

Contents

1	Benefits of Machine Learning and Artificial Intelligence	1
1.1	Artificial Intelligence	3
1.2	Human Intelligence	3
1.3	Rational Decisions	5
1.4	Decision Automation	7
1.5	Decision-Making Process in Humans and Machines	8
1.6	Speed of Decision Making	10
1.7	Data Basis for Decisions	11
1.8	Business Intelligence	13
1.9	Innovation	15
	References	16
2	Machine Learning	17
2.1	Supervised Learning	20
2.2	Unsupervised Learning	21
2.3	Reinforcement Learning	21
	Reference	22
3	Best Practices	23
3.1	Data Quality	25
3.2	Visualization	28
3.3	Development Environment	30
3.4	Programming	30
	References	32
4	Setting up the R Development Environment	33
4.1	Installing R	33
4.2	Installing Rtools	40
4.3	Installing RStudio	43
4.4	User Interface	47

4.5	Advanced Functions.	51
4.6	The machinelearning Package	54
	References.	55
5	Basics of the R Programming Language.	57
5.1	Data Types	59
5.2	Data Structures.	63
5.2.1	Vector	63
5.2.2	Factor Variable	66
5.2.3	List	70
5.2.4	Matrix and Array	71
5.2.5	Dataframe and Tibble.	75
5.2.6	Overview Table	82
5.3	Operations	82
5.3.1	Arithmetic Operations	82
5.3.2	Logical Operations.	85
5.3.3	Selective Operations.	87
5.3.4	Text Operations	89
5.3.5	Date Operations	91
5.4	Functions	97
5.4.1	Calling Functions.	97
5.4.2	Creating Functions.	100
5.5	Control Structures	105
6	Basics of Machine Learning with R	111
6.1	Import and Export of Data.	111
6.1.1	Data Import	113
6.1.2	Data Export	114
6.2	Preprocessing.	118
6.2.1	Selection.	118
6.2.2	Validation	126
6.2.3	Merging	131
6.2.4	Transformation.	138
6.2.5	Addition	144
6.2.6	Summary	150
6.2.7	Best Practices Case Study	156
6.3	Exploratory Data Analysis.	165
6.3.1	Analysis Methods for Base Tables.	167
6.3.2	Statistical Analysis Methods	178
6.3.3	Data Visualization	206
6.4	Modeling and Evaluation.	253
6.4.1	Modeling	255
6.4.2	Samples	279

6.4.3	Training and Test Data	288
6.4.4	Evaluation	291
6.5	The Tidymodels Package	297
6.5.1	Model Selection, Model Training, and Model Application	297
6.5.2	Resampling.	316
6.5.3	Preprocessing	332
6.5.4	Comparison of Models.	359
6.5.5	Parameter Tuning	371
6.5.6	Conclusion	395
	References.	395
7	Application of Machine Learning with R	397
7.1	Supervised Learning	397
7.1.1	Regression	397
7.1.2	Classification	444
7.2	Unsupervised Learning	455
7.2.1	Clustering	455
7.2.2	Association Rules.	471
	References.	488
8	Outlook.	489
	References.	492
	References	493