Advanced Machine Learning

Fundamentals and algorithms

Dr. Amit Kumar Tyagi

Dr. Khushboo Tripathi

Dr. Avinash Kumar Sharma

www.bpbonline.com

First Edition 2024

ISBN: 978-93-55516-343

LIMITS OF LIABILITY AND DISCLAIMER OF WARRANTY

The information contained in this book is true to correct and the best of author's and publisher's knowledge. The author has made every effort to ensure the accuracy of these publications, but publisher cannot be held responsible for any loss or damage arising from any information in this book.

All trademarks referred to in the book are acknowledged as properties of their respective owners but BPB Publications cannot guarantee the accuracy of this information.

Dedicated to

My Mom ***Anita Tyagi****, Dad* ***Devraj Singh Tyagi***

My beloved daughter – ***Vernika Tyagi***

And My beloved Son- ***Apoorv Tyagi***

Book Technical Reviewers

Dr. Aswani Kumar Cherukuri
(Professor, Vellore Institute of Technology, Tamil Nadu, India)

My Mentors and Guides

Dr. G Aghila *(Director, National Institute of Technology, Tiruchirappalli)*

Dr. N Sreenath *(Professor, Puducherry Technological University, Puducherry)*

My Product and Competency Team

My Investors and Partners in Technology

Publishing Team at BPB

– Dr. Amit Kumar Tyagi

About the Authors

- **Dr. Amit Kumar Tyagi** is working as an Assistant Professor, at the National Institute of Fashion Technology, 110016, New Delhi, India. Previously, he has worked as Assistant Professor (Senior Grade 2), and Senior Researcher at Vellore Institute of Technology (VIT), Chennai Campus, 600127, Chennai, Tamilandu, India for the period of 2019-2022. He received his Ph.D. Degree (Full-Time) in 2018 from Pondicherry Central University, 605014, Puducherry, India. About his academic experience, he joined the Lord Krishna College of Engineering, Ghaziabad (LKCE) for the periods of 2009-2010, and 2012-2013. He was an Assistant Professor and Head of Research at Lingaya's Vidyapeeth (formerly known as Lingaya's University), Faridabad, Haryana, India, for the period of 2018-2019. His supervision experience includes more than 10 Master dissertations and one PhD thesis. He has contributed to several projects such as AARIN and P3- Block to address some of the open issues related to privacy breaches in Vehicular Applications (such as Parking) and Medical Cyber-Physical Systems (MCPS). He has published over 200 papers in refereed high-impact journals, conferences, and books, and some of his articles have been awarded as best paper awards. Also, he has filed more than 25 patents (Nationally and Internationally) in the areas of Deep Learning, the Internet of Things, Cyber-Physical Systems, and Computer Vision. He has edited more than 30 books for IET, Elsevier, Springer, CRC Press, etc. Also, he has authored 4 Books on Intelligent Transportation Systems, Vehicular Ad-hoc Networks, Machine Learning, and the Internet of Things, with IET UK, Springer Germany, and BPB India publisher. He is a Winner of the Faculty Research Award for the Year of 2020, 2021, and 2022 (three consecutive years) given by Vellore Institute of Technology, Chennai, India. Recently, he has been awarded the best paper award for a paper titled A Novel Feature Extractor Based on the Modified Approach of Histogram of Oriented Gradient, in ICCSA 2020, Italy (Europe). His current research focuses on Next Generation Machine Based Communications, Blockchain Technology, Smart and Secure Computing, and Privacy. He is a regular member of the ACM, IEEE, MIRLabs, Ramanujan Mathematical Society, Cryptology Research Society, Universal Scientific Education and Research Network, CSI, and ISTE.

- **Dr. Khushboo Tripathi** received her Ph. D. degree in computer science from the University of Allahabad, Prayagraj. She has completed her M. Tech in Computer Science and Engineering from KNIT Sultanpur, M. Sc, and B.Sc. from the University

of Allahabad, Prayagraj. She has more than seventeen years of experience in teaching and research. Her area of interest is Wireless Ad Hoc Networks, particularly, MANET and SENSOR networks, Secure Routing Protocols, SDN, Advanced Networking, Network and Cyber Security, and Artificial Intelligence. She has supervised many Ph.D., M. Tech., MCA, and B. Tech students in their thesis and projects. She has published various papers, and book chapters in International and National reputed journals and conferences in India and Abroad. She is the editor of the book Concepts of Artificial Intelligence and its Application in Modern Healthcare Systems. She is an editor, senior member, and reviewer of many professional organizations. Currently, she is serving as an Associate Professor in the Department of Computer Science and Engineering at Amity University Haryana Gurgaon.

- **Dr. Avinash Kumar Sharma** is currently working as an Associate Professor at the Department of Computer Science & Engineering, Sharda School of Engineering and Technology (SSET), Sharda University, Greater Noida. Dr. Avinash Kumar Sharma has completed his Ph.D at Uttarakhand Technical University, Dehradun (A State Govt. University) in Cloud Computing. His research areas are Cloud Computing, Machine Learning, Smart Agriculture, and Artificial Intelligence. He has more than 17 years of teaching experience. He has published about 30 research articles in national/international conferences, journals, and book chapters. Dr. Avinash Kumar Sharma edited 04 books with IGI Global. He has also published 03 patents, including 01 design patent.

About the Reviewers

- **Ankit Virmani** is an ethical AI and data leader with over a decade of progressive work experience in designing scalable data and Machine Learning platforms. He writes for Forbes being o the Forbes Technology Council and regularly speaks at national and international conferences.

  Ankit has worked for companies like Google, Amazon, CVS, DEloitte and Infosys and is passionate about how AI and data can change the society for the better.

- **Komal Lamba** is a gold-medalist data scientist with a strong background in engineering, statistics, and Machine Learning. Her expertise includes audio analysis, NLP, deep learning, graph neural networks, and user recommendations, with a focus on using GDS-neo4j for feed recommendations and ranking. Komal excels in connecting ideas across disciplines and delivering high-performing AI solutions to solve business problems. She is skilled at communicating complex concepts to non-technical stakeholders.

  Komal is proficient in Python (scikit-learn, NumPy, SciPy, Pandas, and Librosa), TensorFlow, and Excel, and knows SQL. During her postgraduate studies, Komal took on leadership roles, mentoring interns, coordinating large-scale research events, and leading journal clubs.

  Currently, she is working as a Data Scientist at Cygnet One in Ahmedabad, Gujarat, where she works on projects involving Traditional AI and GenAI. She continues to drive innovative, data-driven solutions at Cygnet Infotech as a Data Scientist.

Acknowledgement

There are a few people we want to thank for the continued and ongoing support they have given to us during the writing of this book. First of all, we would like to extend our gratitude to our family members, friends, and supervisors, who stood with us as an advisor in completing this book. Also, we would like to thank our almighty God, who made us write this book. We would also like to thank BPB Publishers (who have provided their continuous support during post COVID-19 Pandemic) and our colleagues with whom we have worked together inside the college/university and others outside premises who have provided their continuous support towards completing this book on a hot topic like Machine Learning.

Also, we would like to thank our Respected Madam, Prof. G Aghila, and our Respected Sir Prof. N Sreenath for giving their valuable inputs and helping us in completing this book.

Finally, we forward our thanks to BPB Publications for giving us this opportunity to write our second book with them.

Preface

In the past years, we have seen a drastic change in the use of smart devices in several sectors. These smart devices produce huge data, called big data, which is used to learn the footprint/ habits of customers/ consumers. Today's companies use this data to launch a new product, understand market needs and scenarios, and take action accordingly. Today, Machine Learning has become a hot topic in data analysis. For this, we have written this book. This book begins with an overview of the fundamental concepts and principles of Machine Learning (which also discusses various types of learning algorithms, such as supervised, unsupervised, and reinforcement learning, and explains the underlying mathematical foundations). Then, we discuss several supervised Machine Learning algorithms, like linear regression, logistic regression, etc., in detail. Further, this book discusses feature engineering, which plays an essential role in enhancing the performance of Machine Learning models. The book discusses different methods for feature extraction, transformation, and selection, enabling readers to effectively pre-process their data and derive informative features. This book also covers various evaluation metrics, including accuracy, precision, recall, and F1-score, to evaluate the learning models.

It also presents techniques like cross-validation and grid search for model selection and hyperparameter tuning. This book also discusses popular ensemble techniques, such as bagging, boosting, and stacking, and elucidates their strengths and limitations. It also discusses ensemble pruning and model combination strategies. The book examines several unsupervised Machine Learning algorithms like clustering algorithms, such as k-means, hierarchical clustering, and density-based methods. Further, it also explores dimensionality reduction techniques like **Principal Component Analysis** (**PCA**) and t-SNE. In the last, applications of learning algorithms have been explained in detail. In the end, this book introduces deep neural networks, including Convolutional Neural Networks (CNNs) for image analysis and **Recurrent Neural Networks** (**RNNs**) for sequence data. It covers advanced topics like transfer learning, **Generative Adversarial Networks** (**GANs**), and deep reinforcement learning. In summary, we provide our readers with the knowledge and skills that are necessary to tackle complex Machine Learning problems. We hope that this book will be helpful to those who are eager to learn more.

Over the 17 chapters, you will learn the following topics:

Chapter 1: Introduction to Machine Learning - This chapter introduces Machine Learning and its types with an explanation of its importance in different sectors/ applications.

Chapter 2: Statistical Analysis - This chapter introduces essential parameters, or statistical analysis parameters, that are used to evaluate the learning models.

Chapter 3: Liner Regression - This chapter introduces linear regression and its explanation in different programming languages with a real-world example.

Chapter 4: Logistic Regression - This chapter introduces logistic regression and its explanation in different programming languages with a real-world example.

Chapter 5: Decision Trees - This chapter introduces the decision tree and its explanation in different programming languages with a real-world example.

Chpter 6: Random Forest - This chapter introduces the random forest and its explanation in different programming languages with a real-world example.

Chapter 7: Rule-Based Classifiers - This chapter introduces rule-based classifiers and their explanation in different programming languages with a real-world example.

Chapter 8: Naïve Bayesian Classifiers - This chapter introduces the naïve Bayes classifier and its explanation in different programming languages with a real-world example.

Chapter 9: K-Nearest neighbors Classifiers - This chapter introduces k-NN and its explanation in different programming languages with a real-world example.

Chapter 10: Support Vector Machine - This chapter introduces SVM and its explanation in different programming languages with a real-world example.

Chapter 11: K-Means Clustering - This chapter introduces k-means clustering (with its importance in different sectors) and its explanation in different programming languages (with a real-world example).

Chapter 12: Dimensionality Reduction - This chapter introduces DR, Principal Component Analysis (PCA) (with its use as a statistical tool), and its explanation in different programming languages with a real-world example.

Chapter 13: Association Rules Mining and FP Growth - This chapter introduces ARM and FP growth (with its importance in different/ useful applications) and its explanation in different programming languages with a real-world example.

Chapter 14: Reinforcement Learning - This chapter introduces **Reinforcement Learning** (with its importance in different sectors) and its explanation in different programming languages with a real-world example.

Chapter 15: Applications of ML Algorithms - This chapter introduces different applications of Machine Learning in detail.

Chapter 16: Applications of Deep Learning - This chapter introduces different applications of deep learning in detail.

Chapter 17: Advanced Topics and Future Directions - This chapter introduces several topics (advanced topics for the future) towards making the next-generation society more secure and safer.

Code Bundle and Coloured Images

Please follow the link to download the
Code Bundle and the ***Coloured Images*** of the book:

https://rebrand.ly/zb5hljw

The code bundle for the book is also hosted on GitHub at
https://github.com/bpbpublications/Advanced-Machine-Learning.
In case there's an update to the code, it will be updated on the existing GitHub repository.

We have code bundles from our rich catalogue of books and videos available at **https://github.com/bpbpublications**. Check them out!

Errata

We take immense pride in our work at BPB Publications and follow best practices to ensure the accuracy of our content to provide with an indulging reading experience to our subscribers. Our readers are our mirrors, and we use their inputs to reflect and improve upon human errors, if any, that may have occurred during the publishing processes involved. To let us maintain the quality and help us reach out to any readers who might be having difficulties due to any unforeseen errors, please write to us at :

errata@bpbonline.com

Your support, suggestions and feedbacks are highly appreciated by the BPB Publications' Family.

Piracy

If you come across any illegal copies of our works in any form on the internet, we would be grateful if you would provide us with the location address or website name. Please contact us at **business@bpbonline.com** with a link to the material.

If you are interested in becoming an author

If there is a topic that you have expertise in, and you are interested in either writing or contributing to a book, please visit **www.bpbonline.com**. We have worked with thousands of developers and tech professionals, just like you, to help them share their insights with the global tech community. You can make a general application, apply for a specific hot topic that we are recruiting an author for, or submit your own idea.

Reviews

Please leave a review. Once you have read and used this book, why not leave a review on the site that you purchased it from? Potential readers can then see and use your unbiased opinion to make purchase decisions. We at BPB can understand what you think about our products, and our authors can see your feedback on their book. Thank you!

For more information about BPB, please visit **www.bpbonline.com**.

Join our book's Discord space

Join the book's Discord Workspace for Latest updates, Offers, Tech happenings around the world, New Release and Sessions with the Authors:

https://discord.bpbonline.com

Table of Contents

1. Introduction to Machine Learning 1

Introduction 1

Structure 2

Objective 2

Overview of AI 3

Artificial narrow intelligence *4*

Artificial general intelligence *4*

Artificial superintelligence *4*

Machine Learning 8

Preliminaries required to understand *9*

Machine Learning: Where several disciplines meet together *10*

Supervised learning 12

Applications *14*

Unsupervised learning 17

Applications *18*

Semi-supervised learning 20

Applications *21*

Reinforcement learning *24*

Difference between semi-supervised learning and reinforcement learning *26*

Validation and evaluation of Machine Learning techniques *27*

Applications of Machine Learning algorithms in today's smart era 29

Automatic recognition of handwritten postal codes *30*

Computer-aided diagnosis *31*

Computer vision *32*

Role of Machine Learning in driverless cars *34*

Role of Machine Learning in face recognition and security *35*

Role of Machine Learning in speech recognition *36*

Role of Machine Learning in text mining *37*

Where text and image data can be used together 38
Role of Machine Learning in energy saving 38
The present and the future of Machine Learning 40
Machine Learning future as thinking machines 41
Smart machines 42
Automated and intelligent machine 44
Automated intelligent and interactive machine 45
Deep blue 46
IBM's Watson 47
Google now 48
Apple's Siri 48
Microsoft's cortana 49
ChatGPT 50
Computer vision vs. Artificial Intelligence vs. expert system vs. robotics 51
Conclusion 52
Points to remember 53
Questions 53
Key terms 54
References 54

2. Statistical Analysis 55

Introduction 55
Structure 55
Objectives 56
Introduction to statistical analysis 57
Gaussian mixture model 58
Learning the concept by example 59
Hidden markov model 60
Example 61
Linear discriminant analysis 63
Example 64
MATLAB code 65

Conditional probability 66
Mean, mode and median 68
Correlation factor 69
Information gain and information Loss 70
Bias, variance and co-variance 72
Metrices 73
FP measure *73*
Confusion matrix *73*
The area under the ROC curve *74*
Recall *74*
Learning the concept by example 74
Gaussian distribution *74*
Example *75*
Skew distribution *76*
Standard deviation *78*
Role of standard deviation, conditional probability, gaussian distribution, skew distribution in ML *80*
Type of data 81
z-test, t-tests, ANOVA and chi-square tests 82
Z-Test *82*
T-tests *83*
ANOVA *83*
Chi-square test *84*
T-Test versus Chi-Square *85*
Conclusion 85
Questions 86
Key terms 87
References 88

3. Linear Regression 89

Introduction 89
Structure 89

Objectives 90
Introduction to linear regression 90
Linear regression 91
Entropy 93
Understanding the concept of several bits *95*
Mean squared error 96
Learning the concept by example *97*
Benefits and limitations of linear regression 105
Open issues and challenges with linear regression 107
Future possibilities 108
Conclusion 110
Questions 110
Key terms 111
References 112

4. Logistic Regression 113

Introduction 113
Structure 113
Objectives 113
Introduction to logistic regression 114
Metrices *116*
Accuracy *116*
F1 score *117*
ROC-AUC *117*
PR-AUC *118*
Learning the concept by example *118*
In language MATLAB *119*
In language R *120*
Python 124
Benefits and limitations with logistic regression *125*
Issues and challenges with logistic regression *127*
Future possibilities and opportunities 128

Conclusion 129
Key terms 130
Questions 130
References 131

5. Decision Trees 133

Introduction 133
Structure 133
Objectives 134
Introduction to decision trees 134
Applications of decision trees 137
Entropy in decision tree 138
Attribute selection measure 139
Information Gain of ID3 140
The problem with information gain 141
Implementation in MATLAB 142
The gain ratio of C4.5 and C5.0 142
Implementation of the concept by example 142
Benefits and limitations of decision tree 145
Issues and challenges with decision tree 147
Future possibilities and opportunities 148
Conclusion 149
Key terms 150
Questions 150
References 151

6. Random Forest 153

Introduction 153
Structure 153
Objectives 154
Introduction to Random Forest 154
Applications of Random Forest 158
Decision Tree versus Random Forest 159

Underfitting versus overfitting 162
Bagging versus boosting 163
Bagging example *164*
Boosting example *165*
Implementation and learning the concept by example 166
In MATLAB *167*
In R language *168*
In Python language *169*
Benefits and limitations of Random Forest 170
Issues and challenges with Random Forest 171
Future possibilities and opportunities for Random Forest 173
Conclusion 174
Questions 175
Key term 176
References 176

7. Rule-Based Classifiers 177

Introduction 177
Structure 177
Objectives 178
Introduction to rule-based classifiers 178
Sequential covering algorithm 180
Algorithm *181*
Visualization 182
Ripper 182
Algorithm *182*
Understanding the rule growing process *183*
Information gain *184*
Pruning *185*
Optimization *186*
Learning the concept by example 187
MATLAB simulator *187*

R language *188*
Python programming languages *190*
Benefits and limitations of Rule based classifier 191
Issues and challenges with Rule based classifier 193
Future possibilities and opportunities 195
Conclusion 196
Key terms 197
References 197
Questions 198

8. Naïve Bayesian Classifier 201

Introduction 201
Structure 201
Objectives 202
Introduction to Naïve Bayes Classifier 202
Example *203*
Prior probability 205
Posterior probability 205
Laplace estimator 207
Likelihood 208
Learning the concept by example *210*
Benefits and limitations with Naïve Bayesian Classifier 213
Issues and challenges with Naïve Bayesian Classifier *215*
Future possibilities and opportunities 217
Conclusion 218
Questions 218
Key terms 220
References 220

9. K-Nearest neighbors Classifiers 221

Introduction 221
Structure 221
Objectives 222

Introduction to k-nearest neighbors 222
Example of k-nearest neighbors *223*
Implementation of k-nearest neighbors in different languages 225
In MATLAB Simulator *225*
In language R *227*
In language Python *228*
Benefits and limitations with k-nearest neighbors classifiers 229
Issues and challenges with k-NN classifiers 230
Future possibilities/ opportunities towards k-NN classifiers 231
Conclusion 233
Questions 234
Key terms 235
References 235

10. Support Vector Machine 237

Introduction 237
Structure 237
Objectives 237
Overview 238
Defining support vector machine 240
Design of the SVM *241*
The case of nonlinear kernel *242*
Learning the concept by example 243
SVM implementation in MATLAB *243*
In 'R' Language *244*
In Python language 249
Benefits and limitations of support vector machine 251
Issues and challenges *253*
Future possibilities and opportunities *254*
Conclusion 256
Keywords 256
References 256
Questions 257

11. K-Means Clustering 259

Introduction 259

Structure 259

Objectives 260

Overview of clustering 260

Types of clusters 261

Key-process 263

The k-means clustering algorithm 264

Learning the concept by example 265

K-means clustering in MATLAB 265

In the language 'R' 266

Benefits and limitations of k-means algorithm 270

Limitations with k-means algorithm 272

Future possibilities toward the k-means algorithm 273

Conclusion 275

Keywords 275

References 275

Questions 276

12. Dimensionality Reduction 279

Introduction 279

Structure 279

Objectives 280

Basic keys 281

Introduction to principal component analysis 282

Explanation of the problem 283

The idea behind the principal component analysis 285

The SVD and dimensionality reduction 286

The PCA implementation 287

Number of principal components for PCA 288

Data reconstruction error in PCA 289

The PCA implementation in MATLAB 290

PCA methods implementation in Weka 292
Polymorphic worm detection using PCA 293
Introduction to SEA, MKMP, and PCA 294
Motivation towards string matching 294
The KMP Algorithm 296
A modified Principal Component Analysis 298
Clustering method for polymorphic worms 299
Signature generation algorithms Pseudo-Codes 301
Signature generation process 301
Learning the PCA concept by example 303
MATLAB 303
R language 304
Python language 305
Benefits and limitations with dimensionality reduction and PCA 306
Challenges with dimensionality reduction and PCA 308
Future possibilities towards dimensionality reduction and PCA 310
Conclusion 311
Keywords 312
References 312
Questions 313

13. Association Rules Mining and FP Growth 315

Introduction 315
Structure 315
Objective 316
An outline 316
Learning the concept by example 322
In MATLAB 322
In R language 325
In Python language 326
Benefits and limitations with ARM and FP growth algorithm 327
Issues with ARM and FP growth algorithm 329

Future possibilities with ARM and FP growth algorithm 330
Conclusion 332
Keywords 332
References 333
Questions 333

14. Reinforcement Learning 335

Introduction 335
Structure 335
Objectives 336
Key features 336
Deep learning 337
Deep learning vs Machine Learning 340
Algorithms 341
Neural network 341
Perceptron neural network 342
MATLAB implementation of the perceptron training and testing algorithms 343
Multilayer perceptron networks 344
The backpropagation algorithm 346
ANN 347
CNN 350
RNN 352
LSTM 354
Other popular reinforcement algorithms 356
Q-learning 356
SARSA 358
DQN 361
DDPG 363
Simulator used 365
Tensor flow 365
PyTorch 367
Learning the neural network concept by example 368

Benefits and limitations of RL algorithm 372
Challenges with a RL algorithm 374
Future opportunities for a RL algorithm 376
Conclusion 377
Keywords 378
References 378
Questions 379

15. Applications of ML Algorithms 381

Introduction 381
Structure 381
Objective 382
Key features of ML 383
Agriculture 384
Bio-medical 385
Healthcare 387
Education 388
Retail 390
Transportation 391
Supply chain management and logistics 392
Churn prediction 393
Manufacturing 394
Media and entertainment 396
Weather prediction 397
Financial sector 398
Recommender systems 400
Textile industry 401
Energy sector 402
Benefits, open issues and critical challenges 404
Future research opportunities towards ML techniques 406
Conclusion 407
Keywords 407
References 408

16. Applications of Deep Learning 409

Introduction 409

Structure 409

Objectives 410

Basic fundamentals 410

Computer vision 412

Natural language processing 415

Sentimental analysis 418

Information gained 419

Customer buying behavior 421

Lesson learned 422

Recommendation systems 424

Information acquired 425

Medical diagnostics 427

Understanding 428

Autonomous vehicles 430

Fraud detection 433

Financial forecasting 434

Drug discovery 435

Benefits, open issues and critical challenges 437

Future research opportunities for DL techniques 438

Conclusion 440

Keywords 440

References 440

17. Advance Topics and Future Directions 443

Introduction 443

Structure 443

Objectives 444

Next-generation machine-based communications 444

Blockchain technology 446

Example 1 447

Digital twin technology 449
Example 2 ..... *450*
Internet of Things 453
Example 3 ..... *454*
Cloud, fog and edge computing 456
Example 4 ..... *457*
Dew computing 459
Example 5 ..... *460*
Parallel computing 461
Example 6 ..... *463*
Soft computing 464
Example 7 ..... *465*
Grid computing 466
Example 8 ..... *467*
Quantum computing 470
Example 9 ..... *471*
6G 472
Example 10 ..... *474*
6G vs. 5G 475
Software-defined networking 476
Example 11 ..... *477*
Popular sectors for the future 480
Popular applications for the future 481
Conclusion 482
Keywords 483
References 483

Index 485-494

Chapter 1
Introduction to Machine Learning

Introduction

This book will discuss **Machine Learning** (**ML**) and its role in the smart era/ many applications with a detailed explanation (with the implementation of code). In general, ML is a subset of **Artificial Intelligence** (**AI**) that focuses on developing algorithms and models that allow computers to learn from and make predictions or decisions based on data without being explicitly programmed. In this smart era, Machine Learning has become a fundamental technology with a wide range of applications that are transforming industries and our daily lives. Using ML techniques, we learn from Data. ML algorithms are designed to analyze and find patterns within large datasets, allowing computers to make predictions or take actions based on that data. Now there are a few types of Machine Learning, added as:

Supervised learning, unsupervised learning, reinforcement learning, semi-supervised learning and self-supervised learning.

We will also learn key algorithms and models, for example, linear regression and logistic regression, decision trees and random forests, neural networks and deep learning.

We will understand the use of Machine Learning in the smart era, personalized recommendations, healthcare, autonomous vehicles, natural language processing (NLP), finance, manufacturing and industry 4.0, energy efficiency, smart cities, environmental monitoring, security, and so on.

Machine Learning has the potential to transform industries and improve the efficiency and convenience of our daily lives in various ways. As data continues to grow and as algorithms become more important, Machine Learning's impact in the smart era is expected to expand even further. This book will cover all such types of ML algorithms and will explain their use in different sectors in the concluding chapters.

Structure

In this chapter, we will discuss the following topics:

- Overview of AI
- Machine Learning
- Supervised learning
- Unsupervised learning
- Semi-supervised learning
- Applications of Machine Learning algorithms in today's smart era
- Where text and image data can be used together
- The present and the future of Machine Learning
- Computer vision vs Artificial Intelligence vs expert system vs robotics

Objective

The objective of this chapter is to provide readers and future researchers with a fundamental understanding of ML. This introductory chapter serves as a stepping stone for those who may be new to the subject or in need of a refresher. It covers various objectives, including defining Machine Learning, explaining its types and applications, offering a historical context, and providing motivation for its importance in multiple domains such as healthcare, finance, and autonomous vehicles.

Basic concepts of Machine Learning, including data, features, labels, algorithms, models, and predictions, are explained. The distinction between supervised and unsupervised learning is clarified. The pivotal role of data in Machine Learning is emphasized, stressing the importance of high-quality, diverse, and well-labeled data for training effective models.

The chapter outlines the typical workflow of a Machine Learning project, encompassing data collection, data preprocessing, model training, evaluation, and deployment. It also explains about the challenges and limitations associated with Machine Learning, covering issues like overfitting, bias, data quality, and ethical issues.

Additionally, essential terminologies related to Machine Learning, such as computer vision and robotics, are introduced. The chapter aims to prevent readers from conflating

these terms and encourages them to distinguish between various concepts they might encounter throughout the book.

In last, a few self-assessment questions are included at the end of the chapter. These questions help readers gauge their grasp of the introductory concepts presented.

Overview of AI

Today AI's rapid growth and powerful capabilities have made people paranoid about the inevitability and proximity of AI growth. Also, the transformation brought about by AI in different industries has made business leaders and the mainstream public think that we are close to achieving the peak of AI research and maxing out AI's potential. However, understanding the types of AI that are possible and the types that exist now will give a clearer picture of existing AI capabilities and the long road ahead for AI research. Since AI research purports to make machines emulate human-like functioning, the degree to which an AI system can replicate human capabilities is used as the criterion for determining the types of AI. Thus, depending on how a machine compares to humans in terms of versatility and performance, AI can be classified as one among the multiple types of AI. Under such a system, an AI that can perform more human-like functions with equivalent levels of proficiency will be considered a more evolved type of AI. In contrast, an AI that has limited functionality and performance would be considered a simple one.

Based on this criterion, there are two ways in which AI is generally classified. One type is based on classifying AI and AI-enabled machines based on their likeness to the human mind and their ability to *think* and perhaps even feel like humans. According to this system of classification, there are four types of AI or AI-based systems: reactive machines, limited memory machines, theory of mind, and self-aware AI, as illustrated in the following figure:

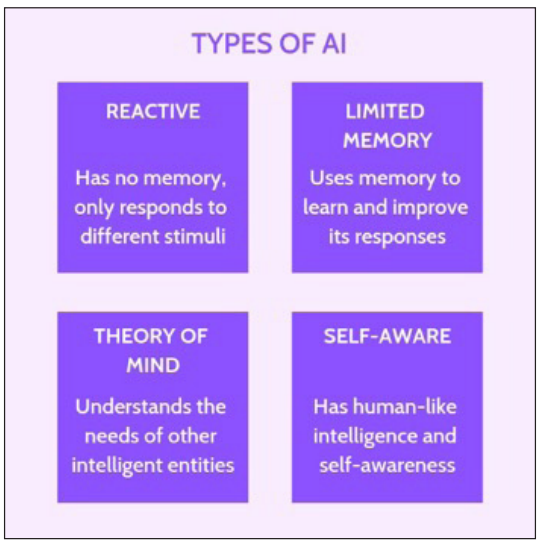

Figure 1.1: *Types of Artificial Intelligence*

Now, let us look at some examples of each:

- **Reactive machines**: Beat at chess by IBM's supercomputer, and Netflix recommendations.
- **Limited memory machines**: Self-driving cars.
- **Theory of mind**: In the future, theory of mind AI machines could be able to understand intentions and predict behavior, as if to simulate human relationships.
- **Self-awareness:** The final evolution of AI would be to design systems that have a sense of self, a conscious understanding of their existence. However, this type of AI does not exist yet.

The alternate system of classification that is more generally used in tech parlance is the classification of the technology into **Artificial Narrow Intelligence** (**ANI**), **Artificial General Intelligence** (**AGI**), and **Artificial Superintelligence** (**ASI**). Let us understand each one by one:

Artificial narrow intelligence

This type of Artificial Intelligence represents all the existing AI, including even the most complicated and capable AI that has ever been created to date. Artificial narrow intelligence refers to AI systems that can only perform a specific task autonomously using human-like capabilities. These machines can do nothing more than what they are programmed to do and thus have a very limited or narrow range of competencies. According to the system of classification, these systems correspond to all the reactive and limited memory AI. Even the most complex AI that uses Machine Learning and deep learning to teach itself falls under ANI.

Artificial general intelligence

Artificial general intelligence is the ability of an AI agent to learn, perceive, understand, and function completely like a human being. These systems will be able to independently build multiple competencies and form connections and generalizations across domains, massively cutting down on the time needed for training. This will make AI systems as capable as humans by replicating our multi-functional capabilities.

Artificial superintelligence

The development of artificial superintelligence will probably mark the maximum use of AI research, as AGI will become by far the most capable form of intelligence on earth. ASI, in addition to replicating the multi-faceted intelligence of human beings, will be exceedingly better at everything they do because of overwhelmingly greater memory, faster data processing and analysis, and decision-making capabilities. The development of AGI and

ASI will lead to a scenario/application most popularly referred to as the singularity. And while the potential of having such powerful machines at our disposal seems appealing, these machines may also threaten our existence or, at the very least, our way of life. Note that AI refers to the development of intelligent machines or systems that can perform tasks that typically require human intelligence. It involves the creation of algorithms and models that enable machines to perceive and understand their environment, reason, learn, and make decisions. AI encompasses various techniques, including ML and DL, as well as areas such as natural language processing, computer vision, expert systems, and robotics. Let us briefly understand:

- **Machine Learning**: ML is a subset of AI that focuses on the development of algorithms and models that allow computer systems to automatically learn from data and improve their performance without explicit programming. ML algorithms can identify patterns, extract information, and make predictions or decisions based on the input data. ML techniques include supervised learning, unsupervised learning, reinforcement learning, and semi-supervised learning.
- **Deep learning**: DL is a subset of ML that specifically focuses on training artificial neural networks with multiple layers (deep neural networks) to learn hierarchical representations of data. DL architectures, often referred to as deep neural networks or deep neural nets, are capable of automatically learning and extracting complex features from raw input data. DL has achieved essential breakthroughs in computer vision, natural language processing, speech recognition, and recommendation systems.

In summary, AI is the broader field that encompasses the development of intelligent systems. At the same time, ML is a subset of AI that focuses on algorithms and models that enable machines to learn from data. DL is a further specialization within ML that adds deep neural networks to learn complex representations. ML and DL are key components of AI, and their advancements have driven many recent breakthroughs in Artificial Intelligence applications. Take a look at the following figure: