

Table 1: Cross-validation accuracy. Ibsen corpus

Features	SVC	Random Forest	Logistic Regression
bigrams	0.646739	0.892935	0.987319
bigrams pos	0.787319	0.792029	0.817391
bigrams pos punct	0.974638	0.995833	1.000000
bigrams pos punct tf	0.974638	1.000000	0.544565
bigrams pos punct tfidf	0.974638	1.000000	0.544565
bigrams pos tf	0.804529	0.974094	0.544565
bigrams pos tfidf	0.804529	0.969928	0.544565
bigrams punct	0.838043	0.970471	1.000000
bigrams punct tf	0.842210	0.982971	0.544565
bigrams punct tfidf	0.842210	0.982790	0.544565
bigrams tf	0.715036	0.973913	0.544565
bigrams tfidf	0.715036	0.969928	0.544565
cohesive	0.694384	0.761232	0.740580
cohesive punctuation	0.838043	0.893659	0.919203
cohesive punctuation tf	0.838043	0.982790	0.544565
cohesive punctuation tfidf	0.838043	0.987138	0.544565
cohesive tf	0.677536	0.960870	0.544565
cohesive tfidf	0.677536	0.952174	0.544565
syntactic n2	0.803442	0.978623	0.995652
syntactic n2 tf	0.709239	0.991486	0.539855
syntactic n2 tfidf	0.709239	0.991304	0.539855
syntactic n3	0.485145	0.952717	0.982971
syntactic n3 tf	0.485145	0.982971	0.539855
syntactic n3 tfidf	0.485145	0.974638	0.539855
trigrams	0.540399	0.651449	0.872283
trigrams pos	0.764855	0.732428	0.817391
trigrams pos punct	0.961232	0.983333	0.995833
trigrams pos punct tf	0.969746	0.995652	0.544565
trigrams pos punct tfidf	0.969746	0.995652	0.544565
trigrams pos tf	0.808333	0.982609	0.544565
trigrams pos tfidf	0.808333	0.982609	0.544565
trigrams punct	0.561594	0.931884	0.970290
trigrams punct tf	0.561594	0.936051	0.544565
trigrams punct tfidf	0.561594	0.940580	0.544565
trigrams tf	0.544565	0.812138	0.544565
trigrams tfidf	0.544565	0.834239	0.544565
unigrams	0.846558	0.949094	0.982790
unigrams punct	0.863587	0.982790	1.000000
unigrams punct tf	0.893116	0.991304	0.544565
unigrams punct tfidf	0.893116	0.991304	0.544565
unigrams tf	0.867935	0.982609	0.544565
unigrams tfidf	0.867935	0.986957	0.544565

Table 2: Cross-validation accuracy. Quixote corpus

Features	SVC	Random Forest	Logistic Regression
bigrams	0.629587	0.970982	0.984211
bigrams pos	0.825462	0.616074	0.862233
bigrams pos punct	0.986771	0.981508	0.992034
bigrams pos punct tf	0.986771	0.973613	0.269915
bigrams pos punct tfidf	0.986771	0.978876	0.269915
bigrams pos tf	0.825462	0.745590	0.269915
bigrams pos tfidf	0.825462	0.700498	0.269915
bigrams punct	0.812091	1.000000	0.986771
bigrams punct tf	0.812091	0.992105	0.269915
bigrams punct tfidf	0.812091	0.997297	0.269915
bigrams tf	0.632290	0.973684	0.269915
bigrams tfidf	0.632290	0.994737	0.269915
cohesive	0.862233	0.918065	0.923257
cohesive punctuation	0.965789	0.960313	0.978805
cohesive punctuation tf	0.963158	0.992105	0.269915
cohesive punctuation tfidf	0.963158	0.984211	0.269915
cohesive tf	0.870128	0.984211	0.269915
cohesive tfidf	0.870128	0.984211	0.269915
syntactic n2	0.674680	0.986842	0.973613
syntactic n2 tf	0.666785	0.970982	0.269915
syntactic n2 tfidf	0.666785	0.994737	0.269915
syntactic n3	0.216856	0.939331	0.920697
syntactic n3 tf	0.211451	0.926245	0.269915
syntactic n3 tfidf	0.211451	0.952489	0.269915
trigrams	0.124253	0.936629	0.889047
trigrams pos	0.841323	0.658321	0.867639
trigrams pos punct	0.994737	0.973613	0.994737
trigrams pos punct tf	0.994737	0.976316	0.269915
trigrams pos punct tfidf	0.994737	0.981508	0.269915
trigrams pos tf	0.849218	0.785277	0.269915
trigrams pos tfidf	0.849218	0.795306	0.269915
trigrams punct	0.471266	0.978876	0.976316
trigrams punct tf	0.476529	0.984211	0.269915
trigrams punct tfidf	0.476529	0.986771	0.269915
trigrams tf	0.126956	0.941892	0.269915
trigrams tfidf	0.126956	0.904836	0.269915
unigrams	0.960384	0.992034	0.989474
unigrams punct	0.960384	0.989474	0.986700
unigrams punct tf	0.960384	0.994737	0.269915
unigrams punct tfidf	0.960384	0.994737	0.269915
unigrams tf	0.960384	1.000000	0.269915
unigrams tfidf	0.960384	0.992105	0.269915