

# Contents

## Chapter 1

### Introduction to Machine Learning — 1

- 1.1 What Is Machine Learning? — 1
  - 1.1.1 Definition of Machine Learning — 1
  - 1.1.2 Historical Background — 2
  - 1.1.3 Types of Machine Learning — 4
  - 1.1.4 Applications of Machine Learning — 7
  - 1.1.5 Challenges in Machine Learning — 8
- 1.2 Python in the Machine Learning Landscape — 11
  - 1.2.1 Why Python for Machine Learning? — 11
  - 1.2.2 Popular Python Libraries for ML — 12
  - 1.2.3 Python vs Other Programming Languages in ML — 13
  - 1.2.4 Community and Resources for Python ML — 14
- 1.3 Setting Up Your Python Environment — 15
  - 1.3.1 Installing Python — 15
  - 1.3.2 Virtual Environments and Why They're Important — 16
  - 1.3.3 Essential Python Libraries for ML — 17
  - 1.3.4 Using Package Managers: Pip and Conda — 19
  - 1.3.5 Setting Up Jupyter Notebook — 20
  - 1.3.6 Best Practices for Managing ML Projects in Python — 22
- Summary — 25
- Exercise (MCQs) — 26
- Answers — 27
- Fill in the Blanks — 27
- Answers — 28
- Descriptive Questions — 28

## Chapter 2

### Basics of Python Programming — 30

- 2.1 Why Python? — 30
  - 2.1.1 Drawbacks of Python — 30
  - 2.1.2 History of Python — 31
  - 2.1.3 Major Features of Python — 31
  - 2.1.4 Market Demand — 32
  - 2.1.5 Why Python in Mobile App Development? — 33
  - 2.1.6 Python Versions — 36
- 2.2 Python Syntax and Structure — 36
  - 2.2.1 Indentation and Whitespace — 36
  - 2.2.2 Comments and Documentation — 37
- 2.3 Data Types and Variables — 38

|       |                                         |           |
|-------|-----------------------------------------|-----------|
| 2.3.1 | Primitive Data Types —                  | <b>39</b> |
| 2.3.2 | Lists and Tuples —                      | <b>39</b> |
| 2.3.3 | Dictionaries and Sets —                 | <b>50</b> |
| 2.4   | Control Structures —                    | <b>60</b> |
| 2.4.1 | Conditional Statements —                | <b>60</b> |
| 2.4.2 | Loops —                                 | <b>62</b> |
| 2.5   | Functions and Modules —                 | <b>63</b> |
| 2.5.1 | Defining Functions —                    | <b>63</b> |
| 2.5.2 | Lambda Functions —                      | <b>65</b> |
| 2.5.3 | Modules and Packages —                  | <b>66</b> |
| 2.6   | Working with Files —                    | <b>68</b> |
| 2.6.1 | File I/O —                              | <b>68</b> |
| 2.7   | Working with Libraries —                | <b>72</b> |
| 2.7.1 | Importing Libraries —                   | <b>72</b> |
| 2.7.2 | Popular Python Libraries Overview —     | <b>73</b> |
| 2.8   | Object-Oriented Programming in Python — | <b>77</b> |
| 2.8.1 | Classes and Objects —                   | <b>77</b> |
| 2.8.2 | Inheritance and Polymorphism —          | <b>80</b> |
|       | Summary —                               | <b>85</b> |
|       | Exercise (MCQs) —                       | <b>85</b> |
|       | Answers —                               | <b>87</b> |
|       | Fill in the Blanks —                    | <b>87</b> |
|       | Answers —                               | <b>88</b> |
|       | Descriptive Questions —                 | <b>88</b> |

## **Chapter 3**

### **Data Preprocessing in Python — 90**

|       |                                                            |            |
|-------|------------------------------------------------------------|------------|
| 3.1   | Numerical and Scientific Computing Using NumPy and SciPy — | <b>90</b>  |
| 3.1.1 | Numerical Computations with NumPy —                        | <b>90</b>  |
| 3.1.2 | Scientific Computations with SciPy —                       | <b>105</b> |
| 3.2   | Loading Data with Pandas —                                 | <b>109</b> |
| 3.2.1 | DataFrame and Series —                                     | <b>109</b> |
| 3.2.2 | Data Manipulation with Pandas —                            | <b>114</b> |
| 3.3   | Data Cleaning and Transformation —                         | <b>120</b> |
| 3.3.1 | Handling Missing Data —                                    | <b>121</b> |
| 3.3.2 | Data Type Conversions —                                    | <b>122</b> |
| 3.4   | Feature Engineering —                                      | <b>123</b> |
| 3.4.1 | Encoding Categorical Variables —                           | <b>124</b> |
| 3.4.2 | Feature Scaling —                                          | <b>126</b> |
| 3.5   | Data Visualization with Matplotlib and Seaborn —           | <b>129</b> |
| 3.5.1 | Basic Plotting with Matplotlib —                           | <b>130</b> |
| 3.5.2 | Advanced Visualizations with Seaborn —                     | <b>140</b> |

|       |                             |     |
|-------|-----------------------------|-----|
| 3.6   | Handling Imbalanced Data —  | 143 |
| 3.6.1 | Resampling Techniques —     | 144 |
| 3.6.2 | Synthetic Data Generation — | 145 |
|       | Summary —                   | 147 |
|       | Exercise (MCQs) —           | 147 |
|       | Answers —                   | 149 |
|       | Fill in the Blanks —        | 149 |
|       | Answers —                   | 149 |
|       | Descriptive Questions —     | 150 |

## Chapter 4

### Foundations of Machine Learning — 152

|       |                                       |     |
|-------|---------------------------------------|-----|
| 4.1   | Supervised vs Unsupervised Learning — | 153 |
| 4.1.1 | Classification vs Regression —        | 154 |
| 4.1.2 | Clustering vs Association —           | 156 |
| 4.2   | Overfitting and Regularization —      | 159 |
| 4.2.1 | Bias-Variance Trade-Off —             | 159 |
| 4.2.2 | L1 and L2 Regularization —            | 163 |
| 4.3   | Evaluation Metrics —                  | 167 |
| 4.3.1 | Metrics for Classification —          | 168 |
| 4.3.2 | Metrics for Regression —              | 171 |
| 4.4   | Cross-Validation —                    | 175 |
| 4.4.1 | k-Fold Cross-Validation —             | 176 |
| 4.4.2 | Leave-One-Out and Stratified K-Fold — | 178 |
|       | Summary —                             | 182 |
|       | Exercise (MCQs) —                     | 183 |
|       | Answers —                             | 185 |
|       | Fill in the Blanks —                  | 186 |
|       | Answers —                             | 186 |
|       | Descriptive Questions —               | 187 |

## Chapter 5

### Classic Machine Learning Algorithms — 188

|       |                                         |     |
|-------|-----------------------------------------|-----|
| 5.1   | Linear Regression —                     | 189 |
| 5.1.1 | Simple Linear Regression —              | 190 |
| 5.1.2 | Multiple Linear Regression —            | 197 |
| 5.1.3 | Polynomial Regression —                 | 203 |
| 5.2   | Logistic Regression —                   | 206 |
| 5.2.1 | Binary Classification —                 | 210 |
| 5.2.2 | Multiclass Classification —             | 214 |
| 5.2.3 | Regularization in Logistic Regression — | 217 |
| 5.3   | Decision Trees and Random Forests —     | 222 |

|       |                                       |     |
|-------|---------------------------------------|-----|
| 5.3.1 | Building Decision Trees —             | 224 |
| 5.3.2 | Entropy and Information Gain —        | 230 |
| 5.3.3 | Random Forests and Bagging —          | 234 |
| 5.4   | Support Vector Machines —             | 239 |
| 5.4.1 | Linear SVM —                          | 242 |
| 5.4.2 | Kernel SVM —                          | 250 |
| 5.4.3 | Hyperparameter Tuning in SVM —        | 255 |
| 5.5   | K-Means Clustering —                  | 260 |
| 5.5.1 | Clustering Basics —                   | 261 |
| 5.5.2 | Selecting the Number of Clusters —    | 262 |
| 5.6   | Principal Component Analysis —        | 270 |
| 5.6.1 | Dimensionality Reduction —            | 273 |
| 5.6.2 | Eigenvectors and Eigenvalues —        | 275 |
| 5.7   | Naive Bayes —                         | 281 |
| 5.7.1 | Gaussian Naive Bayes —                | 285 |
| 5.7.2 | Multinomial Naive Bayes —             | 293 |
| 5.8   | Ensemble Methods: Boosting, Bagging — | 297 |
| 5.8.1 | Bagging and Boosting Overview —       | 298 |
| 5.8.2 | AdaBoost and Gradient Boosting —      | 303 |
|       | Summary —                             | 310 |
|       | Exercise (MCQs) —                     | 311 |
|       | Answers —                             | 313 |
|       | Fill in the Blanks —                  | 313 |
|       | Answers —                             | 314 |
|       | Descriptive Questions —               | 314 |

## Chapter 6

### Advanced Machine Learning Techniques — 316

|       |                                                |     |
|-------|------------------------------------------------|-----|
| 6.1   | Gradient Boosted Trees: XGBoost and LightGBM — | 316 |
| 6.1.1 | XGBoost Algorithm —                            | 319 |
| 6.1.2 | LightGBM Versus XGBoost —                      | 327 |
| 6.2   | Kernel Methods —                               | 331 |
| 6.2.1 | Kernel Tricks —                                | 332 |
| 6.2.2 | Radial Basis Function (RBF) Kernel —           | 333 |
| 6.2.3 | Clustering Techniques Beyond k-Means —         | 336 |
| 6.3   | Anomaly Detection —                            | 347 |
| 6.3.1 | Statistical Methods —                          | 347 |
| 6.3.2 | Distance-Based Methods —                       | 350 |
| 6.4   | Clustering —                                   | 351 |
| 6.4.1 | Isolation Forest —                             | 352 |
| 6.4.2 | Density-Based Methods —                        | 352 |
| 6.4.3 | Other Techniques —                             | 353 |

|                            |     |
|----------------------------|-----|
| Summary                    | 354 |
| Exercise (MCQs)            | 356 |
| Descriptive Type Questions | 358 |
| Fill in the Blanks         | 359 |
| Answers                    | 359 |
| True and False             | 359 |

## Chapter 7

### Neural Networks and Deep Learning — 360

|        |                                                  |     |
|--------|--------------------------------------------------|-----|
| 7.1    | Introduction to Neural Networks                  | 360 |
| 7.2    | Perceptron                                       | 362 |
| 7.2.1  | Structure of Perceptron                          | 362 |
| 7.2.2  | Function of Perceptron                           | 363 |
| 7.2.3  | Where to Use Perceptron                          | 363 |
| 7.2.4  | Where to Use Activation Function                 | 363 |
| 7.3    | TensorFlow                                       | 367 |
| 7.3.1  | Computational Graph                              | 367 |
| 7.3.2  | Eager Execution                                  | 368 |
| 7.3.3  | Keras                                            | 369 |
| 7.3.4  | Sessions                                         | 370 |
| 7.3.5  | Common Operations                                | 370 |
| 7.4    | Implementing Neural Network Using TensorFlow     | 370 |
| 7.5    | Building a Neural Network Using Keras Framework  | 373 |
| 7.5.1  | Difference between Keras and Tensorflow          | 373 |
| 7.6    | Convolutional Neural Network (CNN)               | 374 |
| 7.6.1  | Model: "sequential_2"                            | 386 |
| 7.6.2  | CNN Architecture                                 | 389 |
| 7.7    | Dropout Layers                                   | 390 |
| 7.8    | Recurrent Neural Networks (RNNs)                 | 391 |
| 7.8.1  | Types of RNNs                                    | 391 |
| 7.9    | Sequence-to-Sequence Models                      | 398 |
| 7.10   | Transfer Learning                                | 400 |
| 7.11   | Using Pretrained Models                          | 406 |
| 7.12   | Fine-Tuning and Feature Extraction               | 407 |
| 7.12.1 | Choosing the Right Approach to Fine-Tune         | 407 |
| 7.13   | Generative Adversarial Networks (GANs)           | 408 |
| 7.13.1 | Architecture of GANs                             | 408 |
| 7.13.2 | Best Practices of GANs                           | 409 |
| 7.13.3 | Application of GANs                              | 409 |
| 7.14   | Regularization and Optimization in Deep Learning | 409 |
| 7.15   | Batch Normalization                              | 413 |
| 7.15.1 | How BatchNorm Works?                             | 414 |

|        |                                   |     |
|--------|-----------------------------------|-----|
| 7.15.2 | Best Practices of BatchNorm —     | 414 |
| 7.16   | Gradient Clipping —               | 414 |
| 7.16.1 | Advantages of Gradient Clipping — | 415 |
|        | Summary —                         | 419 |
|        | Exercise (MCQs) —                 | 421 |
|        | Answer Key —                      | 422 |
|        | Some Basic Question —             | 423 |
|        | Fill in the Blanks —              | 423 |
|        | Answers —                         | 423 |

## **Chapter 8**

### **Specialized Applications and Case Studies — 424**

|       |                                                           |     |
|-------|-----------------------------------------------------------|-----|
| 8.1   | Introduction to Natural Language Processing (NLP) —       | 424 |
| 8.1.1 | Tokenization —                                            | 428 |
| 8.1.2 | Word Embeddings —                                         | 429 |
| 8.1.3 | Sequence Modeling for NLP —                               | 431 |
| 8.2   | Time-Series Forecasting —                                 | 433 |
| 8.2.1 | Autoregressive Integrated Moving Average (ARIMA) Models — | 434 |
| 8.2.2 | Prophet and Neural Networks for Time Series —             | 436 |
| 8.3   | Recommender Systems —                                     | 438 |
| 8.4   | Computer Vision Applications —                            | 443 |
| 8.4.1 | Object Detection and Segmentation —                       | 445 |
| 8.5   | Reinforcement Learning —                                  | 447 |
| 8.5.1 | Core Elements of Reinforcement Learning —                 | 447 |
| 8.5.2 | Learning Process —                                        | 448 |
| 8.6   | Application of Reinforcement Learning —                   | 455 |
| 8.7   | Challenges of Reinforcement Learning —                    | 455 |
| 8.8   | Q-learning —                                              | 455 |
| 8.9   | Deep Q Networks —                                         | 457 |
| 8.10  | Policy Gradient Methods —                                 | 459 |
|       | Summary —                                                 | 464 |
|       | Exercise (MCQs) —                                         | 465 |
|       | Answers Key —                                             | 467 |
|       | Some Basic Question —                                     | 468 |
|       | Fill in the Blanks —                                      | 469 |
|       | Answers —                                                 | 469 |

### **References — 471**

### **Index — 475**