

Model\_Name : Model\_ANCOR\_balanced\_ANAPHORE.model

Train Dataset : balanced\_ANAPHORE\_Train.xlsx

Test Dataset : window\_ANAPHORE\_Test.xlsx

{'precision\_score\_pos': 0.20553232550659375, 'recall\_score\_pos': 0.9494799405646359,  
'f1\_score\_pos': 0.3379164463246959, 'precision\_score\_neg': 0.997540865036887,  
'recall\_score\_neg': 0.8481121633255442, 'f1\_score\_neg': 0.916777452805105}

MUC: R=0.974 P=0.240138067061144 F<sub>1</sub>=0.3852848101265823

B<sup>3</sup>: R=0.9920292747837658 P=0.2687463503376037 F<sub>1</sub>=0.42292100468000715

CEAF<sub>m</sub>: R=0.32974051896207585 P=0.33092948717948717 F<sub>1</sub>=0.3303339332133573

CEAF<sub>e</sub>: R=0.13256522254429753 P=0.5679343401737533 F<sub>1</sub>=0.21495614330878815

BLANC: R=0.9780063095684741 P=0.5367818033639943 F<sub>1</sub>=0.5657784473690689

CoNLL-2012 average score: 0.3410539860384592

Model\_Name : Model\_ANCOR\_balanced\_DIRECTE.model

Train Dataset : balanced\_DIRECTE\_Train.xlsx

Test Dataset : window\_DIRECTE\_Test.xlsx

{'precision\_score\_pos': 0.16417910447761194, 'recall\_score\_pos': 0.9506172839506173,  
'f1\_score\_pos': 0.28, 'precision\_score\_neg': 0.999227202472952, 'recall\_score\_neg':  
0.9295470884255931, 'f1\_score\_neg': 0.9631284916201117}

MUC: R=0.9814814814814815 P=0.19099099099099098 F<sub>1</sub>=0.31975867269984914

B<sup>3</sup>: R=0.9964301651048639 P=0.3341916982332532 F<sub>1</sub>=0.5005158839218905

CEAF<sub>m</sub>: R=0.37617135207496655 P=0.37667560321715815 F<sub>1</sub>=0.37642330877428

CEAF<sub>e</sub>: R=0.1796593481901408 P=0.6010592853062825 F<sub>1</sub>=0.2766321048036144

BLANC: R=0.9860363779233123 P=0.537030920093191 F<sub>1</sub>=0.5674381536752174

CoNLL-2012 average score: 0.36563555380845136

Model\_Name : Model\_ANCOR\_balanced\_Full.model

Train Dataset : balanced\_Full\_Train.xlsx

Test Dataset : window\_Full\_Test.xlsx

{'precision\_score\_pos': 0.24341021416803954, 'recall\_score\_pos': 0.9248826291079812,  
'f1\_score\_pos': 0.3853928920769482, 'precision\_score\_neg': 0.9920523222121036,  
'recall\_score\_neg': 0.7653445743118095, 'f1\_score\_neg': 0.8640755696567638}

MUC: R=0.9656804733727811 P=0.35202761000862814 F<sub>1</sub>=0.5159658552007588  
B<sup>3</sup>: R=0.9792180301332337 P=0.22358316960066843 F<sub>1</sub>=0.3640446500315214  
CEAF<sub>m</sub>: R=0.28950317519611507 P=0.29102515959444236 F<sub>1</sub>=0.2902621722846442  
CEAF<sub>e</sub>: R=0.1137015747618428 P=0.6037718404744812 F<sub>1</sub>=0.19136544323720348  
BLANC: R=0.9433639115629608 P=0.5239049062212046 F<sub>1</sub>=0.5404814870582241  
CoNLL-2012 average score: 0.35712531615649457

Model\_Name : Model\_ANCOR\_balanced\_INDIRECTE.model

Train Dataset : balanced\_INDIRECTE\_Train.xlsx

Test Dataset : window\_INDIRECTE\_Test.xlsx

{'precision\_score\_pos': 0.026585312704292874, 'recall\_score\_pos': 0.9312977099236641,  
'f1\_score\_pos': 0.051694915254237285, 'precision\_score\_neg': 0.9992710189535072,  
'recall\_score\_neg': 0.7341704356105689, 'f1\_score\_neg': 0.8464493996569469}

MUC: R=0.9672131147540983 P=0.05175438596491228 F<sub>1</sub>=0.09825145711906745

B<sup>3</sup>: R=0.9975560081466395 P=0.07754737411805192 F<sub>1</sub>=0.14390773993195816

CEAF<sub>m</sub>: R=0.0989816700610998 P=0.09914320685434516 F<sub>1</sub>=0.09906237260497351

CEAF<sub>e</sub>: R=0.026444546657574307 P=0.3607902184334553 F<sub>1</sub>=0.04927725826846714

BLANC: R=0.9736254947303481 P=0.5011001559876569 F<sub>1</sub>=0.49646213163180414

CoNLL-2012 average score: 0.09714548510649758

Model\_Name : Model\_ANCOR\_representative\_ANAPHORE.model

Train Dataset : representative\_ANAPHORE\_Train.xlsx

Test Dataset : window\_ANAPHORE\_Test.xlsx

{'precision\_score\_pos': 0.38080495356037153, 'recall\_score\_pos': 0.9138187221396731,  
'f1\_score\_pos': 0.5375874125874126, 'precision\_score\_neg': 0.9962140992167102,  
'recall\_score\_neg': 0.9385069487147952, 'f1\_score\_neg': 0.9664999050091824}

MUC: R=0.936 P=0.4073107049608355 F<sub>1</sub>=0.5676167374166161

B<sup>3</sup>: R=0.9668783068783068 P=0.4370399834917009 F<sub>1</sub>=0.6019787364764805

CEAF<sub>m</sub>: R=0.5028571428571429 P=0.5132858068697342 F<sub>1</sub>=0.5080179602309173

CEAF<sub>e</sub>: R=0.2563599159618311 P=0.699459161570986 F<sub>1</sub>=0.3752034168263696

BLANC: R=0.9400014618159629 P=0.6005105422588264 F<sub>1</sub>=0.6541826947814775

CoNLL-2012 average score: 0.5149329635731554

Model\_Name : Model\_ANCOR\_representative\_DIRECTE.model

Train Dataset : representative\_DIRECTE\_Train.xlsx

Test Dataset : window\_DIRECTE\_Test.xlsx

{'precision\_score\_pos': 0.375, 'recall\_score\_pos': 0.9135802469135802, 'f1\_score\_pos': 0.5317365269461077, 'precision\_score\_neg': 0.9987150461971487, 'recall\_score\_neg': 0.9778336927869639, 'f1\_score\_neg': 0.98816406841229}

MUC: R=0.9212962962962963 P=0.40447154471544716 F<sub>1</sub>=0.5621468926553672

B<sup>3</sup>: R=0.966646390916464 P=0.5959941884599419 F<sub>1</sub>=0.7373616670212121

CEAF<sub>m</sub>: R=0.6094890510948905 P=0.6239103362391034 F<sub>1</sub>=0.6166153846153846

CEAF<sub>e</sub>: R=0.39146164616461643 P=0.7627837864172269 F<sub>1</sub>=0.5173953273189914

BLANC: R=0.9279735724979613 P=0.6481676182174997 F<sub>1</sub>=0.7110017236434789

CoNLL-2012 average score: 0.6056346289985236

Model\_Name : Model\_ANCOR\_representative\_Full.model

Train Dataset : representative\_Full\_Train.xlsx

Test Dataset : window\_Full\_Test.xlsx

{'precision\_score\_pos': 0.538135593220339, 'recall\_score\_pos': 0.7949921752738655,  
'f1\_score\_pos': 0.6418193303853443, 'precision\_score\_neg': 0.9825878912740081,  
'recall\_score\_neg': 0.9443060611866897, 'f1\_score\_neg': 0.9630667014069828}

MUC: R=0.8710059171597633 P=0.565284178187404 F<sub>1</sub>=0.6856078248719143  
B<sup>3</sup>: R=0.8956161697770648 P=0.5643373893373893 F<sub>1</sub>=0.6923914640228871  
CEAF<sub>m</sub>: R=0.5855087906233352 P=0.6188063063063063 F<sub>1</sub>=0.6016972351491925  
CEAF<sub>e</sub>: R=0.34741262837061176 P=0.7563920516423446 F<sub>1</sub>=0.47613523569518107  
BLANC: R=0.8204225635770281 P=0.6761141312353848 F<sub>1</sub>=0.7113711568655718  
CoNLL-2012 average score: 0.6180448415299942

Model\_Name : Model\_ANCOR\_representative\_INDIRECTE.model

Train Dataset : representative\_INDIRECTE\_Train.xlsx

Test Dataset : window\_INDIRECTE\_Test.xlsx

{'precision\_score\_pos': 0.050728277247614265, 'recall\_score\_pos': 0.7709923664122137,  
'f1\_score\_pos': 0.09519321394910461, 'precision\_score\_neg': 0.9979925053533191,  
'recall\_score\_neg': 0.8875267793382527, 'f1\_score\_neg': 0.9395237495275293}

MUC: R=0.8442622950819673 P=0.06675307841866494 F<sub>1</sub>=0.12372372372372373

B<sup>3</sup>: R=0.9854187860291624 P=0.22841554261775257 F<sub>1</sub>=0.37086604226704445

CEAF<sub>m</sub>: R=0.24974567650050863 P=0.25218284540318436 F<sub>1</sub>=0.2509583439815998

CEAF<sub>e</sub>: R=0.10723978041016852 P=0.4894805818721553 F<sub>1</sub>=0.1759343016693512

BLANC: R=0.9022048516709338 P=0.5081737324256863 F<sub>1</sub>=0.5103027308985852

CoNLL-2012 average score: 0.22350802255337313



Model\_Name : Model\_ANCOR\_window\_ANAPHORE.model

Train Dataset : window\_ANAPHORE\_Train.xlsx

Test Dataset : window\_ANAPHORE\_Test.xlsx

{'precision\_score\_pos': 0.7664670658682635, 'recall\_score\_pos': 0.5705794947994056,  
'f1\_score\_pos': 0.6541737649063032, 'precision\_score\_neg': 0.982414506510892,  
'recall\_score\_neg': 0.992805312999631, 'f1\_score\_neg': 0.987582578908735}

MUC: R=0.654 P=0.7712264150943396 F<sub>1</sub>=0.7077922077922079

B<sup>3</sup>: R=0.683632571635836 P=0.8070337178682713 F<sub>1</sub>=0.7402254144039875

CEAF<sub>m</sub>: R=0.6126224156692056 P=0.8243045387994143 F<sub>1</sub>=0.7028714107365793

CEAF<sub>e</sub>: R=0.5335518003723877 P=0.8631590901777237 F<sub>1</sub>=0.6594637296638066

BLANC: R=0.5659224734579851 P=0.8082865164126218 F<sub>1</sub>=0.6544058787856044

CoNLL-2012 average score: 0.702493783953334

Model\_Name : Model\_ANCOR\_window\_DIRECTE.model

Train Dataset : window\_DIRECTE\_Train.xlsx

Test Dataset : window\_DIRECTE\_Test.xlsx

{'precision\_score\_pos': 0.770949720670391, 'recall\_score\_pos': 0.5679012345679012,  
'f1\_score\_pos': 0.6540284360189572, 'precision\_score\_neg': 0.9937335879684889,  
'recall\_score\_neg': 0.9975437335250419, 'f1\_score\_neg': 0.99563501554652}

MUC: R=0.5972222222222222 P=0.7678571428571429 F<sub>1</sub>=0.671875  
B<sup>3</sup>: R=0.6560539752005835 P=0.868895966029724 F<sub>1</sub>=0.7476214623670197  
CEAF<sub>m</sub>: R=0.5973741794310722 P=0.8694267515923567 F<sub>1</sub>=0.708171206225681  
CEAF<sub>e</sub>: R=0.5386880063228611 P=0.8892041748206132 F<sub>1</sub>=0.6709240802264058  
BLANC: R=0.5049848937789225 P=0.8736704796835795 F<sub>1</sub>=0.6333274899293821  
CoNLL-2012 average score: 0.6968068475311419

Model\_Name : Model\_ANCOR\_window\_Full.model

Train Dataset : window\_Full\_Train.xlsx

Test Dataset : window\_Full\_Test.xlsx

{'precision\_score\_pos': 0.7782805429864253, 'recall\_score\_pos': 0.5383411580594679,  
'f1\_score\_pos': 0.6364477335800185, 'precision\_score\_neg': 0.963242165597159,  
'recall\_score\_neg': 0.9874816376061826, 'f1\_score\_neg': 0.9752113031411631}

MUC: R=0.6437869822485207 P=0.7838616714697406 F<sub>1</sub>=0.7069525666016894

B<sup>3</sup>: R=0.6455773350064008 P=0.8085921555080433 F<sub>1</sub>=0.7179476288906211

CEAF<sub>m</sub>: R=0.5910034602076124 P=0.7981308411214953 F<sub>1</sub>=0.6791252485089463

CEAF<sub>e</sub>: R=0.5098983101452675 P=0.8136675161892567 F<sub>1</sub>=0.6269241518179519

BLANC: R=0.5248421504410351 P=0.8102135460588703 F<sub>1</sub>=0.631366888187501

CoNLL-2012 average score: 0.6839414491034208

Model\_Name : Model\_ANCOR\_window\_INDIRECTE.model

Train Dataset : window\_INDIRECTE\_Train.xlsx

Test Dataset : window\_INDIRECTE\_Test.xlsx

{'precision\_score\_pos': 0.7777777777777778, 'recall\_score\_pos': 0.05343511450381679,  
'f1\_score\_pos': 0.09999999999999999, 'precision\_score\_neg': 0.9926739926739927,  
'recall\_score\_neg': 0.9998809807188764, 'f1\_score\_neg': 0.9962644530091906}

MUC: R=0.05737704918032787 P=0.7777777777777778 F<sub>1</sub>=0.1068702290076336

B<sup>3</sup>: R=0.07415254237288135 P=0.8888888888888888 F<sub>1</sub>=0.1368858469811782

CEAF<sub>m</sub>: R=0.06779661016949153 P=0.8888888888888888 F<sub>1</sub>=0.12598425196850394

CEAF<sub>e</sub>: R=0.07309941520467836 P=0.9259259259259258 F<sub>1</sub>=0.13550135501355012

BLANC: R=0.028924768710557783 P=0.8888888888888888 F<sub>1</sub>=0.05448664736336261

CoNLL-2012 average score: 0.12641914366745396

