### Contents

1 Deep Learning on Graphs: An Introduction  
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

PART ONE FOUNDATIONS  

2 Foundations of Graphs 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
Contents

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

3 Foundations of Deep Learning 43

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
PART TWO  METHODS

4  Graph Embedding  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

5  Graph Neural Networks  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

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

10 Graph Neural Networks in Natural Language Processing 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

11 Graph Neural Networks in Computer Vision 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

12 Graph Neural Networks in Data Mining 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 Traffic Prediction
- 12.3.1 Traffic Prediction

## 12.4 Air Quality Forecasting
- 12.4.1 Air Quality Forecasting

## 12.4 Cybersecurity Data Mining
- 12.4.1 Malicious Account Detection
- 12.4.2 Fake News Detection

## 12.5 Conclusion

## 12.6 Further Reading

### 13 Graph Neural Networks in Biochemistry and Healthcare

#### 13.1 Introduction

#### 13.2 Drug Development and Discovery
- 13.2.1 Molecule Representation Learning
- 13.2.2 Protein Interface Prediction
- 13.2.3 Drug-Target Binding Affinity Prediction

#### 13.3 Drug Similarity Integration

#### 13.4 Polypharmacy Side Effect Prediction

#### 13.5 Disease Prediction

#### 13.6 Conclusion

#### 13.7 Further Reading

### PART FOUR ADVANCES

#### 14 Advanced Topics in Graph Neural Networks

#### 14.1 Introduction

#### 14.2 Deeper Graph Neural Networks
- 14.2.1 Jumping Knowledge
- 14.2.2 DropEdge
- 14.2.3 Pairnorm

#### 14.3 Exploring Unlabeled Data via Self-supervised Learning
- 14.3.1 Node-focused Tasks
- 14.3.2 Graph-focused Tasks

#### 14.4 Expressiveness of Graph Neural Networks
- 14.4.1 Weisfeiler-Lehman Test
- 14.4.2 Expressiveness

#### 14.5 Conclusion

#### 14.6 Further Reading

#### 15 Advanced Applications in Graph Neural Networks

#### 15.1 Introduction

#### 15.2 Combinatorial Optimization on Graphs

#### 15.3 Learning Program Representations
Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.4</td>
<td>Reasoning Interacting Dynamical Systems in Physics</td>
<td>282</td>
</tr>
<tr>
<td>15.5</td>
<td>Conclusion</td>
<td>283</td>
</tr>
<tr>
<td>15.6</td>
<td>Further Reading</td>
<td>283</td>
</tr>
</tbody>
</table>

Bibliography  285
Index  311