Aufgabe 3

Prozess

Ergebnisse

ExampleSet (4000 examples, 1 special attribute, 147 regular attributes) Filter (4,000 / 4,000 examples): all

Row No.	label	HitMultip...	HitMultip...	HitMultip...	HitStatis...	HitStatis...	HitStatis...	HitStatis...	HitStatis...	HitStatis...	HitStatis...	HitStatis...	HitStatis...
1	1	9	22	17	55.942	9399.048	12421.504	17.329	85.550	280.220	182.938	56.429	52.255
2	1	15	50	41	45.717	9837.945	12119.147	4.361	-504.300	-280.550	-443.619	49.701	40.734
3	1	30	140	97	93.834	9882.003	14738.903	57.306	-503.460	52.600	-296.950	103.831	116.446
4	1	9	27	15	106.099	9883.112	12503.293	4.954	-504.150	-145.320	-340.967	97.697	86.889
5	1	13	38	20	120.876	9878.854	11714.215	8.385	-415.030	88.450	-248.253	124.032	49.674
6	1	4	11	9	88.561	9872.640	11345.416	1.918	58	279.320	160.112	85.011	85.299
7	1	9	16	12	41.738	9893.411	11754.055	10.251	4.110	175.030	66.529	49.984	23.136
8	1	8	19	15	92.522	9583.481	12029.004	4.632	-491.480	-148.240	-303.198	97.902	87.155
9	1	12	34	27	225.991	9564.350	13440.005	5.803	-470.670	191.230	-180.130	207.729	224.101
10	1	9	22	16	135.321	9874.513	13671.294	4.481	-281.370	125.560	-77.915	147.280	148.420
11	1	19	108	43	51.667	9874.675	13748.153	18.921	24.250	366.540	196.910	74.036	9.336
12	1	18	68	47	275.385	9872.384	13060.420	26.783	-506.970	189.660	-311.388	218.793	199.982
13	1	6	14	12	82.951	9880.571	11160.141	2.580	-506.440	-263.660	-396.564	78.865	89.166
14	1	5	19	9	121.160	9885.377	11404.301	6.978	140.170	501.480	294.214	131.098	81.893
15	1	10	36	21	345.903	9336.105	13391.496	29.866	-487.550	512.950	-46.808	318.351	313.358
16	1	4	21	14	235.982	9579.292	12316.274	11.293	-367.980	463.490	107.078	273.183	413.129
17	1	5	13	5	41.991	9870.368	10988.441	3.729	161.930	310.600	245.631	47.066	65.824
18	1	5	13	8	116.382	9872.085	11756.348	3.195	-316.930	25.760	-132.977	119.881	137.432
19	1	12	23	18	56.028	9731.615	12414.241	3.824	-504.390	-298.120	-428.749	59.518	54.237
20	1	10	36	27	61.477	9876.188	11931.686	6.669	-350.200	-145.810	-214.708	51.918	51.940
21	1	12	40	30	54.055	9877.487	11979.889	7.280	-385.680	-128.930	-241.837	60.113	3.702
22	1	4	24	16	178.091	9879.500	11997.765	14.379	-506.440	6.520	-260.668	168.534	265.305
23	1	14	38	21	41.996	9875.678	13044.820	6.963	-40.570	180.880	49.093	52.794	8.398
24	1	5	16	11	30.449	9876.812	12085.666	2.756	-506.510	-387.360	-468.335	34.452	11.746
25	1	5	13	8	27.233	9876.938	11768.276	5.526	229.940	332.070	299.312	34.336	-3.020
26	1	9	37	22	110.367	9891.458	11809.221	4.912	-6.210	380.920	139.022	108.908	59.775
27	1	6	22	15	123.262	9887.093	12721.432	5.686	-29.100	448.590	157.895	129.767	136.083
28	1	6	16	7	26.920	9877.361	12097.622	23.405	-210.450	-94.750	-152.376	35.391	12.848
29	1	4	11	7	55.892	10114.906	11758.688	3.209	145.350	328.340	229.637	63.254	66.226

ExampleSet (4000 examples, 1 special attribute, 30 regular attributes) Filter (4,000 / 4,000 examples): all

Row No.	label	HitStatis...	LineFit_T...	LineFit_T...	LineFit_T...	LineFit_T...	MPEFitHi...	MPEFitHi...	MPEFitPa...	MPEFitPa...	MPEFitPa...	MPEFit_T...	MuEXAng...
1	1	52.255	1.769	0.310	0.121	0.061	1.770	0	1.763	0.019	0.009	8.009	1.761
2	1	40.734	1.934	0.273	-0.056	0.097	1.937	0	1.936	0.035	0.018	7.046	2.002
3	1	116.446	2.013	0.302	-0.135	0.129	2.009	0	2.009	0.002	0.002	6.547	2.010
4	1	86.889	2.201	0.281	0.036	0.165	2.210	0	2.202	0.023	0.012	8.818	2.198
5	1	49.674	3.130	0.282	-0.003	0.282	3.016	0	3.016	0.017	0.009	7.647	2.981
6	1	85.299	2.538	0.257	-0.141	0.211	2.255	0	2.255	0.038	0.035	10.778	2.252
7	1	23.136	1.823	0.255	0.121	0.064	1.855	0	1.900	0.106	0.028	9.583	1.838
8	1	87.155	2.620	0.308	-0.015	0.267	2.581	0	2.588	0.054	0.049	8.479	2.648
9	1	224.101	2.210	0.298	0.040	0.178	2.202	0	2.201	0.012	0.003	7.438	2.201
10	1	148.420	1.968	0	0	0	1.956	0	1.961	0.009	0.005	7.873	1.956
11	1	9.336	1.642	0.314	-0.304	0.022	1.609	0	1.599	0.020	0.008	7.232	1.610
12	1	199.982	2.431	0.300	-0.157	0.227	2.440	0	2.442	0.003	0.003	6.721	2.441
13	1	89.166	2.664	0.353	0.102	0.313	2.512	0	2.505	0.070	0.046	9.525	2.527
14	1	81.893	2.475	0.317	0.031	0.249	2.452	0	2.504	0.078	0.043	9.037	2.448
15	1	313.358	2.673	0.307	0.139	0.274	2.676	0	2.677	0.004	0.004	6.714	2.679
16	1	413.129	2.802	0.308	0.101	0.290	2.793	0	2.792	0.011	0.006	6.935	2.787
17	1	65.824	1.851	0.346	0.058	0.096	1.798	0	1.795	0.069	0.041	8.927	1.806
18	1	137.432	2.465	0.298	0.029	0.233	2.521	0	2.494	0.038	0.009	9.058	2.502
19	1	54.237	1.863	0.298	0.168	0.086	1.864	0	1.864	0.028	0.021	7.983	1.887
20	1	51.940	2.159	0.225	-0.021	0.125	2.233	0	2.233	0.011	0.008	6.715	2.220
21	1	3.702	1.929	0.223	0.177	0.078	1.856	0	1.858	0.037	0.019	7.479	1.861
22	1	265.305	2.872	0.317	0.014	0.306	2.874	0	2.873	0.017	0.008	7.184	2.880
23	1	8.398	1.675	0.274	0.104	0.028	1.682	0	1.680	0.066	0.027	7.605	1.693
24	1	11.746	1.817	0.216	0.207	0.053	1.731	0	1.726	0.069	0.021	9.124	1.776
25	1	-3.020	1.507	0.254	-0.042	-0.016	1.546	0	1.546	0.036	0.026	8.863	1.535
26	1	59.775	2.246	0.331	-0.143	0.207	2.229	0	2.231	0.010	0.007	7.572	2.231
27	1	136.083	2.199	0.278	-0.216	0.163	2.172	0	2.176	0.024	0.021	7.881	2.186
28	1	12.848	1.668	0.302	-0.208	0.029	1.708	0	1.705	0.014	0.007	8.186	1.701
29	1	66.226	2.178	0.231	0.165	0.132	2.270	0	2.269	0.037	0.015	8.845	2.241

ExampleSet (30 examples, 0 special attributes, 4 regular attributes)

Filter (30 / 30 examples): all

Row No.	Criterion	Value	Standard...	Variance
1	Robustness	1	?	?
2	Robustness	0.437	?	?
3	Robustness	0.329	?	?
4	Robustness	0.379	?	?
5	Robustness	0.246	?	?
6	Robustness	0.219	?	?
7	Robustness	0.280	?	?
8	Robustness	0.281	?	?
9	Robustness	0.332	?	?
10	Robustness	0.351	?	?
11	Robustness	0.341	?	?
12	Robustness	0.523	?	?
13	Robustness	0.460	?	?
14	Robustness	0.556	?	?
15	Robustness	0.466	?	?
16	Robustness	0.529	?	?
17	Robustness	0.560	?	?
18	Robustness	0.563	?	?
19	Robustness	0.562	?	?
20	Robustness	0.659	?	?
21	Robustness	0.641	?	?
22	Robustness	0.640	?	?
23	Robustness	0.662	?	?
24	Robustness	0.677	?	?
25	Robustness	0.716	?	?
26	Robustness	0.647	?	?
27	Robustness	0.693	?	?
28	Robustness	0.705	?	?
29	Robustness	0.771	?	?

Robustheit

Bei 0 Parametern ist die Unterscheidung zwischen Untgergrund und Signal immer stabil (es gibt ja nur das Label). Ab dem Minimum bei 5 wird die *feature selection* immer stabiler, wie zu erwarten war (ein Attribut 'auf der Kippe' beeinflusst die Stabilität immer weniger, je mehr Attribute es gibt).
Scheinbar ist die Wahl von weniger als 6 Attributen sehr instabil. Wie viele Attribute man für das spezifische Problem behalten sollte hängt von der Weiterverarbeitung ab; mehr sind jedoch (für >5 Attribute) tendenziell stabiler.

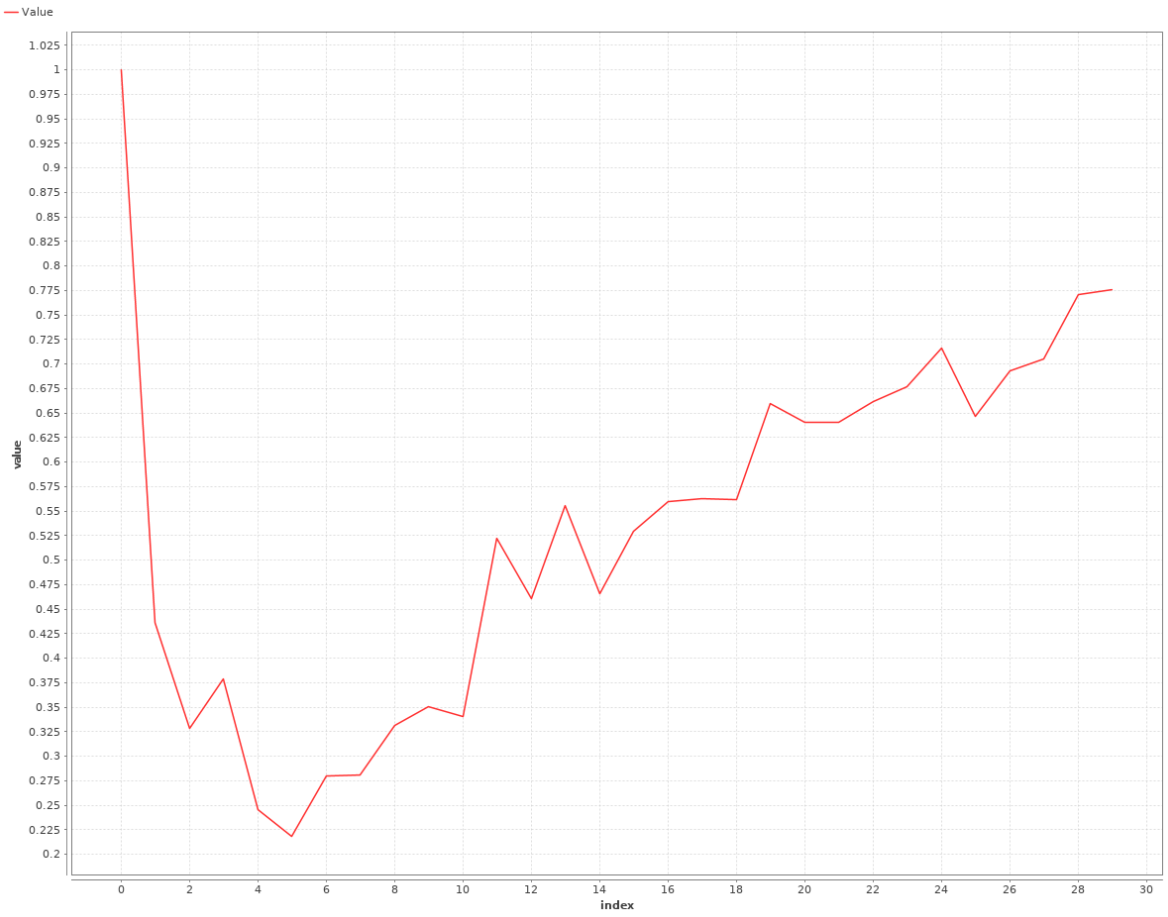

Abbildung 1: Robustheit der MRMR Feature Selection für verschiedene Anzahl von Parametern.