

Contents

Part I Overview of Learning to Rank

1 Introduction	3
1.1 Overview	3
1.2 Ranking in Information Retrieval	7
1.2.1 Conventional Ranking Models	7
1.2.2 Query-Level Position-Based Evaluations	11
1.3 Learning to Rank	15
1.3.1 Machine Learning Framework	16
1.3.2 Definition of Learning to Rank	17
1.3.3 Learning-to-Rank Framework	18
1.4 Book Overview	23
1.5 Exercises	24
References	25

Part II Major Approaches to Learning to Rank

2 The Pointwise Approach	33
2.1 Overview	33
2.2 Regression-Based Algorithms	33
2.2.1 Subset Ranking with Regression	34
2.3 Classification-Based Algorithms	35
2.3.1 Binary Classification for Ranking	35
2.3.2 Multi-class Classification for Ranking	37
2.4 Ordinal Regression-Based Algorithms	39
2.4.1 Perceptron-Based Ranking (PRanking)	39
2.4.2 Ranking with Large Margin Principles	40
2.4.3 Ordinal Regression with Threshold-Based Loss Functions	42
2.5 Discussions	42
2.5.1 Relationship with Relevance Feedback	43
2.5.2 Problems with the Pointwise Approach	44

2.5.3	Improved Algorithms	44
2.6	Summary	45
2.7	Exercises	45
	References	46
3	The Pairwise Approach	49
3.1	Overview	49
3.2	Example Algorithms	49
3.2.1	Ordering with Preference Function	49
3.2.2	SortNet: Neural Network-Based Sorting Algorithm	51
3.2.3	RankNet: Learning to Rank with Gradient Descent	52
3.2.4	FRank: Ranking with a Fidelity Loss	53
3.2.5	RankBoost	54
3.2.6	Ranking SVM	56
3.2.7	GBRank	58
3.3	Improved Algorithms	59
3.3.1	Multiple Hyperplane Ranker	59
3.3.2	Magnitude-Preserving Ranking	60
3.3.3	IR-SVM	61
3.3.4	Robust Pairwise Ranking with Sigmoid Functions	62
3.3.5	P-norm Push	63
3.3.6	Ordered Weighted Average for Ranking	64
3.3.7	LambdaRank	65
3.3.8	Robust Sparse Ranker	66
3.4	Summary	67
3.5	Exercises	67
	References	68
4	The Listwise Approach	71
4.1	Overview	71
4.2	Minimization of Measure-Specific Loss	72
4.2.1	Measure Approximation	72
4.2.2	Bound Optimization	77
4.2.3	Non-smooth Optimization	78
4.2.4	Discussions	80
4.3	Minimization of Non-measure-Specific Loss	80
4.3.1	ListNet	81
4.3.2	ListMLE	82
4.3.3	Ranking Using Cumulative Distribution Networks	83
4.3.4	BoltzRank	84
4.4	Summary	85
4.5	Exercises	86
	References	87
5	Analysis of the Approaches	89
5.1	Overview	89
5.2	The Pointwise Approach	89

- 5.3 The Pairwise Approach 91
- 5.4 The Listwise Approach 94
 - 5.4.1 Non-measure-Specific Loss 94
 - 5.4.2 Measure-Specific Loss 95
- 5.5 Summary 98
- 5.6 Exercises 98
- References 99

Part III Advanced Topics in Learning to Rank

- 6 Relational Ranking 103**
 - 6.1 General Relational Ranking Framework 104
 - 6.1.1 Relational Ranking SVM 104
 - 6.1.2 Continuous Conditional Random Fields 106
 - 6.2 Learning Diverse Ranking 107
 - 6.3 Discussions 110
 - References 111
- 7 Query-Dependent Ranking 113**
 - 7.1 Query-Dependent Loss Function 113
 - 7.2 Query-Dependent Ranking Function 115
 - 7.2.1 Query Classification-Based Approach 115
 - 7.2.2 K Nearest Neighbor-Based Approach 116
 - 7.2.3 Query Clustering-Based Approach 118
 - 7.2.4 Two-Layer Learning Approach 119
 - 7.3 Discussions 120
 - References 121
- 8 Semi-supervised Ranking 123**
 - 8.1 Inductive Approach 123
 - 8.2 Transductive Approach 124
 - 8.3 Discussions 125
 - References 125
- 9 Transfer Ranking 127**
 - 9.1 Feature-Level Transfer Ranking 128
 - 9.2 Instance-Level Transfer Ranking 128
 - 9.3 Discussions 130
 - References 130

Part IV Benchmark Datasets for Learning to Rank

- 10 The LETOR Datasets 133**
 - 10.1 Overview 133
 - 10.2 Document Corpora 133
 - 10.2.1 The “Gov” Corpus and Six Query Sets 134
 - 10.2.2 The OHSUMED Corpus 134
 - 10.2.3 The “Gov2” Corpus and Two Query Sets 135

- 10.3 Document Sampling 135
- 10.4 Feature Extraction 136
- 10.5 Meta Information 136
- 10.6 Learning Tasks 138
- 10.7 Discussions 142
- References 142
- 11 Experimental Results on LETOR 145**
 - 11.1 Experimental Settings 145
 - 11.2 Experimental Results on LETOR 3.0 146
 - 11.3 Experimental Results on LETOR 4.0 149
 - 11.4 Discussions 150
 - 11.5 Exercises 151
 - References 151
- 12 Other Datasets 153**
 - 12.1 Yahoo! Learning-to-Rank Challenge Datasets 153
 - 12.2 Microsoft Learning-to-Rank Datasets 154
 - 12.3 Discussions 155
 - References 155
- Part V Practical Issues in Learning to Rank**
- 13 Data Preprocessing for Learning to Rank 159**
 - 13.1 Overview 159
 - 13.2 Ground Truth Mining from Logs 160
 - 13.2.1 User Click Models 160
 - 13.2.2 Click Data Enhancement 166
 - 13.3 Training Data Selection 168
 - 13.3.1 Document and Query Selection for Labeling 169
 - 13.3.2 Document and Query Selection for Training 171
 - 13.3.3 Feature Selection for Training 175
 - 13.4 Summary 176
 - 13.5 Exercises 176
 - References 177
- 14 Applications of Learning to Rank 181**
 - 14.1 Overview 181
 - 14.2 Question Answering 181
 - 14.2.1 Definitional QA 182
 - 14.2.2 Quantity Consensus QA 183
 - 14.2.3 Non-factoid QA 184
 - 14.2.4 Why QA 185
 - 14.3 Multimedia Retrieval 186
 - 14.4 Text Summarization 187
 - 14.5 Online Advertising 188
 - 14.6 Summary 189

14.7 Exercises 190
 References 190

Part VI Theories in Learning to Rank

15 Statistical Learning Theory for Ranking 195
 15.1 Overview 195
 15.2 Statistical Learning Theory 195
 15.3 Learning Theory for Ranking 197
 15.3.1 Statistical Ranking Framework 197
 15.3.2 Generalization Analysis for Ranking 198
 15.3.3 Statistical Consistency for Ranking 198
 15.4 Exercises 199
 References 199

16 Statistical Ranking Framework 201
 16.1 Document Ranking Framework 202
 16.1.1 The Pointwise Approach 202
 16.1.2 The Pairwise Approach 202
 16.1.3 The Listwise Approach 204
 16.2 Subset Ranking Framework 204
 16.2.1 The Pointwise Approach 205
 16.2.2 The Pairwise Approach 205
 16.2.3 The Listwise Approach 206
 16.3 Two-Layer Ranking Framework 206
 16.3.1 The Pointwise Approach 206
 16.3.2 The Pairwise Approach 207
 16.3.3 The Listwise Approach 208
 16.4 Summary 208
 16.5 Exercises 208
 References 209

17 Generalization Analysis for Ranking 211
 17.1 Overview 211
 17.2 Uniform Generalization Bounds for Ranking 212
 17.2.1 For Document Ranking 212
 17.2.2 For Subset Ranking 214
 17.2.3 For Two-Layer Ranking 216
 17.3 Algorithm-Dependent Generalization Bound 217
 17.3.1 For Document Ranking 218
 17.3.2 For Subset Ranking 219
 17.3.3 For Two-Layer Ranking 220
 17.4 Summary 220
 17.5 Exercises 221
 References 221

18	Statistical Consistency for Ranking	223
18.1	Overview	223
18.2	Consistency Analysis for Document Ranking	224
18.2.1	Regarding Pairwise 0–1 Loss	224
18.3	Consistency Analysis for Subset Ranking	224
18.3.1	Regarding DCG-Based Ranking Error	225
18.3.2	Regarding Permutation-Level 0–1 Loss	225
18.3.3	Regarding Top- k True Loss	226
18.3.4	Regarding Weighted Kendall’s τ	227
18.4	Consistency Analysis for Two-Layer Ranking	229
18.5	Summary	229
18.6	Exercises	230
	References	230

Part VII Summary and Outlook

19	Summary	235
	References	238
20	Future Work	241
20.1	Sample Selection Bias	241
20.2	Direct Learning from Logs	242
20.3	Feature Engineering	243
20.4	Advanced Ranking Models	243
20.5	Large-Scale Learning to Rank	244
20.6	Online Complexity Versus Accuracy	245
20.7	Robust Learning to Rank	245
20.8	Online Learning to Rank	246
20.9	Beyond Ranking	247
	References	247

Part VIII Appendix

21	Mathematical Background	251
21.1	Probability Theory	251
21.1.1	Probability Space and Random Variables	251
21.1.2	Probability Distributions	252
21.1.3	Expectations and Variances	254
21.2	Linear Algebra and Matrix Computation	255
21.2.1	Notations	255
21.2.2	Basic Matrix Operations and Properties	256
21.2.3	Eigenvalues and Eigenvectors	261
21.3	Convex Optimization	262
21.3.1	Convex Set and Convex Function	262
21.3.2	Conditions for Convexity	263
21.3.3	Convex Optimization Problem	263
21.3.4	Lagrangian Duality	264

- 21.3.5 KKT Conditions 265
- References 266
- 22 Machine Learning 267**
- 22.1 Regression 267
- 22.1.1 Linear Regression 267
- 22.1.2 Probabilistic Explanation 268
- 22.2 Classification 269
- 22.2.1 Neural Networks 270
- 22.2.2 Support Vector Machines 271
- 22.2.3 Boosting 273
- 22.2.4 K Nearest Neighbor (KNN) 274
- 22.3 Statistical Learning Theory 274
- 22.3.1 Formalization 275
- 22.3.2 Bounds for $|R(g) - \hat{R}(g)|$ 277
- References 282
- Index 283**



<http://www.springer.com/978-3-642-14266-6>

Learning to Rank for Information Retrieval

Liu, T.-Y.

2011, XVII, 285 p., Hardcover

ISBN: 978-3-642-14266-6