

论文摘要

随着社交应用的大规模发展, 通过跨网络用户对齐实现用户信息连接, 提升用户理解, 增强推荐效果, 精准投放个性化广告成为研究热点。目前大多数的跨网络用户对齐模型利用单一类型的信息作为推断依据, 采用表示学习和用户匹配两个阶段实现用户对齐。由于缺乏对多类型数据的有效融合, 并且独立地实现表示学习和用户匹配流程, 导致这些跨网络用户对齐方法匹配精度不足。针对以上问题本文提出了一个基于图注意力网络的跨网络用户对齐模型, 融合用户属性和网络结构信息, 并将表示学习环节与用户对齐环节串联实现了端到端的预测。在用户对齐部分, 本文还分别讨论了基于二分类和基于距离度量的方法实现。在实验验证阶段, 将论文提出的模型应用于引文网络以及真实社交网络数据集, 并与领域的基准方法进行对比, 结果表明新模型在多项指标上均有明显提升。同时, 通过对比本文还发现基于距离度量的模型能够获得更好的对齐效果。

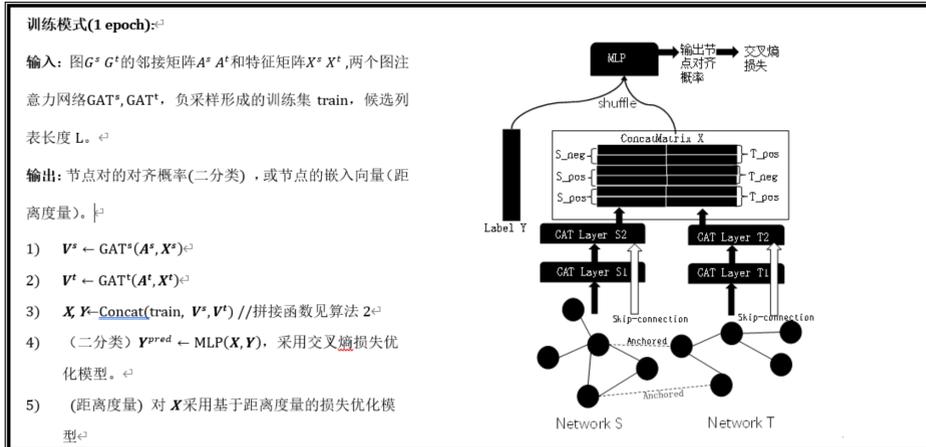
系统模型

图注意力网络

论文简介

1. 使用图注意力网络, 同时学习节点属性、网络结构信息和不同邻居的重要程度, 获得跨网络表示结果, 降低由用户跨网络属性不一致所产生的噪声干扰。
2. 整合表示学习流程和用户对齐流程, 实现端到端的预测, 提升跨网络用户对齐模型学习效率。
3. 在实验中与其他基准模型对比, 验证了本文所提模型的有效性, 并总结了本文所提模型的最优结构。

算法原理



实验仿真

	Citeseer ^{s,t}	Facebook ^{s,t}
MNA	0.2753	0.2601
SVM	0.3949	0.3787
PALE	0.5941	0.6186
ABNE	0.6510	0.6674
GAT-b	0.6594	0.6653
GAT-d	0.7072	0.6806

	citeseer ^{s,t}	cora ^{s,t}	pubmed ^{s,t}	Facebook ^{s,t}
MNA	0.3891	0.3713	0.3424	0.2947
PALE	0.6017	0.6097	0.5805	0.6236
ABNE	0.6715	0.6561	0.6592	0.6883
GCN-d	0.6734	0.6622	0.6415	0.6553
GAT-d	0.7188	0.7039	0.6971	0.6892

论文结论

本文提出基于图注意力网络的模型, 可以实现端到端解决跨网络用户对齐问题, 并比以往工作取得了明显的性能提升。在表示学习阶段采用两层图注意力网络, 在用户对齐阶段使用基于距离度量的损失指导模型的优化。该方法可以较好地去除用户属性信息的噪声, 并整合用户属性与网络的局部特征、全局特征。下一步的工作是对齐节点对在隐空间彼此距离最近为目标, 设计一种能够融合两个网络信息且能更好地学习网络结构的算法, 并在大规模数据集上实现快速计算。