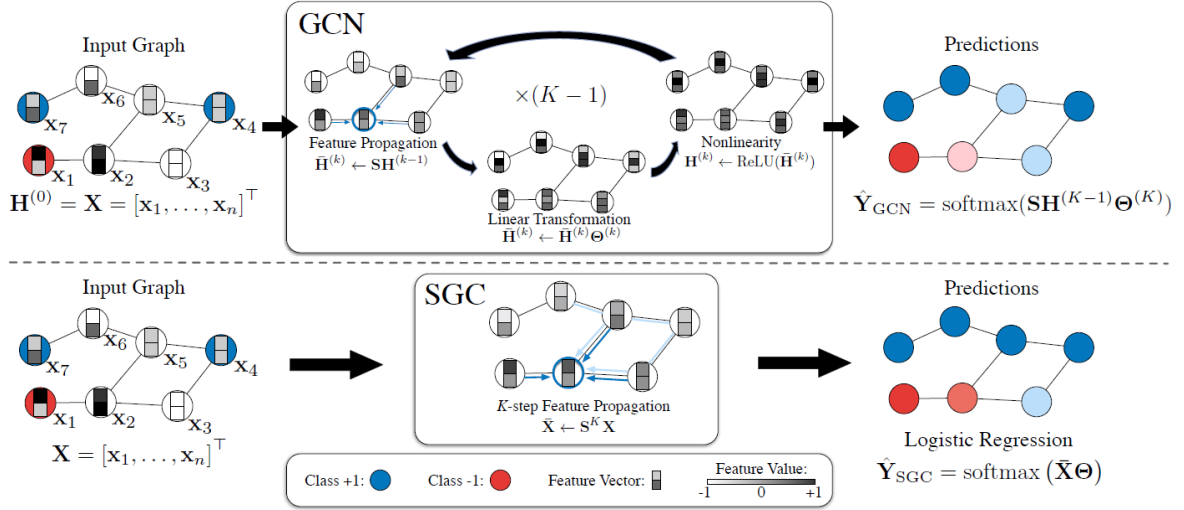


解决的问题

图卷积网络中可能引入了一些不必要的复杂性及冗余的计算

做法及创新



1. 移除图卷积网络各层之间的非线性关系，合并各层之间的权重矩阵

原始图卷积网络

对于一个输入的图，图卷积网络利用多层网络为每个顶点的特征 x_i 学习一个新的特征表示，随即输入一个线性分类器。对第 k 层网络，输入为 $H^{(k-1)}$ ，输出为 $H^{(k)}$ ，其中 $H^{(0)} = X$ 。一个 K 层的图卷积网络等价于对图中每个顶点的特征向量 x_i 应用一个 K 层感知机，不同之处在于顶点的隐层表示local averaging:

$$h_i^{(k)} \leftarrow \frac{1}{d_i + 1} h_i^{(k-1)} + \sum_{j=1}^n \frac{a_{ij}}{\sqrt{(d_i + 1)(d_j + 1)}} h_j^{(k-1)}$$

矩阵形式:

$$S = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$$

其中 $A = A + I$ ，则隐层表示用矩阵的形式表示为:

$$H^{(k)} \leftarrow S H^{(k-1)}$$

Local averaging: this step smoothes the hidden representations locally along the edges of the graph and ultimately encourages similar predictions among locally connected nodes

$\Theta^{(k)}$ 为第 K 层网络的权重矩阵:

$$H^{(k)} \leftarrow \text{ReLU}(H^{(k)} \Theta^{(k)})$$

$Y \in \mathbb{R}^{n \times C}$, y_{ic} 表示第 i 个顶点属于类别 C 的概率

$$Y_{GCN} = \text{softmax}(S H^{(K-1)} \Theta^{(K)})$$

简化图卷积网络

在传统的多层感知机中，多层网络可以提高模型的表现力，是因为这样引入了特征之间的层级关系，例如第二层网络的特征是以第一层网络为基础构建的。而在图卷积网络中，这还有另外一层含义，在每一层中顶点的隐层表示都是以一跳的邻居进行平均，经过 K 层之后，一个顶点就能获得 K 跳邻居的特征信息。这类似于在卷积网路中网络的深度提升了特征的receptive field。

保留local averaging，移除了非线性激活函数：

$$Y = \text{softmax}(S^K X \Theta^{(1)} \dots \Theta^{(K)})$$

其中 S^K 可以预先进行计算，大大减少了模型的训练时间

论文中证明了简化后的图卷积网络等价于谱空间的一个低通滤波器，它通过的低频信号对应于图中平滑后的特征

数据集

Cora、Citeseer、Pubmed、Reddit