Data Mining sheet2

Christian Ley, Christian Stricker, Jonas Grab

Borders

Histogramme

Negative Borders

Threshold	Dm1.csv	Dm2.csv	Dm3.csv	Dm4.csv
0.4	dm1.csv_threshol	dm1.csv_threshol	dm1.csv_threshol	dm1.csv_threshol
	d0.4_negative	d0.5_negative	d0.6_negative	d0.7_negative
0.5	dm1.csv_threshol	dm1.csv_threshol	dm2.csv_threshol	dm2.csv_threshol
	d0.8_negative	d0.9_negative	d0.4_negative	d0.5_negative
0.6	dm2.csv_threshol	dm2.csv_threshol	dm2.csv_threshol	dm2.csv_threshol
	d0.6_negative	d0.7_negative	d0.8_negative	d0.9_negative
0.7	dm3.csv_threshol	dm3.csv_threshol	dm3.csv_threshol	dm3.csv_threshol
	d0.4_negative	d0.5_negative	d0.6_negative	d0.7_negative
0.8	dm3.csv_threshol	dm3.csv_threshol	dm4.csv_threshol	dm4.csv_threshol
	d0.8_negative	d0.9_negative	d0.4_negative	d0.5_negative
	dm4.csv_threshol	dm4.csv_threshol	dm4.csv_threshol	dm4.csv_threshol
0.9	d0.6_negative	d0.7_negative	d0.8_negative	d0.9_negative

Borders

Histogramme

Positive Borders

Dm1.csv

Threshold

0.4	No. of No.	Single Streets	Street Street	Signal Street.
	dm1.csv_threshol d0.4_positive	dm1.csv_threshol d0.5_positive	dm1.csv_threshol d0.6_positive	dm1.csv_threshol d0.7_positive
0.5	The second secon	Total book	j.	Fig. 1
	dm1.csv_threshol d0.8_positive	dm1.csv_threshol d0.9_positive	dm2.csv_threshol d0.4_positive	dm2.csv_threshol d0.5_positive
0.6	dm2.csv_threshol	dm2.csv_threshol	dm2.csv_threshol	dm2.csv_threshol
	d0.6_positive	d0.7_positive	d0.8_positive	d0.9_positive
0.7		Segar brown		
	dm3.csv_threshol d0.4_positive	dm3.csv_threshol d0.5_positive	dm3.csv_threshol d0.6_positive	dm3.csv_threshol d0.7_positive
0.8	Toward brown			
	dm3.csv_threshol d0.8_positive	dm3.csv_threshol d0.9_positive	dm4.csv_threshol d0.4_positive	dm4.csv_threshol d0.5_positive
0.9		The state of the s	The state of the s	The state of the s
	dm4.csv_threshol d0.6_positive	dm4.csv_threshol d0.7_positive	dm4.csv_threshol d0.8_positive	dm4.csv_threshol d0.9_positive

Dm2.csv

Dm3.csv

Dm4.csv

Vergleich Outputs Task 1 & 2

Der Output von Task 2 (positive/negative Border) sind im Vergleich zum Output aus Task 1 (alle frequent Itemssets) viel kleiner. Durch die positive/negative Border kann man jedoch das Ergebnis aus Task 1 reproduzieren, wodurch beim Berechnen der Borders Speicherplatz gespart wird.

Runtimes:

[["	0.4	0.5	0.6	0.7	0.8	0.9]
['dm1.csv'	0.002815000000000012	0.00109499999999402	0.001100000000000989	0.0014569999999984873	0.0014409999999998035	0.0013019999999990262]
['dm2.csv'	0.0014889999999994075	0.001452000000000453	0.001278999999966994	0.001274999999999693	0.0007870000000025357	0.001353999999991892]
['dm3.csv'	0.008100000000005991	0.0032840000000007308	0.0048160000000052605	0.0031370000000094933	0.001385999999965567	0.0014070000000003802]
['dm4.csv'	20.2912880000001	1.9859560000000016	0.5599670000000003	0.10043600000000197	0.053963000000102	0.02220900000003698]