

第 3 章

图神经网络

Lingfei Wu、Peng Cui、Jian Pei、Liang Zhao 和 Le Song^①

摘要

深度学习已经成为当今人工智能研究的主要方法之一。尽管传统的深度学习技术在图像等欧氏数据或文本和信号等序列数据上取得了巨大的成功，但仍有大量的应用可以自然地或最优地用图结构来表征。这一差距推动了对图深度学习的研究热潮，其中，图神经网络（GNN）在应对大量应用领域的各种学习任务方面非常成功。在本章中，我们将沿着三条轴线系统地梳理现有的 GNN 研究——基础、前沿和应用。首先，我们将介绍 GNN 的基本方法，从主流的模型及其表达能力到 GNN 的可扩展性、可解释性和鲁棒性。接下来，我们将讨论各种前沿研究，从图分类和链接预测到图生成、图转换、图匹配和图结构学习。在此基础上，我们将进一步总结在大量应用中充分利用各种 GNN 方法的基本流程。最后，我们将展示本书的组织结构并总结 GNN 的各种研究课题的路线图。

3.1 导读

传统的深度学习技术，如循环神经网络（Schuster and Paliwal, 1997）和卷积神经网络（Krizhevsky et al, 2012），已经在图像等欧氏数据或文本和信号等序列数据上取得巨大的成功。然而，在丰富的科学领域，现实世界中许多重要的对象和问题可以自然地或最优地用复杂的图结构来表达，如社交网络、推荐系统、药物发现和程序分析中的图或流形结构。一方面，这些图结构的数据可以编码复杂的点对关系，以学习更丰富的信息表征；另一方面，原始数据（图像或连续文本）的结构和语义信息中纳入的特定领域知识可以捕捉数据之间更细粒度的关系。

近年来，图深度学习引发了研究界的广泛兴趣（Cui et al, 2018; Wu et al, 2019e; Zhang et al, 2020e）。其中，图神经网络是非常成功的学习框架，可以应对大量应用中的各种任务。新提出的图结构数据的神经网络架构（Kipf and Welling, 2017a; Petar et al, 2018; Hamilton

^① Lingfei Wu

Pinterest, E-mail: lwu@email.wm.edu

Peng Cui

Department of Computer Science, Tsinghua University, E-mail: cuip@tsinghua.edu.cn

Jian Pei

Department of Computer Science, Simon Fraser University, E-mail: jpei@cs.sfu.ca

Liang Zhao

Department of Computer Science, Emory University, E-mail: liang.zhao@emory.edu

Le Song

Mohamed bin Zayed University of Artificial Intelligence, E-mail: dasongle@gmail.com

et al, 2017b) 在一些著名的领域, 如社交网络和生物信息学等, 已经取得令人瞩目的成果。它们还渗透到其他科学研究领域, 包括推荐系统 (Wang et al, 2019j)、计算机视觉 (Yang et al, 2019g)、自然语言处理 (Chen et al, 2020o)、程序分析 (Allamanis et al, 2018b)、软件挖掘 (LeClair et al, 2020)、药物发现 (Ma et al, 2018)、异常检测 (Markovitz et al, 2020) 以及智慧城市 (Yu et al, 2018a) 等。

尽管现有的研究已经取得一些成就, 但是当将 GNN 用于为随时间演化、多关系和多模态的高度结构化数据建模时, 仍然面临许多挑战。要在图和其他高度结构化的数据 (如序列、树和图) 之间建立映射模型也非常困难。图结构数据面临的一个挑战是, 它们的空间局部性和结构不像图像或文本数据那么强。因此, 图结构数据自然不适合高度规则化的神经网络结构, 如卷积神经网络和循环神经网络。

更重要的是, 现实世界中新出现的 GNN 应用领域为 GNN 带来了巨大的挑战。图提供了一种强大的抽象, 可以用来编码任意类型的数据, 如多维数据。例如, 相似性图、核矩阵和协同过滤矩阵也可视为图结构的特例。因此, 一个成功的图的建模过程很可能包含许多应用, 这些应用通常是与专门的和手动设计的方法一起使用的。

3.2 图神经网络概述

在本节中, 我们将从三个重要方面总结图神经网络的发展: (1) 图神经网络基础; (2) 图神经网络前沿; (3) 图神经网络应用。我们将首先讨论 GNN 在前两个维度下的重要研究子领域, 并简要说明每个研究子领域目前的进展和面临的挑战。接下来, 我们将对如何把 GNN 用于丰富的应用进行概括性总结。

3.2.1 图神经网络基础

从概念上讲, 我们可以将图神经网络的基本学习任务分为 5 个不同的方向: (1) 图神经网络方法; (2) 图神经网络的理论理解; (3) 图神经网络的可扩展性; (4) 图神经网络的可解释性; (5) 图神经网络的对抗鲁棒性。

图神经网络方法。图神经网络是专门设计的用于在图结构数据上进行操作的神经网络架构。图神经网络的目标是通过聚合邻居节点的表征及其在前一次迭代中的表征来迭代更新节点表征。目前已有多种图神经网络被提出 (Kipf and Welling, 2017a; Petar et al, 2018; Hamilton et al, 2017b; Gilmer et al, 2017; Xu et al, 2019d; Veličković et al, 2019d; Veličković et al, 2019; Kipf and Welling, 2016), 它们可以进一步划分为有监督的 GNN 和无监督的 GNN。学习到节点表征之后, GNN 的一个基本任务就是进行节点分类, 也就是将节点分类到一些预定义的类别中。尽管各种 GNN 已经取得巨大的成功, 但我们在训练深度图神经网络时仍面临一个严重的问题——过平滑问题 (Li et al, 2018b), 其中所有的节点都有类似的表征。最近有许多研究提出了不同的补救措施来解决过平滑问题。

图神经网络的理论理解。GNN 算法的快速发展引起了人们对 GNN 理论分析的极大兴趣。特别地, 为了描述 GNN 与传统图算法 (如基于图核的方法) 相比表达能力如何, 以及如何构建更强大的 GNN 以克服 GNN 的一些限制, 人们做出了很多努力。具体来说, Xu et al (2019d) 证明了目前的 GNN 方法能够达到一维 Weisfeiler-Lehman 测试 (Weisfeiler and Leman, 1968) 的表达能力, 这是传统图核领域广泛使用的方法 (Shervashidze et al, 2011b)。

最近的许多研究进一步提出了一系列的设计策略，以进一步超越一维 Weisfeiler-Lehman 测试的表达能力，包括附加随机属性、距离属性和利用高阶结构等。

图神经网络的可扩展性。随着图神经网络日益普及，许多人尝试将各种图神经网络方法用于现实世界中的应用，其中图的大小可以有大约 1 亿个节点和 10 亿条边。遗憾的是，因为需要大量的内存，大多数 GNN 方法不能直接应用于这些大规模的图结构数据 (Hu et al, 2020b)。具体来说，这是因为大多数 GNN 需要在内存中存储整个邻接矩阵和中间层的特征矩阵，这对计算机内存消耗和计算成本都是巨大的挑战。为了解决这些问题，最近的许多研究提出了各种抽样策略，如节点抽样 (Hamilton et al, 2017b; Chen et al, 2018d)、层抽样 (Chen and Bansal, 2018; Huang, 2018) 和图抽样 (Chiang et al, 2019; Zeng et al, 2020a)。

图神经网络的可解释性。为了使机器学习过程可以被人类理解，可解释的人工智能正变得越来越流行，特别是由于深度学习技术的黑盒问题。因此，人们对提高 GNN 的可解释性同样深感兴趣。一般来说，GNN 的解释结果可以是重要的节点、边，也可以是节点或边的重要特征。从技术上讲，基于白盒近似的方法 (Baldassarre and Azizpour, 2019; Sanchez-Lengeling et al, 2020) 利用模型内部的信息 (如梯度、中间特征和模型参数) 来提供解释。与之相对，基于黑盒近似的方法 (Huang et al, 2020c; Zhang et al, 2020a; Vu and Thai, 2020) 则放弃了对复杂模型内部信息的使用，而是利用内在可解释的简单模型 (如线性回归和决策树) 来适应复杂模型。然而，大多数现有的工作很耗时，这就造成处理大规模的图成为瓶颈。为此，人们最近做出了很多努力，以便在不影响解释准确性的情况下开发更有效的方法。

图神经网络的对抗鲁棒性。值得信赖的机器学习最近吸引了大量的关注。这是因为现有的研究表明，深度学习模型可以被故意愚弄、逃避、误导和窃取 (Goodfellow et al, 2015)。因此，在计算机视觉和自然语言处理等领域，已有一系列工作广泛地研究了模型的鲁棒性，这也启发了对 GNN 鲁棒性的类似研究。从技术上讲，研究 GNN 鲁棒性的标准方法 (通过对抗性例子) 是构造输入图数据的一个微小变化，然后观察是否导致预测结果产生较大变化 (如节点分类准确性)。目前，越来越多的人开始研究对抗性攻击 (Dai et al, 2018a; Wang and Gong, 2019; Wu et al, 2019b; Zügner et al, 2018; Zügner et al, 2020) 和对抗性训练 (Xu et al, 2019c; Feng et al, 2019b; Chen et al, 2020i; Jin and Zhang, 2019)。最近的许多努力致力于在对抗性训练以及可认证的鲁棒性 (certified robustness) 方面提供理论保证和新算法开发。

3.2.2 图神经网络前沿

在上述 GNN 基本技术的基础上，在处理各种与图有关的研究问题方面，最近的成就增长快速。在本节中，我们将全面介绍这些研究前沿，它们要么是长期存在的图学习问题与新的 GNN 解决方案，要么是最近出现的 GNN 学习问题。

图神经网络——图分类和链接预测。由于 GNN 模型中的每一层都只产生节点级表征，因此需要图池化层来进一步计算基于节点级表征的图级表征。图级表征总结了输入图结构的关键特征，是图分类的关键组成部分。根据图池化层的学习技术，这些方法一般可以分为 4 类——简单的平面池化 (Duvenaud et al, 2015a; Mesquita et al, 2020)、基于注意力的池化 (Lee et al, 2019d; Huang et al, 2019d)、基于聚类的池化 (Ying et al, 2018c)，以及其他类型的池化 (Zhang et al, 2018f; Bianchi et al, 2020; Morris et al, 2020b)。除图分类

以外，另一个长期存在的图学习问题是链接预测任务，其目的是预测任何一对节点之间现在缺失或未来可能形成的链接。由于 GNN 可以从图结构和辅助信息（如节点特征和边特征）中共同学习，因此与其他传统的图学习方法相比，GNN 在链接预测方面具有巨大的优势。基于 GNN 进行链接预测的常见方法有两种——基于节点的方法（Kipf and Welling, 2016）和基于子图的方法（Zhang and Chen, 2018a, 2020）。

图神经网络——图生成和图转换。基于图建立概率模型的图生成问题是一个处于概率论和图论交叉点上的经典研究问题。近年来，人们对建立在现代图深度学习技术（如 GNN）基础上的深度图生成模型的兴趣越来越大。事实证明，这些深度学习模型在成功捕捉图数据中的复杂依赖关系和生成更真实的图方面更有优势。在变分自编码器（Variational AutoEncoder, VAE）（Kingma and Welling, 2013）和生成对抗网络（Generative Adversarial Network）（Goodfellow et al, 2014a; Goodfellow et al, 2014b）的启发下，用于图生成的基于 GNN 的代表性学习范式有三种，分别是 GraphVAE 方法（Jin et al, 2018b; Simonovsky and Komodakis, 2018; Grover et al, 2019）、GraphGAN 方法（De Cao and Kipf, 2018; You et al, 2018a）和深度自回归方法（Li et al, 2018d; You et al, 2018b; Liao et al, 2019a）。图转换问题可以表述为条件图生成概率，其目标是学习输入源图和输出目标图之间的转译映射（Guo et al, 2018b）。这样的学习问题经常出现在其他领域，如自然语言处理领域的机器翻译问题和计算机视觉领域的图像风格转换问题等。根据被转换的图信息，这个问题一般可以分为 4 类，分别是节点级转换（Battaglia et al, 2016; Yu et al, 2018a; Li et al, 2018e）、边级转换（Guo et al, 2018b; Zhu et al, 2017; Do et al, 2019）、节点-边共同转换（Maziarka et al, 2020a; Kaluza et al, 2018; Guo et al, 2019c）以及涉及图的转换（Bastings et al, 2017; Xu et al, 2018c; Li et al, 2020f）。

图神经网络——图匹配和图结构学习。图匹配指的是寻找两个输入图之间的对应关系，这是一个已在各个领域得到广泛研究的问题。通常情况下，图匹配问题是已知的 NP 难问题（Loiola et al, 2007），这使得该问题在现实世界中大规模问题上的精确解和最优解在计算上不可行。由于 GNN 的表达能力，人们越来越关注开发基于 GNN 的各种图匹配方法，以提高匹配的准确性和效率（Zanfir and Sminchisescu, 2018; Rolínek et al, 2020; Li et al, 2019h; Ling et al, 2020）。图匹配问题旨在衡量两个图结构之间的相似性，而是不改变它们。相比之下，图结构学习的目的是通过联合学习隐含的图结构和图节点表征来产生优化的图结构（Chen et al, 2020m; Franceschi et al, 2019; Veličković et al, 2020）。与经常带有噪声或不完整的固有图（intrinsic graph）相比，我们往往可以将学习到的图结构视为一种转变。即便没有提供初始图，但只要提供显示数据点之间相关性的矩阵，就可以使用图结构学习。

动态图神经网络和异质图神经网络。在现实世界的应用中，图的节点（实体）和图的边（关系）经常会随着时间的推移而发生变化，这就自然地产生了动态图。由于图的演化建模对于做出准确的预测至关重要，因此各种 GNN 不能直接应用于动态图。一种简单而有效的方法是将动态图转换为静态图，但这可能导致信息丢失。根据动态图的类型，基于 GNN 的方法有两大类，分别是用于离散时间动态图的 GNN（Seo et al, 2018; Manessi et al, 2020）和用于持续时间动态图的 GNN（Kazemi et al, 2019; Xu et al, 2020a）。独立地看，实际应用中另一种流行的图是异质图——异质图由不同类型的图节点和边组成。用于同质图的各种 GNN 难以充分利用异质图中的这些信息。因此，一个新的研究方向是开发各种异质图神

神经网络，方法有三种，分别是基于消息传递的方法（Wang et al, 2019l; Fu et al, 2020; Hong et al, 2020b）、基于编码器-解码器的方法（Tu et al, 2018; Zhang et al, 2019b）和基于对抗的方法（Wang et al, 2018a; Hu et al, 2018a）。

图神经网络——AutoML 和自监督学习。自动机器学习（AutoML）最近引起学术界和工业界的极大关注，其目的是应对人工调参过程中耗时巨大这一挑战，特别是对于复杂的深度学习模型而言。AutoML 的这一波研究也影响了自动识别优化 GNN 模型架构和训练超参数的研究工作。现有的研究大多集中于架构搜索空间（Gao et al, 2020b; Zhou et al, 2019a）或训练超参数搜索空间（You et al, 2020a; Shi et al, 2020）。GNN 的另一个重要研究方向是解决大多数深度学习模型的局限性——需要大量的有标注的数据集。因此，目前人们已提出自监督学习，目的是基于无标注数据设计和利用领域特定的辅助任务以预训练一个 GNN 模型。为了研究 GNN 中自监督学习的能力，有相当多的文献系统地设计和比较了 GNN 中不同的自监督代理辅助任务（Hu et al, 2020c; Jin et al, 2020d; You et al, 2020c）。

3.2.3 图神经网络应用

由于图神经网络能够对各种具有复杂结构的数据进行建模，因此图神经网络已经被广泛用于多种应用和领域，如现代推荐系统、计算机视觉（Computer Vision, CV）、自然语言处理（Natural Language Processing, NLP）、程序分析、软件挖掘、生物信息学、异常检测和智慧城市等。尽管在不同的应用中 GNN 被用来解决不同的任务，但它们都包括两个重要步骤——图构建和图表征学习。图构建旨在将输入数据转换或表示为结构化数据。在图的基础上，图表征学习则针对下游任务，利用 GNN 来学习节点嵌入或图嵌入。接下来针对不同的应用，我们将简要介绍这两个步骤涉及的技术。

3.2.3.1 图构建

图构建对于捕捉输入数据中对象之间的依赖关系非常重要。鉴于输入数据的不同格式，不同的应用有不同的图构建技术，其中，有些任务需要预先定义节点和边的语义，以充分表达输入数据的结构信息。

具有显式图结构的输入数据。一些应用自然而然地在数据内部存在图结构，而不需要预先定义节点及其之间的边或关系。例如，在推荐系统中，用户与物品的相互作用自然地形成了一个图，其中用户与物品的偏好被视为用户和物品的节点之间的边；在药物开发的任务中，分子也被自然地表示为一个图，其中的每个节点表示一个原子，每条边表示连接两个原子的键；在蛋白质功能和相互作用的任务中，图也可以很容易地适用于蛋白质，其中的每个节点代表一个氨基酸，每条边代表氨基酸之间的相互作用。

有些图是用节点和边的属性构建的。例如，在处理智慧城市交通时，交通网络可以形式化为一个无向图来预测交通状态。具体来说，节点是交通传感位置，如传感器站、路段，边是连接这些交通传感位置的交叉口。一些城市交通网络可以建模为具有预测交通速度属性的有向图，其中的节点是路段，边是交叉口。路段的宽度、长度和方向被表示为节点的属性，交叉口的类型、是否有交通灯或收费站被表示为边的属性。

具有隐式图结构的输入数据。对于许多天然不存在结构化数据的任务，图构建变得非常具有挑战性。选择最佳的表征方法是很重要的，由此节点和边才能捕捉到所有重要的信息。例如，计算机视觉任务有三种图构建方式。第一种是将图像或视频帧分割成规则的

网格，每个网格可作为视觉图的一个顶点。第二种是先得到预处理的结构，再直接借用顶点表征，如场景图的生成。第三种是利用语义信息来表示视觉顶点，比如将具有相似特征的像素分配给同一个顶点。视觉图像中的边缘可以捕捉到两种信息。第一种是空间信息。例如，对于静态方法，在生成场景图 (Xu et al, 2017a) 和人类骨架 (Jain et al, 2016a) 时，自然会选择视觉图中节点之间的边来表示它们的位置连接。第二种是时间信息。例如，为了表示视频，模型不仅要在帧的内部建立空间关系，也要捕捉相邻帧之间的时间联系。

在自然语言处理任务中，根据文本数据构建的图可以分为 5 类——文本图、句法图、语义图、知识图谱和混合图。下面介绍其中的 4 类。文本图通常将单词、句子、段落或文件视为节点，并通过单词共现、位置或文本相似性来构建边。句法图（或树）强调一条句子中单词之间的语法依赖关系，如依赖图和成分图。知识图谱是数据图，旨在积累和传达现实世界的知识。混合图包含多种类型的节点和边，以整合异质信息。在程序分析的任务中，对程序的图表征的表述包括语法树、控制流、数据流、程序依赖性和调用图，其中的每个图都提供了程序的不同视图。在更高的层面上，程序可以认为是一组异质的实体，它们通过各种关系相互关联。这种观点直接将程序映射为一个异质有向图，其中的每个实体被表示为一个节点，每种类型的关系则被表示为一条边。

3.2.3.2 图表征学习

在得到输入数据的图表示后，下一步是应用 GNN 来学习图表征。有些研究直接利用了典型的 GNN，如 GCN (Kipf and Welling, 2017a)、GAT (Petar et al, 2018)、GGNN (Li et al, 2016a) 和 GraphSage (Hamilton et al, 2017b)，而且能够推广到不同的应用任务。不过，一些特殊的任务需要在 GNN 架构上进行额外的设计，以更好地处理具体问题。例如，针对推荐系统中的任务，人们提出了 PinSage (Ying et al, 2018a)，旨在将一个节点的前 k 个计数节点作为其感受野并进行加权聚合。PinSage 可以扩展到具有数百万用户和物品的网络规模的推荐系统中。KGCN (Wang et al, 2019d) 旨在通过在知识图谱中聚合对应的实体邻域来提高物品的表征。KGAT (Wang et al, 2019j) 与 KGCN 的思路基本相似，前者只是在知识图谱的重建中加入了一个辅助损失。例如，在 KB-对齐的 NLP 任务中，Xu et al (2019f) 将其表述为一个图匹配问题，并提出了一种基于图注意力的方法：首先匹配两个知识图谱中的所有实体，然后根据局部匹配信息进行联合建模，进而得到图级匹配向量。我们在后续内容中将详细介绍各种应用的 GNN 技术。

3.2.4 本书组织结构

本书的组织结构见图 3.1。本书分为四部分，读者可以根据需要选择性阅读。第一部分介绍图神经网络的基本概念；第二部分讨论图神经网络成熟的方法；第三部分介绍图神经网络典型的前沿领域；第四部分描述可能对图神经网络未来研究比较重要和有前途的方法与应用的进展情况。

- **第一部分 引言** 这一部分提供从不同数据类型的表征学习到图表征学习的一般介绍，此外还将介绍用于图表征学习的图神经网络的基本思想和典型变体。
- **第二部分 基础** 这一部分通过介绍图神经网络的特性以及这一领域的几个基本问题来描述图神经网络的基础。具体来说，这一部分介绍图的如下基本问题：节点分类、图神经网络的表达能力、图神经网络的可解释性和可扩展性问题，以及

图神经网络的对抗鲁棒性。

- **第三部分 前沿** 这一部分提出图神经网络领域的一些前沿或高级问题。具体来说，这一部分包括关于图分类、链接预测、图生成、图转换、图匹配、图结构学习等技术的介绍。此外，这一部分还将介绍针对不同类型图的 GNN 的几种变体，如针对动态图、异质图的 GNN。这一部分的最后则介绍 GNN 的自动机器学习和自监督学习。
- **第四部分 广泛和新兴的应用** 这一部分介绍涉及 GNN 的广泛和新兴的应用。具体来说，这些基于 GNN 的应用包括现代推荐系统、计算机视觉和自然语言处理、程序分析、软件挖掘、用于药物开发的生物医学知识图谱挖掘、蛋白质和功能相互作用预测、异常检测和智慧城市等。

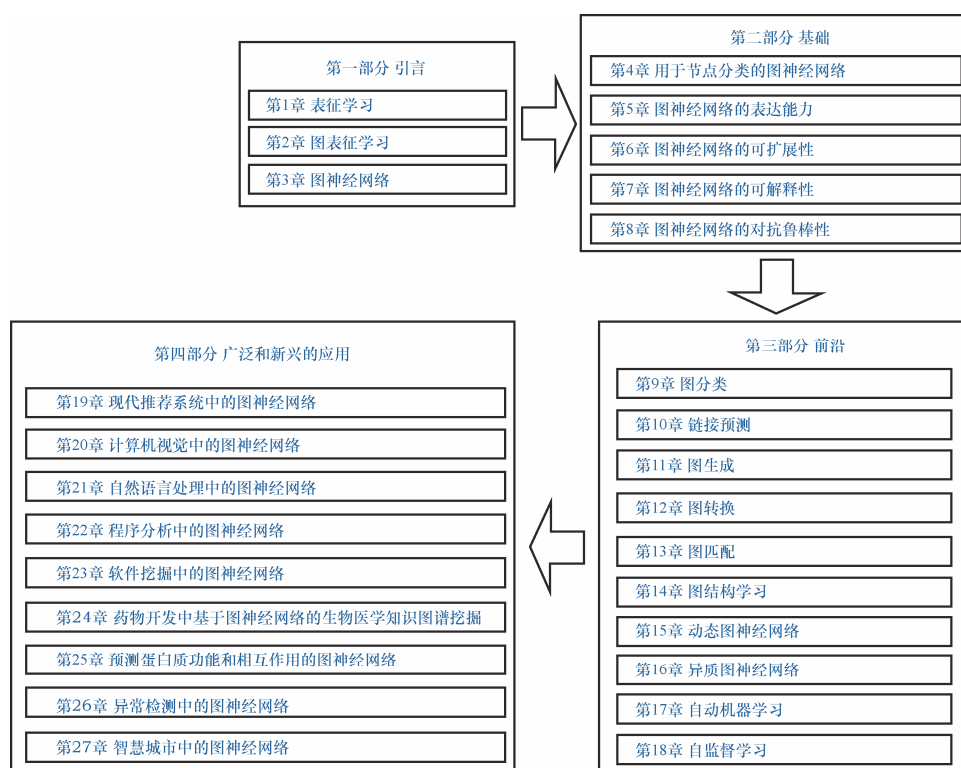


图 3.1 本书的组织结构

3.3 小结

GNN 已经迅速崛起并用于处理图结构数据。传统的深度学习技术由于是为图像和文本等欧几里得数据（又称欧氏数据）设计的，因此不能直接建模图数据。有多种应用可以自然地或最优地建模为图结构，并且已经被各种 GNN 方法成功处理。

在本章中，我们系统地介绍了 GNN 的发展和概况，涵盖对 GNN 基础、前沿和应用的介绍。具体来说，我们介绍了 GNN 的基础理论和方法，从现有的典型 GNN 方法及其表达能力到 GNN 的可扩展性、可解释性和对抗鲁棒性。这些方面促使人们更好地理解 and 利用

GNN。从所回顾的 GNN 的基础方法来看，人们处理图相关研究问题的兴趣正在激增，我们称之为 GNN 的前沿。我们讨论了基于 GNN 的各种前沿研究，从图分类和链接预测到图生成、图转换、图匹配和图结构学习。GNN 由于具有对各种复杂结构的数据的建模能力，目前已被广泛应用于许多应用和领域，如现代推荐系统、计算机视觉、自然语言处理、程序分析、软件挖掘、生物信息学、异常检测和智慧城市。这些任务大多由两个重要步骤组成——图构建和图表征学习。因此，我们对这两个步骤涉及的技术进行了介绍，涵盖不同的应用。最后，我们在本章的末尾提供了本书的组织结构。