Irony Detection in Twitter with Imbalanced Scenarios

Abstract

Irony detection is a not trivial problem and can help to improve natural language processing tasks. When dealing with social media data in real scenarios, an important issue to address is data skew, i.e. the imbalance between ironic and non-ironic samples available. In this work, the main objective is to address irony detection in Twitter considering various degrees of imbalance distribution between classes. We rely on the emotIDM irony detection model. We evaluated it against both benchmark corpora and skewed Twitter datasets collected to simulate a realistic distribution of ironic tweets. We carry out a set of classification experiments aimed to determine the impact of class imbalance for detecting irony and we evaluate the performance of irony detection when different scenarios are considered. We experiment with a set of classifiers applying class imbalance techniques to compensate class distribution. Our results indicate that by using such techniques it is possible to improve the performance of irony detection in imbalanced scenarios.

AUC

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8914	0.8917	0.8662	0.9172	0.8780
	NB	0.8198	0.8341	0.8203	0.8198	0.8595
	RF	0.9724	0.9693	0.9722	0.9711	0.9718
	SVM	0.8601	0.8810	0.8819	0.8511	0.8821
TwIronyBarbieri2014	J48	0.8807	0.8648	0.8490	0.9074	0.8658
	NB	0.8522	0.8490	0.8519	0.8522	0.8502
	RF	0.9598	0.9569	0.9618	0.9603	0.9609
	SVM	0.8250	0.8435	0.8447	0.7401	0.8476
TwMohammad2015	J48	0.5344	0.5710	0.5168	0.5555	0.5355
	NB	0.6364	0.6424	0.6330	0.6364	0.6099
	RF	0.6561	0.6479	0.6397	0.6497	0.6521
	SVM	0.5075	0.6046	0.5984	0.5000	0.5680
TwSarcasmBarbieri2014	J48	0.9072	0.9020	0.8759	0.9304	0.8936
	NB	0.9097	0.9113	0.9099	0.9097	0.8864
	RF	0.9755	0.9732	0.9766	0.9754	0.9758
	SVM	0.8784	0.8943	0.8969	0.8599	0.8981
TwRiloff2013	J48	0.6194	0.6565	0.6067	0.7319	0.6006
	NB	0.7471	0.7446	0.7478	0.7471	0.7331
	RF	0.8006	0.8020	0.8125	0.8078	0.8063
	SVM	0.5485	0.7326	0.7339	0.6230	0.7317
TwPtáček2014	J48	0.7127	0.7376	0.6843	0.7918	0.7234
	NB	0.7390	0.7378	0.7386	0.7390	0.7660
	RF	0.8885	0.8844	0.8886	0.8778	0.8878
	SVM	0.6647	0.7450	0.7448	0.5000	0.7450

TABLE 1: Benchmark Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.5079	0.6385	0.5469	0.7199	0.6562
	NB	0.6620	0.6605	0.6622	0.6620	0.6416
	RF	0.7590	0.7805	0.7809	0.7956	0.7748
	SVM	0.5000	0.6640	0.6759	0.5000	0.6646
TwImbData09112016	J48	0.7749	0.7863	0.6383	0.8597	0.7226
	NB	0.7815	0.7759	0.7809	0.7815	0.8164
	RF	0.9105	0.9194	0.9235	0.9260	0.9203
	SVM	0.5000	0.8114	0.8315	0.7090	0.8264
TwImbData10112016	J48	0.7440	0.8245	0.6766	0.8882	0.7405
	NB	0.7973	0.7948	0.7970	0.7973	0.8492
	RF	0.9231	0.9267	0.9346	0.9350	0.9320
	SVM	0.5000	0.8360	0.8593	0.7052	0.8529
TwImbData11112016	J48	0.7937	0.7882	0.6592	0.8701	0.6856
	NB	0.7977	0.7965	0.7978	0.7977	0.8380
	RF	0.9024	0.9081	0.9188	0.9211	0.9089
	SVM	0.5000	0.8134	0.8477	0.6563	0.8310
TwImbData12112016	J48	0.7784	0.8174	0.6549	0.8596	0.7226
	NB	0.7961	0.7933	0.7961	0.7960	0.8224
	RF	0.8937	0.9118	0.9206	0.9225	0.9103
	SVM	0.5000	0.8227	0.8427	0.7043	0.8313
TwImbData13112016	J48	0.7734	0.7945	0.6573	0.8628	0.7171
	NB	0.7885	0.7910	0.7888	0.7885	0.8253
	RF	0.9074	0.9151	0.9239	0.9241	0.9186
	SVM	0.5000	0.8150	0.8431	0.6963	0.8313
TwImbData14112016	J48	0.7662	0.7908	0.6495	0.8675	0.7154
	NB	0.7875	0.7820	0.7876	0.7875	0.8145
	RF	0.9116	0.9129	0.9226	0.9203	0.9183
	SVM	0.5000	0.8196	0.8399	0.6996	0.8323
TwImbData15112016	J48	0.7785	0.7907	0.6292	0.8608	0.6821
	NB	0.7884	0.7851	0.7883	0.7884	0.8234
	RF	0.8973	0.9059	0.9189	0.9142	0.9165
	SVM	0.5000	0.8069	0.8315	0.6783	0.8272
TwImbData16112016	J48	0.7820	0.7860	0.6553	0.8652	0.7175
	NB	0.7901	0.7848	0.7894	0.7901	0.8026
	RF	0.9036	0.9109	0.9212	0.9180	0.9152
	SVM	0.5000	0.8149	0.8442	0.7024	0.8332
TwImbData17112016	J48	0.7435	0.7936	0.6443	0.8680	0.7236
	NB	0.7917	0.7878	0.7917	0.7917	0.8372
	RF	0.9003	0.9084	0.9127	0.9179	0.9184
	SVM	0.5000	0.8146	0.8405	0.6792	0.8342
TwImbData18112016	J48	0.7484	0.7803	0.6393	0.8505	0.7068
	NB	0.7874	0.7904	0.7871	0.7874	0.8319
	RF	0.8859	0.9027	0.9109	0.9135	0.9080
	SVM	0.5000	0.8116	0.8351	0.6981	0.8274

TABLE 2: New Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8914	0.8917	0.8662	0.9172	0.8780
	NB	0.8198	0.8341	0.8203	0.8198	0.8595
	RF	0.9724	0.9693	0.9722	0.9706	0.9718
	SVM	0.8601	0.8810	0.8819	0.8511	0.8821
TwIronyBarbieri2014	J48	0.8807	0.8648	0.8490	0.9074	0.8658
	NB	0.8522	0.8490	0.8519	0.8522	0.8503
	RF	0.8972	0.9239	0.9400	0.8309	0.9272
	SVM	0.8250	0.8435	0.8447	0.7401	0.8476
TwMohammad2015	J48	0.5344	0.5710	0.5168	0.5555	0.5355
	NB	0.6364	0.6424	0.6330	0.6364	0.6101
	RF	0.6531	0.6478	0.6359	0.6221	0.6509
	SVM	0.5075	0.6046	0.5984	0.5000	0.5680
TwSarcasmBarbieri2014	J48	0.9072	0.9020	0.8759	0.9304	0.8936
	NB	0.9097	0.9113	0.9099	0.9097	0.8866
	RF	0.9145	0.9371	0.9301	0.8831	0.9292
	SVM	0.8784	0.8943	0.8969	0.8599	0.8981
TwRiloff2013	J48	0.6194	0.6565	0.6067	0.7319	0.6006
	NB	0.7471	0.7446	0.7478	0.7471	0.7325
	RF	0.7586	0.7925	0.7866	0.6654	0.7913
	SVM	0.5485	0.7326	0.7339	0.6230	0.7317
TwPtáček2014	J48	0.2873	0.7376	0.6843	0.2082	0.7234
	NB	0.2610	0.7378	0.7386	0.2610	0.7659
	RF	0.1823	0.8354	0.8379	0.1222	0.8443
	SVM	0.3353	0.7450	0.7448	0.5000	0.7450

TABLE 3: Benchmark Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.5079	0.6385	0.5469	0.7199	0.6562
	NB	0.6620	0.6605	0.6622	0.6620	0.6415
	RF	0.7515	0.7753	0.7809	0.5125	0.7681
	SVM	0.5000	0.6640	0.6759	0.5000	0.6646
TwImbData09112016	J48	0.7749	0.7863	0.6383	0.8597	0.7226
	NB	0.7814	0.7758	0.7807	0.7814	0.8105
	RF	0.7997	0.9100	0.8904	0.5237	0.8454
	SVM	0.5000	0.8114	0.8315	0.7090	0.8264
TwImbData10112016	J48	0.7440	0.8245	0.6766	0.8882	0.7405
	NB	0.7973	0.7947	0.7969	0.7973	0.8435
	RF	0.8426	0.9159	0.9170	0.5223	0.8663
	SVM	0.5000	0.8360	0.8593	0.7052	0.8529
TwImbData11112016	J48	0.7937	0.7882	0.6592	0.8701	0.6856
	NB	0.7975	0.7965	0.7977	0.7976	0.8327
	RF	0.8952	0.8964	0.9188	0.5285	0.8703
	SVM	0.5000	0.8134	0.8477	0.6563	0.8310
TwImbData12112016	J48	0.7784	0.8174	0.6549	0.8596	0.7226
	NB	0.7960	0.7932	0.7960	0.7960	0.8166
	RF	0.8350	0.9054	0.9052	0.5217	0.8727
	SVM	0.5000	0.8227	0.8427	0.7043	0.8313
TwImbData13112016	J48	0.7734	0.7945	0.6573	0.8628	0.7171
	NB	0.7884	0.7909	0.7887	0.7884	0.8214
	RF	0.8780	0.9101	0.9192	0.5244	0.8956
	SVM	0.5000	0.8150	0.8431	0.6963	0.8313
TwImbData14112016	J48	0.7662	0.7908	0.6495	0.8675	0.7154
	NB	0.7872	0.7818	0.7873	0.7872	0.8053
	RF	0.8973	0.9075	0.9226	0.5306	0.8863
	SVM	0.5000	0.8196	0.8399	0.6996	0.8323
TwImbData15112016	J48	0.7785	0.7907	0.6292	0.8608	0.6821
	NB	0.7883	0.7849	0.7882	0.7883	0.8166
	RF	0.8682	0.8988	0.9146	0.5248	0.8798
	SVM	0.5000	0.8069	0.8315	0.6783	0.8272
TwImbData16112016	J48	0.7820	0.7860	0.6553	0.8652	0.7175
	NB	0.7899	0.7845	0.7892	0.7899	0.7942
	RF	0.8607	0.9015	0.9201	0.5255	0.8997
	SVM	0.5000	0.8149	0.8442	0.7024	0.8332
TwImbData17112016	J48	0.7435	0.7936	0.6443	0.8680	0.7236
	NB	0.7914	0.7876	0.7915	0.7914	0.8313
	RF	0.8678	0.9008	0.9006	0.5153	0.8855
	SVM	0.5000	0.8146	0.8405	0.6792	0.8342
TwImbData18112016	J48	0.7484	0.7803	0.6393	0.8505	0.7068
	NB	0.7872	0.7902	0.7869	0.7872	0.8254
	RF	0.8607	0.8944	0.8840	0.5186	0.8954
	SVM	0.5000	0.8116	0.8351	0.6981	0.8274

TABLE 4: New Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8914	0.8917	0.8662	0.9172	0.8780
	NB	0.8198	0.8341	0.8203	0.8198	0.8595
	RF	0.9724	0.9693	0.9722	0.9711	0.9718
	SVM	0.8601	0.8810	0.8819	0.8511	0.8821
TwIronyBarbieri2014	J48	0.8807	0.8648	0.8490	0.9074	0.8658
	NB	0.8522	0.8490	0.8519	0.8522	0.8502
	RF	0.9598	0.9569	0.9618	0.9603	0.9609
	SVM	0.8250	0.8435	0.8447	0.7401	0.8476
TwMohammad2015	J48	0.5344	0.5710	0.5168	0.5555	0.5355
	NB	0.6364	0.6424	0.6330	0.6364	0.6099
	RF	0.6561	0.6479	0.6397	0.6497	0.6521
	SVM	0.5075	0.6046	0.5984	0.5000	0.5680
TwSarcasmBarbieri2014	J48	0.9072	0.9020	0.8759	0.9304	0.8936
	NB	0.9097	0.9113	0.9099	0.9097	0.8864
	RF	0.9755	0.9732	0.9766	0.9754	0.9758
	SVM	0.8784	0.8943	0.8969	0.8599	0.8981
TwRiloff2013	J48	0.6194	0.6565	0.6067	0.7319	0.6006
	NB	0.7471	0.7446	0.7478	0.7471	0.7331
	RF	0.8006	0.8020	0.8125	0.8078	0.8063
	SVM	0.5485	0.7326	0.7339	0.6230	0.7317
TwPtáček2014	J48	0.7127	0.7376	0.6843	0.7918	0.7234
	NB	0.7390	0.7378	0.7386	0.7390	0.7660
	RF	0.8885	0.8844	0.8886	0.8778	0.8878
	SVM	0.6647	0.7450	0.7448	0.5000	0.7450

TABLE 5: Benchmark Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.5079	0.6385	0.5469	0.7199	0.6562
	NB	0.6620	0.6605	0.6622	0.6620	0.6416
	RF	0.7590	0.7805	0.7809	0.7956	0.7748
	SVM	0.5000	0.6640	0.6759	0.5000	0.6646
TwImbData09112016	J48	0.7749	0.7863	0.6383	0.8597	0.7226
	NB	0.7815	0.7759	0.7809	0.7815	0.8164
	RF	0.9105	0.9194	0.9235	0.9260	0.9203
	SVM	0.5000	0.8114	0.8315	0.7090	0.8264
TwImbData10112016	J48	0.7440	0.8245	0.6766	0.8882	0.7405
	NB	0.7973	0.7948	0.7970	0.7973	0.8492
	RF	0.9231	0.9267	0.9346	0.9350	0.9320
	SVM	0.5000	0.8360	0.8593	0.7052	0.8529
TwImbData11112016	J48	0.7937	0.7882	0.6592	0.8701	0.6856
	NB	0.7977	0.7965	0.7978	0.7977	0.8380
	RF	0.9024	0.9081	0.9188	0.9211	0.9089
	SVM	0.5000	0.8134	0.8477	0.6563	0.8310
TwImbData12112016	J48	0.7784	0.8174	0.6549	0.8596	0.7226
	NB	0.7961	0.7933	0.7961	0.7960	0.8224
	RF	0.8937	0.9118	0.9206	0.9225	0.9103
	SVM	0.5000	0.8227	0.8427	0.7043	0.8313
TwImbData13112016	J48	0.7734	0.7945	0.6573	0.8628	0.7171
	NB	0.7885	0.7910	0.7888	0.7885	0.8253
	RF	0.9074	0.9151	0.9239	0.9241	0.9186
	SVM	0.5000	0.8150	0.8431	0.6963	0.8313
TwImbData14112016	J48	0.7662	0.7908	0.6495	0.8675	0.7154
	NB	0.7875	0.7820	0.7876	0.7875	0.8145
	RF	0.9116	0.9129	0.9226	0.9203	0.9183
	SVM	0.5000	0.8196	0.8399	0.6996	0.8323
TwImbData15112016	J48	0.7785	0.7907	0.6292	0.8608	0.6821
	NB	0.7884	0.7851	0.7883	0.7884	0.8234
	RF	0.8973	0.9059	0.9189	0.9142	0.9165
	SVM	0.5000	0.8069	0.8315	0.6783	0.8272
TwImbData16112016	J48	0.7820	0.7860	0.6553	0.8652	0.7175
	NB	0.7901	0.7848	0.7894	0.7901	0.8026
	RF	0.9036	0.9109	0.9212	0.9180	0.9152
	SVM	0.5000	0.8149	0.8442	0.7024	0.8332
TwImbData17112016	J48	0.7435	0.7936	0.6443	0.8680	0.7236
	NB	0.7917	0.7878	0.7917	0.7917	0.8372
	RF	0.9003	0.9084	0.9127	0.9179	0.9184
	SVM	0.5000	0.8146	0.8405	0.6792	0.8342
TwImbData18112016	J48	0.7484	0.7803	0.6393	0.8505	0.7068
	NB	0.7874	0.7904	0.7871	0.7874	0.8319
	RF	0.8859	0.9027	0.9109	0.9135	0.9080
	SVM	0.5000	0.8116	0.8351	0.6981	0.8274

TABLE 6: New Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8914	0.8917	0.8662	0.9172	0.8780
	NB	0.8198	0.8341	0.8203	0.8198	0.8595
	RF	0.9724	0.9693	0.9722	0.9711	0.9718
	SVM	0.8601	0.8810	0.8819	0.8511	0.8821
TwIronyBarbieri2014	J48	0.8807	0.8648	0.8490	0.9074	0.8658
	NB	0.8522	0.8490	0.8519	0.8522	0.8502
	RF	0.9598	0.9569	0.9618	0.9603	0.9609
	SVM	0.8250	0.8435	0.8447	0.7401	0.8476
TwMohammad2015	J48	0.5344	0.5710	0.5168	0.5555	0.5355
	NB	0.6364	0.6424	0.6330	0.6364	0.6099
	RF	0.6561	0.6479	0.6397	0.6497	0.6521
	SVM	0.5075	0.6046	0.5984	0.5000	0.5680
TwSarcasmBarbieri2014	J48	0.9072	0.9020	0.8759	0.9304	0.8936
	NB	0.9097	0.9113	0.9099	0.9097	0.8864
	RF	0.9755	0.9732	0.9766	0.9754	0.9758
	SVM	0.8784	0.8943	0.8969	0.8599	0.8981
TwRiloff2013	J48	0.6194	0.6565	0.6067	0.7319	0.6006
	NB	0.7471	0.7446	0.7478	0.7471	0.7331
	RF	0.8006	0.8020	0.8125	0.8078	0.8063
	SVM	0.5485	0.7326	0.7339	0.6230	0.7317
TwPtáček2014	J48	0.7127	0.7376	0.6843	0.7918	0.7234
	NB	0.7390	0.7378	0.7386	0.7390	0.7660
	RF	0.8885	0.8844	0.8886	0.8778	0.8878
	SVM	0.6647	0.7450	0.7448	0.5000	0.7450

TABLE 7: Benchmark Corpora - Youden Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.5079	0.6385	0.5469	0.7199	0.6562
	NB	0.6620	0.6605	0.6622	0.6620	0.6416
	RF	0.7590	0.7805	0.7809	0.7956	0.7748
	SVM	0.5000	0.6640	0.6759	0.5000	0.6646
TwImbData09112016	J48	0.7749	0.7863	0.6383	0.8597	0.7226
	NB	0.7815	0.7759	0.7809	0.7815	0.8164
	RF	0.9105	0.9194	0.9235	0.9260	0.9203
	SVM	0.5000	0.8114	0.8315	0.7090	0.8264
TwImbData10112016	J48	0.7440	0.8245	0.6766	0.8882	0.7405
	NB	0.7973	0.7948	0.7970	0.7973	0.8492
	RF	0.9231	0.9267	0.9346	0.9350	0.9320
	SVM	0.5000	0.8360	0.8593	0.7052	0.8529
TwImbData11112016	J48	0.7937	0.7882	0.6592	0.8701	0.6856
	NB	0.7977	0.7965	0.7978	0.7977	0.8380
	RF	0.9024	0.9081	0.9188	0.9211	0.9089
	SVM	0.5000	0.8134	0.8477	0.6563	0.8310
TwImbData12112016	J48	0.7784	0.8174	0.6549	0.8596	0.7226
	NB	0.7961	0.7933	0.7961	0.7960	0.8224
	RF	0.8937	0.9118	0.9206	0.9225	0.9103
	SVM	0.5000	0.8227	0.8427	0.7043	0.8313
TwImbData13112016	J48	0.7734	0.7945	0.6573	0.8628	0.7171
	NB	0.7885	0.7910	0.7888	0.7885	0.8253
	RF	0.9074	0.9151	0.9239	0.9241	0.9186
	SVM	0.5000	0.8150	0.8431	0.6963	0.8313
TwImbData14112016	J48	0.7662	0.7908	0.6495	0.8675	0.7154
	NB	0.7875	0.7820	0.7876	0.7875	0.8145
	RF	0.9116	0.9129	0.9226	0.9203	0.9183
	SVM	0.5000	0.8196	0.8399	0.6996	0.8323
TwImbData15112016	J48	0.7785	0.7907	0.6292	0.8608	0.6821
	NB	0.7884	0.7851	0.7883	0.7884	0.8234
	RF	0.8973	0.9059	0.9189	0.9142	0.9165
	SVM	0.5000	0.8069	0.8315	0.6783	0.8272
TwImbData16112016	J48	0.7820	0.7860	0.6553	0.8652	0.7175
	NB	0.7901	0.7848	0.7894	0.7901	0.8026
	RF	0.9036	0.9109	0.9212	0.9180	0.9152
	SVM	0.5000	0.8149	0.8442	0.7024	0.8332
TwImbData17112016	J48	0.7435	0.7936	0.6443	0.8680	0.7236
	NB	0.7917	0.7878	0.7917	0.7917	0.8372
	RF	0.9003	0.9084	0.9127	0.9179	0.9184
	SVM	0.5000	0.8146	0.8405	0.6792	0.8342
TwImbData18112016	J48	0.7484	0.7803	0.6393	0.8505	0.7068
	NB	0.7874	0.7904	0.7871	0.7874	0.8319
	RF	0.8859	0.9027	0.9109	0.9135	0.9080
	SVM	0.5000	0.8116	0.8351	0.6981	0.8274

TABLE 8: New Corpora - Youden Thresholding

AUPR

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.6427	0.6002	0.6099	0.7554	0.5288
	NB	0.5460	0.5596	0.5441	0.5422	0.6335
	RF	0.9007	0.8538	0.8838	0.8399	0.8713
	SVM	0.1253	0.0587	0.0587	0.0245	0.0668
TwIronyBarbieri2014	J48	0.6251	0.4738	0.5592	0.7519	0.4958
	NB	0.5276	0.5176	0.5259	0.5262	0.4324
	RF	0.8869	0.8240	0.8684	0.8045	0.8578
	SVM	0.1550	0.0866	0.0883	0.0129	0.0923
TwMohammad2015	J48	0.2516	0.2143	0.2314	0.2983	0.2234
	NB	0.3814	0.3899	0.3762	0.3814	0.3727
	RF	0.4029	0.3960	0.3858	0.4042	0.3947
	SVM	0.2198	0.1449	0.1511	0.0000	0.2189
TwSarcasmBarbieri2014	J48	0.6734	0.5535	0.6176	0.7938	0.5478
	NB	0.6331	0.6311	0.6339	0.6322	0.4461
	RF	0.9140	0.8449	0.8951	0.8285	0.8756
	SVM	0.1064	0.0516	0.0513	0.0167	0.0586
TwRiloff2013	J48	0.2605	0.2308	0.2333	0.3470	0.2243
	NB	0.3971	0.3932	0.3968	0.3971	0.4504
	RF	0.5245	0.5279	0.5249	0.5251	0.5354
	SVM	0.4096	0.0772	0.0859	0.0133	0.0861
TwPtáček2014	J48	0.3655	0.3623	0.3444	0.5934	0.3091
	NB	0.4273	0.4250	0.4268	0.4273	0.4745
	RF	0.7368	0.7220	0.7331	0.7289	0.7293
	SVM	0.2795	0.0615	0.0625	0.0000	0.0722

TABLE 9: Benchmark Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0046	0.0217	0.0501	0.0410	0.0375
	NB	0.0308	0.0309	0.0308	0.0308	0.0351
	RF	0.1162	0.0847	0.1233	0.0911	0.0991
	SVM	0.0000	0.0108	0.0107	0.0000	0.0108
TwImbData09112016	J48	0.1329	0.0491	0.1342	0.0878	0.0828
	NB	0.0481	0.0466	0.0479	0.0479	0.0667
	RF	0.3457	0.2477	0.3533	0.2770	0.3056
	SVM	0.0000	0.0067	0.0066	0.0000	0.0080
TwImbData10112016	J48	0.1431	0.0585	0.1809	0.1211	0.1016
	NB	0.0538	0.0529	0.0534	0.0530	0.0805
	RF	0.3880	0.2747	0.4019	0.3114	0.3545
	SVM	0.0000	0.0090	0.0077	0.0000	0.0100
TwImbData11112016	J48	0.1444	0.0469	0.1649	0.1193	0.0702
	NB	0.0535	0.0518	0.0534	0.0524	0.0880
	RF	0.3346	0.2111	0.3309	0.2409	0.2663
	SVM	0.0000	0.0094	0.0081	0.0000	0.0130
TwImbData12112016	J48	0.1455	0.0512	0.1570	0.0883	0.0805
	NB	0.0533	0.0520	0.0524	0.0519	0.0716
	RF	0.3568	0.2211	0.3643	0.2711	0.3207
	SVM	0.0000	0.0093	0.0091	0.0000	0.0126
TwImbData13112016	J48	0.1363	0.0480	0.1555	0.0949	0.0784
	NB	0.0509	0.0508	0.0508	0.0502	0.0782
	RF	0.3513	0.2493	0.3776	0.2738	0.3024
	SVM	0.0000	0.0097	0.0086	0.0000	0.0113
TwImbData14112016	J48	0.1201	0.0452	0.1516	0.1061	0.0791
	NB	0.0496	0.0484	0.0496	0.0492	0.0728
	RF	0.3422	0.2182	0.3381	0.2451	0.2890
	SVM	0.0000	0.0078	0.0077	0.0000	0.0103
TwImbData15112016	J48	0.1264	0.0470	0.1353	0.0857	0.0688
	NB	0.0513	0.0496	0.0510	0.0506	0.0702
	RF	0.3209	0.2050	0.3340	0.2423	0.3031
	SVM	0.0000	0.0089	0.0083	0.0000	0.0090
TwImbData16112016	J48	0.1380	0.0464	0.1616	0.0888	0.0763
	NB	0.0498	0.0484	0.0496	0.0496	0.0735
	RF	0.3394	0.2232	0.3491	0.2503	0.3013
	SVM	0.0000	0.0085	0.0077	0.0000	0.0107
TwImbData17112016	J48	0.1152	0.0505	0.1444	0.1088	0.0760
	NB	0.0513	0.0508	0.0509	0.0507	0.0829
	RF	0.3270	0.2061	0.3371	0.2532	0.3002
	SVM	0.0000	0.0094	0.0083	0.0001	0.0095
TwImbData18112016	J48	0.1052	0.0448	0.1341	0.0863	0.0755
	NB	0.0502	0.0506	0.0499	0.0497	0.0747
	RF	0.2959	0.2049	0.3178	0.2262	0.2928
	SVM	0.0000	0.0076	0.0070	0.0001	0.0085

TABLE 10: New Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.6427	0.6002	0.6099	0.7554	0.5288
	NB	0.5452	0.5585	0.5435	0.5414	0.6321
	RF	0.9007	0.8538	0.8838	0.8398	0.8713
	SVM	0.1253	0.0587	0.0587	0.0245	0.0668
TwIronyBarbieri2014	J48	0.6251	0.4738	0.5592	0.7519	0.4958
	NB	0.5275	0.5173	0.5258	0.5262	0.4324
	RF	0.2854	0.1613	0.6618	0.6573	0.4983
	SVM	0.1550	0.0866	0.0883	0.0129	0.0923
TwMohammad2015	J48	0.2516	0.2143	0.2314	0.2983	0.2234
	NB	0.3814	0.3896	0.3762	0.3811	0.3727
	RF	0.4006	0.3207	0.3816	0.4009	0.3872
	SVM	0.2198	0.1449	0.1511	0.0000	0.2189
TwSarcasmBarbieri2014	J48	0.6734	0.5535	0.6176	0.7938	0.5478
	NB	0.6331	0.6310	0.6337	0.6322	0.4461
	RF	0.1271	0.0703	0.1429	0.7322	0.1146
	SVM	0.1064	0.0516	0.0513	0.0167	0.0586
TwRiloff2013	J48	0.2605	0.2308	0.2333	0.3468	0.2243
	NB	0.3971	0.3932	0.3968	0.3971	0.4503
	RF	0.4712	0.2641	0.4620	0.5004	0.4094
	SVM	0.4096	0.0772	0.0859	0.0133	0.0861
TwPtáček2014	J48	0.1577	0.3623	0.3444	0.1637	0.3091
	NB	0.1773	0.4250	0.4268	0.1773	0.4745
	RF	0.1414	0.1141	0.3265	0.1566	0.2572
	SVM	0.0926	0.0615	0.0625	0.0000	0.0722

TABLE 11: Benchmark Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0046	0.0217	0.0501	0.0410	0.0375
	NB	0.0308	0.0308	0.0308	0.0307	0.0351
	RF	0.1074	0.0294	0.1233	0.3053	0.0934
	SVM	0.0000	0.0108	0.0107	0.0000	0.0108
TwImbData09112016	J48	0.1329	0.0491	0.1342	0.0877	0.0828
	NB	0.0480	0.0466	0.0479	0.0478	0.0665
	RF	0.3355	0.0442	0.3542	0.4499	0.2871
	SVM	0.0000	0.0067	0.0066	0.0000	0.0080
TwImbData10112016	J48	0.1431	0.0585	0.1809	0.1211	0.1016
	NB	0.0537	0.0528	0.0533	0.0528	0.0803
	RF	0.3888	0.0431	0.4015	0.4376	0.3443
	SVM	0.0000	0.0090	0.0077	0.0000	0.0100
TwImbData11112016	J48	0.1444	0.0469	0.1649	0.1193	0.0702
	NB	0.0534	0.0518	0.0534	0.0524	0.0878
	RF	0.3344	0.0542	0.3309	0.3179	0.2607
	SVM	0.0000	0.0094	0.0081	0.0000	0.0130
TwImbData12112016	J48	0.1455	0.0512	0.1570	0.0883	0.0805
	NB	0.0527	0.0520	0.0524	0.0514	0.0712
	RF	0.3509	0.0572	0.3640	0.4091	0.3085
	SVM	0.0000	0.0093	0.0091	0.0000	0.0126
TwImbData13112016	J48	0.1363	0.0480	0.1555	0.0949	0.0784
	NB	0.0508	0.0506	0.0507	0.0500	0.0781
	RF	0.3349	0.0705	0.3774	0.3871	0.2813
	SVM	0.0000	0.0097	0.0086	0.0000	0.0113
TwImbData14112016	J48	0.1201	0.0452	0.1516	0.1061	0.0791
	NB	0.0496	0.0484	0.0496	0.0490	0.0724
	RF	0.3263	0.0559	0.3381	0.4056	0.2690
	SVM	0.0000	0.0078	0.0077	0.0000	0.0103
TwImbData15112016	J48	0.1264	0.0470	0.1353	0.0857	0.0688
	NB	0.0510	0.0496	0.0510	0.0505	0.0700
	RF	0.3036	0.0553	0.3338	0.4624	0.2939
	SVM	0.0000	0.0089	0.0083	0.0000	0.0090
TwImbData16112016	J48	0.1380	0.0464	0.1616	0.0888	0.0763
	NB	0.0498	0.0484	0.0496	0.0493	0.0732
	RF	0.3333	0.0531	0.3491	0.4117	0.2917
	SVM	0.0000	0.0085	0.0077	0.0000	0.0107
TwImbData17112016	J48	0.1152	0.0505	0.1444	0.1088	0.0760
	NB	0.0513	0.0507	0.0509	0.0506	0.0827
	RF	0.3241	0.0501	0.3370	0.4144	0.2954
	SVM	0.0000	0.0094	0.0083	0.0001	0.0095
TwImbData18112016	J48	0.1052	0.0448	0.1341	0.0863	0.0755
	NB	0.0502	0.0506	0.0499	0.0495	0.0744
	RF	0.2941	0.0482	0.3199	0.3063	0.2887
	SVM	0.0000	0.0076	0.0070	0.0001	0.0085

TABLE 12: New Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.6427	0.6002	0.6099	0.7554	0.5288
	NB	0.5460	0.5596	0.5441	0.5422	0.6335
	RF	0.9007	0.8538	0.8838	0.8399	0.8713
	SVM	0.1253	0.0587	0.0587	0.0245	0.0668
TwIronyBarbieri2014	J48	0.6251	0.4738	0.5592	0.7519	0.4958
	NB	0.5276	0.5176	0.5259	0.5262	0.4324
	RF	0.8869	0.8240	0.8684	0.8045	0.8578
	SVM	0.1550	0.0866	0.0883	0.0129	0.0923
TwMohammad2015	J48	0.2516	0.2143	0.2314	0.2983	0.2234
	NB	0.3814	0.3899	0.3762	0.3814	0.3727
	RF	0.4029	0.3960	0.3858	0.4042	0.3947
	SVM	0.2198	0.1449	0.1511	0.0000	0.2189
TwSarcasmBarbieri2014	J48	0.6734	0.5535	0.6176	0.7938	0.5478
	NB	0.6331	0.6311	0.6339	0.6322	0.4461
	RF	0.9140	0.8449	0.8951	0.8285	0.8756
	SVM	0.1064	0.0516	0.0513	0.0167	0.0586
TwRiloff2013	J48	0.2605	0.2308	0.2333	0.3470	0.2243
	NB	0.3971	0.3932	0.3968	0.3971	0.4504
	RF	0.5245	0.5279	0.5249	0.5251	0.5354
	SVM	0.4096	0.0772	0.0859	0.0133	0.0861
TwPtáček2014	J48	0.3655	0.3623	0.3444	0.5934	0.3091
	NB	0.4273	0.4250	0.4268	0.4273	0.4745
	RF	0.7368	0.7220	0.7331	0.7289	0.7293
	SVM	0.2795	0.0615	0.0625	0.0000	0.0722

TABLE 13: Benchmark Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0046	0.0217	0.0501	0.0410	0.0375
	NB	0.0308	0.0309	0.0308	0.0308	0.0351
	RF	0.1162	0.0847	0.1233	0.0911	0.0991
	SVM	0.0000	0.0108	0.0107	0.0000	0.0108
TwImbData09112016	J48	0.1329	0.0491	0.1342	0.0878	0.0828
	NB	0.0481	0.0466	0.0479	0.0479	0.0667
	RF	0.3457	0.2477	0.3533	0.2770	0.3056
	SVM	0.0000	0.0067	0.0066	0.0000	0.0080
TwImbData10112016	J48	0.1431	0.0585	0.1809	0.1211	0.1016
	NB	0.0538	0.0529	0.0534	0.0530	0.0805
	RF	0.3880	0.2747	0.4019	0.3114	0.3545
	SVM	0.0000	0.0090	0.0077	0.0000	0.0100
TwImbData11112016	J48	0.1444	0.0469	0.1649	0.1193	0.0702
	NB	0.0535	0.0518	0.0534	0.0524	0.0880
	RF	0.3346	0.2111	0.3309	0.2409	0.2663
	SVM	0.0000	0.0094	0.0081	0.0000	0.0130
TwImbData12112016	J48	0.1455	0.0512	0.1570	0.0883	0.0805
	NB	0.0533	0.0520	0.0524	0.0519	0.0716
	RF	0.3568	0.2211	0.3643	0.2711	0.3207
	SVM	0.0000	0.0093	0.0091	0.0000	0.0126
TwImbData13112016	J48	0.1363	0.0480	0.1555	0.0949	0.0784
	NB	0.0509	0.0508	0.0508	0.0502	0.0782
	RF	0.3513	0.2493	0.3776	0.2738	0.3024
	SVM	0.0000	0.0097	0.0086	0.0000	0.0113
TwImbData14112016	J48	0.1201	0.0452	0.1516	0.1061	0.0791
	NB	0.0496	0.0484	0.0496	0.0492	0.0728
	RF	0.3422	0.2182	0.3381	0.2451	0.2890
	SVM	0.0000	0.0078	0.0077	0.0000	0.0103
TwImbData15112016	J48	0.1264	0.0470	0.1353	0.0857	0.0688
	NB	0.0513	0.0496	0.0510	0.0506	0.0702
	RF	0.3209	0.2050	0.3340	0.2423	0.3031
	SVM	0.0000	0.0089	0.0083	0.0000	0.0090
TwImbData16112016	J48	0.1380	0.0464	0.1616	0.0888	0.0763
	NB	0.0498	0.0484	0.0496	0.0496	0.0735
	RF	0.3394	0.2232	0.3491	0.2503	0.3013
	SVM	0.0000	0.0085	0.0077	0.0000	0.0107
TwImbData17112016	J48	0.1152	0.0505	0.1444	0.1088	0.0760
	NB	0.0513	0.0508	0.0509	0.0507	0.0829
	RF	0.3270	0.2061	0.3371	0.2532	0.3002
	SVM	0.0000	0.0094	0.0083	0.0001	0.0095
TwImbData18112016	J48	0.1052	0.0448	0.1341	0.0863	0.0755
	NB	0.0502	0.0506	0.0499	0.0497	0.0747
	RF	0.2959	0.2049	0.3178	0.2262	0.2928
	SVM	0.0000	0.0076	0.0070	0.0001	0.0085

TABLE 14: New Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.6427	0.6002	0.6099	0.7554	0.5288
	NB	0.5460	0.5596	0.5441	0.5422	0.6335
	RF	0.9007	0.8538	0.8838	0.8399	0.8713
	SVM	0.1253	0.0587	0.0587	0.0245	0.0668
TwIronyBarbieri2014	J48	0.6251	0.4738	0.5592	0.7519	0.4958
	NB	0.5276	0.5176	0.5259	0.5262	0.4324
	RF	0.8869	0.8240	0.8684	0.8045	0.8578
	SVM	0.1550	0.0866	0.0883	0.0129	0.0923
TwMohammad2015	J48	0.2516	0.2143	0.2314	0.2983	0.2234
	NB	0.3814	0.3899	0.3762	0.3814	0.3727
	RF	0.4029	0.3960	0.3858	0.4042	0.3947
	SVM	0.2198	0.1449	0.1511	0.0000	0.2189
TwSarcasmBarbieri2014	J48	0.6734	0.5535	0.6176	0.7938	0.5478
	NB	0.6331	0.6311	0.6339	0.6322	0.4461
	RF	0.9140	0.8449	0.8951	0.8285	0.8756
	SVM	0.1064	0.0516	0.0513	0.0167	0.0586
TwRiloff2013	J48	0.2605	0.2308	0.2333	0.3470	0.2243
	NB	0.3971	0.3932	0.3968	0.3971	0.4504
	RF	0.5245	0.5279	0.5249	0.5251	0.5354
	SVM	0.4096	0.0772	0.0859	0.0133	0.0861
TwPtáček2014	J48	0.3655	0.3623	0.3444	0.5934	0.3091
	NB	0.4273	0.4250	0.4268	0.4273	0.4745
	RF	0.7368	0.7220	0.7331	0.7289	0.7293
	SVM	0.2795	0.0615	0.0625	0.0000	0.0722

TABLE 15: Benchmark Corpora - Youden Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0046	0.0217	0.0501	0.0410	0.0375
	NB	0.0308	0.0309	0.0308	0.0308	0.0351
	RF	0.1162	0.0847	0.1233	0.0911	0.0991
	SVM	0.0000	0.0108	0.0107	0.0000	0.0108
TwImbData09112016	J48	0.1329	0.0491	0.1342	0.0878	0.0828
	NB	0.0481	0.0466	0.0479	0.0479	0.0667
	RF	0.3457	0.2477	0.3533	0.2770	0.3056
	SVM	0.0000	0.0067	0.0066	0.0000	0.0080
TwImbData10112016	J48	0.1431	0.0585	0.1809	0.1211	0.1016
	NB	0.0538	0.0529	0.0534	0.0530	0.0805
	RF	0.3880	0.2747	0.4019	0.3114	0.3545
	SVM	0.0000	0.0090	0.0077	0.0000	0.0100
TwImbData11112016	J48	0.1444	0.0469	0.1649	0.1193	0.0702
	NB	0.0535	0.0518	0.0534	0.0524	0.0880
	RF	0.3346	0.2111	0.3309	0.2409	0.2663
	SVM	0.0000	0.0094	0.0081	0.0000	0.0130
TwImbData12112016	J48	0.1455	0.0512	0.1570	0.0883	0.0805
	NB	0.0533	0.0520	0.0524	0.0519	0.0716
	RF	0.3568	0.2211	0.3643	0.2711	0.3207
	SVM	0.0000	0.0093	0.0091	0.0000	0.0126
TwImbData13112016	J48	0.1363	0.0480	0.1555	0.0949	0.0784
	NB	0.0509	0.0508	0.0508	0.0502	0.0782
	RF	0.3513	0.2493	0.3776	0.2738	0.3024
	SVM	0.0000	0.0097	0.0086	0.0000	0.0113
TwImbData14112016	J48	0.1201	0.0452	0.1516	0.1061	0.0791
	NB	0.0496	0.0484	0.0496	0.0492	0.0728
	RF	0.3422	0.2182	0.3381	0.2451	0.2890
	SVM	0.0000	0.0078	0.0077	0.0000	0.0103
TwImbData15112016	J48	0.1264	0.0470	0.1353	0.0857	0.0688
	NB	0.0513	0.0496	0.0510	0.0506	0.0702
	RF	0.3209	0.2050	0.3340	0.2423	0.3031
	SVM	0.0000	0.0089	0.0083	0.0000	0.0090
TwImbData16112016	J48	0.1380	0.0464	0.1616	0.0888	0.0763
	NB	0.0498	0.0484	0.0496	0.0496	0.0735
	RF	0.3394	0.2232	0.3491	0.2503	0.3013
	SVM	0.0000	0.0085	0.0077	0.0000	0.0107
TwImbData17112016	J48	0.1152	0.0505	0.1444	0.1088	0.0760
	NB	0.0513	0.0508	0.0509	0.0507	0.0829
	RF	0.3270	0.2061	0.3371	0.2532	0.3002
	SVM	0.0000	0.0094	0.0083	0.0001	0.0095
TwImbData18112016	J48	0.1052	0.0448	0.1341	0.0863	0.0755
	NB	0.0502	0.0506	0.0499	0.0497	0.0747
	RF	0.2959	0.2049	0.3178	0.2262	0.2928
	SVM	0.0000	0.0076	0.0070	0.0001	0.0085

TABLE 16: New Corpora - Youden Thresholding

Balanced Accuracy

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8311	0.8091	0.8164	0.8054	0.8160
	NB	0.5669	0.5981	0.5802	0.5880	0.6673
	RF	0.8540	0.8306	0.8600	0.8320	0.8606
	SVM	0.8025	0.7887	0.7903	0.7086	0.7981
TwIronyBarbieri2014	J48	0.7996	0.7289	0.7631	0.7288	0.7707
	NB	0.5989	0.5918	0.5957	0.5710	0.6014
	RF	0.8165	0.7982	0.8229	0.8005	0.8256
	SVM	0.7443	0.7137	0.7180	0.4960	0.7289
TwMohammad2015	J48	0.3238	0.4165	0.3358	0.4162	0.3546
	NB	0.4155	0.4393	0.4383	0.4441	0.4240
	RF	0.1294	0.4455	0.2282	0.4667	0.3338
	SVM	0.0384	0.4508	0.4405	0.4323	0.3601
TwSarcasmBarbieri2014	J48	0.8413	0.7948	0.8140	0.7871	0.8239
	NB	0.7028	0.6924	0.6860	0.6534	0.6741
	RF	0.8659	0.8360	0.8708	0.8415	0.8709
	SVM	0.8163	0.7798	0.7853	0.6606	0.7993
TwRiloff2013	J48	0.4118	0.4632	0.3866	0.5011	0.4202
	NB	0.4953	0.4979	0.4976	0.4977	0.4792
	RF	0.3333	0.5318	0.4713	0.5404	0.5372
	SVM	0.1913	0.5411	0.5471	0.4267	0.5438
TwPtáček2014	J48	0.5904	0.6190	0.5790	0.2193	0.5930
	NB	0.5186	0.5337	0.5339	0.4608	0.5632
	RF	0.6239	0.6689	0.6673	0.1485	0.6785
	SVM	0.5026	0.6018	0.6018	0.0000	0.6042

TABLE 17: Benchmark Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0052	0.0685	0.0979	0.0697	0.0696
	NB	0.0625	0.0615	0.0605	0.0562	0.0627
	RF	0.0375	0.0810	0.0526	0.0915	0.0497
	SVM	0.0000	0.0748	0.0785	0.0385	0.0749
TwImbData09112016	J48	0.2690	0.1302	0.2556	0.1245	0.2277
	NB	0.0948	0.0870	0.0889	0.0761	0.0804
	RF	0.0629	0.1356	0.0985	0.1475	0.1497
	SVM	0.0000	0.1287	0.1443	0.0644	0.1489
TwImbData10112016	J48	0.3064	0.1504	0.3308	0.1534	0.2839
	NB	0.1056	0.0990	0.0988	0.0871	0.0941
	RF	0.0443	0.1486	0.1033	0.1626	0.2083
	SVM	0.0000	0.1655	0.1821	0.0636	0.1930
TwImbData11112016	J48	0.2959	0.1310	0.2977	0.1322	0.2160
	NB	0.1062	0.1004	0.1008	0.0886	0.0959
	RF	0.0570	0.1354	0.0902	0.1513	0.1072
	SVM	0.0000	0.1467	0.1722	0.0552	0.1849
TwImbData12112016	J48	0.2888	0.1349	0.2930	0.1290	0.2333
	NB	0.1045	0.0986	0.0996	0.0882	0.0952
	RF	0.0455	0.1401	0.0981	0.1532	0.1481
	SVM	0.0000	0.1546	0.1741	0.0634	0.1844
TwImbData13112016	J48	0.2828	0.1318	0.2875	0.1266	0.2234
	NB	0.1013	0.0947	0.0944	0.0857	0.0885
	RF	0.0716	0.1343	0.0983	0.1522	0.1386
	SVM	0.0000	0.1501	0.1705	0.0618	0.1756
TwImbData14112016	J48	0.2742	0.1246	0.2840	0.1253	0.2284
	NB	0.0974	0.0927	0.0939	0.0808	0.0822
	RF	0.0765	0.1307	0.1136	0.1439	0.1448
	SVM	0.0000	0.1421	0.1600	0.0625	0.1703
TwImbData15112016	J48	0.2825	0.1264	0.2521	0.1237	0.2016
	NB	0.0955	0.0950	0.0959	0.0846	0.0807
	RF	0.0632	0.1286	0.0811	0.1429	0.1150
	SVM	0.0000	0.1386	0.1561	0.0586	0.1568
TwImbData16112016	J48	0.2868	0.1274	0.2920	0.1278	0.2178
	NB	0.0959	0.0944	0.0949	0.0815	0.0796
	RF	0.0516	0.1326	0.0819	0.1462	0.0915
	SVM	0.0000	0.1427	0.1644	0.0630	0.1743
TwImbData17112016	J48	0.2651	0.1282	0.2735	0.1359	0.2160
	NB	0.1003	0.0946	0.0980	0.0841	0.0861
	RF	0.0301	0.1319	0.0610	0.1445	0.1249
	SVM	0.0000	0.1479	0.1654	0.0588	0.1681
TwImbData18112016	J48	0.2158	0.1203	0.2594	0.1226	0.2116
	NB	0.1006	0.0953	0.0947	0.0825	0.0848
	RF	0.0366	0.1259	0.0699	0.1414	0.0921
	SVM	0.0000	0.1342	0.1506	0.0622	0.1535

TABLE 18: New Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8311	0.8092	0.8164	0.8219	0.8160
	NB	0.5408	0.6008	0.5831	0.5686	0.6693
	RF	0.8572	0.8249	0.8572	0.8142	0.8599
	SVM	0.8025	0.7887	0.7903	0.7086	0.7981
TwIronyBarbieri2014	J48	0.7996	0.7289	0.7631	0.7793	0.7707
	NB	0.5854	0.5914	0.5953	0.5213	0.6055
	RF	0.8170	0.7998	0.8200	0.6991	0.8264
	SVM	0.7443	0.7137	0.7180	0.0000	0.7289
TwMohammad2015	J48	0.3238	0.4165	0.3358	0.3652	0.3546
	NB	0.0000	0.4459	0.4365	0.0000	0.4222
	RF	0.2875	0.4627	0.2851	0.0941	0.4009
	SVM	0.0384	0.4508	0.4405	0.0000	0.3601
TwSarcasmBarbieri2014	J48	0.8413	0.7948	0.8140	0.8259	0.8239
	NB	0.7038	0.6931	0.6880	0.7012	0.6774
	RF	0.8676	0.8348	0.8704	0.7753	0.8706
	SVM	0.8163	0.7798	0.7853	0.6606	0.7993
TwRiloff2013	J48	0.4118	0.4632	0.3866	0.5262	0.4202
	NB	0.0000	0.4950	0.4972	0.0000	0.4810
	RF	0.4299	0.5375	0.4659	0.1865	0.5393
	SVM	0.1913	0.5411	0.5471	0.0000	0.5438
TwPtáček2014	J48	0.2253	0.6190	0.5790	0.3217	0.5930
	NB	0.0000	0.5335	0.5334	0.3699	0.5624
	RF	0.2028	0.6688	0.6706	0.1574	0.6789
	SVM	0.2897	0.6018	0.6018	0.0000	0.6042

TABLE 19: Benchmark Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0052	0.0685	0.0979	0.0000	0.0696
	NB	0.0000	0.0612	0.0609	0.0000	0.0630
	RF	0.0586	0.0785	0.0527	0.0313	0.0667
	SVM	0.0000	0.0748	0.0785	0.0000	0.0749
TwImbData09112016	J48	0.2690	0.1302	0.2556	0.0000	0.2277
	NB	0.0000	0.0866	0.0885	0.0000	0.0813
	RF	0.0934	0.1411	0.0822	0.0458	0.2147
	SVM	0.0000	0.1287	0.1443	0.0000	0.1489
TwImbData10112016	J48	0.3064	0.1506	0.3308	0.0000	0.2839
	NB	0.0000	0.0983	0.0984	0.0000	0.0972
	RF	0.1089	0.1538	0.0726	0.0260	0.2654
	SVM	0.0000	0.1655	0.1821	0.0000	0.1930
TwImbData11112016	J48	0.2959	0.1310	0.2977	0.0000	0.2160
	NB	0.0000	0.0993	0.1001	0.0000	0.0980
	RF	0.0927	0.1382	0.0820	0.0375	0.1595
	SVM	0.0000	0.1467	0.1722	0.0000	0.1849
TwImbData12112016	J48	0.2888	0.1349	0.2930	0.0000	0.2333
	NB	0.0000	0.0977	0.0989	0.0000	0.0972
	RF	0.1052	0.1444	0.0758	0.0260	0.2093
	SVM	0.0000	0.1546	0.1741	0.0000	0.1844
TwImbData13112016	J48	0.2828	0.1320	0.2875	0.0000	0.2234
	NB	0.0000	0.0942	0.0936	0.0000	0.0910
	RF	0.1154	0.1429	0.0894	0.0591	0.1857
	SVM	0.0000	0.1501	0.1705	0.0000	0.1756
TwImbData14112016	J48	0.2742	0.1246	0.2840	0.0000	0.2284
	NB	0.0000	0.0919	0.0945	0.0000	0.0840
	RF	0.1165	0.1376	0.1073	0.0637	0.1883
	SVM	0.0000	0.1421	0.1600	0.0000	0.1703
TwImbData15112016	J48	0.2825	0.1264	0.2521	0.0000	0.2016
	NB	0.0000	0.0935	0.0953	0.0000	0.0842
	RF	0.0923	0.1369	0.0728	0.0517	0.1961
	SVM	0.0000	0.1386	0.1561	0.0000	0.1568
TwImbData16112016	J48	0.2868	0.1274	0.2920	0.0000	0.2178
	NB	0.0000	0.0934	0.0942	0.0000	0.0817
	RF	0.0799	0.1359	0.0686	0.0411	0.1695
	SVM	0.0000	0.1427	0.1644	0.0000	0.1743
TwImbData17112016	J48	0.2655	0.1284	0.2735	0.0000	0.2160
	NB	0.0000	0.0934	0.0963	0.0000	0.0902
	RF	0.0769	0.1383	0.0501	0.0227	0.1864
	SVM	0.0000	0.1479	0.1654	0.0000	0.1681
TwImbData18112016	J48	0.2158	0.1203	0.2594	0.0000	0.2116
	NB	0.0000	0.0944	0.0934	0.0000	0.0878
	RF	0.0696	0.1306	0.0534	0.0310	0.1832
	SVM	0.0000	0.1342	0.1506	0.0000	0.1535

TABLE 20: New Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8325	0.8111	0.8187	0.8365	0.8165
	NB	0.5600	0.6080	0.5921	0.5600	0.6713
	RF	0.8594	0.8095	0.8594	0.7762	0.8531
	SVM	0.8025	0.7887	0.7903	0.7086	0.7981
TwIronyBarbieri2014	J48	0.7881	0.7187	0.7676	0.7976	0.7675
	NB	0.5856	0.5901	0.5944	0.5856	0.6116
	RF	0.8191	0.7871	0.8213	0.7396	0.8285
	SVM	0.7443	0.7137	0.7180	0.0000	0.7289
TwMohammad2015	J48	0.3382	0.4135	0.3278	0.2967	0.3541
	NB	0.3852	0.4625	0.4470	0.3852	0.4109
	RF	0.4564	0.4698	0.4564	0.3662	0.4701
	SVM	0.0384	0.4508	0.4405	0.0000	0.3601
TwSarcasmBarbieri2014	J48	0.8390	0.7870	0.8187	0.8341	0.8238
	NB	0.6940	0.6964	0.6941	0.6940	0.6884
	RF	0.8683	0.8249	0.8700	0.8110	0.8699
	SVM	0.8163	0.7798	0.7853	0.6606	0.7993
TwRiloff2013	J48	0.4497	0.4669	0.3851	0.4681	0.4247
	NB	0.4266	0.4915	0.4944	0.4266	0.4866
	RF	0.5249	0.4926	0.5539	0.3639	0.5426
	SVM	0.1913	0.5411	0.5471	0.0000	0.5438
TwPtáček2014	J48	0.6007	0.6208	0.5796	0.5001	0.5977
	NB	0.4521	0.5288	0.5286	0.4521	0.5566
	RF	0.6691	0.6633	0.6805	0.6711	0.6828
	SVM	0.5026	0.6018	0.6018	0.0000	0.6042

TABLE 21: Benchmark Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0107	0.0688	0.0982	0.0209	0.0828
	NB	0.0257	0.0594	0.0599	0.0257	0.0637
	RF	0.1040	0.0622	0.1053	0.0390	0.1428
	SVM	0.0000	0.0748	0.0785	0.0000	0.0749
TwImbData09112016	J48	0.2798	0.1344	0.2561	0.0235	0.2294
	NB	0.0483	0.0857	0.0883	0.0483	0.1030
	RF	0.2565	0.1101	0.2917	0.0527	0.3415
	SVM	0.0000	0.1287	0.1443	0.0000	0.1489
TwImbData10112016	J48	0.3263	0.1524	0.3320	0.1201	0.2851
	NB	0.0645	0.0924	0.0927	0.0645	0.1156
	RF	0.3264	0.1148	0.3525	0.0333	0.3723
	SVM	0.0000	0.1655	0.1821	0.0000	0.1930
TwImbData11112016	J48	0.3069	0.1324	0.2983	0.1970	0.2177
	NB	0.0631	0.0950	0.0960	0.0631	0.1161
	RF	0.3524	0.0983	0.3343	0.0540	0.2977
	SVM	0.0000	0.1467	0.1722	0.0000	0.1849
TwImbData12112016	J48	0.3002	0.1373	0.2945	0.0539	0.2295
	NB	0.0755	0.0896	0.0897	0.0756	0.1096
	RF	0.3043	0.1015	0.3255	0.0336	0.3575
	SVM	0.0000	0.1546	0.1741	0.0000	0.1844
TwImbData13112016	J48	0.3052	0.1331	0.2891	0.0337	0.2253
	NB	0.0604	0.0894	0.0886	0.0604	0.1080
	RF	0.3350	0.1014	0.3757	0.0627	0.3444
	SVM	0.0000	0.1501	0.1705	0.0000	0.1756
TwImbData14112016	J48	0.2806	0.1279	0.2852	0.0759	0.2308
	NB	0.0512	0.0890	0.0901	0.0512	0.1036
	RF	0.2991	0.1012	0.3304	0.0717	0.3172
	SVM	0.0000	0.1421	0.1600	0.0000	0.1703
TwImbData15112016	J48	0.2877	0.1287	0.2540	0.0170	0.2016
	NB	0.0563	0.0872	0.0892	0.0563	0.1047
	RF	0.2568	0.1019	0.2879	0.0583	0.3321
	SVM	0.0000	0.1386	0.1561	0.0000	0.1568
TwImbData16112016	J48	0.2963	0.1312	0.2927	0.0540	0.2179
	NB	0.0441	0.0894	0.0917	0.0441	0.0976
	RF	0.3046	0.1026	0.3226	0.0496	0.3308
	SVM	0.0000	0.1427	0.1644	0.0000	0.1743
TwImbData17112016	J48	0.2811	0.1314	0.2735	0.0160	0.2149
	NB	0.0568	0.0885	0.0896	0.0568	0.1114
	RF	0.2940	0.1050	0.3004	0.0282	0.3578
	SVM	0.0000	0.1479	0.1654	0.0000	0.1681
TwImbData18112016	J48	0.2454	0.1220	0.2607	0.0881	0.2140
	NB	0.0504	0.0902	0.0894	0.0504	0.1081
	RF	0.2818	0.0978	0.2611	0.0384	0.3318
	SVM	0.0000	0.1342	0.1506	0.0000	0.1535

TABLE 22: New Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8320	0.8092	0.8164	0.8088	0.8160
	NB	0.5880	0.6033	0.5874	0.5880	0.6625
	RF	0.8607	0.8283	0.8605	0.8207	0.8580
	SVM	0.8025	0.7887	0.7903	0.7086	0.7981
TwIronyBarbieri2014	J48	0.8003	0.7289	0.7631	0.7311	0.7707
	NB	0.5898	0.5841	0.5882	0.5898	0.6447
	RF	0.8196	0.7879	0.8247	0.7689	0.8282
	SVM	0.7443	0.7137	0.7180	0.4960	0.7289
TwMohammad2015	J48	0.3325	0.4165	0.3358	0.4168	0.3546
	NB	0.4341	0.4409	0.4452	0.4341	0.3879
	RF	0.4564	0.4698	0.4564	0.3662	0.4701
	SVM	0.0384	0.4508	0.4405	0.0000	0.3601
TwSarcasmBarbieri2014	J48	0.8425	0.7948	0.8140	0.7900	0.8239
	NB	0.6902	0.6930	0.6906	0.6902	0.7108
	RF	0.8683	0.8249	0.8700	0.8110	0.8699
	SVM	0.8163	0.7798	0.7853	0.6606	0.7993
TwRiloff2013	J48	0.4296	0.4632	0.3866	0.5016	0.4202
	NB	0.4961	0.4870	0.4926	0.4961	0.4831
	RF	0.5249	0.4926	0.5539	0.3639	0.5426
	SVM	0.1913	0.5411	0.5471	0.4267	0.5438
TwPtáček2014	J48	0.5941	0.6189	0.5790	0.6102	0.5930
	NB	0.5484	0.5470	0.5483	0.5484	0.5641
	RF	0.6691	0.6633	0.6805	0.6719	0.6828
	SVM	0.5026	0.6018	0.6018	0.0000	0.6042

TABLE 23: Benchmark Corpora - Youden Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0079	0.0686	0.0979	0.0712	0.0696
	NB	0.0569	0.0561	0.0564	0.0569	0.0645
	RF	0.1040	0.0622	0.1053	0.0390	0.1428
	SVM	0.0000	0.0748	0.0785	0.0000	0.0749
TwImbData09112016	J48	0.1046	0.1302	0.2556	0.1274	0.2277
	NB	0.0783	0.0763	0.0777	0.0783	0.0933
	RF	0.2565	0.1101	0.2917	0.0527	0.3415
	SVM	0.0000	0.1287	0.1443	0.0644	0.1489
TwImbData10112016	J48	0.1801	0.1506	0.3308	0.1583	0.2839
	NB	0.0833	0.0832	0.0834	0.0833	0.1020
	RF	0.3264	0.1148	0.3525	0.0333	0.3723
	SVM	0.0000	0.1655	0.1821	0.0636	0.1930
TwImbData11112016	J48	0.1552	0.1310	0.2977	0.1359	0.2160
	NB	0.0806	0.0856	0.0819	0.0806	0.1000
	RF	0.3524	0.0983	0.3343	0.0540	0.2977
	SVM	0.0000	0.1467	0.1722	0.0552	0.1849
TwImbData12112016	J48	0.1293	0.1349	0.2930	0.1330	0.2333
	NB	0.0819	0.0791	0.0821	0.0819	0.0923
	RF	0.3043	0.1015	0.3255	0.0336	0.3575
	SVM	0.0000	0.1546	0.1741	0.0634	0.1844
TwImbData13112016	J48	0.1326	0.1320	0.2875	0.1317	0.2234
	NB	0.0764	0.0801	0.0793	0.0764	0.0939
	RF	0.3350	0.1014	0.3757	0.0627	0.3444
	SVM	0.0000	0.1501	0.1705	0.0618	0.1756
TwImbData14112016	J48	0.1517	0.1246	0.2840	0.1286	0.2284
	NB	0.0820	0.0807	0.0825	0.0820	0.0915
	RF	0.2991	0.1012	0.3304	0.0717	0.3172
	SVM	0.0000	0.1421	0.1600	0.0625	0.1703
TwImbData15112016	J48	0.1482	0.1264	0.2521	0.1267	0.2016
	NB	0.0815	0.0798	0.0817	0.0815	0.0886
	RF	0.2568	0.1019	0.2879	0.0583	0.3321
	SVM	0.0000	0.1386	0.1561	0.0586	0.1568
TwImbData16112016	J48	0.1167	0.1274	0.2920	0.1312	0.2178
	NB	0.0805	0.0803	0.0797	0.0805	0.0904
	RF	0.3046	0.1026	0.3226	0.0496	0.3308
	SVM	0.0000	0.1427	0.1644	0.0630	0.1743
TwImbData17112016	J48	0.1630	0.1284	0.2735	0.1404	0.2160
	NB	0.0799	0.0781	0.0801	0.0799	0.0968
	RF	0.2940	0.1050	0.3004	0.0282	0.3578
	SVM	0.0000	0.1479	0.1654	0.0588	0.1681
TwImbData18112016	J48	0.1013	0.1203	0.2594	0.1251	0.2116
	NB	0.0762	0.0770	0.0763	0.0762	0.0977
	RF	0.2818	0.0978	0.2611	0.0384	0.3318
	SVM	0.0000	0.1342	0.1506	0.0622	0.1535

TABLE 24: New Corpora - Youden Thresholding

PPOS

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.2409	0.3059	0.2593	0.3137	0.2586
	NB	0.2605	0.2893	0.2902	0.3226	0.2912
	RF	0.2280	0.2990	0.2422	0.3014	0.2490
	SVM	0.2296	0.3068	0.3059	0.4134	0.2920
TwIronyBarbiery2014	J48	0.1830	0.2700	0.2115	0.2755	0.2061
	NB	0.2599	0.3071	0.3177	0.3901	0.3037
	RF	0.1582	0.2257	0.1649	0.2298	0.1678
	SVM	0.1717	0.2459	0.2421	0.5677	0.2332
TwMohammad2015	J48	0.2431	0.4841	0.2799	0.4624	0.3167
	NB	0.3136	0.3773	0.3691	0.4324	0.3764
	RF	0.0425	0.4971	0.1026	0.5521	0.1612
	SVM	0.0098	0.4049	0.3914	1.0000	0.2395
TwSarcasmBarbieri2014	J48	0.1895	0.2511	0.2106	0.2615	0.2030
	NB	0.2614	0.2971	0.3041	0.3516	0.2802
	RF	0.1788	0.2326	0.1871	0.2332	0.1903
	SVM	0.1896	0.2539	0.2516	0.3769	0.2377
TwRiloff2013	J48	0.1822	0.4272	0.2492	0.4355	0.2797
	NB	0.3223	0.3833	0.3704	0.4254	0.4212
	RF	0.0795	0.3972	0.1387	0.4170	0.2173
	SVM	0.0361	0.3962	0.3731	0.7532	0.3745
TwPtáček2014	J48	0.2534	0.3888	0.2755	0.0374	0.3088
	NB	0.3874	0.4373	0.4388	0.2814	0.4519
	RF	0.1951	0.4236	0.2510	0.0224	0.2948
	SVM	0.1619	0.4752	0.4729	0.0000	0.4493

TABLE 25: Benchmark Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0005	0.3520	0.0282	0.4154	0.0225
	NB	0.2010	0.3210	0.3283	0.5041	0.3287
	RF	0.0004	0.3448	0.0006	0.2924	0.0005
	SVM	0.0000	0.3138	0.3050	1.0000	0.3136
TwImbData09112016	J48	0.0087	0.2235	0.0259	0.2608	0.0308
	NB	0.1492	0.2666	0.2711	0.4327	0.4105
	RF	0.0007	0.2438	0.0011	0.2198	0.0023
	SVM	0.0000	0.2402	0.2158	0.5883	0.2023
TwImbData10112016	J48	0.0109	0.1960	0.0244	0.2109	0.0330
	NB	0.1470	0.2284	0.2255	0.3349	0.3416
	RF	0.0004	0.2185	0.0012	0.1985	0.0033
	SVM	0.0000	0.1791	0.1684	0.5963	0.1518
TwImbData11112016	J48	0.0100	0.2212	0.0248	0.2414	0.0356
	NB	0.1476	0.2304	0.2254	0.3280	0.3113
	RF	0.0006	0.2420	0.0009	0.2123	0.0014
	SVM	0.0000	0.1986	0.1760	0.6921	0.1540
TwImbData12112016	J48	0.0102	0.2192	0.0242	0.2478	0.0310
	NB	0.1446	0.2079	0.2072	0.3019	0.2965
	RF	0.0005	0.2295	0.0010	0.2053	0.0020
	SVM	0.0000	0.1891	0.1701	0.5985	0.1539
TwImbData13112016	J48	0.0101	0.2217	0.0260	0.2550	0.0347
	NB	0.1545	0.2241	0.2271	0.3334	0.3470
	RF	0.0007	0.2418	0.0011	0.2084	0.0018
	SVM	0.0000	0.1920	0.1754	0.6141	0.1621
TwImbData14112016	J48	0.0099	0.2311	0.0242	0.2638	0.0312
	NB	0.1405	0.2374	0.2356	0.3828	0.3896
	RF	0.0008	0.2542	0.0012	0.2262	0.0018
	SVM	0.0000	0.2119	0.1904	0.6078	0.1731
TwImbData15112016	J48	0.0094	0.2288	0.0233	0.2622	0.0355
	NB	0.1501	0.2347	0.2297	0.3482	0.4154
	RF	0.0006	0.2551	0.0009	0.2235	0.0014
	SVM	0.0000	0.2099	0.1920	0.6504	0.1878
TwImbData16112016	J48	0.0104	0.2276	0.0244	0.2582	0.0325
	NB	0.1526	0.2443	0.2461	0.3806	0.3848
	RF	0.0005	0.2490	0.0009	0.2192	0.0011
	SVM	0.0000	0.2073	0.1857	0.6022	0.1669
TwImbData17112016	J48	0.0108	0.2242	0.0243	0.2330	0.0350
	NB	0.1377	0.2380	0.2257	0.3526	0.3803
	RF	0.0003	0.2437	0.0007	0.2162	0.0018
	SVM	0.0000	0.1960	0.1817	0.6458	0.1738
TwImbData18112016	J48	0.0089	0.2378	0.0253	0.2643	0.0332
	NB	0.1521	0.2377	0.2354	0.3569	0.3975
	RF	0.0004	0.2576	0.0007	0.2272	0.0012
	SVM	0.0000	0.2253	0.2050	0.6095	0.1943

TABLE 26: New Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.2409	0.3058	0.2593	0.2933	0.2586
	NB	0.2227	0.3010	0.3027	0.2635	0.2957
	RF	0.2361	0.3072	0.2354	0.1949	0.2575
	SVM	0.2296	0.3068	0.3059	0.4134	0.2920
TwIronyBarbieri2014	J48	0.1830	0.2700	0.2115	0.2248	0.2061
	NB	0.2003	0.3134	0.3198	0.1673	0.2983
	RF	0.1591	0.2221	0.1618	0.1148	0.1709
	SVM	0.1717	0.2459	0.2421	0.0000	0.2332
TwMohammad2015	J48	0.2431	0.4841	0.2799	0.3623	0.3167
	NB	0.0000	0.3955	0.3883	0.0000	0.3614
	RF	0.1254	0.5448	0.1405	0.0228	0.2696
	SVM	0.0098	0.4049	0.3914	0.0000	0.2395
TwSarcasmBarbieri2014	J48	0.1895	0.2511	0.2106	0.2307	0.2030
	NB	0.2254	0.2959	0.3011	0.2717	0.2763
	RF	0.1803	0.2339	0.1857	0.1351	0.1908
	SVM	0.1896	0.2539	0.2516	0.3769	0.2377
TwRiloff2013	J48	0.1822	0.4272	0.2492	0.2973	0.2797
	NB	0.0000	0.3852	0.3727	0.0000	0.4073
	RF	0.1170	0.3815	0.1392	0.0365	0.2228
	SVM	0.0361	0.3962	0.3731	0.0000	0.3745
TwPtáček2014	J48	0.7466	0.3888	0.2755	0.8898	0.3088
	NB	1.0000	0.4334	0.4343	0.8758	0.4348
	RF	0.7689	0.4195	0.2581	0.6947	0.3002
	SVM	0.8381	0.4752	0.4729	1.0000	0.4493

TABLE 27: Benchmark Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0005	0.3520	0.0282	0.0000	0.0225
	NB	0.0000	0.3323	0.3381	0.0000	0.3254
	RF	0.0006	0.3704	0.0005	0.0003	0.0010
	SVM	0.0000	0.3138	0.3050	0.0000	0.3136
TwImbData09112016	J48	0.0087	0.2235	0.0259	0.0000	0.0308
	NB	0.0000	0.2742	0.2770	0.0000	0.3968
	RF	0.0010	0.2301	0.0009	0.0005	0.0048
	SVM	0.0000	0.2402	0.2158	0.0000	0.2023
TwImbData10112016	J48	0.0109	0.1958	0.0244	0.0000	0.0330
	NB	0.0000	0.2372	0.2338	0.0000	0.3220
	RF	0.0012	0.2088	0.0008	0.0003	0.0050
	SVM	0.0000	0.1791	0.1684	0.0000	0.1518
TwImbData11112016	J48	0.0100	0.2212	0.0248	0.0000	0.0356
	NB	0.0000	0.2378	0.2338	0.0000	0.2923
	RF	0.0010	0.2406	0.0008	0.0004	0.0029
	SVM	0.0000	0.1986	0.1760	0.0000	0.1540
TwImbData12112016	J48	0.0102	0.2192	0.0242	0.0000	0.0310
	NB	0.0000	0.2187	0.2175	0.0000	0.2832
	RF	0.0012	0.2174	0.0008	0.0003	0.0035
	SVM	0.0000	0.1891	0.1701	0.0000	0.1539
TwImbData13112016	J48	0.0101	0.2215	0.0260	0.0000	0.0347
	NB	0.0000	0.2335	0.2359	0.0000	0.3235
	RF	0.0013	0.2212	0.0010	0.0006	0.0034
	SVM	0.0000	0.1920	0.1754	0.0000	0.1621
TwImbData14112016	J48	0.0099	0.2311	0.0242	0.0000	0.0312
	NB	0.0000	0.2478	0.2460	0.0000	0.3742
	RF	0.0013	0.2359	0.0012	0.0006	0.0033
	SVM	0.0000	0.2119	0.1904	0.0000	0.1731
TwImbData15112016	J48	0.0094	0.2288	0.0233	0.0000	0.0355
	NB	0.0000	0.2443	0.2391	0.0000	0.3881
	RF	0.0010	0.2311	0.0008	0.0005	0.0030
	SVM	0.0000	0.2099	0.1920	0.0000	0.1878
TwImbData16112016	J48	0.0104	0.2276	0.0244	0.0000	0.0325
	NB	0.0000	0.2526	0.2528	0.0000	0.3658
	RF	0.0009	0.2396	0.0008	0.0004	0.0029
	SVM	0.0000	0.2073	0.1857	0.0000	0.1669
TwImbData17112016	J48	0.0108	0.2237	0.0243	0.0000	0.0350
	NB	0.0000	0.2477	0.2363	0.0000	0.3523
	RF	0.0008	0.2284	0.0005	0.0002	0.0034
	SVM	0.0000	0.1960	0.1817	0.0000	0.1738
TwImbData18112016	J48	0.0089	0.2378	0.0253	0.0000	0.0332
	NB	0.0000	0.2469	0.2440	0.0000	0.3727
	RF	0.0008	0.2443	0.0006	0.0003	0.0033
	SVM	0.0000	0.2253	0.2050	0.0000	0.1943

TABLE 28: New Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.2490	0.3039	0.2551	0.2516	0.2638
	NB	0.2506	0.3725	0.3780	0.2506	0.3276
	RF	0.2428	0.3265	0.2559	0.1744	0.2718
	SVM	0.2296	0.3068	0.3059	0.4134	0.2920
TwIronyBarbieri2014	J48	0.2003	0.2818	0.2073	0.1995	0.2154
	NB	0.2001	0.3333	0.3285	0.2001	0.2892
	RF	0.1607	0.2390	0.1630	0.1286	0.1765
	SVM	0.1717	0.2459	0.2421	0.0000	0.2332
TwMohammad2015	J48	0.2778	0.4826	0.2742	0.2773	0.3297
	NB	0.2732	0.4593	0.4469	0.2732	0.3225
	RF	0.4810	0.7911	0.5219	0.2073	0.6008
	SVM	0.0098	0.4049	0.3914	0.0000	0.2395
TwSarcasmBarbieri2014	J48	0.1998	0.2585	0.2069	0.2006	0.2101
	NB	0.2004	0.2891	0.2893	0.2004	0.2604
	RF	0.1832	0.2433	0.1926	0.1494	0.1964
	SVM	0.1896	0.2539	0.2516	0.3769	0.2377
TwRiloff2013	J48	0.2164	0.4193	0.2460	0.2270	0.2834
	NB	0.2201	0.3935	0.3852	0.2201	0.3306
	RF	0.2289	0.5543	0.2816	0.0855	0.3499
	SVM	0.0361	0.3962	0.3731	0.0000	0.3745
TwPtáček2014	J48	0.2696	0.3809	0.2692	0.2100	0.3172
	NB	0.2680	0.4185	0.4177	0.2680	0.3802
	RF	0.2622	0.4448	0.3036	0.3072	0.3400
	SVM	0.1619	0.4752	0.4729	0.0000	0.4493

TABLE 29: Benchmark Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0021	0.3539	0.0278	0.0144	0.0330
	NB	0.0197	0.3863	0.3825	0.0197	0.3123
	RF	0.0028	0.5525	0.0029	0.0004	0.0322
	SVM	0.0000	0.3138	0.3050	0.0000	0.3136
TwImbData09112016	J48	0.0184	0.2117	0.0257	0.0054	0.0344
	NB	0.0197	0.3053	0.2958	0.0197	0.2237
	RF	0.0049	0.3238	0.0068	0.0006	0.0198
	SVM	0.0000	0.2402	0.2158	0.0000	0.2023
TwImbData10112016	J48	0.0200	0.1894	0.0242	0.0097	0.0354
	NB	0.0196	0.2869	0.2846	0.0196	0.2210
	RF	0.0077	0.3061	0.0081	0.0004	0.0206
	SVM	0.0000	0.1791	0.1684	0.0000	0.1518
TwImbData11112016	J48	0.0194	0.2173	0.0247	0.0140	0.0384
	NB	0.0195	0.2868	0.2771	0.0195	0.2025
	RF	0.0110	0.3627	0.0131	0.0006	0.0255
	SVM	0.0000	0.1986	0.1760	0.0000	0.1540
TwImbData12112016	J48	0.0188	0.2133	0.0238	0.0064	0.0345
	NB	0.0197	0.2944	0.2920	0.0197	0.2209
	RF	0.0074	0.3496	0.0080	0.0003	0.0228
	SVM	0.0000	0.1891	0.1701	0.0000	0.1539
TwImbData13112016	J48	0.0181	0.2184	0.0257	0.0052	0.0382
	NB	0.0195	0.2935	0.2944	0.0195	0.2243
	RF	0.0089	0.3561	0.0106	0.0007	0.0293
	SVM	0.0000	0.1920	0.1754	0.0000	0.1621
TwImbData14112016	J48	0.0198	0.2211	0.0240	0.0137	0.0350
	NB	0.0196	0.2978	0.2904	0.0196	0.2182
	RF	0.0071	0.3570	0.0102	0.0007	0.0272
	SVM	0.0000	0.2119	0.1904	0.0000	0.1731
TwImbData15112016	J48	0.0189	0.2197	0.0230	0.0040	0.0384
	NB	0.0196	0.2995	0.2933	0.0196	0.2375
	RF	0.0057	0.3489	0.0072	0.0006	0.0243
	SVM	0.0000	0.2099	0.1920	0.0000	0.1878
TwImbData16112016	J48	0.0199	0.2183	0.0242	0.0099	0.0366
	NB	0.0198	0.2947	0.2873	0.0198	0.2186
	RF	0.0074	0.3493	0.0080	0.0006	0.0295
	SVM	0.0000	0.2073	0.1857	0.0000	0.1669
TwImbData17112016	J48	0.0202	0.2159	0.0241	0.0021	0.0375
	NB	0.0196	0.2988	0.2922	0.0196	0.2204
	RF	0.0071	0.3383	0.0077	0.0003	0.0275
	SVM	0.0000	0.1960	0.1817	0.0000	0.1738
TwImbData18112016	J48	0.0194	0.2296	0.0251	0.0165	0.0366
	NB	0.0196	0.2952	0.2939	0.0196	0.2230
	RF	0.0071	0.3661	0.0066	0.0004	0.0328
	SVM	0.0000	0.2253	0.2050	0.0000	0.1943

TABLE 30: New Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.2433	0.3058	0.2593	0.3106	0.2586
	NB	0.4252	0.4359	0.4262	0.4252	0.3726
	RF	0.2444	0.3003	0.2510	0.1988	0.2604
	SVM	0.2296	0.3068	0.3059	0.4134	0.2920
TwIronyBarbieri2014	J48	0.1859	0.2700	0.2115	0.2736	0.2061
	NB	0.3436	0.3543	0.3467	0.3436	0.2338
	RF	0.1615	0.2380	0.1678	0.1373	0.1758
	SVM	0.1717	0.2459	0.2421	0.5677	0.2332
TwMohammad2015	J48	0.2483	0.4841	0.2799	0.4587	0.3167
	NB	0.3868	0.4120	0.4825	0.3868	0.2701
	RF	0.4810	0.7911	0.5219	0.2073	0.6008
	SVM	0.0098	0.4049	0.3914	0.0000	0.2395
TwSarcasmBarbieri2014	J48	0.1921	0.2511	0.2106	0.2597	0.2030
	NB	0.2963	0.2945	0.2963	0.2963	0.2287
	RF	0.1832	0.2433	0.1926	0.1494	0.1964
	SVM	0.1896	0.2539	0.2516	0.3769	0.2377
TwRiloff2013	J48	0.1914	0.4272	0.2492	0.4327	0.2797
	NB	0.4596	0.4601	0.4610	0.4596	0.3430
	RF	0.2289	0.5543	0.2816	0.0855	0.3499
	SVM	0.0361	0.3962	0.3731	0.7532	0.3745
TwPtáček2014	J48	0.2587	0.3888	0.2755	0.4533	0.3088
	NB	0.5460	0.5556	0.5448	0.5460	0.4690
	RF	0.2622	0.4448	0.3036	0.3326	0.3400
	SVM	0.1619	0.4752	0.4729	0.0000	0.4493

TABLE 31: Benchmark Corpora - Youden Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0083	0.3516	0.0282	0.4032	0.0225
	NB	0.4860	0.4822	0.4886	0.4860	0.3037
	RF	0.0028	0.5525	0.0029	0.0004	0.0322
	SVM	0.0000	0.3138	0.3050	0.0000	0.3136
TwImbData09112016	J48	0.2374	0.2235	0.0259	0.2518	0.0308
	NB	0.4061	0.4197	0.4086	0.4061	0.2887
	RF	0.0049	0.3238	0.0068	0.0006	0.0198
	SVM	0.0000	0.2402	0.2158	0.5883	0.2023
TwImbData10112016	J48	0.0955	0.1958	0.0244	0.2031	0.0330
	NB	0.3719	0.3711	0.3686	0.3719	0.2948
	RF	0.0077	0.3061	0.0081	0.0004	0.0206
	SVM	0.0000	0.1791	0.1684	0.5963	0.1518
TwImbData11112016	J48	0.1374	0.2212	0.0248	0.2311	0.0356
	NB	0.3948	0.3465	0.3779	0.3948	0.2729
	RF	0.0110	0.3627	0.0131	0.0006	0.0255
	SVM	0.0000	0.1986	0.1760	0.6921	0.1540
TwImbData12112016	J48	0.1825	0.2192	0.0242	0.2369	0.0310
	NB	0.3812	0.4041	0.3796	0.3812	0.3087
	RF	0.0074	0.3496	0.0080	0.0003	0.0228
	SVM	0.0000	0.1891	0.1701	0.5985	0.1539
TwImbData13112016	J48	0.1537	0.2215	0.0260	0.2426	0.0347
	NB	0.4149	0.3908	0.3916	0.4149	0.3038
	RF	0.0089	0.3561	0.0106	0.0007	0.0293
	SVM	0.0000	0.1920	0.1754	0.6141	0.1621
TwImbData14112016	J48	0.1291	0.2311	0.0242	0.2549	0.0312
	NB	0.3643	0.3767	0.3620	0.3643	0.2873
	RF	0.0071	0.3570	0.0102	0.0007	0.0272
	SVM	0.0000	0.2119	0.1904	0.6078	0.1731
TwImbData15112016	J48	0.1529	0.2288	0.0233	0.2541	0.0355
	NB	0.3851	0.3992	0.3844	0.3851	0.3417
	RF	0.0057	0.3489	0.0072	0.0006	0.0243
	SVM	0.0000	0.2099	0.1920	0.6504	0.1878
TwImbData16112016	J48	0.2187	0.2276	0.0244	0.2493	0.0325
	NB	0.3883	0.3858	0.3921	0.3883	0.2897
	RF	0.0074	0.3493	0.0080	0.0006	0.0295
	SVM	0.0000	0.2073	0.1857	0.6022	0.1669
TwImbData17112016	J48	0.1142	0.2237	0.0243	0.2241	0.0350
	NB	0.4080	0.4146	0.4075	0.4080	0.2994
	RF	0.0071	0.3383	0.0077	0.0003	0.0275
	SVM	0.0000	0.1960	0.1817	0.6458	0.1738
TwImbData18112016	J48	0.2521	0.2378	0.0253	0.2570	0.0332
	NB	0.4243	0.4088	0.4258	0.4243	0.2956
	RF	0.0071	0.3661	0.0066	0.0004	0.0328
	SVM	0.0000	0.2253	0.2050	0.6095	0.1943

TABLE 32: New Corpora - Youden Thresholding

F-score

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8932	0.8496	0.8725	0.8451	0.8726
	NB	0.7067	0.7164	0.7044	0.7007	0.7612
	RF	0.9180	0.8656	0.9117	0.8658	0.9077
	SVM	0.8819	0.8359	0.8373	0.7716	0.8466
TwIronyBarbieri2014	J48	0.8898	0.7977	0.8440	0.7961	0.8523
	NB	0.7229	0.7083	0.7092	0.6901	0.7146
	RF	0.9303	0.8570	0.9254	0.8563	0.9235
	SVM	0.8657	0.7965	0.8006	0.6566	0.8111
TwMohammad2015	J48	0.5467	0.5513	0.5407	0.5525	0.5420
	NB	0.5847	0.5847	0.5853	0.5776	0.5760
	RF	0.6152	0.5713	0.5828	0.5834	0.6090
	SVM	0.3655	0.5871	0.5824	0.0000	0.5748
TwSarcasmBarbieri2014	J48	0.9101	0.8432	0.8760	0.8348	0.8877
	NB	0.7845	0.7699	0.7648	0.7410	0.7623
	RF	0.9367	0.8762	0.9311	0.8792	0.9281
	SVM	0.8941	0.8331	0.8372	0.7452	0.8516
TwRiloff2013	J48	0.6458	0.6167	0.5972	0.6412	0.6094
	NB	0.6461	0.6407	0.6411	0.6377	0.6257
	RF	0.7222	0.6620	0.7290	0.6669	0.7063
	SVM	0.7514	0.6678	0.6732	0.6135	0.6711
TwPtáček2014	J48	0.7274	0.2916	0.2908	0.8266	0.2932
	NB	0.6392	0.3558	0.3558	0.6265	0.3353
	RF	0.7937	0.2589	0.2182	0.8565	0.2308
	SVM	0.7384	0.3071	0.3072	0.0000	0.3055

TABLE 33: Benchmark Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0559	0.5128	0.5325	0.5144	0.5235
	NB	0.5092	0.5096	0.5093	0.5097	0.5101
	RF	0.9904	0.5177	0.9655	0.5206	0.9654
	SVM	0.0000	0.5147	0.5160	0.0000	0.5147
TwImbData09112016	J48	0.7105	0.5330	0.6052	0.5320	0.5861
	NB	0.5200	0.5185	0.5192	0.5178	0.5191
	RF	0.9760	0.5355	0.9581	0.5389	0.8558
	SVM	0.0000	0.5329	0.5377	0.5166	0.5389
TwImbData10112016	J48	0.7076	0.5393	0.6435	0.5407	0.6074
	NB	0.5236	0.5221	0.5220	0.5199	0.5229
	RF	0.9904	0.5392	0.9599	0.5435	0.8569
	SVM	0.0000	0.5440	0.5493	0.5164	0.5527
TwImbData11112016	J48	0.7147	0.5332	0.6277	0.5342	0.5781
	NB	0.5238	0.5227	0.5227	0.5204	0.5230
	RF	0.7924	0.5353	0.9907	0.5400	0.8995
	SVM	0.0000	0.5381	0.5462	0.5141	0.5505
TwImbData12112016	J48	0.7058	0.5345	0.6259	0.5332	0.5906
	NB	0.5232	0.5217	0.5220	0.5196	0.5223
	RF	0.8789	0.5366	0.9836	0.5405	0.9101
	SVM	0.0000	0.5405	0.5467	0.5163	0.5501
TwImbData13112016	J48	0.7036	0.5335	0.6197	0.5326	0.5825
	NB	0.5221	0.5205	0.5205	0.5193	0.5209
	RF	0.9906	0.5349	0.9664	0.5402	0.9002
	SVM	0.0000	0.5391	0.5456	0.5159	0.5472
TwImbData14112016	J48	0.6967	0.5312	0.6223	0.5324	0.5862
	NB	0.5209	0.5200	0.5204	0.5185	0.5194
	RF	0.9906	0.5340	0.9762	0.5379	0.9246
	SVM	0.0000	0.5368	0.5424	0.5161	0.5456
TwImbData15112016	J48	0.7141	0.5318	0.6091	0.5317	0.5710
	NB	0.5202	0.5208	0.5211	0.5192	0.5193
	RF	0.9905	0.5333	0.9844	0.5374	0.9222
	SVM	0.0000	0.5356	0.5411	0.5151	0.5413
TwImbData16112016	J48	0.7009	0.5321	0.6254	0.5331	0.5835
	NB	0.5204	0.5208	0.5210	0.5187	0.5182
	RF	0.9905	0.5346	0.9348	0.5384	0.9212
	SVM	0.0000	0.5369	0.5438	0.5162	0.5468
TwImbData17112016	J48	0.6792	0.5323	0.6169	0.5352	0.5773
	NB	0.5219	0.5207	0.5217	0.5191	0.5207
	RF	0.7923	0.5341	0.9530	0.5378	0.8742
	SVM	0.0000	0.5384	0.5440	0.5151	0.5447
TwImbData18112016	J48	0.6647	0.5298	0.6080	0.5314	0.5771
	NB	0.5219	0.5210	0.5207	0.5186	0.5207
	RF	0.9904	0.5324	0.9780	0.5370	0.8652
	SVM	0.0000	0.5344	0.5396	0.5160	0.5403

TABLE 34: New Corpora - No Post-processing

TwReyes2013	J48	0.8933	0.8497	0.8725	0.8618	0.8726
	NB	0.7085	0.7146	0.7026	0.7065	0.7611
	RF	0.9140	0.8596	0.9146	0.9221	0.9023
	SVM	0.8819	0.8359	0.8373	0.7716	0.8466
TwIronyBarbieri2014	J48	0.8898	0.7977	0.8440	0.8460	0.8523
	NB	0.7407	0.7071	0.7087	0.6929	0.7179
	RF	0.9293	0.8601	0.9277	0.9291	0.9204
	SVM	0.8657	0.7965	0.8006	0.2000	0.8111
TwMohammad2015	J48	0.5467	0.5513	0.5407	0.5398	0.5420
	NB	0.2758	0.5855	0.5799	0.2758	0.5783
	RF	0.6093	0.5820	0.5915	0.6863	0.5928
	SVM	0.4758	0.5871	0.5824	0.2758	0.5748
TwSarcasmBarbieri2014	J48	0.9101	0.8432	0.8760	0.8711	0.8877
	NB	0.7997	0.7705	0.7665	0.7806	0.7652
	RF	0.9362	0.8749	0.9322	0.9404	0.9274
	SVM	0.8941	0.8331	0.8372	0.7452	0.8516
TwRiloff2013	J48	0.6458	0.6167	0.5972	0.6716	0.6094
	NB	0.2191	0.6385	0.6407	0.2191	0.6277
	RF	0.7315	0.6669	0.7270	0.7101	0.7083
	SVM	0.7514	0.6678	0.6732	0.2191	0.6711
TwPtáček2014	J48	0.7274	0.2916	0.2908	0.7683	0.2932
	NB	0.7316	0.3556	0.3558	0.6073	0.3351
	RF	0.7849	0.2587	0.2195	0.7610	0.2321
	SVM	0.7384	0.3071	0.3072	0.7316	0.3055

TABLE 35: Benchmark Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0736	0.5128	0.5325	0.0196	0.5235
	NB	0.0196	0.5096	0.5095	0.0196	0.5102
	RF	0.9662	0.5173	0.9905	0.8933	0.8388
	SVM	0.0196	0.5147	0.5160	0.0196	0.5147
TwImbData09112016	J48	0.7105	0.5330	0.6052	0.0196	0.5861
	NB	0.0196	0.5185	0.5192	0.0196	0.5191
	RF	0.9560	0.5370	0.9757	0.9842	0.7704
	SVM	0.0196	0.5329	0.5377	0.0196	0.5389
TwImbData10112016	J48	0.7076	0.5393	0.6435	0.0196	0.6074
	NB	0.0196	0.5220	0.5220	0.0196	0.5236
	RF	0.9605	0.5408	0.9767	0.7962	0.8213
	SVM	0.0196	0.5440	0.5493	0.0196	0.5527
TwImbData11112016	J48	0.7147	0.5332	0.6277	0.0196	0.5781
	NB	0.0196	0.5224	0.5226	0.0196	0.5233
	RF	0.9615	0.5362	0.9906	0.8433	0.8134
	SVM	0.0196	0.5381	0.5462	0.0196	0.5505
TwImbData12112016	J48	0.7058	0.5345	0.6259	0.0196	0.5906
	NB	0.0196	0.5215	0.5219	0.0196	0.5227
	RF	0.9720	0.5378	0.8835	0.7962	0.8421
	SVM	0.0196	0.5405	0.5467	0.0196	0.5501
TwImbData13112016	J48	0.7036	0.5335	0.6197	0.0196	0.5825
	NB	0.0196	0.5205	0.5203	0.0196	0.5213
	RF	0.9586	0.5373	0.9682	0.9905	0.8086
	SVM	0.0196	0.5391	0.5456	0.0196	0.5472
TwImbData14112016	J48	0.6967	0.5312	0.6223	0.0196	0.5862
	NB	0.0196	0.5199	0.5208	0.0196	0.5197
	RF	0.9558	0.5359	0.9762	0.9905	0.8278
	SVM	0.0196	0.5368	0.5424	0.0196	0.5456
TwImbData15112016	J48	0.7141	0.5318	0.6091	0.0196	0.5710
	NB	0.0196	0.5204	0.5210	0.0196	0.5201
	RF	0.9740	0.5355	0.9843	0.9905	0.8749
	SVM	0.0196	0.5356	0.5411	0.0196	0.5413
TwImbData16112016	J48	0.7009	0.5321	0.6254	0.0196	0.5835
	NB	0.0196	0.5205	0.5208	0.0196	0.5186
	RF	0.9623	0.5354	0.9572	0.9779	0.8293
	SVM	0.0196	0.5369	0.5438	0.0196	0.5468
TwImbData17112016	J48	0.6799	0.5323	0.6169	0.0196	0.5773
	NB	0.0196	0.5205	0.5213	0.0196	0.5216
	RF	0.9639	0.5360	0.8709	0.7962	0.8109
	SVM	0.0196	0.5384	0.5440	0.0196	0.5447
TwImbData18112016	J48	0.6647	0.5298	0.6080	0.0196	0.5771
	NB	0.0196	0.5208	0.5204	0.0196	0.5212
	RF	0.9459	0.5337	0.9780	0.7962	0.8099
	SVM	0.0196	0.5344	0.5396	0.0196	0.5403

TABLE 36: New Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8891	0.8514	0.8763	0.8902	0.8703
	NB	0.7064	0.7056	0.6944	0.7064	0.7542
	RF	0.9109	0.8456	0.9029	0.9219	0.8910
	SVM	0.8819	0.8359	0.8373	0.7716	0.8466
TwIronyBarbieri2014	J48	0.8676	0.7885	0.8495	0.8740	0.8442
	NB	0.7410	0.7041	0.7072	0.7410	0.7232
	RF	0.9286	0.8439	0.9272	0.9299	0.9154
	SVM	0.8657	0.7965	0.8006	0.2000	0.8111
TwMohammad2015	J48	0.5427	0.5489	0.5371	0.5143	0.5380
	NB	0.5776	0.5876	0.5775	0.5776	0.5801
	RF	0.5827	0.5970	0.5780	0.6057	0.5854
	SVM	0.4758	0.5871	0.5824	0.2758	0.5748
TwSarcasmBarbieri2014	J48	0.8996	0.8358	0.8815	0.8959	0.8824
	NB	0.8085	0.7737	0.7723	0.8085	0.7762
	RF	0.9335	0.8645	0.9254	0.9401	0.9219
	SVM	0.8941	0.8331	0.8372	0.7452	0.8516
TwRiloff2013	J48	0.6504	0.6186	0.5972	0.6608	0.6110
	NB	0.6329	0.6358	0.6379	0.6329	0.6393
	RF	0.6948	0.6402	0.6929	0.7465	0.6725
	SVM	0.7514	0.6678	0.6732	0.2191	0.6711
TwPtáček2014	J48	0.2734	0.2894	0.2876	0.2590	0.2925
	NB	0.3743	0.3577	0.3577	0.3743	0.3335
	RF	0.2225	0.2627	0.2320	0.2398	0.2403
	SVM	0.2616	0.3071	0.3072	0.7316	0.3055

TABLE 37: Benchmark Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0719	0.5129	0.5329	0.5023	0.5242
	NB	0.5031	0.5094	0.5096	0.5031	0.5104
	RF	0.7141	0.5141	0.7378	0.9779	0.5519
	SVM	0.0196	0.5147	0.5160	0.0196	0.5147
TwImbData09112016	J48	0.6377	0.5341	0.6058	0.5148	0.5831
	NB	0.5146	0.5187	0.5195	0.5146	0.5234
	RF	0.8282	0.5286	0.7977	0.9749	0.6710
	SVM	0.0196	0.5329	0.5377	0.0196	0.5389
TwImbData10112016	J48	0.6554	0.5398	0.6446	0.5649	0.6051
	NB	0.5230	0.5209	0.5209	0.5230	0.5278
	RF	0.7930	0.5299	0.8071	0.9661	0.6834
	SVM	0.0196	0.5440	0.5493	0.0196	0.5527
TwImbData11112016	J48	0.6492	0.5336	0.6282	0.5918	0.5759
	NB	0.5224	0.5218	0.5220	0.5224	0.5278
	RF	0.7451	0.5252	0.7138	0.8476	0.6310
	SVM	0.0196	0.5381	0.5462	0.0196	0.5505
TwImbData12112016	J48	0.6468	0.5352	0.6275	0.5383	0.5837
	NB	0.5288	0.5200	0.5200	0.5288	0.5258
	RF	0.8059	0.5261	0.7922	0.7963	0.6662
	SVM	0.0196	0.5405	0.5467	0.0196	0.5501
TwImbData13112016	J48	0.6551	0.5339	0.6210	0.5169	0.5792
	NB	0.5211	0.5199	0.5196	0.5211	0.5253
	RF	0.7814	0.5262	0.7716	0.9805	0.6424
	SVM	0.0196	0.5391	0.5456	0.0196	0.5472
TwImbData14112016	J48	0.6327	0.5321	0.6233	0.5530	0.5832
	NB	0.5159	0.5198	0.5201	0.5159	0.5236
	RF	0.7990	0.5262	0.7463	0.9781	0.6440
	SVM	0.0196	0.5368	0.5424	0.0196	0.5456
TwImbData15112016	J48	0.6425	0.5324	0.6108	0.4580	0.5690
	NB	0.5191	0.5192	0.5198	0.5191	0.5243
	RF	0.8027	0.5262	0.7850	0.9905	0.6518
	SVM	0.0196	0.5356	0.5411	0.0196	0.5413
TwImbData16112016	J48	0.6411	0.5332	0.6262	0.5227	0.5776
	NB	0.5122	0.5199	0.5206	0.5122	0.5215
	RF	0.8020	0.5265	0.7800	0.9429	0.6432
	SVM	0.0196	0.5369	0.5438	0.0196	0.5468
TwImbData17112016	J48	0.6321	0.5332	0.6175	0.5250	0.5750
	NB	0.5192	0.5196	0.5199	0.5192	0.5263
	RF	0.7828	0.5271	0.7802	0.7962	0.6539
	SVM	0.0196	0.5384	0.5440	0.0196	0.5447
TwImbData18112016	J48	0.6167	0.5303	0.6090	0.5387	0.5752
	NB	0.5157	0.5202	0.5199	0.5157	0.5252
	RF	0.7681	0.5252	0.7805	0.7613	0.6304
	SVM	0.0196	0.5344	0.5396	0.0196	0.5403

TABLE 38: New Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013	J48	0.8923	0.8497	0.8725	0.8481	0.8726
	NB	0.6891	0.6994	0.6886	0.6891	0.7420
	RF	0.9108	0.8643	0.9064	0.9223	0.8997
	SVM	0.8819	0.8359	0.8373	0.7716	0.8466
TwIronyBarbieri2014	J48	0.8874	0.7977	0.8440	0.7980	0.8523
	NB	0.7031	0.6990	0.7020	0.7031	0.7602
	RF	0.9279	0.8450	0.9230	0.9319	0.9159
	SVM	0.8657	0.7965	0.8006	0.6566	0.8111
TwMohammad2015	J48	0.5507	0.5513	0.5407	0.5536	0.5420
	NB	0.5790	0.5810	0.5743	0.5790	0.5856
	RF	0.5827	0.5970	0.5780	0.6057	0.5854
	SVM	0.4758	0.5871	0.5824	0.2758	0.5748
TwSarcasmBarbieri2014	J48	0.9085	0.8432	0.8760	0.8372	0.8877
	NB	0.7686	0.7707	0.7689	0.7686	0.8023
	RF	0.9335	0.8645	0.9254	0.9401	0.9219
	SVM	0.8941	0.8331	0.8372	0.7452	0.8516
TwRiloff2013	J48	0.6522	0.6167	0.5972	0.6418	0.6094
	NB	0.6366	0.6309	0.6351	0.6366	0.6370
	RF	0.6948	0.6402	0.6929	0.7465	0.6725
	SVM	0.7514	0.6678	0.6732	0.6135	0.6711
TwPtáček2014	J48	0.2728	0.2916	0.2908	0.3011	0.2932
	NB	0.3451	0.3456	0.3453	0.3451	0.3349
	RF	0.2225	0.2627	0.2320	0.2463	0.2403
	SVM	0.2616	0.3071	0.3072	0.7316	0.3055

TABLE 39: Benchmark Corpora - Youden Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwImbData08112016	J48	0.0693	0.5128	0.5325	0.5148	0.5235
	NB	0.5097	0.5093	0.5095	0.5097	0.5106
	RF	0.7141	0.5141	0.7378	0.9779	0.5519
	SVM	0.0196	0.5147	0.5160	0.0196	0.5147
TwImbData09112016	J48	0.5243	0.5330	0.6052	0.5328	0.5861
	NB	0.5180	0.5175	0.5178	0.5180	0.5212
	RF	0.8282	0.5286	0.7977	0.9749	0.6710
	SVM	0.0196	0.5329	0.5377	0.5166	0.5389
TwImbData10112016	J48	0.5500	0.5393	0.6435	0.5421	0.6074
	NB	0.5193	0.5192	0.5193	0.5193	0.5247
	RF	0.7930	0.5299	0.8071	0.9661	0.6834
	SVM	0.0196	0.5440	0.5493	0.5164	0.5527
TwImbData11112016	J48	0.5401	0.5332	0.6277	0.5352	0.5781
	NB	0.5187	0.5196	0.5188	0.5187	0.5238
	RF	0.7451	0.5252	0.7138	0.8476	0.6310
	SVM	0.0196	0.5381	0.5462	0.5141	0.5505
TwImbData12112016	J48	0.5320	0.5345	0.6259	0.5343	0.5906
	NB	0.5190	0.5183	0.5190	0.5190	0.5215
	RF	0.8059	0.5261	0.7922	0.7963	0.6662
	SVM	0.0196	0.5405	0.5467	0.5163	0.5501
TwImbData13112016	J48	0.5327	0.5335	0.6197	0.5340	0.5825
	NB	0.5175	0.5184	0.5181	0.5175	0.5220
	RF	0.7814	0.5262	0.7716	0.9805	0.6424
	SVM	0.0196	0.5391	0.5456	0.5159	0.5472
TwImbData14112016	J48	0.5396	0.5312	0.6223	0.5333	0.5862
	NB	0.5185	0.5183	0.5187	0.5185	0.5207
	RF	0.7990	0.5262	0.7463	0.9781	0.6440
	SVM	0.0196	0.5368	0.5424	0.5161	0.5456
TwImbData15112016	J48	0.5389	0.5318	0.6091	0.5326	0.5710
	NB	0.5189	0.5185	0.5189	0.5189	0.5207
	RF	0.8027	0.5262	0.7850	0.9905	0.6518
	SVM	0.0196	0.5356	0.5411	0.5151	0.5413
TwImbData16112016	J48	0.5281	0.5321	0.6254	0.5341	0.5835
	NB	0.5185	0.5183	0.5183	0.5185	0.5205
	RF	0.8020	0.5265	0.7800	0.9429	0.6432
	SVM	0.0196	0.5369	0.5438	0.5162	0.5468
TwImbData17112016	J48	0.5434	0.5323	0.6169	0.5366	0.5773
	NB	0.5188	0.5182	0.5189	0.5188	0.5228
	RF	0.7828	0.5271	0.7802	0.7962	0.6539
	SVM	0.0196	0.5384	0.5440	0.5151	0.5447
TwImbData18112016	J48	0.5233	0.5298	0.6080	0.5321	0.5771
	NB	0.5176	0.5176	0.5177	0.5176	0.5231
	RF	0.7681	0.5252	0.7805	0.7613	0.6304
	SVM	0.0196	0.5344	0.5396	0.5160	0.5403

TABLE 40: New Corpora - Youden Thresholding