

【动手做】 使用图注意力网络GAT 分类

AchieveFun

讨论内容

0 工具准备

- Python和Anaconda
- DGL-Deep Graph Library

1 知识准备

- 注意力机制
- 图注意力网络

2 案例演示

- 代码运行流程
- 代码演示

3 参考资料

4 动手做系列之图神经网络模型

AchieveFun

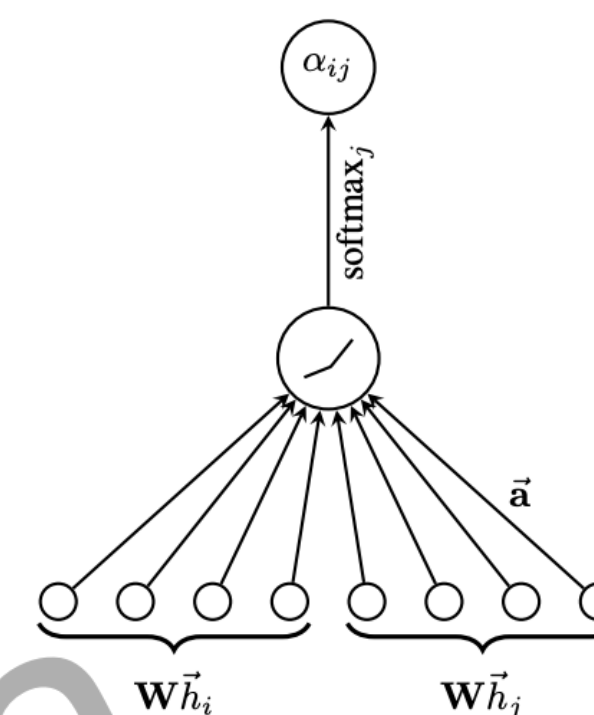
1 知识准备：注意力机制

- 参考幻灯片 <https://github.com/bettermorn/IAICourse/blob/main/courseware/%E5%B7%A5%E4%B8%9A%E6%99%BA%E8%83%BD%E5%AE%9E%E6%88%985-%E7%94%9F%E4%BA%A7%E8%BF%90%E8%90%A5%E4%BC%98%E5%8C%96%E4%B8%8E%E6%9C%BA%E5%99%A8%E7%BF%BB%E8%AF%91.pdf>
- 视频： <https://www.bilibili.com/video/BV1a24y1F73K> 【专题5:生产运营优化与机器翻译原理】 2.排程排产优化问题与Seq2seq 2小时25分左右 注意力机制部分

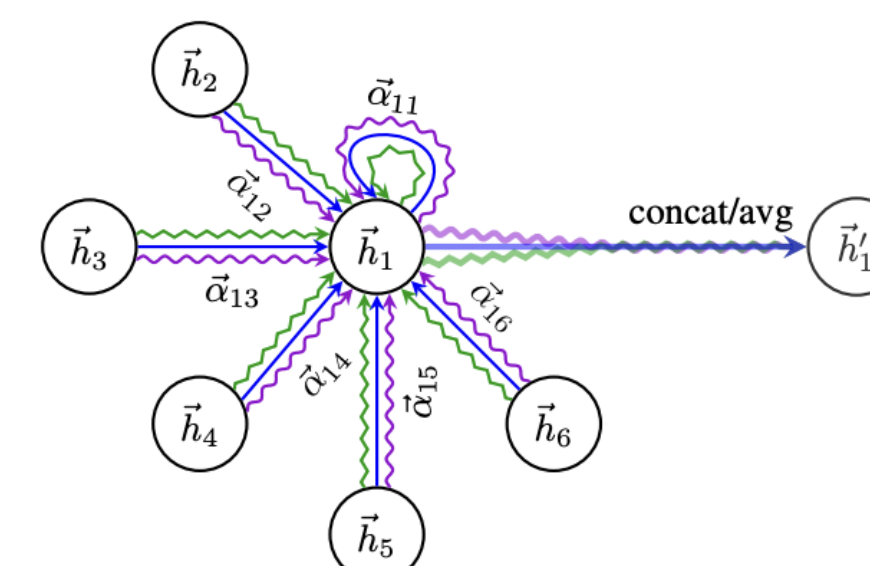
1 知识准备：图注意力网络

Published as a conference paper at ICLR 2018

- 在图卷积神经网络聚合周围邻居特征的基础上，考虑了不同邻居节点对当前节点特征学习的不同重要性。
- 带权重的聚合操作：为某个节点的邻居节点学习不同的权重
- 操作高效：每个“节点-邻居节点对”之间是并行计算
- 模型可泛化到新的图谱上，用归纳的方法分类节点：以共享方式应用于图的所有边，不必事先知道全局图结构或其所有节点的特征。



注意力机制



节点 1 对其邻域的多头关注 (K = 3 头) 示意图

$$\vec{h}'_i = \sigma \left(\sum_{j \in \mathcal{N}_i} \alpha_{ij} \mathbf{W} \vec{h}_j \right)$$

$$\alpha_{ij} = \frac{\exp \left(\text{LeakyReLU} \left(\vec{a}^T [\mathbf{W} \vec{h}_i \| \mathbf{W} \vec{h}_j] \right) \right)}{\sum_{k \in \mathcal{N}_i} \exp \left(\text{LeakyReLU} \left(\vec{a}^T [\mathbf{W} \vec{h}_i \| \mathbf{W} \vec{h}_k] \right) \right)}$$

图片来源: <https://arxiv.org/pdf/1710.10903.pdf>

GRAPH ATTENTION NETWORKS

图注意力网络：区分每个邻居节点的贡献度

- We present graph attention networks (GATs), novel neural network architectures that operate on graph-structured data, leveraging masked self-attentional layers to address the shortcomings of prior methods based on graph convolutions or their approximations. By stacking layers in which nodes are able to attend over their neighborhoods' features, we enable (implicitly) specifying different weights to different nodes in a neighborhood, without requiring any kind of costly matrix operation (such as inversion) or depending on knowing the graph structure upfront. In this way, we address several key challenges of spectral-based graph neural networks simultaneously, and make our model readily applicable to inductive as well as transductive problems. Our GAT models have achieved or matched state-of-the-art results across four established transductive and inductive graph benchmarks: the Cora, Citeseer and Pubmed citation network datasets, as well as a protein-protein interaction dataset (wherein test graphs remain unseen during training).

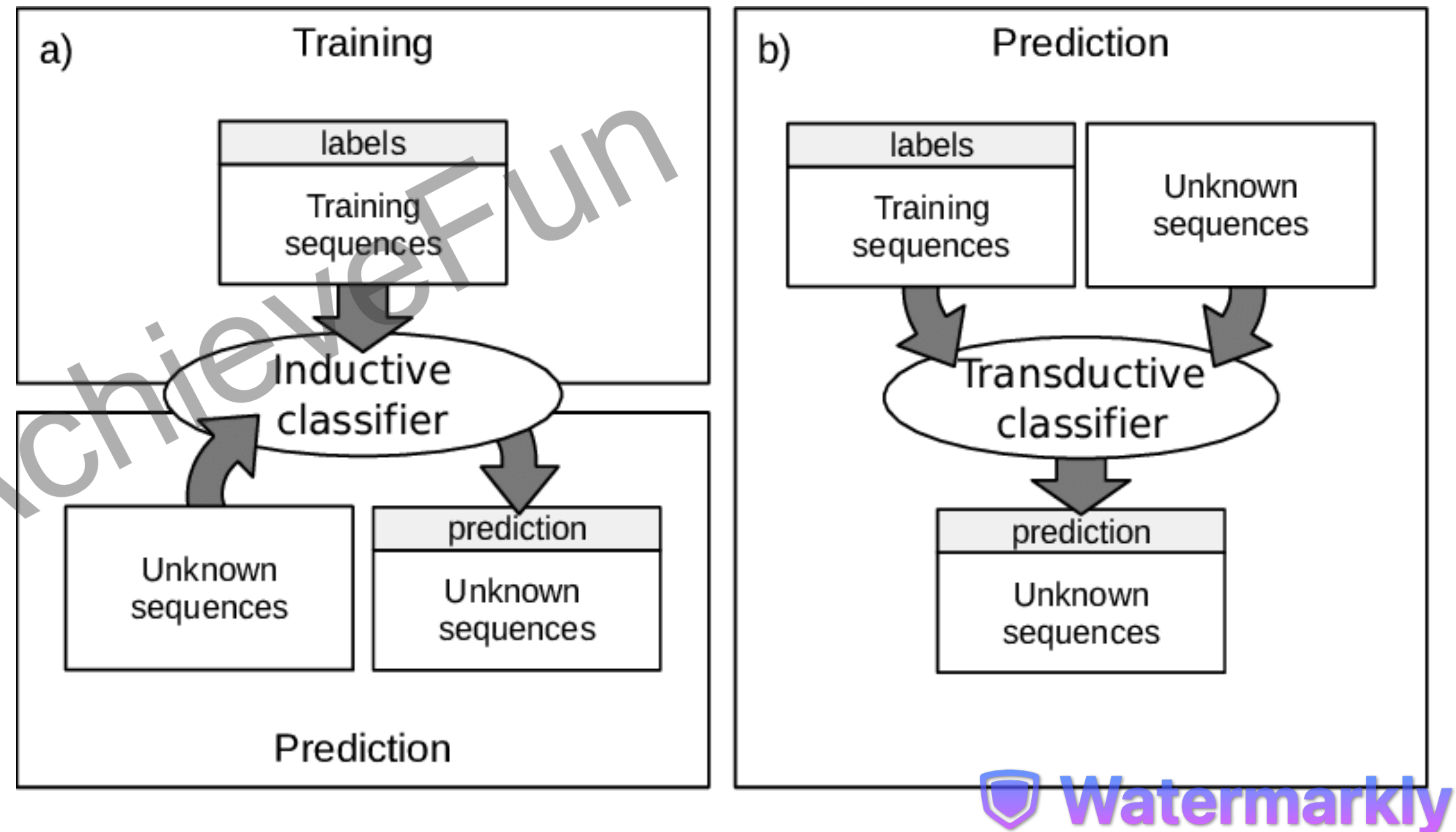
- 来源: <https://arxiv.org/pdf/1710.10903.pdf>

机器翻译：

我们提出了图注意力网络（GATs），这是一种新型的神经网络架构，可对图结构数据进行操作，利用遮蔽的自我注意力层来解决之前基于图卷积或其近似值的方法的不足之处。通过堆叠节点能够关注其邻域特征的层，我们能够（隐含地）为邻域中的不同节点指定不同的权重，而不需要任何形式的高成本矩阵操作（如反转），也不依赖于对图结构的预先了解。通过这种方式，我们同时解决了基于谱的图神经网络所面临的几个关键挑战，并使我们的模型可随时应用于归纳和直推式问题。我们的 GAT 模型在四个已确立的转导和归纳图基准中取得了与一流水平相当的结果：Cora、Citeseer 和 Pubmed 引用网络数据集，以及蛋白质相互作用数据集（其中测试图在训练期间保持未见）。

Transductive learning vs inductive learning

- Transductive learning: 模型在一个图上进行训练，其中一部分节点的标签是已知的，而另一部分节点的标签是未知的。利用图中的结构信息和节点特征，以及已知标签的节点信息，来预测未知标签节点的标签。适合于那些图结构不会发生太大变化的场景，比如社交网络分析、推荐系统等。
- Inductive learning: 指模型通过训练集上的样本学习到一般性的规律或模式，然后将这些学到的规律应用到新的、未见过的数据上进行预测的过程。适用于图数据会不断增长或变化的场景，例如，推荐系统中不断添加新用户和新商品，或者在不同的生物网络中预测蛋白质的功能。



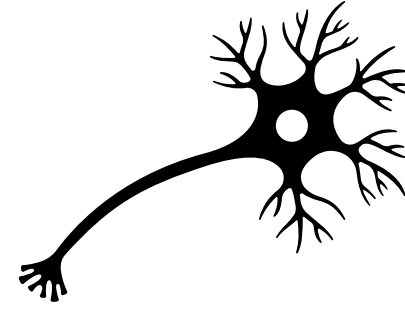
2 案例演示

- 案例来源：
 - <https://paperswithcode.com/paper/graph-attention-networks>
 - https://docs.dgl.ai/tutorials/models/1_gnn/9_gat.html
- 代码
 - <https://github.com/dmlc/dgl/tree/master/examples/pytorch/gat> 包括dropout, skip connections, and hyper-parameter tuning
 - https://docs.dgl.ai/downloads/5debe8ae501b96c82cb090e18a7c302f/9_gat.py
 - https://docs.dgl.ai/downloads/30dde79d1418f8a8224fa315d0f97a95/9_gat.ipynb

学习目标

- 理解GAT 只是一种不同的聚合函数，它关注的是邻域的特征，而不是简单的均值聚合。
- GraphSAGE：简单地对所有相邻节点特征进行平均。
- GCN：聚合方式与结构有关，可能影响通用性。邻居节点特征的归一化总和。
- GAT：引入了注意力机制来替代静态归一化卷积操作。
- 多头注意力：用于丰富模型容量并稳定学习过程。中间层串联法，最终层平均法。
- 可视化和理解学习到的注意力：层数越高，注意力越敏锐。
- GAT 的优势可能更多在于其处理具有更复杂邻域结构的图的能力。
- 来源： https://docs.dgl.ai/tutorials/models/1_gnn/9_gat.html

环境准备



1. 安装Anaconda <https://www.anaconda.com/download/>
2. 在Anaconda建立环境dgl, 根据dgl文档(<https://www.dgl.ai/pages/start.html>)确定**Python版本**
3. 根据代码, 使用pip install -r requirements.txt 安装依赖库并验证库的存在pip list
 - o torch
4. 安装dgl <https://www.dgl.ai/pages/start.html>

例如: conda install -c dgteam dgl

2.1 代码运行流程

- 问题：在一个图（如引文网络）中对节点（如文档）进行分类，在这个图中，只有一小部分节点有标签。
 - Cora数据集由 2708 篇科学出版物组成，分为七类。引用网络由 5429 个链接组成。数据集中的每份出版物都由一个 0/1 值的单词向量来描述，表示字典中相应单词的缺失/存在。字典由 1433 个独特的单词组成。
- 流程：
 - 定义GATLayer(包括edge_attention, reduce_func)
 - 定义MultiHeadGATLayer
 - 定义两层GAT模型
 - 装载数据集Cora
 - 训练模型（与GCN类似）
 - 评估模型

2.2 代码演示

- https://docs.dgl.ai/downloads/5debe8ae501b96c82cb090e18a7c302f/9_gat.py
- https://docs.dgl.ai/downloads/30dde79d1418f8a8224fa315d0f97a95/9_gat.ipynb
- <https://github.com/dmlc/dgl/tree/master/examples/pytorch/gat> 包括dropout, skip connections, and hyper-parameter tuning

3 参考资料

- https://docs.dgl.ai/tutorials/models/1_gnn/9_gat.html
- <https://arxiv.org/pdf/1710.10903.pdf>
- <https://github.com/dmlc/dgl/tree/master/examples/pytorch/gat> 包括dropout, skip connections, and hyper-parameter tuning
- https://docs.dgl.ai/downloads/5debe8ae501b96c82cb090e18a7c302f/9_gat.py
- https://docs.dgl.ai/downloads/30dde79d1418f8a8224fa315d0f97a95/9_gat.ipynb

4 动手做系列之图神经网络模型

- GraphSAGE: 链接预测
- GCN图卷积网络: 分类
- GAT图注意力网络: 分类
- VGAE 图变分自编码模型: 链接预测
- DiffPool: Hierarchical Graph Representation Learning with Differentiable Pooling 利用可变池进行层次图表示学习
- GCC: Graph Contrastive Coding for Graph Neural Network Pre-Training 图对比学习模型 自监督学习
- MVGRL: Contrastive Multi-View Representation Learning on Graphs 图上的对比多视图表征学习
- HGT: Heterogeneous Graph Transformer 异构图转换器
- Text Generation from Knowledge Graphs with Graph Transformers 利用图转换器从知识图谱生成文本
- GPT-GNN: Generative Pre-Training of Graph Neural Networks 生成式图网络预训练框架

总结

0 工具准备

- Python和Anaconda
- DGL-Deep Graph Library

1 知识准备

- 注意力机制
- 图注意力网络

2 案例演示

- 代码运行流程
- 代码演示

3 参考资料

4 动手做系列之图神经网络模型

AchieveFun