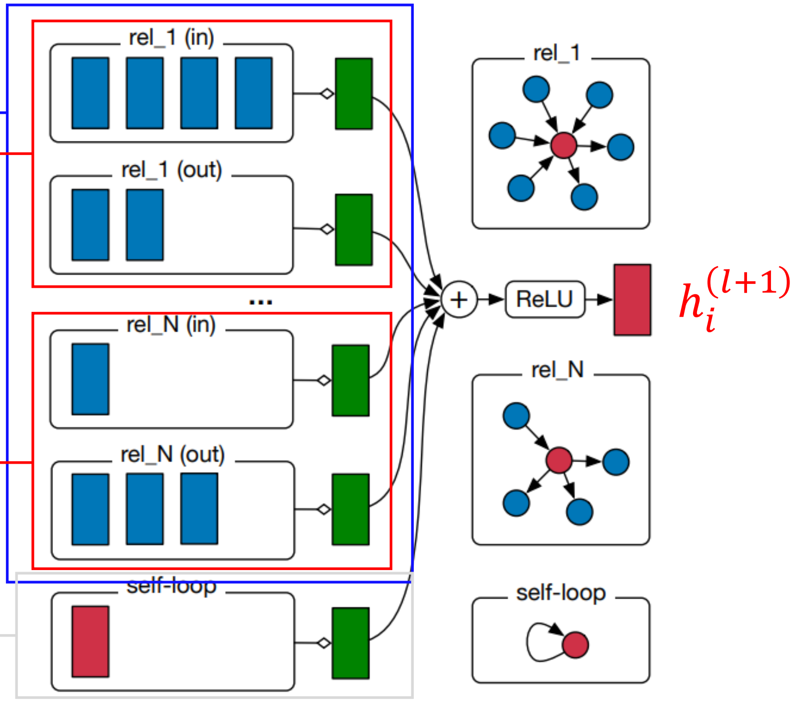
❖ 아래 같은 지식 그래프 데이터가 주어졌을 때, **Relation-specific** Graph Convolutional 연산 수행

- Directed and labeled multi-graphs as $G = (\mathcal{V}, \mathcal{E}, \mathcal{R})$
- v_i(entities) $\in \mathcal{V}$, r(a relation type) $\in \mathcal{R}$, (v_i, r, v_j)(labeled edges,) $\in \mathcal{E}$

$$h_i^{(l+1)} = \sigma\left(\sum_{r\in\mathcal{R}}\sum_{j\in\mathcal{N}_i^r}\frac{1}{c_{i,r}}W_r^{(l)}h_j^{(l)} + W_0^{(l)}h_i^{(l)}\right)$$

$l+1$번째 i 노드의 representation 벡터

l번째 i 노드의 representation 벡터



Proposed Method

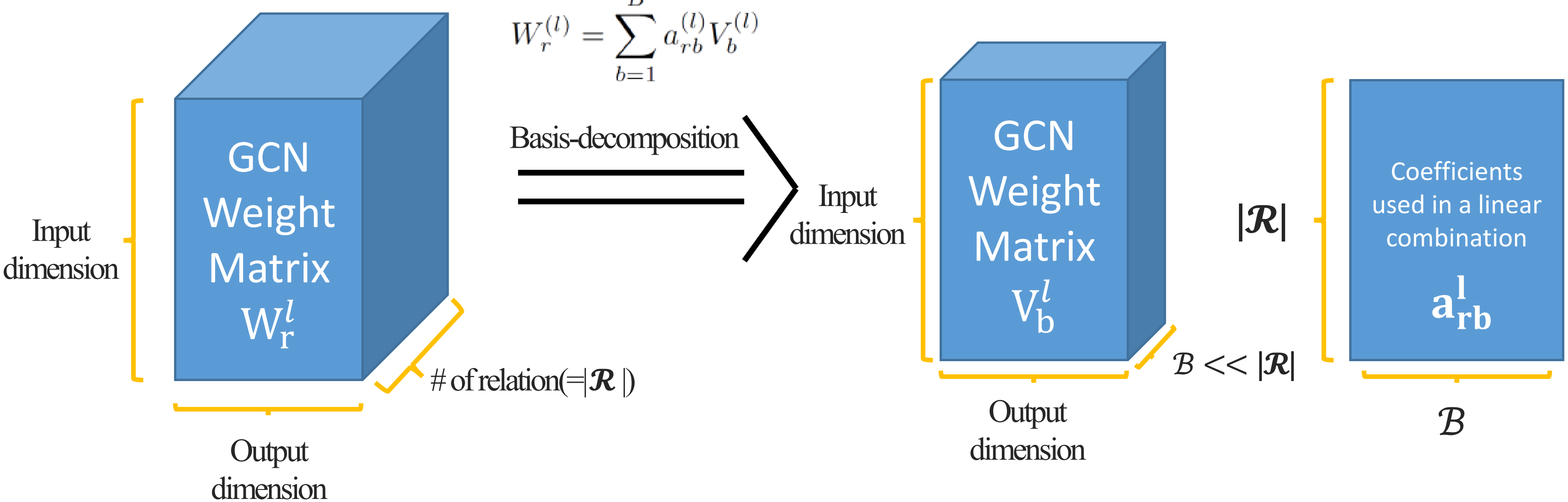
$$h_i^{(l+1)} = \sigma\left(\sum_{\boxed{r\in\mathcal{R}}}\sum_{j\in\mathcal{N}_i^r}\frac{1}{c_{i,r}}\boxed{W_r^{(l)}}h_j^{(l)} + W_0^{(l)}h_i^{(l)}\right)$$

Relational GCNs

❖ Highly **Multi-relational** Data에서 **모든 $r \in \mathcal{R}$**에 대한 **Relation-specific** GCN 연산을 수행하게 되면 모델 복잡도가 많이 상승 (W_r^l=> 多)하므로, 아래와 같은 Regularization method 적용

- basis-decomposition (선형결합을 활용해 학습해야 할 파라미터 수 감소)
- block-diagonal-decomposition

$$W_r^{(l)} = \sum_{b=1}^{B} a_{rb}^{(l)} V_b^{(l)}$$



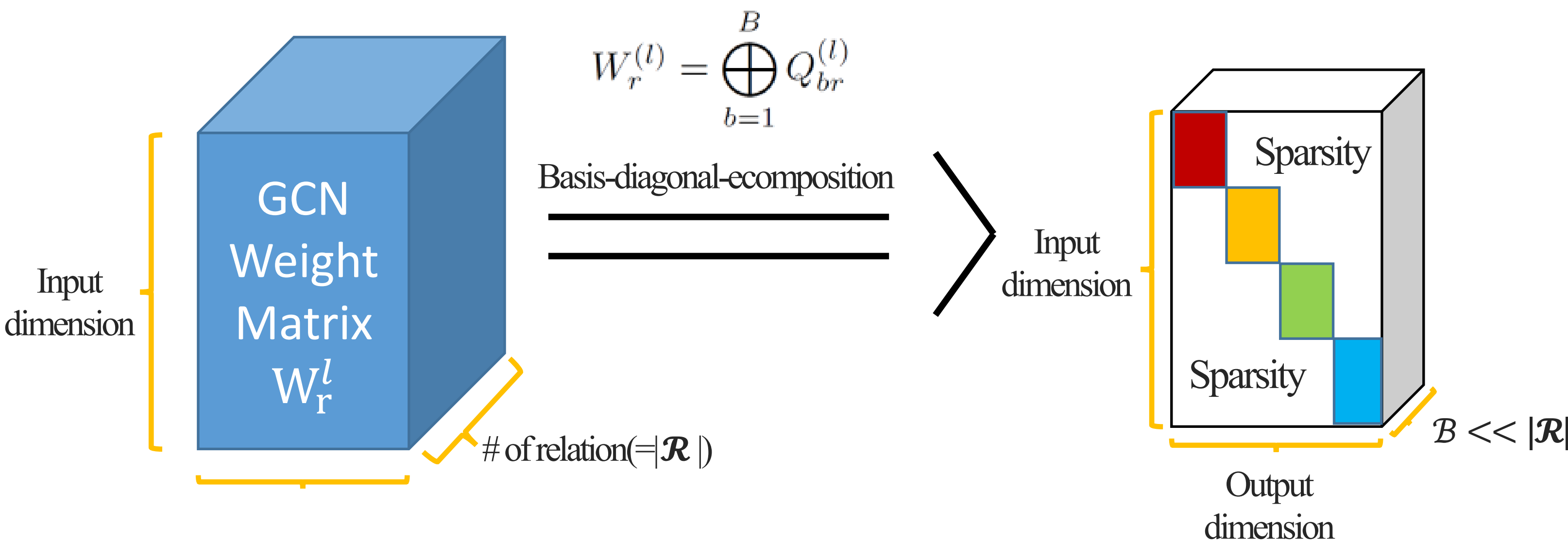
Proposed Method

$$h_i^{(l+1)} = \sigma\left(\sum_{\boxed{r\in\mathcal{R}}}\sum_{j\in\mathcal{N}_i^r}\frac{1}{c_{i,r}}\boxed{W_r^{(l)}}h_j^{(l)} + W_0^{(l)}h_i^{(l)}\right)$$

Relational GCNs

❖ Highly **multi-relational** data에서 **모든 $r \in \mathcal{R}$**에 대한 **relation-specific** GCN 연산을 수행하게 되면 모델 복잡도가 많이 상승 (W_r^l=> 多)하므로, 아래와 같은 Regularization method 적용

- basis-decomposition (선형결합을 활용해 학습해야 할 파라미터 수 감소)
- block-diagonal-decomposition (대각선 block들을 제외한 나머지 학습 파라미터는 모두 0 값)



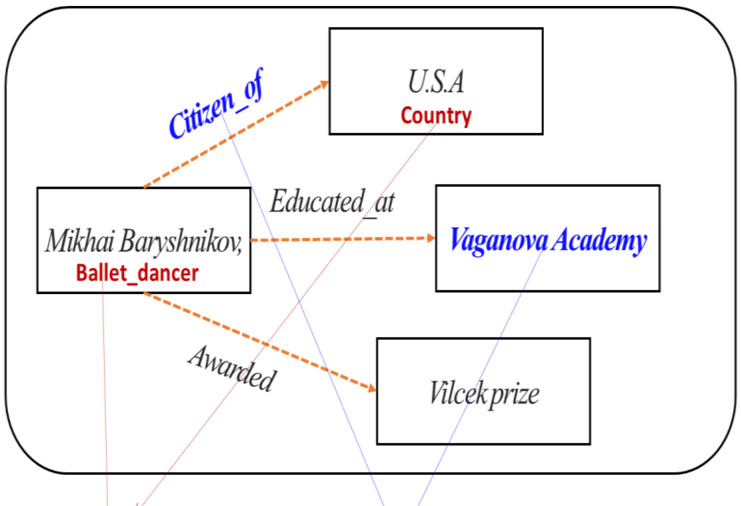
DMQA hcai

Proposed Method

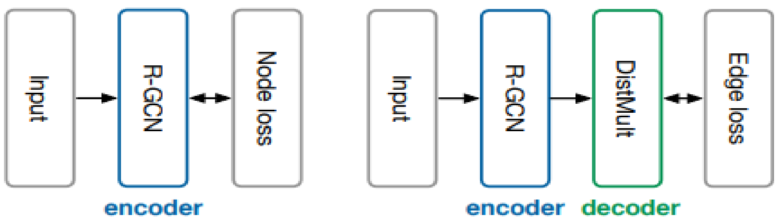
Entity classification and Link Prediction

❖ Entity classification은 R-GCNs로 얻은 노드 임베딩 벡터를 classification layer에 입력해 노드별 클래스 예측을 의미. (일반적인 분류 문제처럼 cross-entropy 나 sigmoid 함수로 학습)

❖ Link Prediction은 R-GCNs의 출력벡터(지식 그래프 구조를 반영한 임베딩 벡터)를 DistMult(기존 Link Prediction 알고리즘)에 입력하여 수행. (+정답 이외의 클래스에 대해서는 Negative Sampling)



[Entity classification and link prediction in Knowledge Graph]



(a) Entity classification (b) Link prediction

Figure 3: (a) Depiction of an R-GCN model for entity classification with a per-node loss function. (b) Link prediction model with an R-GCN encoder (interspersed with fully-connected/dense layers) and a DistMult decoder that takes pairs of hidden node representations and produces a score for every (potential) edge in the graph. The loss is evaluated per edge.

Experiments

Entity classification

❖ (Subject, predicate, object)의 triple 구조를 가진 4가지 Knowledge Graph 데이터셋과 Accuracy로 알고리즘 성능 평가 진행

- Dataset: AIFB, MUTAG, BGS, AM, 층별 16개 히든 유닛을 가진 2층 R-GCNs 모델 구조 선택
- 비교 알고리즘: Feat(hand-designed feature extractors, 2012), WL(Weisfeiler-Lehman kernels, 2015), RDF2Vec(2016)
- AIFB, AM 데이터에서 가장 우수한 성능을 보였고, MUTAG, BGS 데이터셋은 Normalization consant $(1/c_{i,r})$ 고도화를 통해 성능 향상 가능성 기대

Dataset	AIFB	MUTAG	BGS	AM
Entities	8,285	23,644	333,845	1,666,764
Relations	45	23	103	133
Edges	29,043	74,227	916,199	5,988,321
Labeled	176	340	146	1,000
Classes	4	2	2	11

Datasets

Results All results in Table 2 are reported on the train/test benchmark splits from Ristoski, de Vries, and Paulheim (2016). We further set aside 20% of the training set as a validation set for hyperparameter tuning. For R-GCN, we report performance of a 2-layer model with 16 hidden units (10 for AM), basis function decomposition (Eq. 3), and trained with Adam (Kingma and Ba 2014) for 50 epochs using a learning rate of 0.01. The normalization constant is chosen as $c_{i,r} = |\mathcal{N}_i^r|$. Further details on (baseline) models and hyperparameter choices are provided in the supplementary material.

Model	AIFB	MUTAG	BGS	AM
Feat	55.55	77.94	72.41	66.66
WL	80.55	**80.88**	86.20	87.37
RDF2Vec	88.88	67.20	**87.24**	88.33
R-GCN	**95.83**	73.23	83.10	**89.29**

Experiments

Link Prediction

❖ Link Prediction 분야의 벤치마킹 데이터셋(FB 15k, WN18, FB15k-237)과 MRR(mean reciprocal rank), H@n(Hits at n)로 알고리즘 성능 평가 진행

- FB15k—237은 Inverse relation이 있는 경우를 사전에 제거한 데이터 버전을 의미 (Inverse relation이 있는 경우 상대적으로 원래 relation 예측이 쉬워짐)
- FB 15k, WN 18에서는 단층 R-GCNs, FB15k-237에서는 이층 R-GCNs 모델 구조를 사용함
- **R-GCN**이 R-GCNs의 임베딩 벡터와 DistMult 알고리즘 병합 모델을 의미하고, **R-GCN+은** R-GCN과 DistMult의 앙상블 모델을 의미
- <u>R-GCN이 지식 데이터간 관계성을 올바르게 학습했기에, 기존 DistMult 알고리즘보다 우수한 성능을 확보</u>

Dataset	WN18	FB15K	FB15k-237
Entities	40,943	14,951	14,541
Relations	18	1,345	237
Train edges	141,442	483,142	272,115
Val. edges	5,000	50,000	17,535
Test edges	5,000	59,071	20,466

Table 3: Number of entities and relation types along with the number of edges per split for the three datasets.

	FB15k					WN18				
	MRR		Hits @			MRR		Hits @		
Model	Raw	Filtered	1	3	10	Raw	Filtered	1	3	10
LinkFeat		0.779			0.804		0.938			0.939
DistMult	0.248	0.634	0.522	0.718	0.814	0.526	0.813	0.701	0.921	0.943
R-GCN	0.251	0.651	0.541	0.736	0.825	0.553	0.814	0.686	0.928	0.955
R-GCN+	**0.262**	**0.696**	**0.601**	**0.760**	**0.842**	0.561	0.819	0.697	0.929	**0.964**
CP*	0.152	0.326	0.219	0.376	0.532	0.075	0.058	0.049	0.080	0.125
TransE*	0.221	0.380	0.231	0.472	0.641	0.335	0.454	0.089	0.823	0.934
HolE**	0.232	0.524	0.402	0.613	0.739	**0.616**	0.938	0.930	**0.945**	0.949
ComplEx*	0.242	0.692	0.599	0.759	0.840	0.587	**0.941**	**0.936**	**0.945**	0.947

Table 4: Results on the the Freebase and WordNet datasets. Results marked (*) taken from Trouillon et al. (2016). Results marks (**) taken from Nickel, Rosasco, and Poggio (2015). R-GCN+ denotes an ensemble between R-GCN and DistMult – see main text for details.

Conclusion

❖ Conclusion

- Knowledge Graph 분야에서 Knowledge base(노드간 관계성)를 고려한 Entity representation learning을 위해 GCN 모델을 처음으로 적용한 논문
- Knowledge relation별 GCN 연산 아이디어를 제안했으며, 이 때 특정 relation에 대한 오버피팅이 일어나는 것을 방지하고자 모델 파라미터에 대한 정규화 기법을 적용
  - <u>Parameter sharing</u> using basis decomposition <u>between different relation types.</u>
  - <u>Enforce the sparsity constraints</u> using block-diagonal decomposition <u>for each relation type.</u>

Thank You