

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	The Quantified Self	2
1.2	The Goal of this Book	4
1.3	Basic Terminology	5
1.3.1	Data Terminology	5
1.3.2	Machine Learning Terminology	7
1.4	Basic Mathematical Notation	8
1.5	Overview of the Book	10
 <b>Part I Sensory Data and Features</b>		
<b>2</b>	<b>Basics of Sensory Data</b>	<b>15</b>
2.1	Crowdsignals Dataset	15
2.2	Converting the Raw Data to an Aggregated Data Format	17
2.3	Exploring the Dataset	19
2.4	Machine Learning Tasks	23
2.5	Exercises	24
2.5.1	Pen and Paper	24
2.5.2	Coding	24
<b>3</b>	<b>Handling Noise and Missing Values in Sensory Data</b>	<b>25</b>
3.1	Detecting Outliers	27
3.1.1	Distribution-Based Models	28
3.1.2	Distance-Based Models	30
3.2	Imputation of Missing Values	34
3.3	A Combined Approach: The Kalman Filter	35
3.4	Transformation	37
3.4.1	Lowpass Filter	38
3.4.2	Principal Component Analysis	38

3.5	Case Study . . . . .	42
3.5.1	Outlier Detection . . . . .	43
3.5.2	Missing Value Imputation . . . . .	45
3.5.3	Kalman Filter . . . . .	46
3.5.4	Data Transformation . . . . .	47
3.6	Exercises. . . . .	49
3.6.1	Pen and Paper . . . . .	49
3.6.2	Coding. . . . .	50
<b>4</b>	<b>Feature Engineering Based on Sensory Data . . . . .</b>	<b>51</b>
4.1	Time Domain . . . . .	51
4.1.1	Numerical Data . . . . .	52
4.1.2	Categorical Data . . . . .	54
4.1.3	Mixed Data . . . . .	56
4.2	Frequency Domain . . . . .	58
4.2.1	Fourier Transformations . . . . .	58
4.2.2	Features in Frequency Domain . . . . .	60
4.3	Features for Unstructured Data . . . . .	62
4.3.1	Pre-processing Text Data. . . . .	62
4.3.2	Bag of Words . . . . .	63
4.3.3	TF-IDF . . . . .	63
4.3.4	Topic Modeling. . . . .	64
4.4	Case Study . . . . .	65
4.4.1	Time Domain . . . . .	66
4.4.2	Frequency Domain . . . . .	67
4.4.3	New Dataset . . . . .	68
4.5	Exercises. . . . .	69
4.5.1	Pen and Paper . . . . .	69
4.5.2	Coding. . . . .	70
 <b>Part II Learning Based on Sensory Data</b>		
<b>5</b>	<b>Clustering . . . . .</b>	<b>73</b>
5.1	Learning Setup . . . . .	73
5.2	Distance Metrics . . . . .	74
5.2.1	Individual Data Points Distance Metrics . . . . .	74
5.2.2	Person Level Distance Metrics . . . . .	77
5.3	Non-hierarchical Clustering . . . . .	82
5.4	Hierarchical Clustering . . . . .	84
5.4.1	Agglomerative Clustering . . . . .	84
5.4.2	Divisive Clustering . . . . .	87
5.5	Subspace Clustering . . . . .	88
5.6	Datastream Clustering. . . . .	91
5.7	Performance Evaluation . . . . .	93

5.8	Case Study . . . . .	94
5.8.1	Non-hierarchical Clustering . . . . .	94
5.8.2	Hierarchical Clustering . . . . .	98
5.9	Exercises. . . . .	98
5.9.1	Pen and Paper . . . . .	98
5.9.2	Coding. . . . .	100
<b>6</b>	<b>Mathematical Foundations for Supervised Learning . . . . .</b>	<b>101</b>
6.1	Learning Process and Elements . . . . .	101
6.1.1	Unknown Target Function. . . . .	102
6.1.2	Observed Data. . . . .	104
6.1.3	Error Measure . . . . .	105
6.1.4	Hypothesis Set and the Learning Machine. . . . .	107
6.1.5	Model Selection and Evaluation . . . . .	111
6.2	Learning Theory . . . . .	114
6.2.1	PAC Learnability. . . . .	114
6.2.2	VC-Dimension and VC-Bound . . . . .	116
6.2.3	Implications. . . . .	118
6.3	Exercises. . . . .	120
6.3.1	Pen and Paper . . . . .	120
6.3.2	Coding. . . . .	121
<b>7</b>	<b>Predictive Modeling without Notion of Time. . . . .</b>	<b>123</b>
7.1	Learning Setup . . . . .	123
7.2	Feedforward Neural Networks . . . . .	125
7.2.1	Perceptron . . . . .	125
7.2.2	Multi-layer Perceptron. . . . .	128
7.2.3	Convolutional Neural Networks. . . . .	129
7.3	Support Vector Machines . . . . .	131
7.4	K-Nearest Neighbor . . . . .	134
7.5	Decision Trees . . . . .	135
7.6	Naive Bayes . . . . .	139
7.7	Ensembles. . . . .	140
7.7.1	Bagging. . . . .	141
7.7.2	Boosting . . . . .	141
7.8	Predictive Modeling for Data Streams . . . . .	144
7.9	Practical Considerations . . . . .	145
7.9.1	Feature Selection . . . . .	145
7.9.2	Regularization . . . . .	147
7.10	Case Study . . . . .	148
7.10.1	Classification: Predicting the Activity Label . . . . .	149
7.10.2	Regression: Predicting the Heart Rate . . . . .	157

7.11	Exercises. . . . .	163
7.11.1	Pen and Paper . . . . .	163
7.11.2	Coding. . . . .	164
<b>8</b>	<b>Predictive Modeling with Notion of Time . . . . .</b>	<b>167</b>
8.1	Learning Setup . . . . .	167
8.2	Time Series Analysis . . . . .	168
8.2.1	Basic Concepts . . . . .	169
8.2.2	Filtering and Smoothing . . . . .	170
8.2.3	Autoregressive Integrated Moving Average Model—ARIMA . . . . .	173
8.2.4	Estimating and Forecasting Time Series Models . . . . .	176
8.2.5	Example Application . . . . .	177
8.3	Neural Networks. . . . .	181
8.3.1	Recurrent Neural Networks . . . . .	182
8.3.2	Echo State Networks . . . . .	184
8.4	Dynamical Systems Models . . . . .	186
8.4.1	Example Based on Bruce’s Data . . . . .	186
8.4.2	Parameter Optimization . . . . .	188
8.5	Case Study . . . . .	195
8.5.1	Tuning Parameters. . . . .	195
8.5.2	Results. . . . .	197
8.6	Exercises. . . . .	201
8.6.1	Pen and Paper . . . . .	201
8.6.2	Coding. . . . .	201
<b>9</b>	<b>Reinforcement Learning to Provide Feedback and Support . . . . .</b>	<b>203</b>
9.1	Basic Setting. . . . .	203
9.2	One-Step SARSA Temporal Difference Learning . . . . .	208
9.3	Q-Learning . . . . .	210
9.4	SARSA( $\lambda$ ) and Q( $\lambda$ ) . . . . .	211
9.5	Approximate Solutions . . . . .	212
9.6	Discretizing the State Space . . . . .	212
9.7	Exercises. . . . .	213
9.7.1	Pen and Paper . . . . .	213
9.7.2	Coding. . . . .	214

### Part III Discussion

<b>10</b>	<b>Discussion . . . . .</b>	<b>217</b>
10.1	Learning Full Circle . . . . .	217
10.2	Heterogeneity . . . . .	218
10.3	Effective Data Collection and Reuse. . . . .	219
10.4	Data Processing and Storage. . . . .	219

10.5	Better Predictive Modeling and Clustering . . . . .	220
10.6	Validation . . . . .	221
<b>References . . . . .</b>		<b>223</b>
<b>Index . . . . .</b>		<b>229</b>

<http://www.springer.com/978-3-319-66307-4>

Machine Learning for the Quantified Self  
On the Art of Learning from Sensory Data

Hoogendoorn, M.; Funk, B.

2018, XV, 231 p. 89 illus., 72 illus. in color., Hardcover

ISBN: 978-3-319-66307-4