

Counting trees in Random Forests: Predicting symptom severity in psychiatric intake reports - Supplementary

1. Complete results on positive valence severity prediction - iteration over parameters

An overview of results is shown in Table 1. This table shows the individual results for all feature sets with differing parameters. In the first row, experiments
5 are performed with the basic featureset. In the second row, all context cues (family, uncertainty and negation cues) are added. In the third row, we also add bootstrapping and outlier detection.

	Development phase (10-fold CV)		Testing phase	
System	MAE	F-score	MAE	F-score
Random Baseline	50.96 (SD 5.32)	24.44 (SD 5.26)	57.28	30.90
Bag of words	74.65 (SD 5.05)	43.53 (SD 9.15)	75.91	54.31
Bag of UMLS concepts	72.76 (SD 4.42)	42.88 (SD 7.95)	72.88	47.37
DSM+1	74.34 (SD 4.27)	46.97 (SD 7.43)	75.17	51.47
DSM+2	71.9 (SD 3.46)	40.51 (SD 4.84)	76.82	55.21
DSM	67.5 (SD 3.91)	38.79 (SD 8.25)	70.36	46.93
SNOMED+1	72.31 (SD 4.37)	43.66 (SD 7.62)	77.7	58.35
SNOMED	72.86 (SD 3.69)	47.06 (SD 5.93)	76.4	55.44
Added context cues				
Bag of words	76.42 (SD 4.22)	53.0 (SD 8.42)	78.98	59.18
Bag of UMLS concepts	75.49 (SD 3.73)	51.21 (SD 8.6)	79.41	62.42
DSM+1	78.3 (SD 2.65)	57.05 (SD 4.58)	79.52	61.15
DSM+2	76.81 (SD 3.73)	53.62 (SD 5.52)	79.78	61.09
DSM	72.82 (SD 3.22)	47.13 (SD 6.32)	71.91	49.95
SNOMED+1	75.97 (SD 3.5)	52.69 (SD 6.06)	79.73	61.38
SNOMED	74.7 (SD 3.05)	51.12 (SD 5.4)	75.99	56.24
Added Context, Bootstrap and Outlier Detection				
Bag of words	74.1 (SD 3.81)	51.4 (SD 5.45)	79.11	58.48
Bag of UMLS concepts	74.37 (SD 2.19)	50.01 (SD 4.75)	78.51	58.82
DSM+1	78.77 (SD 3.61)	56.7 (SD 6.66)	80.64	63.67
DSM+2	76.18 (SD 4.06)	52.44 (SD 6.81)	79.2	60.37
DSM	73.3 (SD 3.72)	48.89 (SD 6.52)	72.53	50.78
SNOMED+1	74.28 (SD 3.57)	48.32 (SD 6.3)	79.18	60.51
SNOMED	73.55 (SD 3.17)	47.37 (SD 6.05)	75.8	56.91
Differing Systems				
Questionsets	78.01% (SD 2.63)	53.57 (SD 5.24)	79.34	60.46

Table 1: Results achieved (in %) when iterating over the different parameters, i.e. context, and outlier detection and bootstrapping. Models submitted to the challenge are marked in **bold**.