

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abkürzungsverzeichnis	III
Formelverzeichnis	IV
Abbildungsverzeichnis	V
Tabellenverzeichnis	VII
1 Einleitung.....	1
1.1 Motivation	1
1.2 Ziel der Arbeit und Vorgehensweise	3
2 Patente als Träger technischer Informationen.....	5
2.1 Patente als Informationsquelle.....	5
2.1.1 Aufbau einer Patentschrift.....	6
2.1.2 Patentzeichnungen	8
2.1.3 Die Internationale Patentklassifikation	10
2.1.4 Recherchierbare Patentinformationen	11
2.2 Patente als Ideenpool	15
2.3 Bedeutung für die weitere Untersuchung	17
3 Methoden zum Auffinden von Informationen.....	19
3.1 Text Mining: Informationen aus Texten	20
3.1.1 Wortbasierte Merkmalsextraktion aus Texten	21
3.1.2 N-Gramm basierte Merkmalsextraktion aus Texten	22
3.1.3 Merkmalsgewichtung zur Bewertung der Dokumente	23
3.1.4 Merkmalsauswahl für effizientes Klassifikatortraining	24
3.2 Image Mining: Informationen aus Bildern	25
3.2.1 Merkmalsauswahl in Bildern	26
3.3 Automatische Klassifikation zur effizienten Selektion.....	27
3.3.1 Übersicht zu Klassifikationsverfahren	29
3.3.2 Gütekriterien zur Bewertung der Klassifikationsperformance	33
3.4 Softwarebasierte Methodenumsetzung.....	34
3.4.1 RapidMiner	36
3.5 Automatisierte Patentklassifikation in Kinderschuhen	37
3.6 Bedeutung für die Arbeit	39
4 Forschungsansatz: Softwarebasiertes Patent-Mining.....	43
4.1 Anforderungen und Abgrenzung	44
4.2 Forschungshypothesen zu ausgewählten Problemfeldern	46
5 Verfahrensentwurf zur zweistufigen Patentklassifikation	49
5.1 Genereller Ablauf.....	49
5.2 Aufbau der Dokumentenbasis.....	52
5.2.1 Stufe 1: Auswahl der Textdateien	53
5.2.2 Stufe 2: Auswahl der Bilddateien	55
5.3 Ablauf der Merkmalsextraktion.....	56
5.3.1 Stufe 1: Merkmale aus Text	57
5.3.2 Stufe 2: Merkmale aus Bildern	62
5.4 Ablauf der automatischen Klassifikation und Auswahl der Parameter	63
5.4.1 Kritische Betrachtung des Verfahrensaufbaus	68
6 Evaluation des Verfahrens.....	69
6.1 Testumgebung und Festlegung der Dokumentenbasis	69
6.2 Überprüfung der Merkmalsextraktion aus Texten und Bildern.....	72
6.3 Überprüfung der automatischen Patentklassifikation	76
6.3.1 Ergebnisse der Textklassifikation mit Wörtern als Merkmale	78
6.3.2 Ergebnisse der Textklassifikation mit N-Grammen als Merkmale.....	87
6.3.3 Kritische Betrachtung der Ergebnisse der Textklassifikation	92
6.3.4 Ergebnisse der Bildklassifikation.....	93
6.4 Referenztests zur Überprüfung der Praxistauglichkeit des Verfahrens.....	94
6.4.1 Ergebnisse der Referenztests zur technischen Funktion <i>Verriegeln</i>	97
6.4.2 Ergebnisse der Referenztests zur technischen Funktion <i>Sever</i>	100

6.5	Zusammenfassung der Ergebnisse der Evaluation	102
7	Wesentliche Erkenntnisse und Ausblick.....	107
8	Literatur- und Quellenverzeichnis	111
9	Anhang	117
9.1	Extrahierte Merkmale.....	117
9.2	Merkmale zur Bildklassifikation	119
9.3	Vorkommenshäufigkeit von Wörtern	120
9.4	Wortliste	121
9.5	Standardabweichung bei wortbasierter Klassifikation	123
9.6	Preprocessing.....	129
9.7	Einfluss der Vorverarbeitung.....	142
9.8	N-Gramme bestimmen	144
9.9	Beispiel von 3-Grammen	145
9.10	Naive Bayes Klassifikation an Bildern	148
9.11	Referenztests zur technischen Funktion <i>Durchtrennen</i>	149