

# Recommender Systems: The Textbook



Charu C. Aggarwal

# Recommender Systems

The Textbook

Charu C. Aggarwal  
IBM T.J. Watson Research Center  
Yorktown Heights, NY, USA

ISBN 978-3-319-29657-9      ISBN 978-3-319-29659-3 (eBook)  
DOI 10.1007/978-3-319-29659-3

Library of Congress Control Number: 2016931438

Springer Cham Heidelberg New York Dordrecht London

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*To my wife Lata, my daughter Sayani,  
and my late parents Dr. Prem Sarup and Mrs. Pushplata Aggarwal.*



---

---

# Contents

---

---

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>An Introduction to Recommender Systems</b>   | <b>1</b> |
| 1.1      | Introduction . . . . .  | 1        |
| 1.2      | Goals of Recommender Systems . . . . .  | 3        |
| 1.2.1    | The Spectrum of Recommendation Applications . . . . .   | 7        |
| 1.3      | Basic Models of Recommender Systems . . . . .   | 8        |
| 1.3.1    | Collaborative Filtering Models . . . . .  | 8        |
| 1.3.1.1  | Types of Ratings . . . . .  | 10       |
| 1.3.1.2  | Relationship with Missing Value Analysis . . . . .  | 13       |
| 1.3.1.3  | Collaborative Filtering as a Generalization of Classification and Regression Modeling . . . . . | 13       |
| 1.3.2    | Content-Based Recommender Systems . . . . .   | 14       |
| 1.3.3    | Knowledge-Based Recommender Systems . . . . .   | 15       |
| 1.3.3.1  | Utility-Based Recommender Systems . . . . .   | 18       |
| 1.3.4    | Demographic Recommender Systems . . . . .   | 19       |
| 1.3.5    | Hybrid and Ensemble-Based Recommender Systems . . . . .   | 19       |
| 1.3.6    | Evaluation of Recommender Systems . . . . .   | 20       |
| 1.4      | Domain-Specific Challenges in Recommender Systems . . . . .                                     | 20       |
| 1.4.1    | Context-Based Recommender Systems . . . . .   | 20       |
| 1.4.2    | Time-Sensitive Recommender Systems . . . . .  | 21       |
| 1.4.3    | Location-Based Recommender Systems . . . . .  | 21       |
| 1.4.4    | Social Recommender Systems . . . . .  | 22       |
| 1.4.4.1  | Structural Recommendation of Nodes and Links . . . . .  | 22       |
| 1.4.4.2  | Product and Content Recommendations with Social Influence . . . . .                             | 23       |
| 1.4.4.3  | Trustworthy Recommender Systems . . . . .   | 23       |
| 1.4.4.4  | Leveraging Social Tagging Feedback for Recommendations . . . . .                                | 23       |
| 1.5      | Advanced Topics and Applications . . . . .  | 23       |
| 1.5.1    | The Cold-Start Problem in Recommender Systems . . . . .   | 24       |
| 1.5.2    | Attack-Resistant Recommender Systems . . . . .  | 24       |
| 1.5.3    | Group Recommender Systems . . . . .   | 24       |

|          |  |           |
|----------|--|-----------|
| 1.5.4    | Multi-Criteria Recommender Systems . . . . .                     | 24        |
| 1.5.5    | Active Learning in Recommender Systems . . . . .                 | 25        |
| 1.5.6    | Privacy in Recommender Systems . . . . .                         | 25        |
| 1.5.7    | Application Domains . . . . .                                    | 26        |
| 1.6      | Summary . . . . .  | 26        |
| 1.7      | Bibliographic Notes . . . . .                                    | 26        |
| 1.8      | Exercises . . . . .  | 28        |
| <b>2</b> | <b>Neighborhood-Based Collaborative Filtering</b>                | <b>29</b> |
| 2.1      | Introduction . . . . .   | 29        |
| 2.2      | Key Properties of Ratings Matrices . . . . .                     | 31        |
| 2.3      | Predicting Ratings with Neighborhood-Based Methods . . . . .     | 33        |
| 2.3.1    | User-Based Neighborhood Models . . . . .                         | 34        |
| 2.3.1.1  | Similarity Function Variants . . . . .                           | 37        |
| 2.3.1.2  | Variants of the Prediction Function . . . . .                    | 38        |
| 2.3.1.3  | Variations in Filtering Peer Groups . . . . .                    | 39        |
| 2.3.1.4  | Impact of the Long Tail . . . . .                                | 39        |
| 2.3.2    | Item-Based Neighborhood Models . . . . .                         | 40        |
| 2.3.3    | Efficient Implementation and Computational Complexity . . . . .  | 41        |
| 2.3.4    | Comparing User-Based and Item-Based Methods . . . . .            | 42        |
| 2.3.5    | Strengths and Weaknesses of Neighborhood-Based Methods . . . . . | 44        |
| 2.3.6    | A Unified View of User-Based and Item-Based Methods . . . . .    | 44        |
| 2.4      | Clustering and Neighborhood-Based Methods . . . . .              | 45        |
| 2.5      | Dimensionality Reduction and Neighborhood Methods . . . . .      | 47        |
| 2.5.1    | Handling Problems with Bias . . . . .                            | 49        |
| 2.5.1.1  | Maximum Likelihood Estimation . . . . .                          | 49        |
| 2.5.1.2  | Direct Matrix Factorization of Incomplete Data . . . . .         | 50        |
| 2.6      | A Regression Modeling View of Neighborhood Methods . . . . .     | 51        |
| 2.6.1    | User-Based Nearest Neighbor Regression . . . . .                 | 53        |
| 2.6.1.1  | Sparsity and Bias Issues . . . . .                               | 54        |
| 2.6.2    | Item-Based Nearest Neighbor Regression . . . . .                 | 55        |
| 2.6.3    | Combining User-Based and Item-Based Methods . . . . .            | 57        |
| 2.6.4    | Joint Interpolation with Similarity Weighting . . . . .          | 57        |
| 2.6.5    | Sparse Linear Models (SLIM) . . . . .                            | 58        |
| 2.7      | Graph Models for Neighborhood-Based Methods . . . . .            | 60        |
| 2.7.1    | User-Item Graphs . . . . .                                       | 61        |
| 2.7.1.1  | Defining Neighborhoods with Random Walks . . . . .               | 61        |
| 2.7.1.2  | Defining Neighborhoods with the Katz Measure . . . . .           | 62        |
| 2.7.2    | User-User Graphs . . . . .                                       | 63        |
| 2.7.3    | Item-Item Graphs . . . . .                                       | 66        |
| 2.8      | Summary . . . . .  | 67        |
| 2.9      | Bibliographic Notes . . . . .                                    | 67        |
| 2.10     | Exercises . . . . .  | 69        |
| <b>3</b> | <b>Model-Based Collaborative Filtering</b>                       | <b>71</b> |
| 3.1      | Introduction . . . . .   | 71        |
| 3.2      | Decision and Regression Trees . . . . .                          | 74        |
| 3.2.1    | Extending Decision Trees to Collaborative Filtering . . . . .    | 76        |



|          |  |            |
|----------|--|------------|
| 3.3      | Rule-Based Collaborative Filtering . . . . .                           | 77         |
| 3.3.1    | Leveraging Association Rules for Collaborative Filtering . . . . .     | 79         |
| 3.3.2    | Item-Wise Models versus User-Wise Models . . . . .                     | 80         |
| 3.4      | Naive Bayes Collaborative Filtering . . . . .                          | 82         |
| 3.4.1    | Handling Overfitting . . . . .   | 84         |
| 3.4.2    | Example of the Bayes Method with Binary Ratings . . . . .              | 85         |
| 3.5      | Using an Arbitrary Classification Model as a Black-Box . . . . .       | 86         |
| 3.5.1    | Example: Using a Neural Network as a Black-Box . . . . .               | 87         |
| 3.6      | Latent Factor Models . . . . .   | 90         |
| 3.6.1    | Geometric Intuition for Latent Factor Models . . . . .                 | 91         |
| 3.6.2    | Low-Rank Intuition for Latent Factor Models . . . . .                  | 93         |
| 3.6.3    | Basic Matrix Factorization Principles . . . . .                        | 94         |
| 3.6.4    | Unconstrained Matrix Factorization . . . . .                           | 96         |
| 3.6.4.1  | Stochastic Gradient Descent . . . . .                                  | 99         |
| 3.6.4.2  | Regularization . . . . .   | 100        |
| 3.6.4.3  | Incremental Latent Component Training . . . . .                        | 103        |
| 3.6.4.4  | Alternating Least Squares and Coordinate Descent . . . . .             | 105        |
| 3.6.4.5  | Incorporating User and Item Biases . . . . .                           | 106        |
| 3.6.4.6  | Incorporating Implicit Feedback . . . . .                              | 109        |
| 3.6.5    | Singular Value Decomposition . . . . .                                 | 113        |
| 3.6.5.1  | A Simple Iterative Approach to SVD . . . . .                           | 114        |
| 3.6.5.2  | An Optimization-Based Approach . . . . .                               | 116        |
| 3.6.5.3  | Out-of-Sample Recommendations . . . . .                                | 116        |
| 3.6.5.4  | Example of Singular Value Decomposition . . . . .                      | 117        |
| 3.6.6    | Non-negative Matrix Factorization . . . . .                            | 119        |
| 3.6.6.1  | Interpretability Advantages . . . . .                                  | 121        |
| 3.6.6.2  | Observations about Factorization with Implicit Feedback . . . . .      | 122        |
| 3.6.6.3  | Computational and Weighting Issues with Implicit<br>Feedback . . . . . | 124        |
| 3.6.6.4  | Ratings with Both Likes and Dislikes . . . . .                         | 124        |
| 3.6.7    | Understanding the Matrix Factorization Family . . . . .                | 126        |
| 3.7      | Integrating Factorization and Neighborhood Models . . . . .            | 128        |
| 3.7.1    | Baseline Estimator: A Non-Personalized Bias-Centric Model . . . . .    | 128        |
| 3.7.2    | Neighborhood Portion of Model . . . . .                                | 129        |
| 3.7.3    | Latent Factor Portion of Model . . . . .                               | 130        |
| 3.7.4    | Integrating the Neighborhood and Latent Factor Portions . . . . .      | 131        |
| 3.7.5    | Solving the Optimization Model . . . . .                               | 131        |
| 3.7.6    | Observations about Accuracy . . . . .                                  | 132        |
| 3.7.7    | Integrating Latent Factor Models with Arbitrary Models . . . . .       | 133        |
| 3.8      | Summary . . . . .  | 134        |
| 3.9      | Bibliographic Notes . . . . .  | 134        |
| 3.10     | Exercises . . . . .  | 136        |
| <b>4</b> | <b>Content-Based Recommender Systems</b> . . . . .                     | <b>139</b> |
| 4.1      | Introduction . . . . .   | 139        |
| 4.2      | Basic Components of Content-Based Systems . . . . .                    | 141        |
| 4.3      | Preprocessing and Feature Extraction . . . . .                         | 142        |
| 4.3.1    | Feature Extraction . . . . .   | 142        |
| 4.3.1.1  | Example of Product Recommendation . . . . .                            | 143        |

|          |  |            |
|----------|--|------------|
| 4.3.1.2  | Example of Web Page Recommendation . . . . .                     | 143        |
| 4.3.1.3  | Example of Music Recommendation . . . . .                        | 144        |
| 4.3.2    | Feature Representation and Cleaning . . . . .                    | 145        |
| 4.3.3    | Collecting User Likes and Dislikes . . . . .                     | 146        |
| 4.3.4    | Supervised Feature Selection and Weighting . . . . .             | 147        |
| 4.3.4.1  | Gini Index . . . . .   | 147        |
| 4.3.4.2  | Entropy . . . . .  | 148        |
| 4.3.4.3  | $\chi^2$ -Statistic . . . . .                                    | 148        |
| 4.3.4.4  | Normalized Deviation . . . . .                                   | 149        |
| 4.3.4.5  | Feature Weighting . . . . .                                      | 150        |
| 4.4      | Learning User Profiles and Filtering . . . . .                   | 150        |
| 4.4.1    | Nearest Neighbor Classification . . . . .                        | 151        |
| 4.4.2    | Connections with Case-Based Recommender Systems . . . . .        | 152        |
| 4.4.3    | Bayes Classifier . . . . .                                       | 153        |
| 4.4.3.1  | Estimating Intermediate Probabilities . . . . .                  | 154        |
| 4.4.3.2  | Example of Bayes Model . . . . .                                 | 155        |
| 4.4.4    | Rule-based Classifiers . . . . .                                 | 156        |
| 4.4.4.1  | Example of Rule-based Methods . . . . .                          | 157        |
| 4.4.5    | Regression-Based Models . . . . .                                | 158        |
| 4.4.6    | Other Learning Models and Comparative Overview . . . . .         | 159        |
| 4.4.7    | Explanations in Content-Based Systems . . . . .                  | 160        |
| 4.5      | Content-Based Versus Collaborative Recommendations . . . . .     | 161        |
| 4.6      | Using Content-Based Models for Collaborative Filtering . . . . . | 162        |
| 4.6.1    | Leveraging User Profiles . . . . .                               | 163        |
| 4.7      | Summary . . . . .  | 163        |
| 4.8      | Bibliographic Notes . . . . .                                    | 164        |
| 4.9      | Exercises . . . . .  | 165        |
| <b>5</b> | <b>Knowledge-Based Recommender Systems</b> . . . . .             | <b>167</b> |
| 5.1      | Introduction . . . . .   | 167        |
| 5.2      | Constraint-Based Recommender Systems . . . . .                   | 172        |
| 5.2.1    | Returning Relevant Results . . . . .                             | 174        |
| 5.2.2    | Interaction Approach . . . . .                                   | 176        |
| 5.2.3    | Ranking the Matched Items . . . . .                              | 178        |
| 5.2.4    | Handling Unacceptable Results or Empty Sets . . . . .            | 179        |
| 5.2.5    | Adding Constraints . . . . .                                     | 180        |
| 5.3      | Case-Based Recommenders . . . . .                                | 181        |
| 5.3.1    | Similarity Metrics . . . . .                                     | 183        |
| 5.3.1.1  | Incorporating Diversity in Similarity Computation . . . . .      | 187        |
| 5.3.2    | Critiquing Methods . . . . .                                     | 188        |
| 5.3.2.1  | Simple Critiques . . . . .                                       | 188        |
| 5.3.2.2  | Compound Critiques . . . . .                                     | 190        |
| 5.3.2.3  | Dynamic Critiques . . . . .                                      | 192        |
| 5.3.3    | Explanation in Critiques . . . . .                               | 193        |
| 5.4      | Persistent Personalization in Knowledge-Based Systems . . . . .  | 194        |
| 5.5      | Summary . . . . .  | 195        |
| 5.6      | Bibliographic Notes . . . . .                                    | 195        |
| 5.7      | Exercises . . . . .  | 197        |

|          |  |            |
|----------|--|------------|
| <b>6</b> | <b>Ensemble-Based and Hybrid Recommender Systems</b>           | <b>199</b> |
| 6.1      | Introduction . . . . .   | 199        |
| 6.2      | Ensemble Methods from the Classification Perspective . . . . . | 204        |
| 6.3      | Weighted Hybrids . . . . .                                     | 206        |
| 6.3.1    | Various Types of Model Combinations . . . . .                  | 208        |
| 6.3.2    | Adapting Bagging from Classification . . . . .                 | 209        |
| 6.3.3    | Randomness Injection . . . . .                                 | 211        |
| 6.4      | Switching Hybrids . . . . .                                    | 211        |
| 6.4.1    | Switching Mechanisms for Cold-Start Issues . . . . .           | 212        |
| 6.4.2    | Bucket-of-Models . . . . .                                     | 212        |
| 6.5      | Cascade Hybrids . . . . .                                      | 213        |
| 6.5.1    | Successive Refinement of Recommendations . . . . .             | 213        |
| 6.5.2    | Boosting . . . . .   | 213        |
| 6.5.2.1  | Weighted Base Models . . . . .                                 | 214        |
| 6.6      | Feature Augmentation Hybrids . . . . .                         | 215        |
| 6.7      | Meta-Level Hybrids . . . . .                                   | 216        |
| 6.8      | Feature Combination Hybrids . . . . .                          | 217        |
| 6.8.1    | Regression and Matrix Factorization . . . . .                  | 218        |
| 6.8.2    | Meta-level Features . . . . .                                  | 218        |
| 6.9      | Mixed Hybrids . . . . .  | 220        |
| 6.10     | Summary . . . . .  | 221        |
| 6.11     | Bibliographic Notes . . . . .                                  | 222        |
| 6.12     | Exercises . . . . .  | 224        |
| <b>7</b> | <b>Evaluating Recommender Systems</b>                          | <b>225</b> |
| 7.1      | Introduction . . . . .   | 225        |
| 7.2      | Evaluation Paradigms . . . . .                                 | 227        |
| 7.2.1    | User Studies . . . . .   | 227        |
| 7.2.2    | Online Evaluation . . . . .                                    | 227        |
| 7.2.3    | Offline Evaluation with Historical Data Sets . . . . .         | 229        |
| 7.3      | General Goals of Evaluation Design . . . . .                   | 229        |
| 7.3.1    | Accuracy . . . . .   | 229        |
| 7.3.2    | Coverage . . . . .   | 231        |
| 7.3.3    | Confidence and Trust . . . . .                                 | 232        |
| 7.3.4    | Novelty . . . . .  | 233        |
| 7.3.5    | Serendipity . . . . .  | 233        |
| 7.3.6    | Diversity . . . . .  | 234        |
| 7.3.7    | Robustness and Stability . . . . .                             | 235        |
| 7.3.8    | Scalability . . . . .  | 235        |
| 7.4      | Design Issues in Offline Recommender Evaluation . . . . .      | 235        |
| 7.4.1    | Case Study of the Netflix Prize Data Set . . . . .             | 236        |
| 7.4.2    | Segmenting the Ratings for Training and Testing . . . . .      | 238        |
| 7.4.2.1  | Hold-Out . . . . .   | 238        |
| 7.4.2.2  | Cross-Validation . . . . .                                     | 239        |
| 7.4.3    | Comparison with Classification Design . . . . .                | 239        |
| 7.5      | Accuracy Metrics in Offline Evaluation . . . . .               | 240        |
| 7.5.1    | Measuring the Accuracy of Ratings Prediction . . . . .         | 240        |
| 7.5.1.1  | RMSE versus MAE . . . . .                                      | 241        |
| 7.5.1.2  | Impact of the Long Tail . . . . .                              | 241        |

|          |  |            |
|----------|--|------------|
| 7.5.2    | Evaluating Ranking via Correlation . . . . .                           | 242        |
| 7.5.3    | Evaluating Ranking via Utility . . . . .                               | 244        |
| 7.5.4    | Evaluating Ranking via Receiver Operating Characteristic . . . . .     | 247        |
| 7.5.5    | Which Ranking Measure is Best? . . . . .                               | 250        |
| 7.6      | Limitations of Evaluation Measures . . . . .                           | 250        |
| 7.6.1    | Avoiding Evaluation Gaming . . . . .                                   | 252        |
| 7.7      | Summary . . . . .  | 252        |
| 7.8      | Bibliographic Notes . . . . .  | 253        |
| 7.9      | Exercises . . . . .  | 254        |
| <b>8</b> | <b>Context-Sensitive Recommender Systems</b> . . . . .                 | <b>255</b> |
| 8.1      | Introduction . . . . .   | 255        |
| 8.2      | The Multidimensional Approach . . . . .                                | 256        |
| 8.2.1    | The Importance of Hierarchies . . . . .                                | 259        |
| 8.3      | Contextual Pre-filtering: A Reduction-Based Approach . . . . .         | 262        |
| 8.3.1    | Ensemble-Based Improvements . . . . .                                  | 264        |
| 8.3.2    | Multi-level Estimation . . . . .                                       | 265        |
| 8.4      | Post-Filtering Methods . . . . .                                       | 266        |
| 8.5      | Contextual Modeling . . . . .  | 268        |
| 8.5.1    | Neighborhood-Based Methods . . . . .                                   | 268        |
| 8.5.2    | Latent Factor Models . . . . .   | 269        |
| 8.5.2.1  | Factorization Machines . . . . .                                       | 272        |
| 8.5.2.2  | A Generalized View of Second-Order Factorization<br>Machines . . . . . | 275        |
| 8.5.2.3  | Other Applications of Latent Parametrization . . . . .                 | 276        |
| 8.5.3    | Content-Based Models . . . . .   | 277        |
| 8.6      | Summary . . . . .  | 279        |
| 8.7      | Bibliographic Notes . . . . .  | 280        |
| 8.8      | Exercises . . . . .  | 281        |
| <b>9</b> | <b>Time- and Location-Sensitive Recommender Systems</b> . . . . .      | <b>283</b> |
| 9.1      | Introduction . . . . .   | 283        |
| 9.2      | Temporal Collaborative Filtering . . . . .                             | 285        |
| 9.2.1    | Recency-Based Models . . . . .   | 286        |
| 9.2.1.1  | Decay-Based Methods . . . . .  | 286        |
| 9.2.1.2  | Window-Based Methods . . . . .   | 288        |
| 9.2.2    | Handling Periodic Context . . . . .                                    | 288        |
| 9.2.2.1  | Pre-Filtering and Post-Filtering . . . . .                             | 289        |
| 9.2.2.2  | Direct Incorporation of Temporal Context . . . . .                     | 290        |
| 9.2.3    | Modeling Ratings as a Function of Time . . . . .                       | 290        |
| 9.2.3.1  | The Time-SVD++ Model . . . . .   | 291        |
| 9.3      | Discrete Temporal Models . . . . .                                     | 295        |
| 9.3.1    | Markovian Models . . . . .   | 295        |
| 9.3.1.1  | Selective Markov Models . . . . .                                      | 298        |
| 9.3.1.2  | Other Markovian Alternatives . . . . .                                 | 300        |
| 9.3.2    | Sequential Pattern Mining . . . . .                                    | 300        |
| 9.4      | Location-Aware Recommender Systems . . . . .                           | 302        |
| 9.4.1    | Preference Locality . . . . .  | 303        |
| 9.4.2    | Travel Locality . . . . .  | 305        |
| 9.4.3    | Combined Preference and Travel Locality . . . . .                      | 305        |

|           |   |            |
|-----------|---|------------|
| 9.5       | Summary . . . . .   | 305        |
| 9.6       | Bibliographic Notes . . . . .   | 306        |
| 9.7       | Exercises . . . . .   | 308        |
| <b>10</b> | <b>Structural Recommendations in Networks</b>                             | <b>309</b> |
| 10.1      | Introduction . . . . .  | 309        |
| 10.2      | Ranking Algorithms . . . . .  | 311        |
| 10.2.1    | PageRank . . . . .  | 311        |
| 10.2.2    | Personalized PageRank . . . . .   | 314        |
| 10.2.3    | Applications to Neighborhood-Based Methods . . . . .                      | 316        |
| 10.2.3.1  | Social Network Recommendations . . . . .                                  | 317        |
| 10.2.3.2  | Personalization in Heterogeneous Social Media . . . . .                   | 317        |
| 10.2.3.3  | Traditional Collaborative Filtering . . . . .                             | 319        |
| 10.2.4    | SimRank . . . . .   | 321        |
| 10.2.5    | The Relationship Between Search and Recommendation . . . . .              | 322        |
| 10.3      | Recommendations by Collective Classification . . . . .                    | 323        |
| 10.3.1    | Iterative Classification Algorithm . . . . .                              | 324        |
| 10.3.2    | Label Propagation with Random Walks . . . . .                             | 325        |
| 10.3.3    | Applicability to Collaborative Filtering in Social Networks . . . . .     | 326        |
| 10.4      | Recommending Friends: Link Prediction . . . . .                           | 326        |
| 10.4.1    | Neighborhood-Based Measures . . . . .                                     | 327        |
| 10.4.2    | Katz Measure . . . . .  | 328        |
| 10.4.3    | Random Walk-Based Measures . . . . .                                      | 329        |
| 10.4.4    | Link Prediction as a Classification Problem . . . . .                     | 329        |
| 10.4.5    | Matrix Factorization for Link Prediction . . . . .                        | 330        |
| 10.4.5.1  | Symmetric Matrix Factorization . . . . .                                  | 333        |
| 10.4.6    | Connections Between Link Prediction and Collaborative Filtering . . . . . | 335        |
| 10.4.6.1  | Using Link Prediction Algorithms for Collaborative Filtering . . . . .    | 336        |
| 10.4.6.2  | Using Collaborative Filtering Algorithms for Link Prediction . . . . .    | 337        |
| 10.5      | Social Influence Analysis and Viral Marketing . . . . .                   | 337        |
| 10.5.1    | Linear Threshold Model . . . . .  | 339        |
| 10.5.2    | Independent Cascade Model . . . . .                                       | 340        |
| 10.5.3    | Influence Function Evaluation . . . . .                                   | 340        |
| 10.5.4    | Targeted Influence Analysis Models in Social Streams . . . . .            | 341        |
| 10.6      | Summary . . . . .   | 342        |
| 10.7      | Bibliographic Notes . . . . .   | 343        |
| 10.8      | Exercises . . . . .   | 344        |
| <b>11</b> | <b>Social and Trust-Centric Recommender Systems</b>                       | <b>345</b> |
| 11.1      | Introduction . . . . .  | 345        |
| 11.2      | Multidimensional Models for Social Context . . . . .                      | 347        |
| 11.3      | Network-Centric and Trust-Centric Methods . . . . .                       | 349        |
| 11.3.1    | Collecting Data for Building Trust Networks . . . . .                     | 349        |
| 11.3.2    | Trust Propagation and Aggregation . . . . .                               | 351        |
| 11.3.3    | Simple Recommender with No Trust Propagation . . . . .                    | 353        |
| 11.3.4    | TidalTrust Algorithm . . . . .  | 353        |

|           |  |            |
|-----------|--|------------|
| 11.3.5    | MoleTrust Algorithm . . . . .                                    | 356        |
| 11.3.6    | TrustWalker Algorithm . . . . .                                  | 357        |
| 11.3.7    | Link Prediction Methods . . . . .                                | 358        |
| 11.3.8    | Matrix Factorization Methods . . . . .                           | 361        |
| 11.3.8.1  | Enhancements with Logistic Function . . . . .                    | 364        |
| 11.3.8.2  | Variations in the Social Trust Component . . . . .               | 364        |
| 11.3.9    | Merits of Social Recommender Systems . . . . .                   | 365        |
| 11.3.9.1  | Recommendations for Controversial Users and Items . . . . .      | 365        |
| 11.3.9.2  | Usefulness for Cold-Start . . . . .                              | 366        |
| 11.3.9.3  | Attack Resistance . . . . .                                      | 366        |
| 11.4      | User Interaction in Social Recommenders . . . . .                | 366        |
| 11.4.1    | Representing Folksonomies . . . . .                              | 367        |
| 11.4.2    | Collaborative Filtering in Social Tagging Systems . . . . .      | 368        |
| 11.4.3    | Selecting Valuable Tags . . . . .                                | 371        |
| 11.4.4    | Social-Tagging Recommenders with No Ratings Matrix . . . . .     | 372        |
| 11.4.4.1  | Multidimensional Methods for Context-Sensitive Systems . . . . . | 372        |
| 11.4.4.2  | Ranking-Based Methods . . . . .                                  | 373        |
| 11.4.4.3  | Content-Based Methods . . . . .                                  | 374        |
| 11.4.5    | Social-Tagging Recommenders with Ratings Matrix . . . . .        | 377        |
| 11.4.5.1  | Neighborhood-Based Approach . . . . .                            | 378        |
| 11.4.5.2  | Linear Regression . . . . .                                      | 379        |
| 11.4.5.3  | Matrix Factorization . . . . .                                   | 380        |
| 11.4.5.4  | Content-Based Methods . . . . .                                  | 382        |
| 11.5      | Summary . . . . .  | 382        |
| 11.6      | Bibliographic Notes . . . . .                                    | 382        |
| 11.7      | Exercises . . . . .  | 384        |
| <b>12</b> | <b>Attack-Resistant Recommender Systems</b> . . . . .            | <b>385</b> |
| 12.1      | Introduction . . . . .   | 385        |
| 12.2      | Understanding the Trade-Offs in Attack Models . . . . .          | 386        |
| 12.2.1    | Quantifying Attack Impact . . . . .                              | 390        |
| 12.3      | Types of Attacks . . . . .                                       | 392        |
| 12.3.1    | Random Attack . . . . .  | 393        |
| 12.3.2    | Average Attack . . . . .   | 393        |
| 12.3.3    | Bandwagon Attack . . . . .                                       | 394        |
| 12.3.4    | Popular Attack . . . . .   | 395        |
| 12.3.5    | Love/Hate Attack . . . . .                                       | 395        |
| 12.3.6    | Reverse Bandwagon Attack . . . . .                               | 396        |
| 12.3.7    | Probe Attack . . . . .   | 396        |
| 12.3.8    | Segment Attack . . . . .   | 396        |
| 12.3.9    | Effect of Base Recommendation Algorithm . . . . .                | 397        |
| 12.4      | Detecting Attacks on Recommender Systems . . . . .               | 398        |
| 12.4.1    | Individual Attack Profile Detection . . . . .                    | 399        |
| 12.4.2    | Group Attack Profile Detection . . . . .                         | 402        |
| 12.4.2.1  | Preprocessing Methods . . . . .                                  | 402        |
| 12.4.2.2  | Online Methods . . . . .   | 403        |
| 12.5      | Strategies for Robust Recommender Design . . . . .               | 403        |
| 12.5.1    | Preventing Automated Attacks with CAPTCHAs . . . . .             | 403        |
| 12.5.2    | Using Social Trust . . . . .                                     | 404        |

|           |  |            |
|-----------|--|------------|
| 12.5.3    | Designing Robust Recommendation Algorithms . . . . .             | 404        |
| 12.5.3.1  | Incorporating Clustering in Neighborhood Methods . . . . .       | 405        |
| 12.5.3.2  | Fake Profile Detection during Recommendation Time . . . . .      | 405        |
| 12.5.3.3  | Association-Based Algorithms . . . . .                           | 405        |
| 12.5.3.4  | Robust Matrix Factorization . . . . .                            | 405        |
| 12.6      | Summary . . . . .  | 408        |
| 12.7      | Bibliographic Notes . . . . .                                    | 408        |
| 12.8      | Exercises . . . . .  | 410        |
| <b>13</b> | <b>Advanced Topics in Recommender Systems</b> . . . . .          | <b>411</b> |
| 13.1      | Introduction . . . . .   | 411        |
| 13.2      | Learning to Rank . . . . .                                       | 413        |
| 13.2.1    | Pairwise Rank Learning . . . . .                                 | 415        |
| 13.2.2    | Listwise Rank Learning . . . . .                                 | 416        |
| 13.2.3    | Comparison with Rank-Learning Methods in Other Domains . . . . . | 417        |
| 13.3      | Multi-Armed Bandit Algorithms . . . . .                          | 418        |
| 13.3.1    | Naive Algorithm . . . . .  | 419        |
| 13.3.2    | $\epsilon$ -Greedy Algorithm . . . . .                           | 420        |
| 13.3.3    | Upper Bounding Methods . . . . .                                 | 421        |
| 13.4      | Group Recommender Systems . . . . .                              | 423        |
| 13.4.1    | Collaborative and Content-Based Systems . . . . .                | 424        |
| 13.4.2    | Knowledge-Based Systems . . . . .                                | 425        |
| 13.5      | Multi-Criteria Recommender Systems . . . . .                     | 426        |
| 13.5.1    | Neighborhood-Based Methods . . . . .                             | 427        |
| 13.5.2    | Ensemble-Based Methods . . . . .                                 | 428        |
| 13.5.3    | Multi-Criteria Systems without Overall Ratings . . . . .         | 429        |
| 13.6      | Active Learning in Recommender Systems . . . . .                 | 430        |
| 13.6.1    | Heterogeneity-Based Models . . . . .                             | 431        |
| 13.6.2    | Performance-Based Models . . . . .                               | 432        |
| 13.7      | Privacy in Recommender Systems . . . . .                         | 432        |
| 13.7.1    | Condensation-Based Privacy . . . . .                             | 434        |
| 13.7.2    | Challenges for High-Dimensional Data . . . . .                   | 434        |
| 13.8      | Some Interesting Application Domains . . . . .                   | 435        |
| 13.8.1    | Portal Content Personalization . . . . .                         | 435        |
| 13.8.1.1  | Dynamic Profiler . . . . .                                       | 436        |
| 13.8.1.2  | Google News Personalization . . . . .                            | 436        |
| 13.8.2    | Computational Advertising versus Recommender Systems . . . . .   | 438        |
| 13.8.2.1  | Importance of Multi-Armed Bandit Methods . . . . .               | 442        |
| 13.8.3    | Reciprocal Recommender Systems . . . . .                         | 443        |
| 13.8.3.1  | Leveraging Hybrid Methods . . . . .                              | 444        |
| 13.8.3.2  | Leveraging Link Prediction Methods . . . . .                     | 445        |
| 13.9      | Summary . . . . .  | 446        |
| 13.10     | Bibliographic Notes . . . . .                                    | 446        |
|           | <b>Bibliography</b> . . . . .                                    | <b>449</b> |
|           | <b>Index</b> . . . . .   | <b>493</b> |





---

---

# Preface

---

---

*“Nature shows us only the tail of the lion. But I do not doubt that the lion belongs to it even though he cannot at once reveal himself because of his enormous size.”* – Albert Einstein

The topic of recommender systems gained increasing importance in the nineties, as the Web became an important medium for business and e-commerce transactions. It was recognized early on that the Web provided unprecedented opportunities for personalization, which were not available in other channels. In particular, the Web provided ease in data collection and a user interface that could be employed to recommend items in a non-intrusive way.

Recommender systems have grown significantly in terms of public awareness since then. An evidence of this fact is that many conferences and workshops are exclusively devoted to this topic. The *ACM Conference on Recommender Systems* is particularly notable because it regularly contributes many of the cutting-edge results in this topic. The topic of recommender systems is very diverse because it enables the ability to use various types of user-preference and user-requirement data to make recommendations. The most well-known methods in recommender systems include collaborative filtering methods, content-based methods, and knowledge-based methods. These three methods form the fundamental pillars of research in recommender systems. In recent years, specialized methods have been designed for various data domains and contexts, such as time, location and social information. Numerous advancements have been proposed for specialized scenarios, and the methods have been adapted to various application domains, such as query log mining, news recommendations, and computational advertising. The organization of the book reflects these important topics. The chapters of this book can be organized into three categories:

1. *Algorithms and evaluation*: These chapters discuss the fundamental algorithms in recommender systems, including collaborative filtering methods (Chapters 2 and 4), content-based methods (Chapter 4), and knowledge-based methods (Chapter 5). Techniques for hybridizing these methods are discussed in Chapter 6. The evaluation of recommender systems is discussed in Chapter 7.
2. *Recommendations in specific domains and contexts*: The context of a recommender system plays a critical role in providing effective recommendations. For example, a

user looking for a restaurant would want to use their location as additional *context*. The context of a recommendation can be viewed as important side information that affects the recommendation goals. Different types of domains such as temporal data, spatial data, and social data, provide different types of contexts. These methods are discussed in Chapters 8, 9, 10, and 11. Chapter 11 also discusses the issue of using social information to increase the trustworthiness of the recommendation process. Recent topics such as factorization machines and trustworthy recommender systems are also covered in these chapters.

3. *Advanced topics and applications*: In Chapter 12, we discuss various robustness aspects of recommender systems, such as shilling systems, attack models, and their defenses. In addition, recent topics, such as learning to rank, multi-armed bandits, group recommender systems, multi-criteria systems, and active learning systems, are discussed in Chapter 13. An important goal of this chapter is to introduce the reader to the basic ideas and principles underlying recent developments. Although it is impossible to discuss all the recent developments in detail in a single book, it is hoped that the material in the final chapter will play the role of “breaking the ice” for the reader in terms of advanced topics. This chapter also investigates some application settings in which recommendation technology is used, such as news recommendations, query recommendations, and computational advertising. The application section provides an idea of how the methods introduced in earlier chapters apply to these different domains.

Although this book is primarily written as a textbook, it is recognized that a large portion of the audience will comprise industrial practitioners and researchers. Therefore, we have taken pains to write the book in such a way that it is also useful from an applied and reference point of view. Numerous examples and exercises have been provided to enable its use as a textbook. As most courses on recommender systems will teach only the fundamental topics, the chapters on fundamental topics and algorithms are written with a particular emphasis on classroom teaching. On the other hand, advanced industrial practitioners might find the chapters on context-sensitive recommendation useful, because many real-life applications today arise in the domains where a significant amount of contextual side-information is available. The application portion of Chapter 13 is particularly written for industrial practitioners, although instructors might find it useful towards the end of a recommender course.

We conclude with a brief introduction to the notations used in this book. This book consistently uses an  $m \times n$  ratings matrix denoted by  $R$ , where  $m$  is the number of users and  $n$  is the number of items. The matrix  $R$  is typically incomplete because only a subset of entries are observed. The  $(i, j)$ th entry of  $R$  indicates the rating of user  $i$  for item  $j$ , and it is denoted by  $r_{ij}$  when it is actually observed. When the entry  $(i, j)$  is *predicted* by a recommender algorithm (rather than being specified by a user), it is denoted by  $\hat{r}_{ij}$ , with a “hat” symbol (i.e., a circumflex) denoting that it is a predicted value. Vectors are denoted by an “overline,” as in  $\overline{X}$  or  $\overline{y}$ .

---

---

# Acknowledgments

---

---

I would like to thank my wife and daughter for their love and support during the writing of this book. I also owe my late parents a debt of gratitude for instilling in me a love of education, which has played an important inspirational role in my book-writing efforts.

This book has been written with the direct and indirect support of many individuals to whom I am grateful. During the writing of this book, I received feedback from many colleagues. In particular, I received feedback from Xavier Amatriain, Kanishka Bhaduri, Robin Burke, Martin Ester, Bart Goethals, Huan Liu, Xia Ning, Saket Sathe, Jiliang Tang, Alexander Tuzhilin, Koen Versetrepren, and Jieping Ye. I would like to thank them for their constructive feedback and suggestions. Over the years, I have benefited from the insights of numerous collaborators. These insights have influenced this book directly or indirectly. I would first like to thank my long-term collaborator Philip S. Yu for my years of collaboration with him. Other researchers with whom I have had significant collaborations include Tarek F. Abdelzaher, Jing Gao, Quanquan Gu, Manish Gupta, Jiawei Han, Alexander Hinneburg, Thomas Huang, Nan Li, Huan Liu, Ruoming Jin, Daniel Keim, Arijit Khan, Latifur Khan, Mohammad M. Masud, Jian Pei, Magda Procopiuc, Guojun Qi, Chandan Reddy, Saket Sathe, Jaideep Srivastava, Karthik Subbian, Yizhou Sun, Jiliang Tang, Min-Hsuan Tsai, Haixun Wang, Jianyong Wang, Min Wang, Joel Wolf, Xifeng Yan, Mohammed Zaki, ChengXiang Zhai, and Peixiang Zhao. I would also like to thank my advisor James B. Orlin for his guidance during my early years as a researcher.

I would also like to thank my manager Nagui Halim for providing the tremendous support necessary for the writing of this book. His professional support has been instrumental for my many book efforts in the past and present.

Finally, I would like to thank Lata Aggarwal for helping me with some of the figures drawn using Microsoft Powerpoint.



---

## Author Biography

---

**Charu C. Aggarwal** is a Distinguished Research Staff Member (DRSM) at the IBM T. J. Watson Research Center in Yorktown Heights, New York. He completed his B.S. from IIT Kanpur in 1993 and his Ph.D. from the Massachusetts Institute of Technology in 1996.



He has worked extensively in the field of data mining. He has published more than 300 papers in refereed conferences and journals and authored over 80 patents. He is the author or editor of 15 books, including a textbook on data mining and a comprehensive book on outlier analysis. Because of the commercial value of his patents, he has thrice been designated a Master Inventor at IBM. He is a recipient of an IBM Corporate Award (2003) for his work on bio-terrorist threat detection in data streams, a recipient of the IBM Outstanding Innovation Award (2008) for his scientific contributions to privacy technology, a recipient of two IBM Outstanding Technical Achievement Award (2009,

2015) for his work on data streams and high-dimensional data, respectively. He received the EDBT 2014 Test of Time Award for his work on condensation-based privacy-preserving data mining. He is also a recipient of the IEEE ICDM Research Contributions Award (2015), which is one of the two highest awards for influential research contributions in the field of data mining.

He has served as the general co-chair of the IEEE Big Data Conference (2014), program co-chair of the ACM CIKM Conference (2015), IEEE ICDM Conference (2015), and the ACM KDD Conference (2016). He served as an associate editor of the IEEE Transactions on Knowledge and Data Engineering from 2004 to 2008. He is an associate editor of the ACM Transactions on Knowledge Discovery from Data, an associate editor of the IEEE Transactions on Big Data, an action editor of the Data Mining and Knowledge Discovery Journal, editor-in-chief of the ACM SIGKDD Explorations, and an associate editor of the Knowledge and Information Systems Journal. He serves on the advisory board of the Lecture Notes on Social Networks, a publication by Springer. He has served as the vice-president of the SIAM Activity Group on Data Mining. He is a fellow of the SIAM, ACM, and the IEEE, for “contributions to knowledge discovery and data mining algorithms.”