

Inhalt

Vorwort	IX
1 Einleitung	1
1.1 Eigenschaften von Zeitreihen	2
1.2 Daten, Korrelationen und das „Ende der Theorie“	5
1.3 Inhalt	7
1.4 Voraussetzungen, Ressourcen und praktische Hinweise	9
2 Verarbeitung und Vorverarbeitung von Zeitreihen	11
2.1 Mit Datumsangaben arbeiten	12
2.1.1 Daten laden	13
2.1.2 Datumsangaben und Zeitstempel parsen	13
2.1.3 Indexbezogene Datums-Funktionen	15
2.2 Operationen entlang der Zeitachse	17
2.2.1 Mit Lags arbeiten	18
2.2.2 Differenzen erster und höherer Ordnung	18
2.3 Imputation fehlender Werte	20
2.3.1 Fehlende Werte vorbehandeln	20
2.3.2 Einfache Imputationsverfahren	22
2.3.3 Mit gleitenden Mittelwerten arbeiten	24
2.3.4 Zeitfenster imputieren	27
2.4 Daten mit NumPy in Form bringen	31
2.4.1 Pandas und NumPy	31
2.4.2 Slicing	33
2.4.3 Restrukturierungsmaßnahmen	34
3 Grundprinzipien maschinellen Lernens	37
3.1 Lineare Regression	38
3.1.1 Grundlagen	38
3.1.2 Umsetzung mit Scikit-learn	45
3.1.3 Trainings- und Testdaten separieren	51
3.2 Logistische Regression	54
3.2.1 Grundlagen	54
3.2.2 Umsetzung mit Scikit-learn	59

3.3	Softmax-Regression	63
3.3.1	Grundlagen	63
3.3.2	Umsetzung mit Scikit-learn	65
3.4	Feature-Vorverarbeitung	67
3.4.1	One-Hot-Codierung	68
3.4.2	Standardisierung	73
3.4.3	Hauptkomponentenanalyse	75
3.4.4	Vorverarbeitungsmethoden in der Praxis	82
3.5	Zeitreihen mit Standardverfahren verarbeiten	88
3.5.1	Modelle mit Zeitangaben anlernen	89
3.5.2	Mit Interaktionsvariablen arbeiten	92
3.5.3	Nicht-lineare Beziehungen modellieren	96
4	Prognoseverfahren für univariate Zeitreihen:	
	ARIMA & Seasonal ARIMA	101
4.1	Autoregression (AR)	102
4.2	Moving-Averages-Modell (MA)	104
4.3	Stationarität	106
4.3.1	Auf Stationarität testen	106
4.3.2	Saisonale Komponenten entfernen	109
4.3.3	Trends entfernen	111
4.3.4	Warum man Zeitreihen stationär macht	114
4.4	Autokorrelationen und partielle Autokorrelationen	115
4.5	ARIMA-Verfahren	118
4.5.1	Umsetzung mit statsmodels	118
4.5.2	Evaluation des Modells über die Testdaten	121
4.5.3	Modell-Kenngrößen zur Evaluation einsetzen	125
4.6	Zeitreihen mit saisonalen Komponenten modellieren	127
4.6.1	Saisonale Daten analysieren und modellieren	127
4.6.2	Modelle mit der Brut-Force-Methode anlernen	133
4.6.3	Prognosen erstellen	136
4.7	Trends verarbeiten	138
4.7.1	Konstante Trends modellieren	139
4.7.2	Mit komplexen Trends arbeiten	140
4.8	Kontexteffekte integrieren	142
5	Deep-Learning-Verfahren	149
5.1	Arbeitsweise neuronaler Netze	150
5.1.1	Aufbau neuronaler Netze	150
5.1.2	Training eines neuronalen Netzes	154
5.1.3	Neuronale Netze auf Lernaufgaben einstellen	155
5.2	Mit TensorFlow 2/Keras arbeiten	157
5.2.1	Ein einfaches Keras-Modell aufbauen	158

5.2.2	Einen Klassifizierer anlernen	165
5.2.3	Den Anlernprozess steuern	170
5.3	Überanpassung verhindern	175
5.3.1	Regularisierung	177
5.3.2	Dropout	182
5.4	Rekurrente Netze	185
5.4.1	Funktionsweise rekurrenter Layer	185
5.4.2	Long Short Term Memory (LSTM) und Gated Recurrent Unit (GRU)	190
5.4.3	Training eines rekurrenten Layers mit Keras	191
5.4.4	Mit Zeitfenstern arbeiten	195
5.5	Konvolutionale Netze	203
5.5.1	Funktionsweise konvolutionaler Schichten	204
5.5.2	Zeitreihen mit konvolutionalen Layern verarbeiten	207
5.5.3	Training einer Zeitreihe mit einer konvolutionalen Schicht	208
5.6	Mit Zeitfenstern in der Praxis arbeiten	213
5.6.1	Generatoren	213
5.6.2	Zeitfenster mit Generatoren anlernen	220
5.7	Domänenspezifische Lernarchitekturen umsetzen	230
5.7.1	Saisonale Komponenten vorverarbeiten	231
5.7.2	Mit der funktionalen Keras-API arbeiten	235
5.7.3	Zeitreihendaten und zeitkonstante Daten in einem Modell verarbeiten ...	237
5.7.4	Lernarchitektur und Preprocessing	241
5.8	Dimensionsreduktion mit Autoencodern	248
5.8.1	Architektur eines Autoencoders	248
5.8.2	Einen Autoencoder anlernen	251
5.8.3	Dimensionen kategorialer Variablen reduzieren	255
Literaturverzeichnis		259
Stichwortverzeichnis		261