Multimodal Knowledge Graph Embeddings

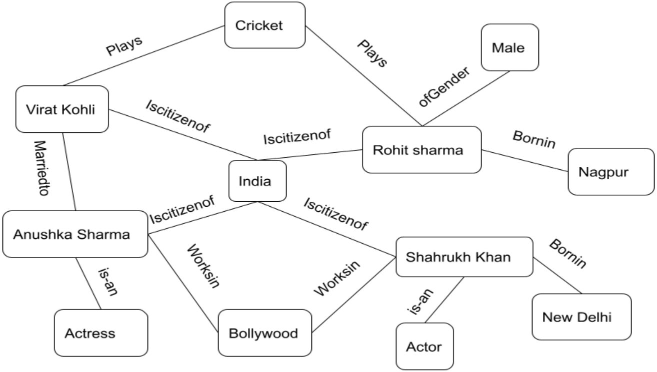
Presented by: Parsa Revanth,
Advisor: Dr.Anand Mishra

Indian Institute of Technology
Jodhpur

June 11, 2022

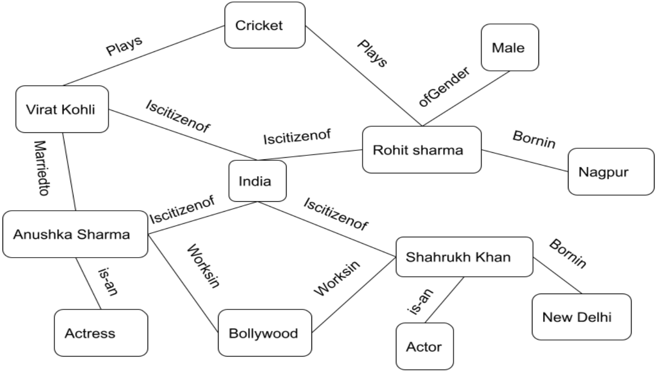
Introduction: What is Knowledge Graph

- Knowledge graph is a network of entities
- Knowledge graph is represented as set of triplets (Subject, Relation, Object)



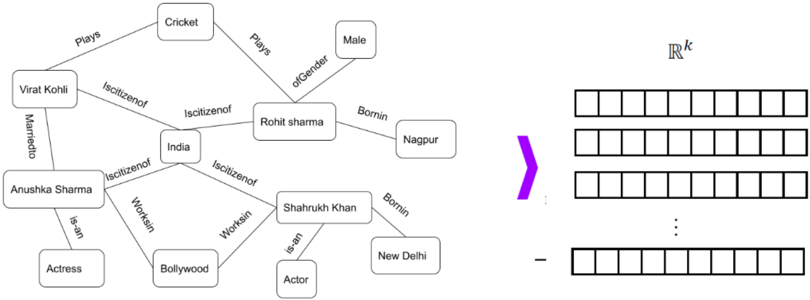
Introduction: Applications of Knowledge Graph

- Question Answering System
- Recommender System
- Information Retrieval



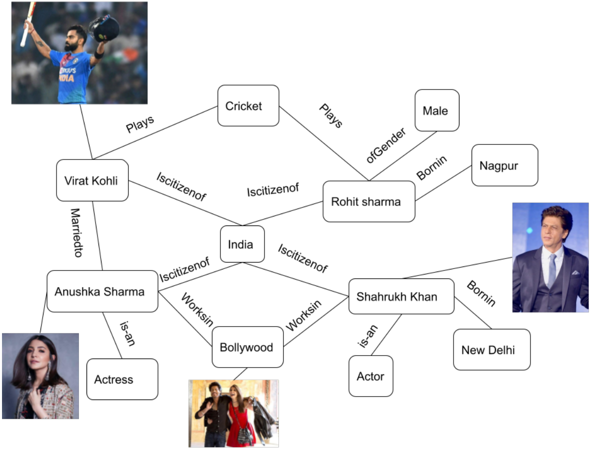
Problem Statement

- One of the task for Knowledge graph completion is link prediction
- Learn the low-dimensional representations of entities and the relations in a knowledge graph



Problem Statement

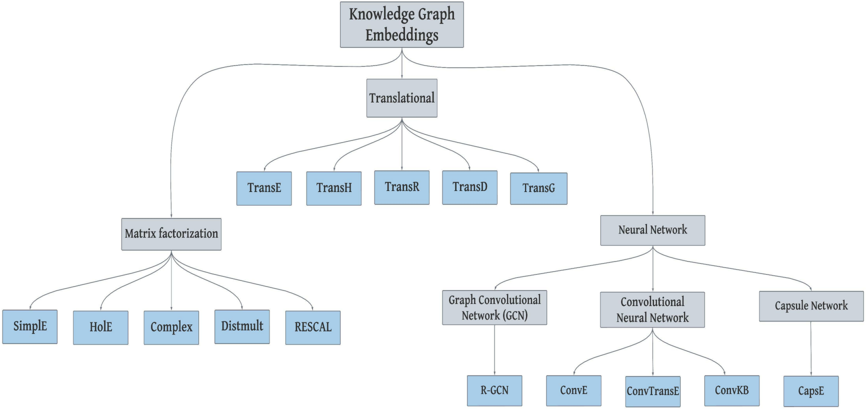
- Multimodality is using image and text information
- Multimodality is a complementary



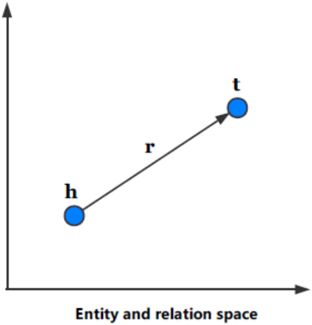
Contributions

1. Detailed survey of knowledge graph embeddings techniques and their implementations
2. Exploring few of knowledge graph embedding techniques for multimodal data on two datasets

Survey of Knowledge Graph Embeddings



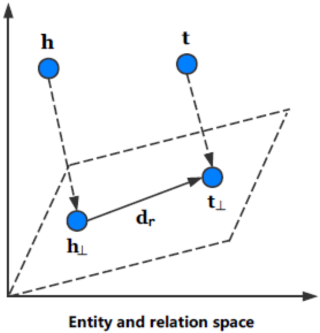
1. TransE [NeurIPS'13]



$$f_r(h, t) = \|h + r - t\|_{l1/l2}. \quad (1)$$

Translational-based models

2. TransH [AAAI'14]

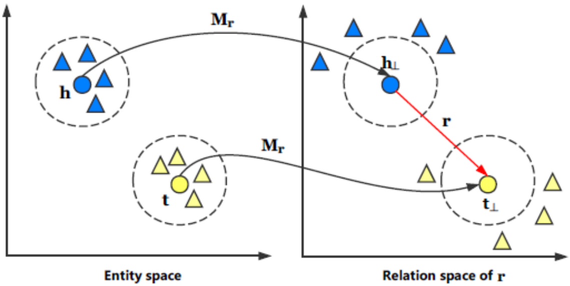


$$\mathbf{h}_\perp = \mathbf{h} - \mathbf{w}_r^\top \mathbf{h} \mathbf{w}_r, \mathbf{t}_\perp = \mathbf{t} - \mathbf{w}_r^\top \mathbf{t} \mathbf{w}_r \tag{2}$$

$$f_r(h, t) = \|\mathbf{h}_\perp + d_r - \mathbf{t}_\perp\|_{l2}. \tag{3}$$

Translational-based models

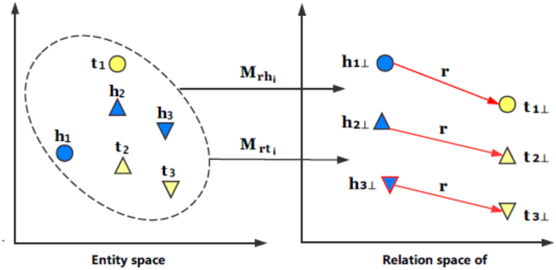
3. TransR [AAAI'15]



$$\mathbf{h}_\perp = \mathbf{M}_r\mathbf{h}, \mathbf{t}_\perp = \mathbf{M}_r\mathbf{t} \tag{4}$$

$$f_r(h,t) = \|\mathbf{h}_\perp + r - \mathbf{t}_\perp\|_{l2}. \tag{5}$$

4. **TransD [IJCNLP'15]**



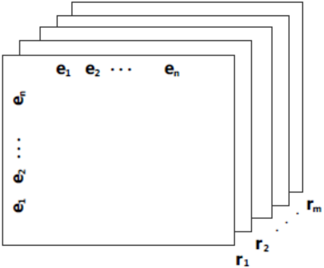
$$M_{rh} = \mathbf{r}^{'}\mathbf{h}^{'\top} + \mathbf{I}, M_{rt} = \mathbf{r}^{'}\mathbf{t}^{'\top} + \mathbf{I}. \tag{6}$$

$$\mathbf{h}_{\perp} = \mathbf{M}_{rh}\mathbf{h}, \mathbf{t}_{\perp} = \mathbf{M}_{rt}\mathbf{t}. \tag{7}$$

$$f_r(h,t) = \|\mathbf{h}_{\perp} + r - \mathbf{t}_{\perp}\|_{l2}. \tag{8}$$

Matrix Factorization-based models

1. RESCAL [ICML'11]



$$\mathbf{Z_r} = \mathbf{A}\mathbf{M_r}\mathbf{A}^\top. \tag{9}$$

$$f_r(h, t) = \mathbf{h}^\top \mathbf{M_r} \mathbf{t}. \tag{10}$$

Matrix Factorization-based models

2. **Distmult [ICLR'15]**

$$f_r(h,t) = \mathbf{h}^\top \mathbf{diag}(\mathbf{r})\mathbf{t}. \tag{11}$$

3. **ComplEx [ICML'16]**

$$f_r(h,t) = \mathbf{Re}(\mathbf{h}^\top \mathbf{diag}(\mathbf{r})\mathbf{t}^{'}). \tag{12}$$

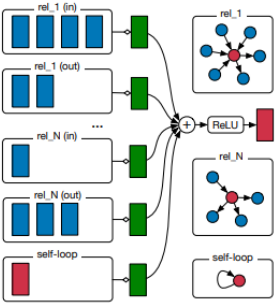
4. **HolE [AAAI'16]**

$$f_r(h,t) = \mathbf{r}^\top (\mathbf{h} * \mathbf{t}). \tag{13}$$

5. **SimplE [NeurIPS'18]**

$$f_r(h,t) = \frac{1}{2}(\mathbf{h} \circ \mathbf{rt} + \mathbf{t} \circ \mathbf{r}^{'}\mathbf{h}). \tag{14}$$

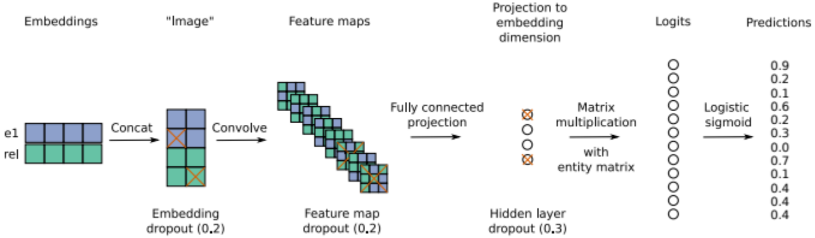
1. **R-GCN [ESWC'18]**



$$q_i^{(l+1)} = \sigma\left(\sum_{r\in\mathcal{R}}\sum_{j\in\mathcal{N}_i^r}\frac{1}{c_{i,r}}W_r^{(l)}q_j^{(l)} + W_0^{(l)}q_i^{(l)}\right). \qquad (15)$$

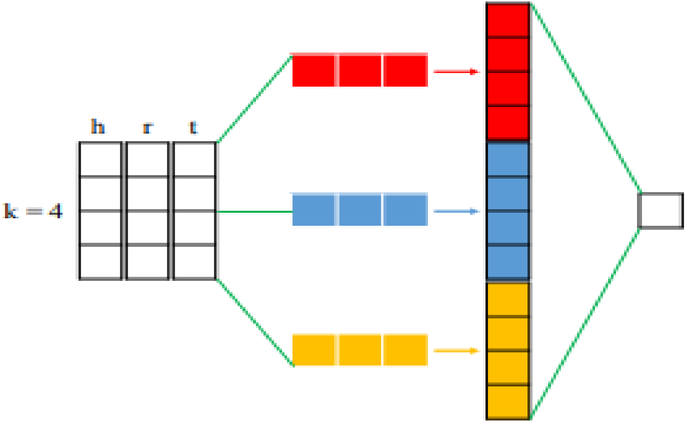
$$f_r(h,t) = \mathbf{h}^\top\mathbf{R}\mathbf{t}. \qquad (16)$$

2. ConvE [AAAI'18]



$$f_r(h,t) = f(vec(f([h;r] * w))W)t. \tag{17}$$

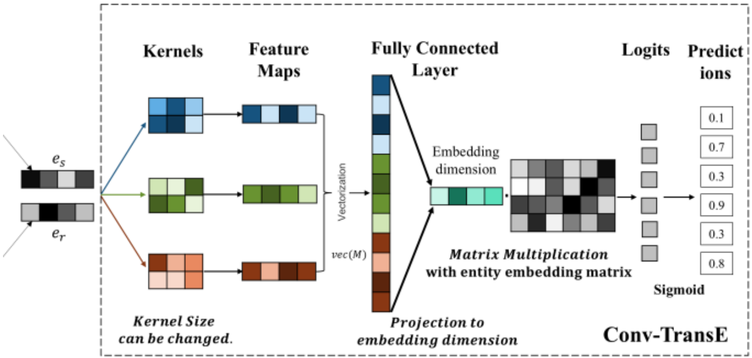
3. ConvKB [NAACL'18]



$$f_r(h, t) = f(vec(f([h; r; t] * w))W). \tag{18}$$

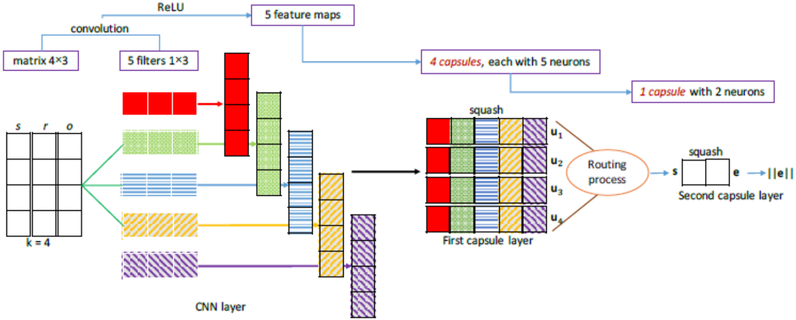
Neural Network-based models

4. Conv-TransE [AAAI'19]



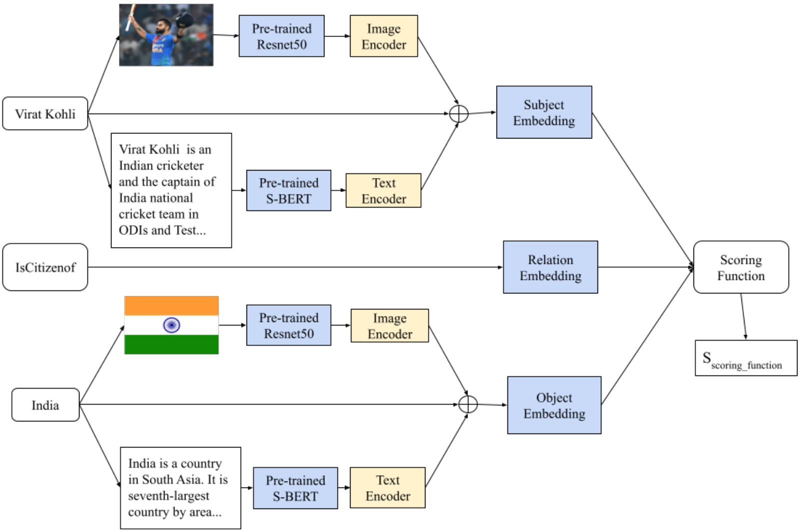
$$f_r(h, t) = f(vec(M_{(h,r)}))W)t. \tag{19}$$

5. CapsE [NAACL'19]



$$f_r(h, t) = ||(capsnet(f[h; r; t] * k))||. \quad (20)$$

Model architecture



Datasets used in the project

S.No	Dataset	Entities	Relations	Triplets			Images
				Train	Test	Valid	
1	FB15k-237	14505	237	272115	17535	20466	12237
2	WN18RR	40943	11	86835	3034	3134	-
3	YAGO3-10	123143	37	1079040	5000	5000	61223

Results

Table 1: Results on WN18RR dataset

S.No	Dataset	Modal	Embedding Technique	Data type	Embedding Dimension	Hits @ (%)		
						10	3	1
1	WN18RR	Translational	TransE	Entity	200	45.51	37.24	0.22
2			TransH			45.42	37.68	0.21
3			TransR			39.52	36.18	0.59
4			TransD			45.42	37.38	0.24
5		Matrix factorization	Distmult	Entity	200	46.09	36.57	23.87
6			ComplEx			**47.52**	**42.17**	33.77
7			SimplE			45.33	39.04	28.65
8		Neural network	ConvE	Entity	200	38.61	28.73	4.20
9			R-GCN		100	43.97	40.03	**34.56**

Results

Table 2: Results on FB15k-237 dataset

S.No	Dataset	Modal	Embedding Technique	Data type	Embedding Dimension	Hits @ (%)		
						10	3	1
1	FB15k-237	Translational	TransE	Entity	200	47.73	32.71	19.28
2				Entity + Image		38.59	25.65	17.12
3			TransH	Entity		48.59	33.55	20.39
4				Entity + Image		39.81	26.98	17.56
5			TransR	Entity		**48.70**	**34.09**	**21.29**
6			TransD			48.36	33.32	19.96
7		Matrix factorization	Distmult	Entity	200	34.31	20.01	9.94
8			ComplEx			40.94	26.82	14.88
9			SimplE			36.12	21.30	10.14
10		Neural network	ConvE	Entity	100	30.20	16.46	8.37
11			R-GCN			41.85	26.07	15.29

Results

Table 3: Results on YAGO3-10 dataset

S.No	Dataset	Modal	Embedding Technique	Data type	Embedding Dimension	Hits @ (%)		
						10	3	1
1	YAGO3-10	Translational	TransE	Entity	300	61.93	46.78	**29.45**
2				Entity + Image	350	24.55	14.31	7.79
3			TransH	Entity	300	**65.07**	**49.92**	16.04
4				Entity + Image		24.03	13.74	8.02
5			TransR	Entity		18.96	10.72	5.28
6			TransD			32.93	20.15	10.03
7		Matrix factorization	Distmult	Entity	300	19.30	7.86	3.06
8			ComplEx			34.06	14.73	3.02
9			SimplE			21.24	8.50	3.15
10		Neural network	ConvE	Entity	100	4.98	3.85	2.5
11			R-GCN			16.04	8.57	4.99

Conclusion

- I did a detailed survey on knowledge graph embedding techniques
- Implemented four translational embedding techniques TransE, TransH, TransR, and TransD for the entity only data
- Implemented three matrix factorization-based embedding techniques Distmult, ComplEx, and SimplE for the entity only data
- Implemented two neural network-based embedding techniques ConvE and R-GCN for the entity only data.
- For the entity and image data, I implemented TransE and TransH embedding techniques.

Thank You