Contents

Dedication v

Preface xi

1 Introduction 1
- 1.1 Examples of Social Learning . . . 2
- 1.2 Building Opinions . . . 6
- 1.3 Book Organization . . . 10
- 1.4 Notation, Symbols, and Conventions . . . 13

2 Bayesian Learning 17
- 2.1 The Bayesian Way . . . 18
- 2.2 Properties of Bayes' Rule . . . 22
- 2.3 Information-Theoretic Interpretations . . . 31
- 2.4 Stochastic-Optimization Interpretation . . . 34

3 From Single-Agent to Social Learning 41
- 3.1 Bayesian versus Non-Bayesian Learning . . . 42
- 3.2 Non-Bayesian Social Learning . . . 45
- 3.3 Information-Theoretic Viewpoint . . . 47
- 3.4 Behavioral Viewpoint . . . 52
- 3.5 Unifying Framework . . . 63

4 Network Models 67
- 4.1 Network Graphs . . . 67
- 4.2 Combination Matrices . . . 73
- 4.3 Strong and Primitive Graphs . . . 77
- 4.4 Stochastic Combination Matrices . . . 80
- 4.5 Weak Graphs . . . 82
- 4.6 Combination Policies . . . 91

5 Social Learning with Geometric Averaging 95
- 5.1 Belief Convergence . . . 96
- 5.2 Learning over Connected Graphs . . . 101

5.3 Objective Evidence . . . 104
5.4 Subjective Evidence . . . 108
5.5 Fake Evidence . . . 112
5.6 Learning over Weak Graphs . . . 114

6 Error Probability Performance **121**
6.1 Useful Statistical Descriptors . . . 122
6.2 Normal Approximation for Large t . . . 127
6.3 Large Deviations for Large t . . . 134

7 Social Learning with Arithmetic Averaging **145**
7.1 Modeling Assumptions . . . 146
7.2 Belief Convergence . . . 147

8 Adaptive Social Learning **157**
8.1 Stubbornness of Agents . . . 157
8.2 Adaptive Update . . . 159
8.3 Learning versus Adaptation . . . 164
8.4 Adaptive Setting . . . 166
8.5 Variation on ASL . . . 168

9 Learning Accuracy under ASL **171**
9.1 Steady-State Analysis . . . 172
9.2 Small-δ Regime . . . 178
9.3 Consistency of Adaptive Social Learning . . . 180
9.4 Normal Approximation for Small δ . . . 187
9.5 Large Deviations for Small δ . . . 192
9.6 Main Performance Characteristics . . . 202

10 Adaptation under ASL **205**
10.1 Qualitative Description of the Transient Phase . . . 205
10.2 Quantitative Transient Analysis . . . 210
10.3 Adaptation Time . . . 216
10.4 Summary: Learning and Adaptation under ASL . . . 232

11 Partial Information Sharing **235**
11.1 Partial Information Framework . . . 236
11.2 Decoding Strategies . . . 237
11.3 Asymptotic Learning Objectives . . . 241

11.4 Memoryless Strategy . . . 242
11.5 Memory in Partial Information . . . 257
11.6 Comparing Strategies . . . 270
11.A Appendix: Preliminary Results . . . 274
11.B Appendix: Proof of Theorem 11.2 . . . 286
11.C Appendix: Proof of Theorem 11.3 . . . 290

12 Social Machine Learning **297**
12.1 Social Machine Learning Model . . . 299
12.2 General Decision Statistics . . . 302
12.3 Training Phase . . . 306
12.4 Performance Guarantees . . . 312
12.5 Sample Complexity . . . 318
12.6 Illustrative Examples . . . 321
12.A Appendix: Notation for Binary Decision Problems . . . 328
12.B Appendix: Bounds for Consistent Learning . . . 330
12.C Appendix: Proof of Theorem 12.1 . . . 338
12.D Appendix: Proof of Theorem 12.2 . . . 341
12.E Appendix: Auxiliary Results . . . 342

13 Extensions and Conclusions **347**
13.1 Non-Bayesian Updates . . . 347
13.2 Censored Beliefs . . . 361
13.3 Learning the Social Graph . . . 368

Appendices **373**

A Convex Functions **375**

B Entropy and KL Divergence **377**

C Probabilistic Inequalities **381**

D Stochastic Convergence **385**
D.1 Types of Stochastic Convergence . . . 385
D.2 Fundamental Asymptotic Results . . . 388
D.3 Convergence of Sums and Recursions . . . 394
D.4 Martingales . . . 397

E Large Deviations **401**
E.1 Empirical Averages 401
E.2 Large Deviation Principle 427

F Random Sums and Series **431**
F.1 Convergent Random Series 431
F.2 Random Sums Relevant to Adaptive Social Learning 432
F.3 Vector Case for Network Behavior 436

G Rademacher Complexity **451**
G.1 General Case 451
G.2 Multilayer Perceptrons 454

References **459**

About the Authors **469**