

# Contents

<b>1</b>	<b>Deep Learning on Graphs: An Introduction</b>	<i>page</i> 1
1.1	Introduction	1
1.2	Why Deep Learning on Graphs?	1
1.3	What Content is Covered?	3
1.4	Who Should Read the Book?	6
1.5	Feature Learning on Graphs: A Brief History	8
	1.5.1 Feature Selection on Graphs	9
	1.5.2 Representation Learning on Graphs	10
1.6	Conclusion	12
1.7	Further Reading	13
	<b>PART ONE FOUNDATIONS</b>	15
<b>2</b>	<b>Foundations of Graphs</b>	17
2.1	Introduction	17
2.2	Graph Representations	18
2.3	Properties and Measures	19
	2.3.1 Degree	19
	2.3.2 Connectivity	21
	2.3.3 Centrality	23
2.4	Spectral Graph Theory	26
	2.4.1 Laplacian Matrix	26
	2.4.2 The Eigenvalues and Eigenvectors of the Laplacian Matrix	28
2.5	Graph Signal Processing	29
	2.5.1 Graph Fourier Transform	30
2.6	Complex Graphs	33

2.6.1	Heterogeneous Graphs	33
2.6.2	Bipartite Graphs	33
2.6.3	Multi-dimensional Graphs	34
2.6.4	Signed Graphs	35
2.6.5	Hypergraphs	36
2.6.6	Dynamic Graphs	37
2.7	Computational Tasks on Graphs	39
2.7.1	Node-focused Tasks	39
2.7.2	Graph-focused Tasks	41
2.8	Conclusion	42
2.9	Further Reading	42
<b>3</b>	<b>Foundations of Deep Learning</b>	<b>43</b>
3.1	Introduction	43
3.2	Feedforward Networks	44
3.2.1	The Architecture	46
3.2.2	Activation Functions	47
3.2.3	Output Layer and Loss Function	50
3.3	Convolutional Neural Networks	51
3.3.1	The Convolution Operation and Convolutional Layer	52
3.3.2	Convolutional Layers in Practice	56
3.3.3	Non-linear Activation Layer	57
3.3.4	Pooling Layer	58
3.3.5	An Overall CNN Framework	58
3.4	Recurrent Neural Networks	59
3.4.1	The Architecture of Traditional RNNs	60
3.4.2	Long Short-Term Memory	61
3.4.3	Gated Recurrent Unit	63
3.5	Autoencoders	63
3.5.1	Undercomplete Autoencoders	65
3.5.2	Regularized Autoencoders	66
3.6	Training Deep Neural Networks	67
3.6.1	Training with Gradient Descent	67
3.6.2	Backpropagation	68
3.6.3	Preventing Overfitting	70
3.7	Conclusion	71
3.8	Further Reading	72

	<b>PART TWO METHODS</b>	73
<b>4</b>	<b>Graph Embedding</b>	75
4.1	Introduction	75
4.2	Graph Embedding on Simple Graphs	77
4.2.1	Preserving Node Co-occurrence	77
4.2.2	Preserving Structural Role	86
4.2.3	Preserving Node Status	89
4.2.4	Preserving Community Structure	91
4.3	Graph Embedding on Complex Graphs	93
4.3.1	Heterogeneous Graph Embedding	94
4.3.2	Bipartite Graph Embedding	96
4.3.3	Multi-dimensional Graph Embedding	97
4.3.4	Signed Graph Embedding	98
4.3.5	Hypergraph Embedding	101
4.3.6	Dynamic Graph Embedding	103
4.4	Conclusion	104
4.5	Further Reading	105
<b>5</b>	<b>Graph Neural Networks</b>	106
5.1	Introduction	106
5.2	The General GNN Frameworks	108
5.2.1	A General Framework for Node-focused Tasks	108
5.2.2	A General Framework for Graph-focused Tasks	109
5.3	Graph Filters	111
5.3.1	Spectral-based Graph Filters	111
5.3.2	Spatial-based Graph Filters	121
5.4	Graph Pooling	127
5.4.1	Flat Graph Pooling	128
5.4.2	Hierarchical Graph Pooling	129
5.5	Parameter Learning for Graph Neural Networks	133
5.5.1	Parameter Learning for Node Classification	133
5.5.2	Parameter Learning for Graph Classification	134
5.6	Conclusion	135
5.7	Further Reading	136
<b>6</b>	<b>Robust Graph Neural Networks</b>	137
6.1	Introduction	137
6.2	Graph Adversarial Attacks	137
6.2.1	Taxonomy of Graph Adversarial Attacks	138
6.2.2	White-box Attack	140
6.2.3	Gray-box Attack	143

6.2.4	Black-box Attack	147
6.3	Graph Adversarial Defenses	150
6.3.1	Graph Adversarial Training	151
6.3.2	Graph Purification	153
6.3.3	Graph Attention	154
6.3.4	Graph Structure Learning	158
6.4	Conclusion	159
6.5	Further Reading	159
<b>7</b>	<b>Scalable Graph Neural Networks</b>	<b>160</b>
7.1	Introduction	160
7.2	Node-wise Sampling Methods	164
7.3	Layer-wise Sampling Methods	166
7.4	Subgraph-wise Sampling Methods	170
7.5	Conclusion	172
7.6	Further Reading	172
<b>8</b>	<b>Graph Neural Networks on Complex Graphs</b>	<b>174</b>
8.1	Introduction	174
8.2	Heterogeneous Graph Neural Networks	174
8.3	Bipartite Graph Neural Networks	176
8.4	Multi-dimensional Graph Neural Networks	177
8.5	Signed Graph Neural Networks	179
8.6	Hypergraph Neural Networks	182
8.7	Dynamic Graph Neural Networks	183
8.8	Conclusion	185
8.9	Further Reading	185
<b>9</b>	<b>Beyond GNNs: More Deep Models on Graphs</b>	<b>186</b>
9.1	Introduction	186
9.2	Autoencoders on Graphs	187
9.3	Recurrent Neural Networks on Graphs	189
9.4	Variational Autoencoders on Graphs	191
9.4.1	Variational Autoencoders for Node Representation Learning	193
9.4.2	Variational Autoencoders for Graph Generation	193
9.5	Generative Adversarial Networks on Graphs	196
9.5.1	Generative Adversarial Networks for Node Representation Learning	197
9.5.2	Generative Adversarial Networks for Graph Generation	199
9.6	Conclusion	200

9.7	Further Reading	200
	<b>PART THREE APPLICATIONS</b>	203
<b>10</b>	<b>Graph Neural Networks in Natural Language Processing</b>	205
10.1	Introduction	205
10.2	Semantic Role Labeling	206
10.3	Neural Machine Translation	208
10.4	Relation Extraction	209
10.5	Question Answering	210
	10.5.1 The Multi-hop QA Task	211
	10.5.2 Entity-GCN	212
10.6	Graph to Sequence Learning	214
10.7	Graph Neural Networks on Knowledge Graphs	215
	10.7.1 Graph Filters for Knowledge Graphs	216
	10.7.2 Transforming Knowledge Graphs to Simple Graphs	217
	10.7.3 Knowledge Graph Completion	217
10.8	Conclusion	218
10.9	Further Reading	218
<b>11</b>	<b>Graph Neural Networks in Computer Vision</b>	220
11.1	Introduction	220
11.2	Visual Question Answering	220
	11.2.1 Images as Graphs	221
	11.2.2 Images and Questions as Graphs	223
11.3	Skeleton-based Action Recognition	225
11.4	Image Classification	227
	11.4.1 Zero-shot Image Classification	228
	11.4.2 Few-shot Image Classification	229
	11.4.3 Multi-label Image Classification	230
11.5	Point Cloud Learning	231
11.6	Conclusion	232
11.7	Further Reading	232
<b>12</b>	<b>Graph Neural Networks in Data Mining</b>	233
12.1	Introduction	233
12.2	Web Data Mining	233
	12.2.1 Social Network Analysis	234
	12.2.2 Recommender Systems	237
12.3	Urban Data Mining	241

12.3.1	Traffic Prediction	241
12.3.2	Air Quality Forecasting	243
12.4	Cybersecurity Data Mining	244
12.4.1	Malicious Account Detection	244
12.4.2	Fake News Detection	246
12.5	Conclusion	247
12.6	Further Reading	247
<b>13</b>	<b>Graph Neural Networks in Biochemistry and Healthcare</b>	<b>249</b>
13.1	Introduction	249
13.2	Drug Development and Discovery	249
13.2.1	Molecule Representation Learning	250
13.2.2	Protein Interface Prediction	251
13.2.3	Drug-Target Binding Affinity Prediction	253
13.3	Drug Similarity Integration	255
13.4	Polypharmacy Side Effect Prediction	256
13.5	Disease Prediction	259
13.6	Conclusion	260
13.7	Further Reading	261
	<b>PART FOUR ADVANCES</b>	<b>263</b>
<b>14</b>	<b>Advanced Topics in Graph Neural Networks</b>	<b>265</b>
14.1	Introduction	265
14.2	Deeper Graph Neural Networks	266
14.2.1	Jumping Knowledge	268
14.2.2	DropEdge	268
14.2.3	Pairnorm	268
14.3	Exploring Unlabeled Data via Self-supervised Learning	268
14.3.1	Node-focused Tasks	269
14.3.2	Graph-focused Tasks	272
14.4	Expressiveness of Graph Neural Networks	273
14.4.1	Weisfeiler-Lehman Test	274
14.4.2	Expressiveness	275
14.5	Conclusion	277
14.6	Further Reading	277
<b>15</b>	<b>Advanced Applications in Graph Neural Networks</b>	<b>278</b>
15.1	Introduction	278
15.2	Combinatorial Optimization on Graphs	278
15.3	Learning Program Representations	280

	<i>Contents</i>	ix
15.4	Reasoning Interacting Dynamical Systems in Physics	282
15.5	Conclusion	283
15.6	Further Reading	283
	<i>Bibliography</i>	285
	<i>Index</i>	311