图数据表示学习综述论文

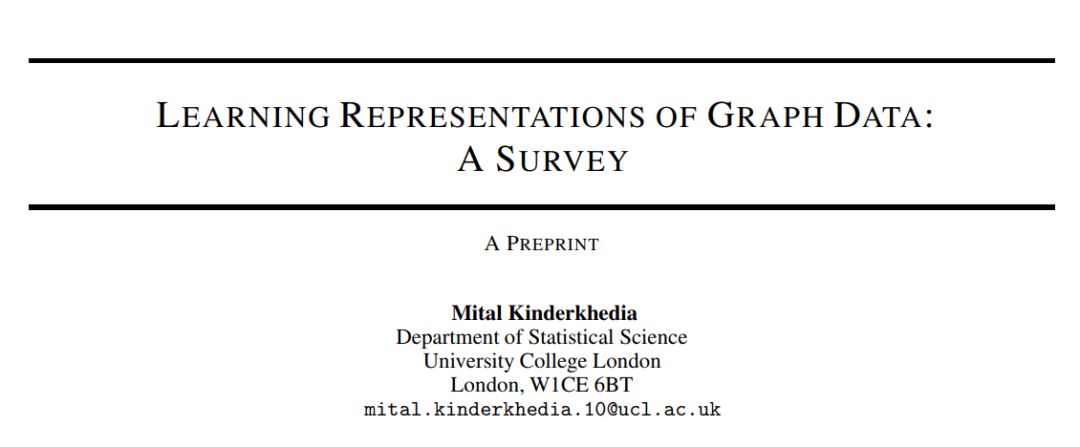
[专知](javascript:void(0);) 今天

导读

本文总结和讨论学习图数据表示方法的最新进展，包括核方法、卷积方法、图神经网络方法、图嵌入方法和概率方法等。

作者 | Mital Kinderkhedia

编译 | Xiaowen



深度神经网络在目标识别、图像分类和自然语言处理等领域取得了巨大的成功。然而，设计能够学习和输出任意图的最优神经网络结构是一个正在研究中的问题。

本综述的目的是总结和讨论学习图数据表示方法的最新进展。我们首先

确定了常用的图数据类型，并回顾了图论的基本知识。然后讨论了

图核方法与神经网络之间的关系。接下来，我们确定了

用于学习图数据表示的主要方法：

核方法、卷积方法、图神经网络方法、图嵌入方法和概率方法。

讨论了每一种方法下的各种变形方法，最后简要讨论了

学习图数据表示的前景。

将数据构造成图表有助于系统地发现复杂的关系和模式。例如，考虑到网页之间的链接频率很高，万维网图表显示了复杂的结构。在自然语言处理中，有时使用树来表示文本，以理解单词之间的联系，从而推断句子的含义。然而，机器学习方面的研究主要集中在以向量形式表示的数据上，真实世界的数据不容易表示为向量。具有复杂图形结构的现实世界场景包括分子和生物网络、计算机网络、传感器网络、社会网络、引用网络、电网和交通网络。使用基于图的表示，可以捕获结构化数据的顺序、拓扑、几何和其他关系特征。

神经网络是通用的函数逼近器。近年来，深度学习模式在语音识别、对象识别检测和学习自然语言处理方面取得了巨大的成功。此外，成分的融合：大型数据集、先进的计算处理能力以及机器学习方法中的新兴研究极大地促进了深度学习研究。机器学习的神经和非神经方法之间的主要区别在于数据的学习表示。数据对象的表示被定义为作为一组提供与给定学习任务相关的学习信息的信息片段。在机器学习术语中，我们使用术语特征，而在表示学习术语中，我们关注的是学习数据的表示，使得更容易使用学习信息来执行诸如预测、分类等任务的学习信息。

学习图表示的思想是：

学习映射，从而将顶点、子图或整体图嵌入到低维向量空间中的点中。然后对这些映射进行优化，使它们能够反映嵌入空间内的几何结构，

然后所学习的嵌入可以用作机器学习任务的矢量输入。

本研究的贡献在于提出了一种主要的学习图表示方法的分类法：核方法、卷积方法、图神经网络方法、图嵌入方法和概率方法。本文对这些方法所使用的技术进行了比较、对比和概述。然而，这项调查并非详尽无遗。