Contents

Preface xix

Section 1: Theoretical Fundamentals 1

1 Mathematical Foundation 3
Afroz and Basharat Hussain

1.1 Concept of Linear Algebra 3
1.1.1 Introduction 3
1.1.2 Vector Spaces 5
1.1.3 Linear Combination 6
1.1.4 Linearly Dependent and Independent Vectors 7
1.1.5 Linear Span, Basis and Subspace 8
1.1.6 Linear Transformation (or Linear Map) 9
1.1.7 Matrix Representation of Linear Transformation 10
1.1.8 Range and Null Space of Linear Transformation 13
1.1.9 Invertible Linear Transformation 15
1.2 Eigenvalues, Eigenvectors, and Eigendecomposition of a Matrix 15
1.2.1 Characteristics Polynomial 16
1.2.1.1 Some Results on Eigenvalue 16
1.2.2 Eigendecomposition 18
1.3 Introduction to Calculus 20
1.3.1 Function 20
1.3.2 Limits of Functions 21
1.3.2.1 Some Properties of Limits 22
1.3.2.2 Infinite Limits 25
1.3.2.3 Limits at Infinity 26
1.3.3 Continuous Functions and Discontinuous Functions 26
1.3.3.1 Discontinuous Functions 27
1.3.3.2 Properties of Continuous Function 27
1.3.4 Differentiation 28
References 29

2 Theory of Probability **31**
Parvaze Ahmad Dar and Afroz
2.1 Introduction 31
2.1.1 Definition 31
2.1.1.1 Statistical Definition of Probability 31
2.1.1.2 Mathematical Definition of Probability 32
2.1.2 Some Basic Terms of Probability 32
2.1.2.1 Trial and Event 32
2.1.2.2 Exhaustive Events (Exhaustive Cases) 33
2.1.2.3 Mutually Exclusive Events 33
2.1.2.4 Equally Likely Events 33
2.1.2.5 Certain Event or Sure Event 33
2.1.2.6 Impossible Event or Null Event (ϕ) 33
2.1.2.7 Sample Space 34
2.1.2.8 Permutation and Combination 34
2.1.2.9 Examples 35
2.2 Independence in Probability 38
2.2.1 Independent Events 38
2.2.2 Examples: Solve the Following Problems 38
2.3 Conditional Probability 41
2.3.1 Definition 41
2.3.2 Mutually Independent Events 42
2.3.3 Examples 42
2.4 Cumulative Distribution Function 43
2.4.1 Properties 44
2.4.2 Example 44
2.5 Baye's Theorem 46
2.5.1 Theorem 46
2.5.1.1 Examples 47
2.6 Multivariate Gaussian Function 50
2.6.1 Definition 50
2.6.1.1 Univariate Gaussian (i.e., One Variable Gaussian) 50
2.6.1.2 Degenerate Univariate Gaussian 51
2.6.1.3 Multivariate Gaussian 51
References 51

3 Correlation and Regression **53**
Mohd. Abdul Haleem Rizwan
3.1 Introduction 53

3.2 Correlation 54
3.2.1 Positive Correlation and Negative Correlation 54
3.2.2 Simple Correlation and Multiple Correlation 54
3.2.3 Partial Correlation and Total Correlation 54
3.2.4 Correlation Coefficient 55
3.3 Regression 57
3.3.1 Linear Regression 64
3.3.2 Logistic Regression 64
3.3.3 Polynomial Regression 65
3.3.4 Stepwise Regression 66
3.3.5 Ridge Regression 67
3.3.6 Lasso Regression 67
3.3.7 Elastic Net Regression 68
3.4 Conclusion 68
References 69

Section 2: Big Data and Pattern Recognition 71

4 Data Preprocess 73
Md. Sharif Hossen
4.1 Introduction 73
4.1.1 Need of Data Preprocessing 74
4.1.2 Main Tasks in Data Preprocessing 75
4.2 Data Cleaning 77
4.2.1 Missing Data 77
4.2.2 Noisy Data 78
4.3 Data Integration 80
4.3.1 $\chi 2$ Correlation Test 82
4.3.2 Correlation Coefficient Test 82
4.3.3 Covariance Test 83
4.4 Data Transformation 83
4.4.1 Normalization 83
4.4.2 Attribute Selection 85
4.4.3 Discretization 86
4.4.4 Concept Hierarchy Generation 86
4.5 Data Reduction 88
4.5.1 Data Cube Aggregation 88
4.5.2 Attribute Subset Selection 90
4.5.3 Numerosity Reduction 91
4.5.4 Dimensionality Reduction 95
4.6 Conclusion 101

Acknowledgements 101
References 101

5 Big Data 105
R. Chinnaiyan
5.1 Introduction 105
5.2 Big Data Evaluation With Its Tools 107
5.3 Architecture of Big Data 107
5.3.1 Big Data Analytics Framework Workflow 107
5.4 Issues and Challenges 109
5.4.1 Volume 109
5.4.2 Variety of Data 110
5.4.3 Velocity 110
5.5 Big Data Analytics Tools 110
5.6 Big Data Use Cases 114
5.6.1 Banking and Finance 114
5.6.2 Fraud Detection 114
5.6.3 Customer Division and Personalized Marketing 114
5.6.4 Customer Support 115
5.6.5 Risk Management 116
5.6.6 Life Time Value Prediction 116
5.6.7 Cyber Security Analytics 117
5.6.8 Insurance Industry 118
5.6.9 Health Care Sector 118
5.6.9.1 Big Data Medical Decision Support 120
5.6.9.2 Big Data–Based Disorder Management 120
5.6.9.3 Big Data–Based Patient Monitoring and Control 120
5.6.9.4 Big Data–Based Human Routine Analytics 120
5.6.10 Internet of Things 121
5.6.11 Weather Forecasting 121
5.7 Where IoT Meets Big Data 122
5.7.1 IoT Platform 122
5.7.2 Sensors or Devices 123
5.7.3 Device Aggregators 123
5.7.4 IoT Gateway 123
5.7.5 Big Data Platform and Tools 124
5.8 Role of Machine Learning For Big Data and IoT 124
5.8.1 Typical Machine Learning Use Cases 125
5.9 Conclusion 126
References 127

6 Pattern Recognition Concepts **131**
Ambeshwar Kumar, R. Manikandan and C. Thaventhiran
6.1 Classifier 132
6.1.1 Introduction 132
6.1.2 Explanation-Based Learning 133
6.1.3 Isomorphism and Clique Method 135
6.1.4 Context-Dependent Classification 138
6.1.5 Summary 139
6.2 Feature Processing 140
6.2.1 Introduction 140
6.2.2 Detection and Extracting Edge With Boundary Line 141
6.2.3 Analyzing the Texture 142
6.2.4 Feature Mapping in Consecutive Moving Frame 143
6.2.5 Summary 145
6.3 Clustering 145
6.3.1 Introduction 145
6.3.2 Types of Clustering Algorithms 146
6.3.2.1 Dynamic Clustering Method 148
6.3.2.2 Model-Based Clustering 148
6.3.3 Application 149
6.3.4 Summary 150
6.4 Conclusion 151
References 151

Section 3: Machine Learning: Algorithms & Applications **153**

7 Machine Learning **155**
Elham Ghanbari and Sara Najafzadeh
7.1 History and Purpose of Machine Learning 155
7.1.1 History of Machine Learning 155
7.1.1.1 What is Machine Learning? 156
7.1.1.2 When the Machine Learning is Needed? 157
7.1.2 Goals and Achievements in Machine Learning 158
7.1.3 Applications of Machine Learning 158
7.1.3.1 Practical Machine Learning Examples 159
7.1.4 Relation to Other Fields 161
7.1.4.1 Data Mining 161
7.1.4.2 Artificial Intelligence 162
7.1.4.3 Computational Statistics 162
7.1.4.4 Probability 163

7.1.5 Limitations of Machine Learning 163
7.2 Concept of Well-Defined Learning Problem 164
7.2.1 Concept Learning 164
7.2.1.1 Concept Representation 166
7.2.1.2 Instance Representation 167
7.2.1.3 The Inductive Learning Hypothesis 167
7.2.2 Concept Learning as Search 167
7.2.2.1 Concept Generality 168
7.3 General-to-Specific Ordering Over Hypotheses 169
7.3.1 Basic Concepts: Hypothesis, Generality 169
7.3.2 Structure of the Hypothesis Space 169
7.3.2.1 Hypothesis Notations 169
7.3.2.2 Hypothesis Evaluations 170
7.3.3 Ordering on Hypotheses: General to Specific 170
7.3.3.1 Most Specific Generalized 171
7.3.3.2 Most General Specialized 173
7.3.3.3 Generalization and Specialization Operators 173
7.3.4 Hypothesis Space Search by Find-S Algorithm 174
7.3.4.1 Properties of the Find-S Algorithm 176
7.3.4.2 Limitations of the Find-S Algorithm 176
7.4 Version Spaces and Candidate Elimination Algorithm 177
7.4.1 Representing Version Spaces 177
7.4.1.1 General Boundary 178
7.4.1.2 Specific Boundary 178
7.4.2 Version Space as Search Strategy 179
7.4.3 The List-Eliminate Method 179
7.4.4 The Candidate-Elimination Method 180
7.4.4.1 Example 181
7.4.4.2 Convergence of Candidate-Elimination Method 183
7.4.4.3 Inductive Bias for Candidate-Elimination 184
7.5 Concepts of Machine Learning Algorithm 185
7.5.1 Types of Learning Algorithms 185
7.5.1.1 Incremental vs. Batch Learning Algorithms 186
7.5.1.2 Offline vs. Online Learning Algorithms 188
7.5.1.3 Inductive vs. Deductive Learning Algorithms 189
7.5.2 A Framework for Machine Learning Algorithms 189
7.5.2.1 Training Data 190
7.5.2.2 Target Function 190
7.5.2.3 Construction Model 191

7.5.2.4 Evaluation 191
7.5.3 Types of Machine Learning Algorithms 194
7.5.3.1 Supervised Learning 196
7.5.3.2 Unsupervised Learning 198
7.5.3.3 Semi Supervised Learning 200
7.5.3.4 Reinforcement Learning 200
7.5.3.5 Deep Learning 202
7.5.4 Types of Machine Learning Problems 203
7.5.4.1 Classification 204
7.5.4.2 Clustering 204
7.5.4.3 Optimization 205
7.5.4.4 Regression 205
Conclusion 205
References 206

8 Performance of Supervised Learning Algorithms on Multi-Variate Datasets 209
Asif Iqbal Hajamydeen and Rabab Alayham Abbas Helmi
8.1 Introduction 209
8.2 Supervised Learning Algorithms 210
8.2.1 Datasets and Experimental Setup 211
8.2.2 Data Treatment/Preprocessing 212
8.3 Classification 212
8.3.1 Support Vector Machines (SVM) 213
8.3.2 Naïve Bayes (NB) Algorithm 214
8.3.3 Bayesian Network (BN) 214
8.3.4 Hidden Markov Model (HMM) 215
8.3.5 K-Nearest Neighbour (KNN) 216
8.3.6 Training Time 216
8.4 Neural Network 217
8.4.1 Artificial Neural Networks Architecture 219
8.4.2 Application Areas 222
8.4.3 Artificial Neural Networks and Time Series 224
8.5 Comparisons and Discussions 225
8.5.1 Comparison of Classification Accuracy 225
8.5.2 Forecasting Efficiency Comparison 226
8.5.3 Recurrent Neural Network (RNN) 226
8.5.4 Backpropagation Neural Network (BPNN) 228
8.5.5 General Regression Neural Network 229
8.6 Summary and Conclusion 230
References 231

9 Unsupervised Learning **233**
M. Kumara Swamy and Tejaswi Puligilla
9.1 Introduction 233
9.2 Related Work 234
9.3 Unsupervised Learning Algorithms 235
9.4 Classification of Unsupervised Learning Algorithms 238
9.4.1 Hierarchical Methods 238
9.4.2 Partitioning Methods 239
9.4.3 Density-Based Methods 242
9.4.4 Grid-Based Methods 245
9.4.5 Constraint-Based Clustering 245
9.5 Unsupervised Learning Algorithms in ML 246
9.5.1 Parametric Algorithms 246
9.5.2 Non-Parametric Algorithms 246
9.5.3 Dirichlet Process Mixture Model 247
9.5.4 X-Means 248
9.6 Summary and Conclusions 248
References 248

10 Semi-Supervised Learning **251**
Manish Devgan, Gaurav Malik and Deepak Kumar Sharma
10.1 Introduction 252
10.1.1 Semi-Supervised Learning 252
10.1.2 Comparison With Other Paradigms 255
10.2 Training Models 257
10.2.1 Self-Training 257
10.2.2 Co-Training 259
10.3 Generative Models—Introduction 261
10.3.1 Image Classification 264
10.3.2 Text Categorization 266
10.3.3 Speech Recognition 268
10.3.4 Baum-Welch Algorithm 268
10.4 S3VMs 270
10.5 Graph-Based Algorithms 274
10.5.1 Mincut 275
10.5.2 Harmonic 276
10.5.3 Manifold Regularization 277
10.6 Multiview Learning 277
10.7 Conclusion 278
References 279

11 Reinforcement Learning **281**
Amandeep Singh Bhatia, Mandeep Kaur Saggi, Amit Sundas and Jatinder Ashta
11.1 Introduction: Reinforcement Learning 281
11.1.1 Elements of Reinforcement Learning 283
11.2 Model-Free RL 284
11.2.1 Q-Learning 285
11.2.2 R-Learning 286
11.3 Model-Based RL 287
11.3.1 SARSA Learning 289
11.3.2 Dyna-Q Learning 290
11.3.3 Temporal Difference 291
11.3.3.1 TD(0) Algorithm 292
11.3.3.2 TD(1) Algorithm 293
11.3.3.3 TD(λ) Algorithm 294
11.3.4 Monte Carlo Method 294
11.3.4.1 Monte Carlo Reinforcement Learning 296
11.3.4.2 Monte Carlo Policy Evaluation 296
11.3.4.3 Monte Carlo Policy Improvement 298
11.4 Conclusion 298
References 299

12 Application of Big Data and Machine Learning **305**
Neha Sharma, Sunil Kumar Gautam, Azriel A. Henry and Abhimanyu Kumar
12.1 Introduction 306
12.2 Motivation 307
12.3 Related Work 308
12.4 Application of Big Data and ML 309
12.4.1 Healthcare 309
12.4.2 Banking and Insurance 312
12.4.3 Transportation 314
12.4.4 Media and Entertainment 316
12.4.5 Education 317
12.4.6 Ecosystem Conservation 319
12.4.7 Manufacturing 321
12.4.8 Agriculture 322
12.5 Issues and Challenges 324
12.6 Conclusion 326
References 326

Section 4: Machine Learning's Next Frontier 335

13 Transfer Learning 337
Riyanshi Gupta, Kartik Krishna Bhardwaj and Deepak Kumar Sharma

13.1 Introduction 338
13.1.1 Motivation, Definition, and Representation 338
13.2 Traditional Learning vs. Transfer Learning 338
13.3 Key Takeaways: Functionality 340
13.4 Transfer Learning Methodologies 341
13.5 Inductive Transfer Learning 342
13.6 Unsupervised Transfer Learning 344
13.7 Transductive Transfer Learning 346
13.8 Categories in Transfer Learning 347
13.9 Instance Transfer 348
13.10 Feature Representation Transfer 349
13.11 Parameter Transfer 349
13.12 Relational Knowledge Transfer 350
13.13 Relationship With Deep Learning 351
13.13.1 Transfer Learning in Deep Learning 351
13.13.2 Types of Deep Transfer Learning 352
13.13.3 Adaptation of Domain 352
13.13.4 Domain Confusion 353
13.13.5 Multitask Learning 354
13.13.6 One-Shot Learning 354
13.13.7 Zero-Shot Learning 355
13.14 Applications: Allied Classical Problems 355
13.14.1 Transfer Learning for Natural Language Processing 356
13.14.2 Transfer Learning for Computer Vision 356
13.14.3 Transfer Learning for Audio and Speech 357
13.15 Further Advancements and Conclusion 357
References 358

Section 5: Hands-On and Case Study 361

14 Hands on MAHOUT—Machine Learning Tool
Uma N. Dulhare and Sheikh Gouse

14.1 Introduction to Mahout 363
14.1.1 Features 366
14.1.2 Advantages 366
14.1.3 Disadvantages 366

14.1.4 Application 366
14.2 Installation Steps of Apache Mahout Using Cloudera 367
14.2.1 Installation of VMware Workstation 367
14.2.2 Installation of Cloudera 368
14.2.3 Installation of Mahout 383
14.2.4 Installation of Maven 384
14.2.5 Testing Mahout 386
14.3 Installation Steps of Apache Mahout Using Windows 10 386
14.3.1 Installation of Java 386
14.3.2 Installation of Hadoop 387
14.3.3 Installation of Mahout 387
14.3.4 Installation of Maven 387
14.3.5 Path Setting 388
14.3.6 Hadoop Configuration 391
14.4 Installation Steps of Apache Mahout Using Eclipse 395
14.4.1 Eclipse Installation 395
14.4.2 Installation of Maven Through Eclipse 396
14.4.3 Maven Setup for Mahout Configuration 399
14.4.4 Building the Path- 402
14.4.5 Modifying the pom.xml File 405
14.4.6 Creating the Data File 407
14.4.7 Adding External Jar Files 408
14.4.8 Creating the New Package and Classes 410
14.4.9 Result 411
14.5 Mahout Algorithms 412
14.5.1 Classification 412
14.5.2 Clustering 413
14.5.3 Recommendation 415
14.6 Conclusion 418
References 418

15 Hands-On H2O Machine Learning Tool 423
Uma N. Dulhare, Azmath Mubeen and Khaleel Ahmed
15.1 Introduction 424
15.2 Installation 425
15.2.1 The Process of Installation 425
15.3 Interfaces 431
15.4 Programming Fundamentals 432
15.4.1 Data Manipulation 432
15.4.1.1 Data Types 432
15.4.1.2 Data Import 435

15.4.2 Models 436
15.4.2.1 Model Training 436
15.4.3 Discovering Aspects 437
15.4.3.1 Converting Data Frames 437
15.4.4 H2O Cluster Actions 438
15.4.4.1 H2O Key Value Retrieval 438
15.4.4.2 H2O Cluster Connection 438
15.4.5 Commands 439
15.4.5.1 Cluster Information 439
15.4.5.2 General Data Operations 441
15.4.5.3 String Manipulation Commands 442
15.5 Machine Learning in H2O 442
15.5.1 Supervised Learning 442
15.5.2 Unsupervised Learning 443
15.6 Applications of H2O 443
15.6.1 Deep Learning 443
15.6.2 K-Fold Cross-Authentication or Validation 448
15.6.3 Stacked Ensemble and Random Forest Estimator 450
15.7 Conclusion 452
References 453

16 Case Study: Intrusion Detection System Using Machine Learning 455
Syeda Hajra Mahin, Fahmina Taranum and Reshma Nikhat
16.1 Introduction 456
16.1.1 Components Used to Design the Scenario Include 456
16.1.1.1 Black Hole 456
16.1.1.2 Intrusion Detection System 457
16.1.1.3 Components Used From MATLAB Simulator 458
16.2 System Design 465
16.2.1 Three Sub-Network Architecture 465
16.2.2 Using Classifiers of MATLAB 465
16.3 Existing Proposals 467
16.4 Approaches Used in Designing the Scenario 469
16.4.1 Algorithm Used in QualNet 469
16.4.2 Algorithm Applied in MATLAB 471
16.5 Result Analysis 471
16.5.1 Results From QualNet 471
16.5.1.1 Deployment 471
16.5.1.2 Detection 472

16.5.1.3 Avoidance 473
16.5.1.4 Validation of Conclusion 473
16.5.2 Applying Results to MATLAB 473
16.5.2.1 K-Nearest Neighbor 475
16.5.2.2 SVM 477
16.5.2.3 Decision Tree 477
16.5.2.4 Naïve Bayes 479
16.5.2.5 Neural Network 479
16.6 Conclusion 484
References 484

17 Inclusion of Security Features for Implications of Electronic Governance Activities 487
Prabal Pratap and Nripendra Dwivedi
17.1 Introduction 487
17.2 Objective of E-Governance 491
17.3 Role of Identity in E-Governance 493
17.3.1 Identity 493
17.3.2 Identity Management and its Buoyancy Against Identity Theft in E-Governance 494
17.4 Status of E-Governance in Other Countries 496
17.4.1 E-Governance Services in Other Countries Like Australia and South Africa 496
17.4.2 Adaptation of Processes and Methodology for Developing Countries 496
17.4.3 Different Programs Related to E-Governance 499
17.5 Pros and Cons of E-Governance 501
17.6 Challenges of E-Governance in Machine Learning 502
17.7 Conclusion 503
References 503

Index 505