

TransE的训练

$$d(sub, rel, obj) = ||s + r - o||_2^2$$

To learn such embeddings, we minimize a margin based ranking criterion over the training set:

$$\mathcal{L} = \sum_{(h,r,t) \in \mathcal{S}_{pos}} \sum_{(h',r',t') \in \mathcal{S}_{neg}} [\gamma + d(h,r,t) - d(h',r',t')]$$

where $[\gamma]_+$ denotes the positive part of x , $\gamma > 0$ is a margin hyperparameter, and $\mathcal{S}_{pos} = \{(h,r,t) | (h,r,t) \in \mathcal{E}\}$

Randomly change an entity

不确定知识图谱推理

- 不确定知识图谱: 为每个三元组添加一个置信度和描述三元组的不确定性
- 将传统知识图谱中三元组 (h, r, t) 拓展为 (h, r, t, s) , 其中 s 代表置信度, r 代表三元组之间的关系, s 代表置信度

1. $(h, r, t, s) \in \mathcal{E}$, $s \in [0, 1]$
2. $(h, r, t, s) \in \mathcal{E}$, $s \in [0, 1]$
3. $(h, r, t, s) \in \mathcal{E}$, $s \in [0, 1]$
4. $(h, r, t, s) \in \mathcal{E}$, $s \in [0, 1]$

- 三元组置信度如何获得的呢?
- 三元组置信度与三元组属性有关
- 三元组置信度与三元组属性有关
- 三元组置信度与三元组属性有关

GTransE

- 专注于学习那些置信度更高的三元组, 降低那些置信度低、置信度较低的三元组对训练及关系最小的贡献
- 利用三元组置信度动态地调整TransE中的Margin
- 置信度越高, 置信度小, 置信度小则Margin小

$$\mathcal{L} = \sum_{(h,r,t,s) \in \mathcal{S}_{pos}} \sum_{(h',r',t',s') \in \mathcal{S}_{neg}} [\gamma + d(h,r,t) - d(h',r',t')]$$



传统知识工程: 依赖专家/干预(自上而下); 知识获取困难; 知识应用困难

语义网络

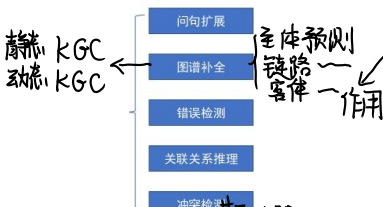
NODE 实体、概念和情况
Arc 节点间关系

多元 $R(x_1, x_2, \dots, x_n)$
二元 $R_2(x_1, x_2) \wedge R_3(x_1, x_3) \wedge \dots \wedge R_n(x_1, x_n) \wedge \dots \wedge R_{n+1}(x_{n+1}, x_n)$

NODE 实体、概念、值
Arc 关系、属性

优势: 尺度大, 语义丰富, 数据量大, 质量高, 结构易于实现

挑战: 高质量模式缺失, 封闭世界假设不再成立, 大规模自动化知识获取有前提



知识图谱

表示

构建

$G = G(V, E)$ 有问何句, 描述现实世界中资源, 路径可到达
RDF: 描述现实世界中资源, 路径可到达

三元组: (主体, 属性, 客体)

模式设计
明确数据来源

词汇挖掘

实体发现

关系发现

词汇挖掘

质量控制

人工干预

$$d(sub, rel, obj) = ||s + r - o||_2^2$$

To learn such embeddings, we minimize a margin based ranking criterion over the training set:

$$\mathcal{L} = \sum_{(h,r,t) \in \mathcal{S}_{pos}} \sum_{(h',r',t') \in \mathcal{S}_{neg}} [\gamma + d(h,r,t) - d(h',r',t')]$$

where $[\gamma]_+$ denotes the positive part of x , $\gamma > 0$ is a margin hyperparameter, and $\mathcal{S}_{pos} = \{(h,r,t) | (h,r,t) \in \mathcal{E}\}$

Randomly change an entity

- 专注于学习那些置信度更高的三元组, 降低那些置信度低、置信度较低的三元组对训练及关系最小的贡献
- 利用三元组的置信度动态地调整TransE中的Margin
- 置信度越高, 置信度小, 置信度小则Margin小

$$\mathcal{L} = \sum_{(h,r,t,s) \in \mathcal{S}_{pos}} \sum_{(h',r',t',s') \in \mathcal{S}_{neg}} [\gamma + d(h,r,t) - d(h',r',t')]$$

