

TransE, TransH, TransR 的比较

Alouette

任务介绍

- 数据格式是(h, t, r)三元组，在FB15K237数据集上都已经转换成了id的格式，数据集中实体有14541个，关系有237个
- 测试集依照实体之间对应的数量关系被分为四类：一对一，一对多，多对一，多对多，分类是为了衡量之后TransH和TransR相对于TransE在后三种情形下的提升
- 模型的目的是学到实体和关系的合理embedding表示，做到给定三元组中任意两个实体，根据embedding向量之间的关系，成功预测出剩下实体

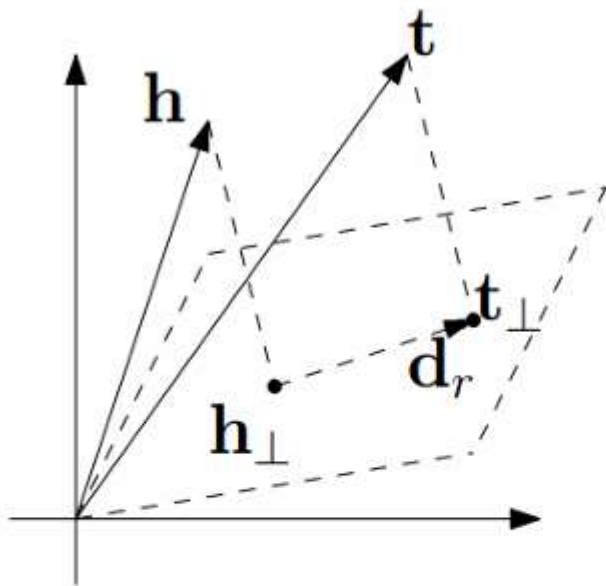
Loss函数

$$f(h, r, t) = \|\mathbf{h} + \mathbf{r} - \mathbf{t}\|$$

$$\sum_{(h,r,t) \in \Delta} \sum_{(h',r,t') \in \Delta'} [\gamma + f(h,r,t) - f(h',r,t')]_+$$

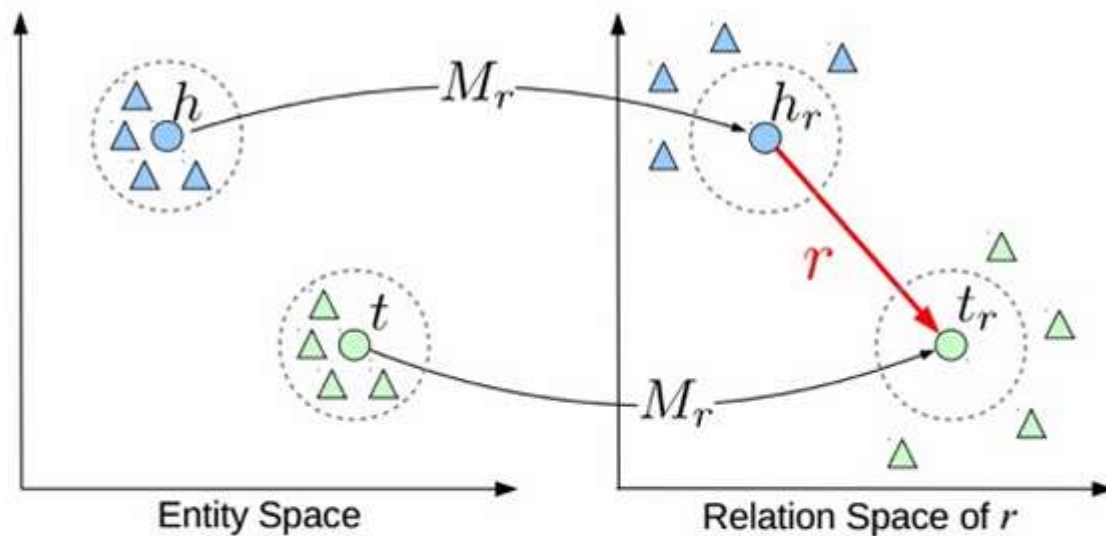
- 理想状态下，正例三元组f的值应该尽量趋于0，负例三元组f的值应该尽量远离0，因此用 γ 来作为边界，尽量区分开正负例
- 类似地，TransH和TransR只是f函数变了，增加了投影操作

模型原理



TransH

$$f_r(h, t) = -\|(\mathbf{h} - \mathbf{w}_r^\top \mathbf{h} \mathbf{w}_r) + \mathbf{r} - (\mathbf{t} - \mathbf{w}_r^\top \mathbf{t} \mathbf{w}_r)\|_2^2$$



TransR

$$\begin{aligned} \mathbf{h}_r &= \mathbf{h} \mathbf{M}_r, \quad \mathbf{t}_r = \mathbf{t} \mathbf{M}_r \\ f_r(h, t) &= \|\mathbf{h}_r + \mathbf{r} - \mathbf{t}_r\|_{L1/L2} \end{aligned}$$

TransE原理

- 实体 e 和关系 r 都在同一个向量空间
- 运算就在原空间进行
- 需要学习的参数只有embedding
- 问题：太过简单直接，表达能力不够，对于一对多、多对一、多对多的情形表现不好

TransH原理

- 实体 e 和关系 r 仍在同一空间
- 运算时实体 e 投影到关系 r 对应的一个超平面上
- 使得遇到不同的 r ，同一个 e 也可能有不同的投影表示
- 每个 r 都需要学习一个超平面的法向量
- 需要学习的参数有embeddings, normal_vector_embedding
- 问题：投影操作太单一，实体和关系仍没有各自不同的空间

TransR原理

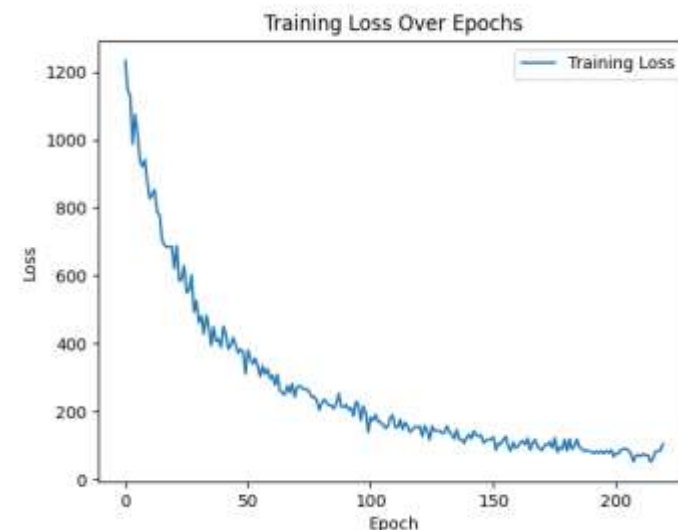
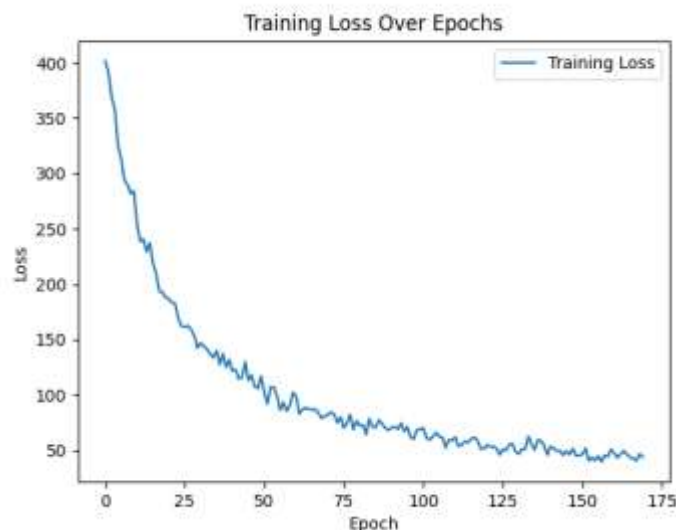
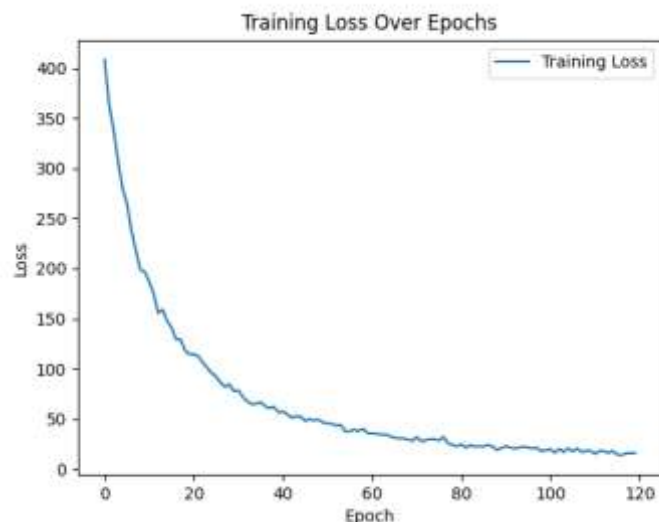
- 实体和关系分别在不同的空间
- 运算时实体 e 投影到关系 r 空间
- 每个 r 都需要学习一个映射矩阵 把 h 和 t 映射到 r 的空间

数据构造和读取

- 需要构造负例，随机选取更换头实体h或者尾实体t
- 需要用到random因此提前固定seed为42
- 每条正例构造三条负例
- 最终传入forward的时候把正例列表复制三遍，然后shuffle，分batch传入

模型训练

- 由于三个模型的参数量递增，收敛所需的epoch也增加，为了尽量让三个模型都收敛，分别选取了150, 200, 250的epoch值
- TransE的一个epoch只需几秒，TransR的一个epoch已需半分钟
- loss图像的绘制去掉了前30个epoch（数量级太大）



评测方式

- 根据验证集和测试集的三元组id，取出对应学习完的embedding，衡量f函数的值是否趋于0
- 设置了一个阈值，低于这个阈值的即视为预测正确
- 用accuracy作为衡量指标

评测结果

- 取threshold为5, 6, 7的对比

----- TransE Model -----	----- TransE Model -----	----- TransE Model -----
Validation Set Accuracy: 0.350499001996008	Validation Set Accuracy: 0.8495580268035358	Validation Set Accuracy: 0.9870544625035643
Test Set Accuracy: 0.34911560637154304	Test Set Accuracy: 0.8442783152545685	Test Set Accuracy: 0.9850483729111698
1-1 Test Set Accuracy: 0.4635416666666667	1-1 Test Set Accuracy: 0.7083333333333334	1-1 Test Set Accuracy: 0.9427083333333334
1-n Test Set Accuracy: 0.2088167053364269	1-n Test Set Accuracy: 0.7509667440061871	1-n Test Set Accuracy: 0.9767981438515081
n-1 Test Set Accuracy: 0.17395459976105138	n-1 Test Set Accuracy: 0.7252090800477897	n-1 Test Set Accuracy: 0.9753882915173238
n-n Test Set Accuracy: 0.40943498242768317	n-n Test Set Accuracy: 0.887875101378751	n-n Test Set Accuracy: 0.9890510948905109
----- TransH Model -----	----- TransH Model -----	----- TransH Model -----
Validation Set Accuracy: 0.16429997148560022	Validation Set Accuracy: 0.5974337040205304	Validation Set Accuracy: 0.920045623039635
Test Set Accuracy: 0.16290432913124206	Test Set Accuracy: 0.5967458223394899	Test Set Accuracy: 0.9151763901104271
1-1 Test Set Accuracy: 0.4322916666666667	1-1 Test Set Accuracy: 0.5833333333333334	1-1 Test Set Accuracy: 0.875
1-n Test Set Accuracy: 0.06419180201082754	1-n Test Set Accuracy: 0.4354215003866976	1-n Test Set Accuracy: 0.8499613302397525
n-1 Test Set Accuracy: 0.07311827956989247	n-1 Test Set Accuracy: 0.40908004778972523	n-1 Test Set Accuracy: 0.822700119474313
n-n Test Set Accuracy: 0.19343065693430658	n-n Test Set Accuracy: 0.6640984049743174	n-n Test Set Accuracy: 0.9475533928088673
----- TransR Model -----	----- TransR Model -----	----- TransR Model -----
Validation Set Accuracy: 0.012432278300541773	Validation Set Accuracy: 0.02566295979469632	Validation Set Accuracy: 0.04950099800399202
Test Set Accuracy: 0.01314375061076908	Test Set Accuracy: 0.02467507084921333	Test Set Accuracy: 0.04827518811687677
1-1 Test Set Accuracy: 0.3854166666666667	1-1 Test Set Accuracy: 0.390625	1-1 Test Set Accuracy: 0.4010416666666667
1-n Test Set Accuracy: 0.007733952049497293	1-n Test Set Accuracy: 0.01160092807424594	1-n Test Set Accuracy: 0.02397525135344161
n-1 Test Set Accuracy: 0.008841099163679809	n-1 Test Set Accuracy: 0.017682198327359618	n-1 Test Set Accuracy: 0.03082437275985663
n-n Test Set Accuracy: 0.010002703433360368	n-n Test Set Accuracy: 0.023046769397134362	n-n Test Set Accuracy: 0.05075696134090295

结果分析

- TransE模型反而是表现最好的一个，而且有时候对n-n的情况表现最好，很反直觉
- TransE收敛的时候loss是十几，TransH最后loss是五十几，TransR最后loss是一百出头，怀疑其实并未收敛，或者模型过于复杂反而起到了反作用
- 训练的充分可能比模型的种类更重要，该数据集的任务应该无需过于复杂的模型