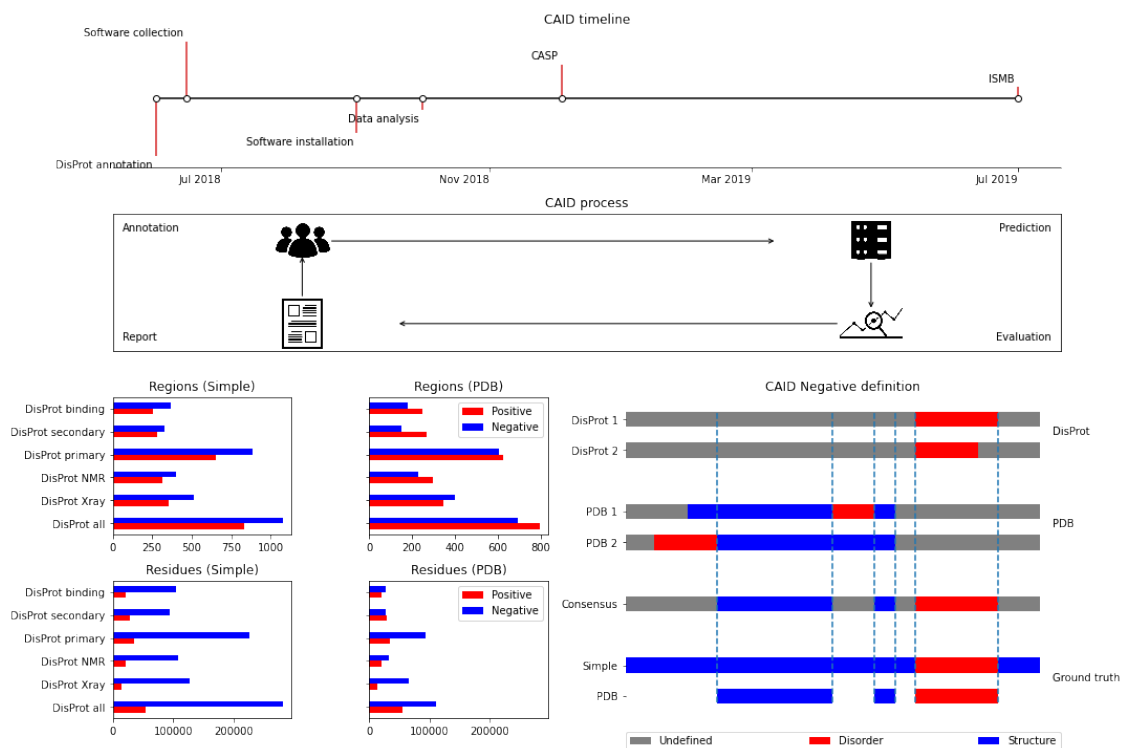


CAID-paper

February 12, 2020

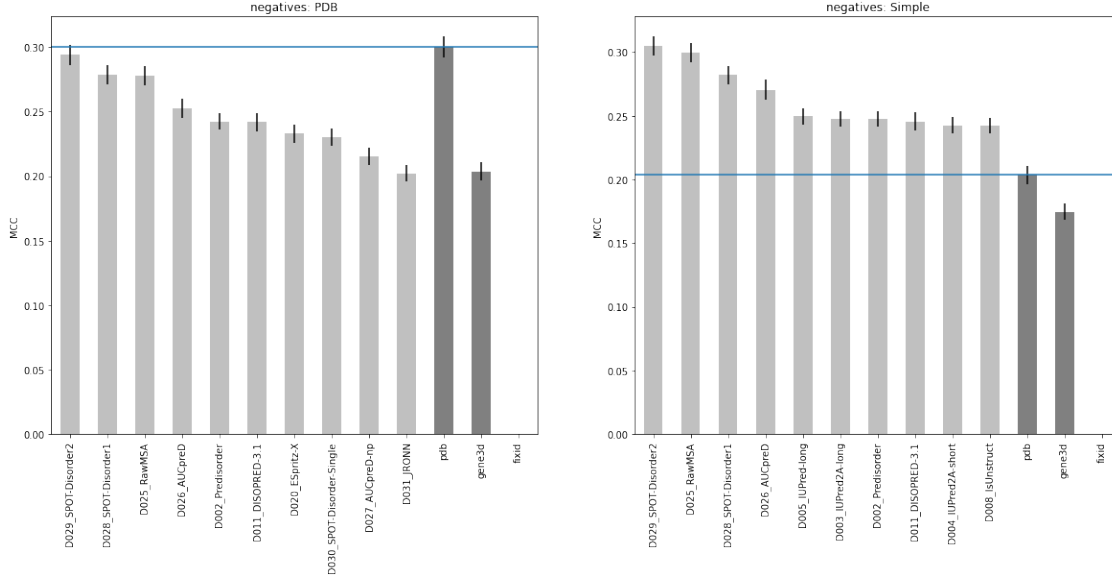
0.1 Display element 1 - CAID

Drafted, refinements needed



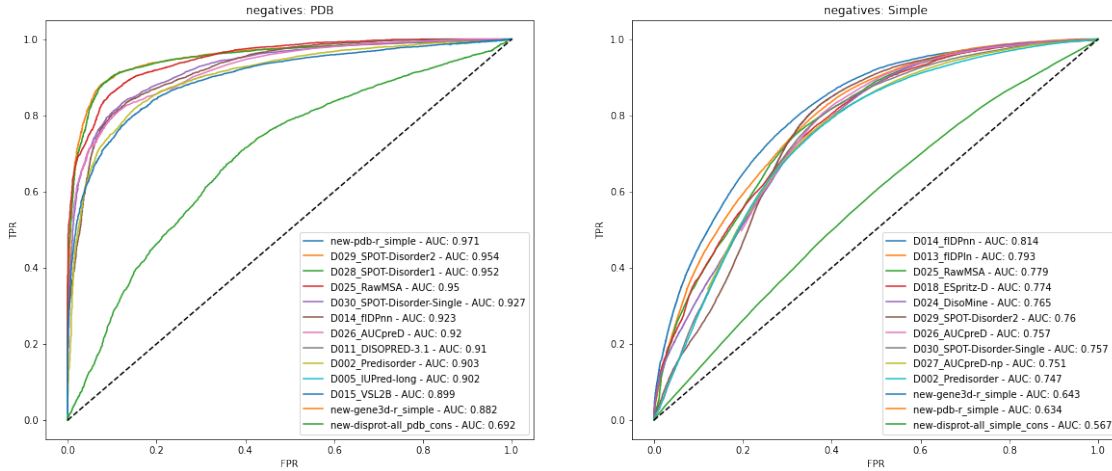
0.2 Display element 2 - MCC

MCC of the first 10 predictors and 3 baselines. MCC is calculated for each target and its mean and standard deviations are plotted as a bar plot with error whiskers.



0.3 Display element 3 - ROC curve

ROC curve of the best 10 ranking methods (ranking is based on ROC AUC) and baselines. X axis is 1-specificity; Y axis is sensitivity.



0.4 Display element 4 - Fully ID table

Performance of predictors (shown as confusion matrix) on their ability to discriminate fully disordered proteins (>95% id content) from non-fully disordered proteins.

Negatives: Simple				
	tn	fp	fn	tp
D016_DisEMBL-HL	529	0	36	4
D032_DFLpred	529	0	40	0
D009_GlobPlot	529	0	40	0
D033_DynaMine	529	0	40	0
D017_DisEMBL-465	529	0	38	2
D006_IUPred-short	528	1	30	10
D004_IUPred2A-short	527	2	30	10
D021_MobiDB-lite	527	2	29	11
D019_ESpritz-N	525	4	28	12
D001_PyHCA	524	5	28	12
D011_DISOPRED-3.1	524	5	34	6
D020_ESpritz-X	523	6	31	9
D005_IUPred-long	523	6	25	15
D003_IUPred2A-long	523	6	27	13
D031_JRONN	523	6	28	12
D030_SPOT-Disorder-Single	522	7	25	15
D027_AUCpreD-np	519	10	30	10
D026_AUCpreD	517	12	23	17
D002_Predisorder	517	12	21	19
D008_IsUnstruct	516	13	24	16
D014_fIDPnn	514	15	14	26
D010_DisPredict-2	514	15	27	13
D025_RawMSA	511	18	14	26
D015_VSL2B	506	23	17	23
D029_SPOT-Disorder2	503	26	19	21
D023_S2D-2	500	29	26	14
D028_SPOT-Disorder1	500	29	18	22
D013_fIDPln	498	31	14	26
D018_ESpritz-D	487	42	19	21
D024_DisoMine	481	48	13	27
D022_S2D-1	477	52	25	15
new-pdb-r_simple	412	117	8	32
new-gene3d-r_simple	394	135	3	37
new-disprot-all_simple_cons	393	136	34	6
D007_FoldUnfold	386	143	9	31

0.5 Display element 5 - Binding

MCC and ROC curve for binding

