

Contents

Acknowledgments — VII

Preface — XV

1 Introduction — 1

- 1.1 Artificial Intelligence — 1
- 1.2 What Is Machine Learning? — 3
- 1.3 Areas of Machine Learning — 4
- 1.4 Machine Learning Workflow — 5

Part I: Machine Learning I

2 Simple Linear Regression – Concept — 9

- 2.1 Bird's Eye View — 9
- 2.2 Fundamental Equation of SLR — 14
- 2.3 Some Assumptions of Simple Linear Regression — 15
- 2.4 Case Study: Is Earning Related to Learning? — 17
- 2.5 Interpreting Regression — 20
- 2.6 Summary — 22

3 Simple Linear Regression – Theory — 23

- 3.1 Some Basics of Measuring Distance — 23
- 3.2 Total Absolute Error (TAE) — 26
- 3.3 Residual Sum of Squares (RSS) — 27
- 3.4 Mean Squared Error — 28
- 3.5 Analytical vs. Numerical Solutions in Machine Learning — 29
- 3.6 Summary — 30

4 Simple Linear Regression – Practice — 31

5 K-Nearest Neighbors (KNN) – Concept — 34

- 5.1 Bird's Eye View — 34
- 5.2 Prediction Structure — 35
- 5.3 Case Study: To Loan or Not to Loan — 38
- 5.4 Summary — 42

6 K-Nearest Neighbors (KNN) – Theory — 43

- 6.1 Distance Metrics — 43
- 6.2 Choosing k in KNN — 50

- 6.3 ML Pipelines, Hyperparameters — 52
- 6.4 Summary — 52
- 7 K-Nearest Neighbors (KNN) – Practice — 53**

Part II: Model Assessment

- 8 Model Assessment – Bias-Variance Tradeoff — 59**
 - 8.1 Train-Test Split — 59
 - 8.2 K-Fold Cross-Validation — 60
 - 8.3 Underfit and Overfit — 60
 - 8.4 Bias-Variance Tradeoff — 64
 - 8.5 Summary — 65
- 9 Model Assessment – Regression — 66**
 - 9.1 Mean Squared Error — 66
 - 9.2 R^2 – Intuition — 67
 - 9.3 R^2 – Computation — 70
 - 9.3.1 Calculating R^2 — 72
 - 9.4 Interpreting R^2 — 73
 - 9.5 Summary — 74
- 10 Model Assessment – Classification — 75**
 - 10.1 Accuracy — 75
 - 10.2 Confusion Matrix — 76
 - 10.3 Precision and Recall — 77
 - 10.4 ROC/AUC Curve — 79
 - 10.5 Summary — 81

Part III: Machine Learning II

- 11 Multiple Linear Regression – Concept — 85**
 - 11.1 Bird's Eye View — 85
 - 11.2 Multiple Regression Workflow — 88
 - 11.3 Case Study: Does Sleep Improve Academic Performance? — 88
 - 11.4 Interpreting the Regression Equation — 95
 - 11.5 Summary — 96
- 12 Multiple Linear Regression – Theory — 97**
 - 12.1 Standardized Coefficients — 97

12.2	Multiple Linear Regression Diagnostics Checklist —	100
12.3	Summary —	106
13	Multiple Linear Regression – Practice —	108
14	Logistic Regression – Concept —	111
14.1	Bird's Eye View —	111
14.2	Probabilities and Decision Making —	112
14.3	Case Study: Credit Card Payments —	114
14.4	Summary —	118
15	Logistic Regression – Theory —	119
15.1	Logistic Regression Function —	119
15.2	Odds, Odds Ratio, and Logit —	122
15.3	Logistic Regression Equation in Logit Form —	123
15.4	Interpreting the Coefficients —	124
15.5	Derivation of the Logistic Regression Equation as Log Odds —	125
15.6	Summary —	126
16	Logistic Regression – Practice —	127
17	<i>K</i>-Means – Concept —	131
17.1	Bird's Eye View —	131
17.2	Case Study: Marketing Segmentation —	133
17.3	Summary —	135
18	<i>K</i>-Means – Theory —	136
18.1	Expectation – Maximization —	136
18.2	Choice of K —	139
18.3	Summary —	142
19	<i>K</i>-Means – Practice —	144

Part IV: Deep Learning

20	Deep Learning – Bird's Eye View —	149
20.1	Foundational Principles —	149
20.2	From Evidence to Judgment —	150
20.3	Summary —	152
21	Neurons —	154

21.1	Neurons as Functions —	154
21.1.1	Activation Function Computation —	158
21.2	Mathematics of Neural Computation —	158
21.3	Linear Algebra for Neural Computation —	159
21.4	Neural Computation in Matrix Form —	162
21.5	Summary —	163
22	Neurons – Practice —	164
23	Network Architecture —	167
23.1	Layers —	167
23.2	Parameters —	168
23.3	Hyperparameters —	171
23.4	Summary —	171
24	Network Architecture – Practice —	172
25	Forward Propagation —	174
25.1	Forward Propagation Flow —	174
25.2	Function Composition —	174
25.3	Forward Propagation Computation —	176
25.4	Computation Worked Example —	178
25.5	Summary —	180
26	Forward Propagation – Practice —	181
27	Loss Function —	185
27.1	A Game of Arrows —	185
27.2	Common Loss Functions in Deep Learning —	186
27.3	Summary —	190
28	Loss Function – Practice —	191
29	Backward Propagation —	194
29.1	Computation as Gears —	194
29.2	Gradient Descent —	195
29.3	Gradient Descent for Simple Linear Regression —	197
29.4	Gradients and Backpropagation —	199
29.5	Summary —	200
30	Backward Propagation – Practice —	201

31 Deep Learning – Practice — 205

List of Figures — 209

List of Tables — 215

About the Authors — 217

Index — 219