

---

# INHALTSVERZEICHNIS

---

1	EINFÜHRUNG	1
1.1	Motivation	2
1.2	Herausforderungen, Zielsetzung und innovative Aspekte	4
1.3	Aktives Lernbeispiel	8
1.4	Eigene Veröffentlichungen	10
1.5	Aufbau der Arbeit	11
2	THEORETISCHE UND METHODISCHE GRUNDLAGEN	13
2.1	Aktives Lernen	15
2.2	Modellierung von Struktur in Daten	21
2.3	Classifier based on Mixture Models (CMM)	26
2.3.1	Shared-Components-Ansatz (CMM <sub>sha</sub> )	28
2.3.2	Separate-Components-Ansatz (CMM <sub>sep</sub> )	29
2.4	Support Vector Machines (SVM)	31
2.4.1	Linearer Fall	32
2.4.2	Nicht-linearer Fall	35
2.5	Evaluation	38
2.5.1	Allgemeine Anmerkungen	38
2.5.2	Bewertungsmaße	39
2.5.3	Datensätze	48
3	NUTZUNG VON STRUKTURINFORMATION ZUR AKTIVEN MUSTER- AUSWAHL	51
3.1	Stand der Technik	52
3.2	Grundanforderungen	60
3.3	Distance-Density-Distribution-Diversity Sampling (4DS)	61
3.4	Experimentelle Untersuchungen	71
3.4.1	Aktive Selektionsstrategien	71
3.4.2	Exemplarische Untersuchung der Selektionsstrategie 4DS	73
3.4.3	Vergleich basierend auf 20 Benchmark-Datensätzen	81
3.5	Abschlussbemerkung	92
4	AKTIVE VERFEINERUNG DER STRUKTURINFORMATION	93
4.1	Stand der Technik	94
4.2	Erweiterung des aktiven Lernzyklus	102
4.3	Transduktiver Lernprozess	105
4.3.1	Schritt 0: Aktualisierung der Klassen-A-posteriori	105
4.3.2	Schritt 1: Bewertung der Modellkomponenten	106

4.3.3	Schritt 2: Mustererfassung . . . . .	108
4.3.4	Schritt 3: Musterbasierter Klassifikator . . . . .	109
4.3.5	Schritt 4: Lokaler Klassifikator $CMM_{sep}$ . . . . .	114
4.3.6	Schritt 5: Fusion der Modellkomponenten . . . . .	114
4.3.7	Schritt 6: Musterklassifikation . . . . .	115
4.4	Erklärung am Beispiel . . . . .	115
4.5	Experimentelle Untersuchungen . . . . .	122
4.5.1	Klassifikatoren . . . . .	122
4.5.2	Aktive Selektionsstrategien . . . . .	124
4.5.3	Vergleich basierend auf 20 Benchmark-Datensätzen . . . . .	125
4.6	Abschlussbemerkung . . . . .	138
5	NUTZUNG VON STRUKTURINFORMATION FÜR DAS AKTIVE TRAINING VON SVM . . . . .	139
5.1	Responsibility-Weighted-Mahalanobis-Kernel . . . . .	143
5.1.1	RWM-Ähnlichkeitsmaß . . . . .	143
5.1.2	RWM-Kernelfunktion . . . . .	149
5.1.3	Erweiterung für kategorische Eingabedimensionen . . . . .	154
5.2	Halb-überwachtes Training einer SVM mit RWM-Kernel . . . . .	155
5.2.1	Stand der Technik . . . . .	156
5.2.2	Experimentelle Untersuchungen . . . . .	158
5.2.3	Abschlussbemerkung . . . . .	170
5.3	Aktives Training einer SVM mit RWM-Kernel . . . . .	172
5.3.1	Stand der Technik . . . . .	173
5.3.2	Experimentelle Untersuchungen . . . . .	175
5.3.3	Abschlussbemerkung . . . . .	208
6	ZUSAMMENFASSUNG UND AUSBLICK . . . . .	211
6.1	Zusammenfassung . . . . .	211
6.2	Ausblick . . . . .	214
A	ANHANG . . . . .	219
A.1	Semi-Metrik-Beweis der RWM-Ähnlichkeit . . . . .	219
A.2	Künstlich erzeugte Datensätze . . . . .	220
	LITERATURVERZEICHNIS . . . . .	221