

Inhaltsverzeichnis

1	Einleitung	1
2	Sentiment Retrieval – Meinungsäußerungen identifizieren	5
2.1	Meinungsäußerungen	5
2.1.1	Meinungsäußerungen vs. Sachinformation	6
2.1.2	Arten von Meinungsäußerungen	8
2.2	Methoden für die Suche nach Meinungsäußerungen	10
2.3	Zusammenfassung	13
2.4	Übungen	14
2.5	Weiterführende Literatur	15
3	Polarität: Dokumente klassifizieren	17
3.1	Die Aufgabe	17
3.2	Vorbereitung der Daten	18
3.3	Wortlistenabgleich	20
3.4	Qualitätssicherung und systematische Evaluation	20
3.5	Klassifikation und Regression	23
3.6	Dokumentklassifikation mit maschinellem Lernen	25
3.6.1	Supervised Learning – Probabilistisches Sprachmodell	25
3.6.2	Supervised Learning mit Features	27
3.6.3	Deep Learning	33
3.7	Zusammenfassung	33
3.8	Übungen	33
3.9	Weiterführende Literatur	35
4	Wörter in der Sentiment-Analyse	37
4.1	Normalisierung der Texte	37
4.2	Einbindung eines existierenden Sentiment-Wörterbuchs	38
4.3	Gewinnung von Sentiment-Wörtern mithilfe von WordNet	40
4.4	Gewinnung von Sentiment-Wörtern aus annotierten Korpora	42
4.5	Gewinnung von Wörtern aus nicht annotierten Korpora	44

4.6	Zusammenfassung	45
4.7	Übungen	46
4.8	Weiterführende Literatur	48
5	Sentiment-Analyse auf Satzebene	49
5.1	Satz-Tokenisierung	50
5.2	Identifikation von Sätzen mit Meinungsäußerungen	51
5.3	Satzanalyse	52
5.4	Zusammenfassung	55
5.5	Übungen	55
5.6	Weiterführende Literatur	57
6	Was bewertet wird: Aspekte identifizieren	59
6.1	Taxonomie der Aspekte	60
6.2	Phrasen-Lexikon der Aspekte	62
6.3	Aspekte im Text identifizieren und interpretieren	64
6.4	Aspektidentifizierung ohne Beschränkung auf eine Domäne	65
6.5	Sentiment-Klassifikation des Aspekts	66
6.6	Zusammenfassung	68
6.7	Übungen	68
6.8	Weiterführende Literatur	69
7	Ironie	71
7.1	Übungen	73
7.2	Weiterführende Literatur	74
8	Analyse politischer Trends	75
8.1	Aufstellung der Datenbasis	76
8.1.1	Tweets mit Meinungen zu Politikerinnen und Politikern	76
8.1.2	ZDF-Politbarometer	77
8.2	Sentiment-Analyse für Tweets	78
8.3	Zusammenfassung	79
8.4	Übungen	79
8.5	Weiterführende Literatur	80
9	Opinion Spam	81
9.1	Gefälschte Bewertungen	82
9.2	Annotierte Korpora für Opinion Spam	83
9.3	Klassifikation von Bewertungen	85
9.3.1	Klassifikation mit Meta-Daten	85
9.3.2	Klassifikation mit linguistischer Information	86
9.4	Beobachtungen über Opinion Spam im deutschsprachigen Amazon-Portal	87
9.5	Maschinelles Lernen für die automatische Klassifikation	89

9.6	Zusammenfassung	91
9.7	Übungen	91
9.8	Weiterführende Literatur	92
10	Erkennung und Klassifikation von Aggression in Meinungsäußerungen	93
10.1	Daten, Daten, Daten	95
10.2	Methoden zur automatischen Klassifikation	97
10.3	Zusammenfassung	100
10.4	Übungen	101
10.5	Weiterführende Literatur	101
11	Sentiment-Analyse im Unternehmenskontext und Softwarelösungen im Markt	103
11.1	Markt für kostenpflichtige Sentiment-Analyse-Tools und -Services	103
11.1.1	Technische Bereitstellung der Lösung	105
11.1.2	Art der Sentiment-Analyse	107
11.1.3	Sprachenabdeckung	108
11.1.4	Leistungsumfang und Funktionen	108
11.2	Tools für deutschsprachige Texte	109
11.2.1	Amazon Comprehend	109
11.2.2	Cogito Intelligence Plattform von Expert System	110
11.2.3	InMap, Insius	111
11.2.4	Monkey Learn	111
11.2.5	OpenText Magellan Text Mining	112
11.2.6	ParallelDots	112
11.2.7	SAS® Visual Text Analytics	114
11.2.8	Sentiment Intelligence in SAP Hana	114
11.2.9	Sentiment Lab von m-result	115
11.2.10	Ubermetrics	116
11.3	Anwendung der Sentiment-Analyse in der Praxis	117
11.4	Zusammenfassung	121
11.5	Übungen	121
11.6	Weiterführende Literatur	122
Literatur	123	
Stichwortverzeichnis	129	