

# CONTENTS IN DETAIL

<b>FOREWORD</b>	<b>xvii</b>
-----------------	-------------

<b>ACKNOWLEDGMENTS</b>	<b>xxi</b>
------------------------	------------

<b>INTRODUCTION</b>	<b>xxiii</b>
---------------------	--------------

Who Is This Book For? .....	xxiv
About This Book .....	xxiv

<b>1</b>	
<b>SETTING THE STAGE</b>	<b>1</b>

Installing the Toolkits .....	2
Linux .....	2
macOS .....	3
Windows .....	3
NumPy .....	4
Defining Arrays .....	5
Data Types .....	5
2D Arrays .....	6
Zeros and Ones .....	7
Advanced Indexing .....	7
Reading and Writing to Disk .....	10
SciPy .....	11
Matplotlib .....	12
Scikit-Learn .....	14
Summary .....	15

<b>2</b>	
<b>PROBABILITY</b>	<b>17</b>

Basic Concepts .....	18
Sample Space and Events .....	18
Random Variables .....	19
Humans Are Bad at Probability .....	19
The Rules of Probability .....	21
Probability of an Event .....	21
Sum Rule .....	24
Product Rule .....	25

Sum Rule Revisited .....	25
The Birthday Paradox .....	26
Conditional Probability .....	30
Total Probability .....	31
Joint and Marginal Probability .....	32
Joint Probability Tables .....	33
Chain Rule for Probability .....	37
Summary .....	39

### **3 MORE PROBABILITY 41**

Probability Distributions .....	41
Histograms and Probabilities .....	42
Discrete Probability Distributions .....	45
Continuous Probability Distributions .....	51
Central Limit Theorem .....	55
The Law of Large Numbers .....	58
Bayes' Theorem .....	59
Cancer or Not Redux .....	60
Updating the Prior .....	61
Bayes' Theorem in Machine Learning .....	62
Summary .....	65

### **4 STATISTICS 67**

Types of Data .....	68
Nominal Data .....	68
Ordinal Data .....	68
Interval Data .....	68
Ratio Data .....	68
Using Nominal Data in Deep Learning .....	69
Summary Statistics .....	70
Means and Median .....	70
Measures of Variation .....	74
Quantiles and Box Plots .....	78
Missing Data .....	83
Correlation .....	85
Pearson Correlation .....	86
Spearman Correlation .....	90
Hypothesis Testing .....	92
Hypotheses .....	93
The t-test .....	95
The Mann-Whitney U Test .....	99
Summary .....	102

**5**  
**LINEAR ALGEBRA** **103**

Scalars, Vectors, Matrices, and Tensors .....	104
Scalars .....	104
Vectors .....	104
Matrices .....	105
Tensors .....	106
Arithmetic with Tensors .....	109
Array Operations .....	109
Vector Operations .....	111
Matrix Multiplication .....	120
Kronecker Product .....	125
Summary .....	126

**6**  
**MORE LINEAR ALGEBRA** **127**

Square Matrices .....	128
Why Square Matrices? .....	128
Transpose, Trace, and Powers .....	129
Special Square Matrices .....	131
The Identity Matrix .....	132
Determinants .....	134
Inverses .....	137
Symmetric, Orthogonal, and Unitary Matrices .....	139
Definiteness of a Symmetric Matrix .....	140
Eigenvectors and Eigenvalues .....	141
Finding Eigenvalues and Eigenvectors .....	141
Vector Norms and Distance Metrics .....	144
L-Norms and Distance Metrics .....	145
Covariance Matrices .....	146
Mahalanobis Distance .....	148
Kullback-Leibler Divergence .....	151
Principal Component Analysis .....	153
Singular Value Decomposition and Pseudoinverse .....	157
SVD in Action .....	158
Two Applications .....	159
Summary .....	161

**7**  
**DIFFERENTIAL CALCULUS** **163**

Slope .....	164
Derivatives .....	165
A Formal Definition .....	165
Basic Rules .....	167
Rules for Trigonometric Functions .....	172

Rules for Exponentials and Logarithms .....	175
Minima and Maxima of Functions .....	177
Partial Derivatives .....	181
Mixed Partial Derivatives .....	183
The Chain Rule for Partial Derivatives .....	184
Gradients .....	186
Calculating the Gradient .....	186
Visualizing the Gradient .....	189
Summary .....	191

## **8**

### **MATRIX CALCULUS** **193**

The Formulas .....	194
A Vector Function by a Scalar Argument .....	194
A Scalar Function by a Vector Argument .....	196
A Vector Function by a Vector .....	197
A Matrix Function by a Scalar .....	198
A Scalar Function by a Matrix .....	198
The Identities .....	199
A Scalar Function by a Vector .....	199
A Vector Function by a Scalar .....	202
A Vector Function by a Vector .....	203
A Scalar Function by a Matrix .....	203
Jacobians and Hessians .....	205
Concerning Jacobians .....	205
Concerning Hessians .....	211
Some Examples of Matrix Calculus Derivatives .....	217
Derivative of Element-Wise Operations .....	217
Derivative of the Activation Function .....	218
Summary .....	220

## **9**

### **DATA FLOW IN NEURAL NETWORKS** **221**

Representing Data .....	222
Traditional Neural Networks .....	222
Deep Convolutional Networks .....	223
Data Flow in Traditional Neural Networks .....	225
Data Flow in Convolutional Neural Networks .....	229
Convolution .....	229
Convolutional Layers .....	234
Pooling Layers .....	237
Fully Connected Layers .....	239
Data Flow Through a Convolutional Neural Network .....	239
Summary .....	242

<b>10</b>		
<b>BACKPROPAGATION</b>		<b>243</b>
What Is Backpropagation? .....	244	
Backpropagation by Hand .....	245	
Calculating the Partial Derivatives .....	246	
Translating into Python.....	249	
Training and Testing the Model.....	253	
Backpropagation for Fully Connected Networks .....	254	
Backpropagating the Error .....	255	
Calculating Partial Derivatives of the Weights and Biases .....	258	
A Python Implementation.....	260	
Using the Implementation .....	264	
Computational Graphs.....	267	
Summary .....	269	

<b>11</b>		
<b>GRADIENT DESCENT</b>		<b>271</b>
The Basic Idea .....	272	
Gradient Descent in One Dimension .....	272	
Gradient Descent in Two Dimensions.....	276	
Stochastic Gradient Descent .....	282	
Momentum .....	284	
What Is Momentum? .....	284	
Momentum in 1D .....	285	
Momentum in 2D .....	287	
Training Models with Momentum .....	289	
Nesterov Momentum .....	294	
Adaptive Gradient Descent.....	297	
RMSprop.....	297	
Adagrad and Adadelta.....	299	
Adam.....	300	
Some Thoughts About Optimizers.....	301	
Summary .....	303	
Epilogue.....	303	

<b>APPENDIX: GOING FURTHER</b>		<b>305</b>
Probability and Statistics .....	305	
Linear Algebra .....	306	
Calculus .....	306	
Deep Learning .....	307	

<b>INDEX</b>		<b>309</b>
--------------	--	------------