卷积神经网络在知识图谱嵌入中的应用

邹长龙

（大连海事大学信息科学技术学院 辽宁大连）

摘要: 人们构建的知识库通常被表示为网络形式，节点代表实体。连边代表实体间的关系。在网络表示形式下，人们需要设计专门的图算法存储和利用知识库。存在费时费力的缺点，并受到数据稀疏问题的困扰。近期，以深度学习为代表的表示学习不仅为机器学习算法提供更好的性能， 同时也为知识表示提供了新思路。而这其中受到关注最多的就是卷积神经网络。本文将重点介绍卷积神经网络在知识表示学习上的应用。

关键字: 深度学习;表示学习;卷积神经网络

**The Application of Convolutional Neural Network in Knowledge Graph Embedding**

ZOU ChangLong

(School of Information Science and Engineering, Dalian Maritime University, Dalian China)

**Abstract:** The knowledge base constructed by people is usually expressed in the form of network, and nodes represent entities. Edges represent the relationship between entities. In the form of network representation, people need to design special graph algorithm to store and utilize knowledge base. It is time-consuming and laborious, and is troubled by the problem of data sparsity. Recently, representation learning represented by deep learning not only provides better performance for machine learning algorithms, but also provides new ideas for knowledge representation. Among them, convolution neural network is the most concerned. This paper will focus on the application of convolution neural network in knowledge representation learning.

Keyword: Deep learning; Representation Learning; Convolution Neural Network

1 引言

随着互联网技术的发展，不论是数据规模还是数据种类都呈现指数级的增长。当代人工智能不断发展时，人们便为互联网构建大量的知识库，这种知识库会将数据依据彼此间的关联性进行分层分类管理，使资料存储、管理及应用更加系统化，进而搜索人工智能学科的发展并开发智能信息服务应用。近年来，国内外互联网公司纷纷推出知识库产品。知识库的主要研究目标就是从无结构或半结构的互联网信息中获得有结构的知识，自动融合构建知识库、服务知识推理等相关应用。知识表示是知识获取与应用的基础，所以在构建整个知识库和应用知识库的过程中知识表示学习都是贯穿其中的重要研究课题。

2 知识表示学习概述

表示学习又称表征学习(Representation learning)，主要是利用机器学习技术自动获取每个实体或者关系的向量化表达， 旨在将描述对象表示为低维稠密向量( 即分布式向量)。分布式表示的对象均被表示成一个低维的稠密、实值向量，利用对象在空间的相对距离，反映它们之间的语义关系。两个对象离得越近，说明关系越紧密，两个对象离得越远，说明它们之间没有太强的关系。

将表示学习应用于知识表示中已经成为了现在知识图谱主流的研究方向，这样的表示学习方法又称之为知识表示学习或知识图嵌入。知识表示学习是将知识库中的实体和关系嵌入到连续的向量空间中，以便简化操作， 同时保持知识图谱的结构。

从形式上知识表示学习主要可以分成两类: 一种是基于结构的表示学习方法， 另一种是基于语义的表示学习方法，前者主要是从三元组的结构出发学习知识图谱中实体和关系的表示，而后者则是通过考虑文本语义来学习实体和关系的表示。

从发展来看， 目前知识表示学习的研究进程主要可以分成两个阶段，以2013年 Borders 等人受Mikolov发现的词向量空间中存在的平移不变性这一有趣现象的启发，从而提出的TransE模型为分割。在TransE之前主要包括结构表示、能量模型、矩阵分解模型等。在TransE之后，人们陆续提出了在其基础上加以改进的新Trans系列模型以及添加卷积的新模型。本文主要介绍添加卷积的新模型。

3 卷积介绍

可以把卷积想象成信息的混合。想象一下，两桶装满信息的东西被倒在一个桶里，然后根据一个特定的规则进行混合。每个信息桶都有自己的配方，它描述了一个桶中的信息如何与另一个桶中的信息混合。所以卷积是一个有序的过程，两个信息源相互交织。卷积也可以用数学来描述，事实上，它是一种像加法、乘法或导数一样的数学运算，虽然这个运算本身很复杂，但它对于简化更复杂的方程是非常有用的。

3.1 卷积的定义

我们称为和的卷积。其连续的定义为：



其离散的定义为：



3.2卷积在深度学习中的作用

从函数（或者说映射、变换）的角度理解：卷积过程是在图像每个位置进行线性变换映射成新值的过程，将卷积核看成权重，若拉成向量记为w，图像对应位置的像素拉成向量记为x，则该位置卷积结果为，即向量内积加上偏置，将x变换为y。从这个角度看，多层卷积是在进行逐层映射，整体构成一个复杂函数，训练过程是在学习每个局部映射所需的权重，训练过程可以看成是函数拟合的过程。

从模版匹配的角度理解：前面我们已经知道，卷积与相关在计算上可以等价，相关运算常用[模板匹配](https://wiki2.org/en/Template_matching)，即认为卷积核定义了某种模式，卷积（相关）运算是在计算每个位置与该模式的相似程度，或者说每个位置具有该模式的分量有多少，当前位置与该模式越像，响应越强。

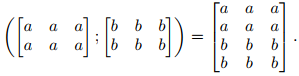
4 基于卷积的知识表示学习方法

本文将三元组表示为(s,r,o)，其中s代表头实体，t代表尾实体，r代表头实体与尾实体之间的关系。同时引进一个新的三元组(o,r\_reverse,s),该三元组为(s,r,o)的相反三元组。r与r\_reverse仅仅是不同的符号，没有特殊的意义，本文称r\_reverse为反向关系。例如小明的爸爸是小强可以表示为：(xiaoming,father\_of,xiaoqiang)，则与之对应的反关系三元组可以表示为(xiaoqiang,father\_of\_reverse,xiaomming)。同样的，我们定义知识库，其中代表知识库中实体的集合，其中包含|E|种不同的实体；代表知识库中关系的集合，包含|R|种不同的关系；S代表知识库种三元组的集合，。

4.1 ConvE

ConvE是将卷积神经网络用于知识图谱嵌入的开山之作。ConvE认为以前关于链接预测的工作集中在可以扩展到大型知识图的浅层快速模型上。然而，这些模型比深层的多层模型学习到的表达功能更少，这可能会限制一些性能。它是一个在嵌入上使2D卷积来预测知识图中缺失链接的模型。ConvE是用于链路预测的最简单的多层卷积体系结构：它由单个卷积层、嵌入维的投影层和内积层定义。

在ConvE中，作者首先介绍了一维卷积与二维卷积在相互作用数上的差别，例如：我们连接两个维数为3的1D嵌入: 如上为一个过滤器大小为k=3的填充1D卷积将能够在连接点周围模拟这两个嵌入之间的交互作用数。如果我们将两行二维嵌入连接起来对于维数m×n，其中m=2和n=3，我们得到以下内容：



卷积核大小为3×3的填充二维卷积将能够模拟整个连接线周围的相互作用数。因此，与1D卷积相比，2D卷积能够提取两个嵌入之间更多的特征交互。

ConvE是一个卷积神经网络链路预测模型，其中输入实体和关系之间的交互是由卷积层和全连接层建模。评分函数如下：



其中是一个取决于r的关系参数，和是和的2D重构，k为嵌入维数。

在前馈过程中,模型在两个嵌入矩阵中做了一个行向量的查找操作，分别是实体和关系的查找，k是实体和关系的嵌入维数，|E| ，|R|代表着实体和关系的数量。然后这个模型链接和，并且使用他们作为一个2D卷积层的输入，卷积核为*w*。该隐藏层返回一个特征映射张量，其中c是尺寸为m和n的二维特征映射的数量，张量随后被重构为向量*vec*(*T*),然后使用矩阵W进行参数化的线性变换将其投影到k维空间。并通过内积与嵌入对象匹配。具体过程图示如图1。

对于训练模型的参数，ConvE应用sigmoid函数去归一化评分，即。损失函数如下：

其中，t是标签向量，其维数为（表示1-1评分），（表示1-N评分）；向量t的元素表示存在的关系，否则为零。

每层的输出要先经过非线性函数ReLU的作用，以便更快地训练，并在每层之后进行批量归一化，以提高融合速率。并且分阶段对模型进行dropout。特别是，作者在嵌入、卷积运算后的特征映射和完全连接层后的隐藏单元上使用了dropout。使用Adam作为优化器，并使用标签平滑来减少由于标签输出非线性饱和而导致的过度拟合。

CovnE的链路预测方式与传统的距离模型如：Trans系列，张量分解模型不同。与其他将实体对和关系作为三元组（s，r，o）并对其进行评分（1-1评分）的其他链接预测模型不同，我们采用一个（s，r）对，同时对所有实体o∈E进行评分（1-N评分）。

作者在这片文章中发现了TransE作者提取的数据集的缺点，并给出了修改之后的数据集。可以看到ConvE的缺点也是十分的明显，无法预测关系(r)。

4.2 InteractE

InteractE认为ConvE可以捕获的交互数量是有限的。InteractE是增加交互的数量是来影响链路预测性能的。InteractE在ConvE提出了3个用于增加交互的观点如下：1.特征排列2.“棋盘”重塑3.循环卷积。IteractE的评分函数如下：

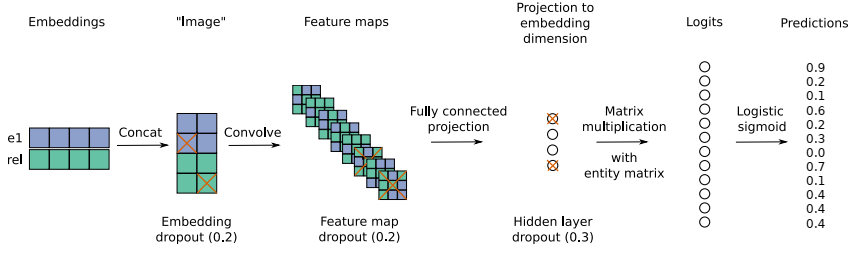


图1



代表循环卷积，*vec*(*·*)表示向量链接，表示宾语实体的嵌入和*W*是一个可学习的权重矩阵。f和g表示ReLU和sigmoid。损失函数采用交叉熵损失函数，和CovnE同样，也采用了标签顺滑。

为了捕获更多的的异构交互，InteractE首先产生*es*和*er*的t随机排列，记。其次进行“棋盘”式的重塑，重塑的图示如图2。

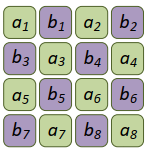


图2

2D重构增强了实体和关系嵌入之间的相互作用，这已被发现有助于学习更好的表示。最后采用循环卷积进行计算，与标准卷积相比，循环卷积允许捕捉更多的特征交互。卷积以深度方式在不同的输入排列上执行。InteractE将每个重新成形的排列堆叠为一个单独的通道。对于卷积排列，我们以深度方式进行循环卷积。虽然不同的滤波

器组可以应用不同的排列，但在实践中，InteractE发现跨通道共享滤波器的效果更好，因为它允许在更多的输入实例上训练一组内核的权重。

IteractE的执行过程图示如图3。

4.3 HypER

HypER是一种超网络体系结构，可以生成简化特定于关系卷积滤波器，它在标准数据集上的性能优于ConvE 和所有以前的方法；且可以被框定为张量因子分解，从而在一个已经建立的因子分解模型中进行链路预测。HypER也认为2D卷积对于从实体和关系嵌入中提取信息是有效的。

HypER是一个使用超网络为每个关系生成卷积滤波器权重的模型。它生成特定于关系的过滤器权重来处理输入实体，并在知识图中实现跨关系的多任务知识共享。HypER的超网络生成特定于关系的过滤器，从而从主实体(s)嵌入中提取特定于关系的特征。HypER实际上属于一大类张量因子分解模型，这在表达能力和要学习的参数数量之间提供了一个很好的折衷。

HypER模型没有向CovnE一样的向量的整形或拼接，因此避免了在嵌入中隐含任何固有的2D结构和将交互限制在拼接边界。取而代之的是，HypER用超网络生成的特定关系卷积滤波器对嵌入的主题实体的每个维度进行卷积。这样，实体和关系嵌入以非线性

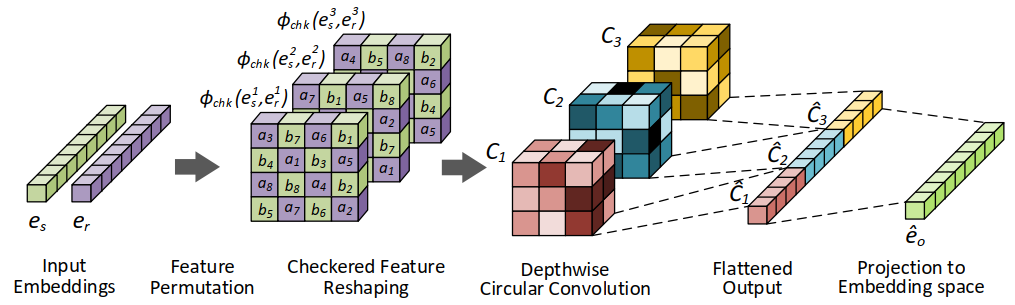


图3

(二次)方式组合，不像ConvE中的线性组合(加权和)。这给了HypER更多的表达能力，同时也减少了参数。

HypER的评分函数为：



 表示将向量重新构造成矩阵，f为ReLU函数。

在前馈过程中，模型从实体和关系嵌入矩阵和 获得输入三元组的嵌入。超网络是一个全连接层 （lf表示滤波器长度，nf表示每个关系的滤波器数量，即卷积的输出信道）应用于关系嵌入。结果被重新整形以生成卷积滤波器矩阵。虽然滤波器组的整体维数是lf\*nf，但是秩被限制为dr，以鼓励关系之间的参数共享。其中de，dr为实体和关系的嵌入维数。

将嵌入的主体实体e1与一组特定关系滤波器Fr卷积，得到二维特征映射，其中lm=de−lf+1为特征映射长度。将特征映射矢量化为，并且利用权矩阵投影到de维空间，在应用ReLU激活函数后，将结果与嵌入的每个对象实体的内积相结合，其中i在数据集中的所有实体上变化（大小为ne），并给出一个分数向量。将sigmoid按元素应用于得分向量，以获得每个预期三元组为真的预测概率。

损失函数采用的交叉熵损失函数：

y是标记向量，并且也在训练的过程中采用了标签平滑。

HypER的执行过程如图4。

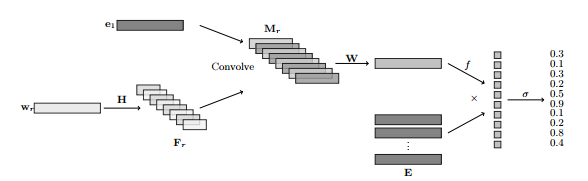


图4

评分过程与CovnE一样也采用的1-N评分。由于HypER在计算的过程中采用了分组卷积，也就是上文中说的超网络，所以HypER的参数数量远小于ConvE。

5 总结

通过以上对卷积在知识图谱嵌入中方法的简要介绍，我们知道知识表示学习对知识库的构建和融合以及提升知识的获取和推理能力具有重大的意义。对于知识图谱来说知识表示学习是最基础且最有研究价值的。可以看到卷积神经网络在知识图谱嵌入上也可以很好的应用前景。

参考文献

[1] Dettmers T , Minervini P , Stenetorp P , et al. Convolutional 2D Knowledge Graph Embeddings[J]. 2017.

[2] Vashishth S , Sanyal S , Nitin V , et al. InteractE: Improving Convolution-Based Knowledge Graph Embeddings by Increasing Feature Interactions[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(3):3009-3016.

[3] Balazevic I , Allen C , Hospedales T M . Hypernetwork Knowledge Graph Embeddings[J]. 2018.

[4] 刘知远, 孙茂松, 林衍凯,等. 知识表示学习研究进展[J]. 计算机研究与发展, 2016, 53(2):247-261.

[5] 王子悦, 陈华辉. 知识表示学习综述[J]. 无线通信技术, 2019, v.28;No.112(04):59-64.