

This is a note for the paper: Diffusion-Convolutional Neural Networks(2016)

概念辨析

adjacency matrix:邻接矩阵， $N \times N$ ，描述了节点间的连接关系

affinity(类同、类似的意思) matrix:关联矩阵， $N \times N$ ，描述了节点间的相似度，用于spectral clustering

incidence(关联，接合的意思) matrix：关联矩阵， $N \times E$ ，描述了节点和边的关系，即节点是否为边的端点

方法

基于一歩转移概率矩阵 P 定义图上的扩散卷积(diffusion convolution)，这里所谓的卷积，就是一种局部运算。 n 步转移概率矩阵就是 P^n 。

$P^n x, n = 1, 2, 3 \dots$ 可以看作不同深度的扩散操作。将这些操作的结果拼在一起，乘上逐元素系数 W_{ij} ，经过激活函数，再通过全连接网络即得到输出。一个式子表达出来就是：

$$Z = f(W \odot P^* X)$$

其中 $X \in R^{N \times d}$ ， N 为节点个数， d 为节点特征维度

$P^* \in R^{N \times H \times N}$ ， H 为扩散深度，不同的下标对应步数不同的转移概率矩阵。

注： W 为可学习参数。前向传播没有特征维度的变化，只有特征的局部融合。 H 决定了融合的深度，即距离节点 H 以内的特征会被融合。

我觉得这个方法可以被归为是基于random walk的。

实验

作者的实验做的非常详细，用了三种baseline，一种采用加正则化的logistic regression,即只采用节点特征信息做分类;一种采用基于图的kernel method,只利用图的结构信息;左后一种采用CRF-LBP，同时利用节点特征和图的结构。

节点分类

采用citation network数据集，在一个图上进行节点分类，图上节点划分为训练、验证、测试。Paper特征为0-1词袋，或tf-idf

H=2的DCNN就表现非常好，超过了其他方法。H到3时性能会饱和。

图分类

采用基于化学分子式进行性质预测的数据集，对化学分子进行分类，原子特征为一个高维向量，具体细节不太了解。

DCNN表现不好，H扩大对性能没有提升。作者因此说：DCNN对提取图的整体信息无能为力。

DCNN缺点

因为涉及概率转移矩阵，DCNN需要的内存太大，扩展性差，N只能为几万到几十万。

DCNN只能提取局部信息，对远程信息无能为力。