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.

\mathbf{AUC}

$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
TwReyes2013 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 RF 0.9598 0.9569 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 NB 0.6364 0.6424 0.6330 0.6364 0.6099 RF 0.6561 0.6479 0.6337 0.6497 0.6521 SVM 0.5075 0.6046 0.5984 0.5000 0.5680 TwSarcasmBarbieri2014 NB 0.9072 0.9020 0.8759 0.9304 0.8936 RF 0.9097 0.9113 0.9099 0.9097 0.8864 RF 0.9755 0.9732 0.9766 0.9754 0.975		J48	0.8914	0.8917	0.8662	0.9172	0.8780
RF 0.9/24 0.9693 0.9722 0.9711 0.9718 SVM 0.8601 0.8810 0.8819 0.8511 0.8821 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 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 NB 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	TD0012	NB	0.8198	0.8341	0.8203	0.8198	0.8595
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wReyes2013	RF	0.9724	0.9693	0.9722	0.9711	0.9718
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.8601	0.8810	0.8819	0.8511	0.8821
TwlronyBarbieri2014 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 MB 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 NB 0.7471 0.7446 0.7478 0.7471 0.7331 RF 0.8006 0.8020 0.8125 0.8078 0.8063 <td< td=""><td></td><td>J48</td><td>0.8807</td><td>0.8648</td><td>0.8490</td><td>0.9074</td><td>0.8658</td></td<>		J48	0.8807	0.8648	0.8490	0.9074	0.8658
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 NB 0.9072 0.9020 0.8759 0.9304 0.8936 RF 0.9755 0.9732 0.9766 0.9754 0.9758 SVM 0.8784 0.8943 0.8969 0.8599 0.8981 TwRiloff2013 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 NB <td>TInonPonkioni2014</td> <td>NB</td> <td>0.8522</td> <td>0.8490</td> <td>0.8519</td> <td>0.8522</td> <td>0.8502</td>	TInonPonkioni2014	NB	0.8522	0.8490	0.8519	0.8522	0.8502
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wirony barbieri 2014	RF	0.9598	0.9569	0.9618	0.9603	0.9609
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.8250	0.8435	0.8447	0.7401	0.8476
TwMohammad2015 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 RF 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 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 NB 0.7127 0.7376 0.6843 0.7918 0.7234 TwPtáček2014 NB 0.7390 0.7378 0.7386 0.7390 0.7660		J48	0.5344	0.5710	0.5168	0.5555	0.5355
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwMohammad2015	NB	0.6364	0.6424	0.6330	0.6364	0.6099
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w Monammad 2015	RF	0.6561	0.6479	0.6397	0.6497	0.6521
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.5075	0.6046	0.5984	0.5000	0.5680
TwSarcasmBarbieri2014 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 RF 0.8006 0.8020 0.8125 0.8078 0.8063 SVM 0.5485 0.7326 0.7339 0.6230 0.7317 J48 0.7127 0.7376 0.6843 0.7918 0.7234 TwPtáček2014 NB 0.7390 0.7378 0.7386 0.7390 0.7660 RF 0.8885 0.8844 0.8886 0.8778 0.8878		J48	0.9072	0.9020	0.8759	0.9304	0.8936
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwSarcacmBarbiori2014	NB	0.9097	0.9113	0.9099	0.9097	0.8864
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w5a1CasiiiDa1bie112014	RF	0.9755	0.9732	0.9766	0.9754	0.9758
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.8784	0.8943	0.8969	0.8599	0.8981
TwRiloff2013 RF 0.8006 0.8020 0.8125 0.8078 0.8063 SVM 0.5485 0.7326 0.7339 0.6230 0.7317 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		J48	0.6194	0.6565	0.6067	0.7319	0.6006
TwPtáček2014 RF 0.8006 0.8020 0.8125 0.8078 0.8063 0.7317 0.7326 0.7339 0.6230 0.7317 0.7376 0.6843 0.7918 0.7234 0.7390 0.7378 0.7386 0.7390 0.7660 0.8885 0.8844 0.8886 0.8778 0.8878	TwPiloff2013	NB	0.7471	0.7446	0.7478	0.7471	0.7331
TwPtáček2014 J48 0.7127 0.7376 0.6843 0.7918 0.7234 0.7390 0.7378 0.7386 0.7390 0.7660 0.8885 0.8844 0.8886 0.8778 0.8878	1 WIGHOREOUS	RF	0.8006	0.8020	0.8125	0.8078	0.8063
TwPtáček2014 NB 0.7390 0.7378 0.7386 0.7390 0.7660 RF 0.8885 0.8844 0.8886 0.8778 0.8878		SVM	0.5485	0.7326	0.7339	0.6230	0.7317
TwPtáček2014 RF 0.8885 0.8844 0.8886 0.8778 0.8878		J48	0.7127	0.7376	0.6843	0.7918	0.7234
RF 0.8885 0.8844 0.8886 0.8778 0.8878	TwPtscale2014	NB	0.7390	0.7378	0.7386	0.7390	0.7660
SVM 0.6647 0.7450 0.7448 0.5000 0.7450	IWF taceK2014	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
	J48	0.5079	0.6385	0.5469	0.7199	0.6562
F. I. I.D	NB	0.6620	0.6605	0.6622	0.6620	0.6416
TwImbData08112016	RF	0.7590	0.7805	0.7809	0.7956	0.7748
	SVM	0.5000	0.6640	0.6759	0.5000	0.6646
	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
TwImbData09112016	SVM	0.5000	0.8114	0.8315	0.7090	0.8264
	J48	0.7440	0.8245	0.6766	0.8882	0.7405
T. I. I.D. 4 10119016	NB	0.7973	0.7948	0.7970	0.7973	0.8492
TwImbData10112016	RF	0.9231	0.9267	0.9346	0.9350	0.9320
	SVM	0.5000	0.8360	0.8593	0.7052	0.8529
	J48	0.7937	0.7882	0.6592	0.8701	0.6856
T I I D / 11110016	NB	0.7977	0.7965	0.7978	0.7977	0.8380
TwImbData11112016	RF	0.9024	0.9081	0.9188	0.9211	0.9089
	SVM	0.5000	0.8134	0.8477	0.6563	0.8310
	J48	0.7784	0.8174	0.6549	0.8596	0.7226
T I I D / 10110016	NB	0.7961	0.7933	0.7961	0.7960	0.8224
TwImbData12112016	RF	0.8937	0.9118	0.9206	0.9225	0.9103
	SVM	0.5000	0.8227	0.8427	0.7043	0.8313
	J48	0.7734	0.7945	0.6573	0.8628	0.7171
T. I. I.D. 4 19119016	NB	0.7885	0.7910	0.7888	0.7885	0.8253
TwImbData13112016	RF	0.9074	0.9151	0.9239	0.9241	0.9186
	SVM	0.5000	0.8150	0.8431	0.6963	0.8313
	J48	0.7662	0.7908	0.6495	0.8675	0.7154
T. I. I.D. 4 14119016	NB	0.7875	0.7820	0.7876	0.7875	0.8145
TwImbData14112016	RF	0.9116	0.9129	0.9226	0.9203	0.9183
	SVM	0.5000	0.8196	0.8399	0.6996	0.8323
	J48	0.7785	0.7907	0.6292	0.8608	0.6821
T. I. I.D. + 15110016	NB	0.7884	0.7851	0.7883	0.7884	0.8234
TwImbData15112016	RF	0.8973	0.9059	0.9189	0.9142	0.9165
	SVM	0.5000	0.8069	0.8315	0.6783	0.8272
-	J48	0.7820	0.7860	0.6553	0.8652	0.7175
T I I D + 10110010	NB	0.7901	0.7848	0.7894	0.7901	0.8026
TwImbData16112016	RF	0.9036	0.9109	0.9212	0.9180	0.9152
	SVM	0.5000	0.8149	0.8442	0.7024	0.8332
	J48	0.7435	0.7936	0.6443	0.8680	0.7236
m I ID : 45440010	NB	0.7917	0.7878	0.7917	0.7917	0.8372
TwImbData17112016	RF	0.9003	0.9084	0.9127	0.9179	0.9184
	SVM	0.5000	0.8146	0.8405	0.6792	0.8342
	J48	0.7484	0.7803	0.6393	0.8505	0.7068
T I I D + 10112012	NB	0.7874	0.7904	0.7871	0.7874	0.8319
TwImbData18112016	RF	0.8859	0.9027	0.9109	0.9135	0.9080
	SVM	0.5000	0.8116	0.8351	0.6981	0.8274
	DID 9. Nor. C					

Table 2: New Corpora - No Post-processing

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
	J48	0.8914	0.8917	0.8662	0.9172	0.8780
Tw.Powog9019	NB	0.8198	0.8341	0.8203	0.8198	0.8595
TwReyes2013	RF	0.9724	0.9693	0.9722	0.9706	0.9718
	SVM	0.8601	0.8810	0.8819	0.8511	0.8821
	J48	0.8807	0.8648	0.8490	0.9074	0.8658
Tw.Inony.Panhioni2014	NB	0.8522	0.8490	0.8519	0.8522	0.8503
TwIronyBarbieri2014	RF	0.8972	0.9239	0.9400	0.8309	0.9272
	SVM	0.8250	0.8435	0.8447	0.7401	0.8476
	J48	0.5344	0.5710	0.5168	0.5555	0.5355
TwMohammad2015	NB	0.6364	0.6424	0.6330	0.6364	0.6101
1 wMonammad2015	RF	0.6531	0.6478	0.6359	0.6221	0.6509
	SVM	0.5075	0.6046	0.5984	0.5000	0.5680
	J48	0.9072	0.9020	0.8759	0.9304	0.8936
TwSarcasmBarbieri2014	NB	0.9097	0.9113	0.9099	0.9097	0.8866
1 wsarcasiiibarbieri2014	RF	0.9145	0.9371	0.9301	0.8831	0.9292
	SVM	0.8784	0.8943	0.8969	0.8599	0.8981
	J48	0.6194	0.6565	0.6067	0.7319	0.6006
TwRiloff2013	NB	0.7471	0.7446	0.7478	0.7471	0.7325
1 WKH0H2013	RF	0.7586	0.7925	0.7866	0.6654	0.7913
	SVM	0.5485	0.7326	0.7339	0.6230	0.7317
	J48	0.2873	0.7376	0.6843	0.2082	0.7234
TwPtáček2014	NB	0.2610	0.7378	0.7386	0.2610	0.7659
I WF tacekZU14	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

TwImbData08112016 NB NB N6620 N6605 N6602 N6802	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
NumbData08112016 RF 0.7515 0.7753 0.7809 0.5125 0.7681		J48	0.5079	0.6385	0.5469	0.7199	0.6562
RF	TII D-+-00110016	NB	0.6620	0.6605	0.6622	0.6620	0.6415
TwImbData10112016 TwImbData09112016 TwImbData09112016 TwImbData09112016 TwImbData09112016 TwImbData09112016 TwImbData09112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData11112016 TwImbData111112016 TwImbData11	TwImbData08112016	RF	0.7515	0.7753	0.7809	0.5125	0.7681
TwImbData09112016 NB RF 0.7814 0.7758 0.7807 0.7814 0.8105 0.8454 SVM 0.5000 0.8114 0.8315 0.7090 0.8264 J48 0.7440 0.8245 0.6766 0.8882 0.7405 TwImbData10112016 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 MB 0.7977 0.7965 0.9777 0.7976 0.8529 TwImbData11112016 NB 0.7975 0.7965 0.7977 0.7976 0.8520 RF 0.8952 0.8964 0.9188 0.5285 0.8703 SVM 0.5000 0.8134 0.8477 0.6563 0.8310 TwImbData12112016 NB 0.7784 0.8174 0.6549 0.8596 0.7226 TwImbData13112016 NB 0.7884 0.7992 0.8427		SVM	0.5000	0.6640	0.6759	0.5000	0.6646
NumbData09112016 RF 0.7997 0.9100 0.8904 0.5237 0.8454		J48	0.7749	0.7863	0.6383	0.8597	0.7226
RF 0.7997 0.9100 0.8904 0.5237 0.8454 SVM 0.5000 0.8114 0.8315 0.7090 0.8264 TwImbData10112016 NB 0.7440 0.8245 0.6766 0.8882 0.7405 RF 0.8426 0.9159 0.9170 0.5223 0.8663 SVM 0.5000 0.8360 0.8593 0.7052 0.8529 MB 0.7937 0.7882 0.6592 0.8701 0.6856 RF 0.8952 0.8964 0.9188 0.5285 0.8301 SVM 0.5000 0.8134 0.8477 0.6563 0.8331 TwImbData12112016 NB 0.7796 0.7960 0.7226 0.7960 0.7960 0.7226 TwImbData13112016 NB 0.7960 0.7932 0.7960 0.7960 0.8166 RF 0.8350 0.9054 0.952 0.5217 0.8727 TwImbData13112016 NB 0.7784 0.7945 0.6573 0.8628	TII D-+-00119016	NB	0.7814	0.7758	0.7807	0.7814	0.8105
TwImbData10112016 TwImbData10112016 RF NB NB N-7973 N-7947 N-7969 N-7973 N-8623 N-8635 N-7940 N-8593 N-7952 N-8522 N-8523 N-8522 N-8523	1 w1mbData09112016	RF	0.7997	0.9100	0.8904	0.5237	0.8454
TwImbData10112016 NB RF 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 0.8523 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 NB 0.7784 0.8174 0.6549 0.8596 0.7226 RF 0.8350 0.9054 0.9052 0.5217 0.8727 SVM 0.5000 0.8227 0.8427 0.7043 0.8313 TwImbData13112016 RF 0.8780 0.9010 0.9952 0.5217 0.8727 TwImbData14112016 RF 0.8784 0.7909 0.7887 0.7884 0.8211		SVM	0.5000	0.8114	0.8315	0.7090	0.8264
Name	-	J48	0.7440	0.8245	0.6766	0.8882	0.7405
NF	TIh Doto 10119016	NB	0.7973	0.7947	0.7969	0.7973	0.8435
TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData13112016	1 wimbData10112016	RF	0.8426	0.9159	0.9170	0.5223	0.8663
TwImbData11112016 NB RF 0.7975 0.7965 0.7967 0.7977 0.7976 0.8327 0.8327 0.8964 0.9188 0.5285 0.8703 SVM 0.5000 0.8134 0.8477 0.6563 0.8310 J48 0.7784 0.8174 0.6549 0.8596 0.7226 NB D.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 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 J48 0.7662 0.7908 0.6495 0.8675 0.7154 RF 0.8973 0.9075 0.9226 0.5306 0.8863 SVM 0.5000 0.8150 0.8431 0.6963 0.8313 J48 0.7662 0.7908 0.6495 0.8675 0.7154 TwImbData14112016 RF 0.8973 0.9075 0.9226 0.5306 0.8863 RF 0.8973 0.9075 0.9226 0.5306 0.8863 SVM 0.5000 0.8196 0.8399 0.6996 0.8323 J48 0.7785 0.7907 0.6292 0.8608 0.6821 TwImbData15112016 RF 0.8682 0.8988 0.9146 0.5248 0.8798 RF 0.8682 0.8988 0.9146 0.5248 0.8798 SVM 0.5000 0.8149 0.8342 0.7892 0.7899 0.7942 TwImbData16112016 RF 0.8667 0.9015 0.9201 0.5255 0.8997 NB 0.7890 0.7845 0.7892 0.7899 0.7942 0.7896 0.7902 0.7899 0.7942 TwImbData17112016 RF 0.8667 0.9015 0.9201 0.5255 0.8997 NB 0.7892 0.7908		SVM	0.5000	0.8360	0.8593	0.7052	0.8529
NumbData11112016 RF 0.8952 0.8964 0.9188 0.5285 0.8703 SVM 0.5000 0.8134 0.8477 0.6563 0.8310 0.8104 0.8174 0.6549 0.8596 0.7226 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 0.8313 0.8313 0.7734 0.7945 0.6573 0.8628 0.7171 0.8727 0.8728 0.8734 0.7945 0.6573 0.8628 0.7171 0.8728 0.8734 0.7945 0.6573 0.8628 0.7171 0.8728 0.8734 0.7945 0.6573 0.8628 0.7171 0.8728 0.8042	-	J48	0.7937	0.7882	0.6592	0.8701	0.6856
NB	TIh Doto 11112016	NB	0.7975	0.7965	0.7977	0.7976	0.8327
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wiiiibData11112010	RF	0.8952	0.8964	0.9188	0.5285	0.8703
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.5000	0.8134	0.8477	0.6563	0.8310
NB		J48	0.7784	0.8174	0.6549	0.8596	0.7226
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData12112016	NB	0.7960	0.7932	0.7960	0.7960	0.8166
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wimbData12112010	RF	0.8350	0.9054	0.9052	0.5217	0.8727
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.5000	0.8227	0.8427	0.7043	0.8313
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.7734	0.7945	0.6573	0.8628	0.7171
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TIh Doto 12112016	NB	0.7884	0.7909	0.7887	0.7884	0.8214
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 willibData15112010	RF	0.8780	0.9101	0.9192	0.5244	0.8956
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.5000	0.8150	0.8431	0.6963	0.8313
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.7662	0.7908	0.6495	0.8675	0.7154
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData14112016	NB	0.7872	0.7818	0.7873	0.7872	0.8053
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wiiiibData14112010	RF	0.8973	0.9075	0.9226	0.5306	0.8863
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$			0.5000			0.6996	0.8323
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.7785		0.6292	0.8608	0.6821
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData15112016		0.7883	0.7849	0.7882	0.7883	0.8166
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 willibData15112010	RF	0.8682	0.8988	0.9146	0.5248	0.8798
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$							
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$			0.7820	0.7860	0.6553		
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData16112016		0.7899	0.7845	0.7892	0.7899	0.7942
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIDData10112010	RF	0.8607	0.9015	0.9201	0.5255	0.8997
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$			0.5000	0.8149	0.8442	0.7024	0.8332
TwImbData17112016 RF 0.8678 0.9008 0.9006 0.5153 0.8855 SVM 0.5000 0.8146 0.8405 0.6792 0.8342 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			0.7435	0.7936			0.7236
RF 0.8678 0.9008 0.9006 0.5153 0.8855 SVM 0.5000 0.8146 0.8405 0.6792 0.8342 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	TwImbData17112016	NB	0.7914	0.7876	0.7915	0.7914	0.8313
TwImbData18112016	1 WIIIDData1 (112010	RF	0.8678	0.9008	0.9006	0.5153	0.8855
TwImbData18112016			0.5000	0.8146	0.8405		
1WImbData18112016 RF 0.8607 0.8944 0.8840 0.5186 0.8954			0.7484		$0.6\overline{393}$	0.8505	
RF $0.8607 0.8944 0.8840 0.5186 0.8954$	TwImbData18119016	NB	0.7872	0.7902	0.7869	0.7872	0.8254
SVM 0.5000 0.8116 0.8351 0.6981 0.8274	1 willionarator12010		0.8607	0.8944	0.8840	0.5186	
		SVM	0.5000	0.8116	0.8351	0.6981	0.8274

Table 4: New Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
	J48	0.8914	0.8917	0.8662	0.9172	0.8780
Tw.Powo2012	NB	0.8198	0.8341	0.8203	0.8198	0.8595
TwReyes2013	RF	0.9724	0.9693	0.9722	0.9711	0.9718
	SVM	0.8601	0.8810	0.8819	0.8511	0.8821
	J48	0.8807	0.8648	0.8490	0.9074	0.8658
Tw.Inony.Panhioni2014	NB	0.8522	0.8490	0.8519	0.8522	0.8502
TwIronyBarbieri2014	RF	0.9598	0.9569	0.9618	0.9603	0.9609
	SVM	0.8250	0.8435	0.8447	0.7401	0.8476
	J48	0.5344	0.5710	0.5168	0.5555	0.5355
TwMohammad2015	NB	0.6364	0.6424	0.6330	0.6364	0.6099
1 wMohammad2015	RF	0.6561	0.6479	0.6397	0.6497	0.6521
	SVM	0.5075	0.6046	0.5984	0.5000	0.5680
	J48	0.9072	0.9020	0.8759	0.9304	0.8936
TwSarcasmBarbieri2014	NB	0.9097	0.9113	0.9099	0.9097	0.8864
1 wsarcasin barbieri 2014	RF	0.9755	0.9732	0.9766	0.9754	0.9758
	SVM	0.8784	0.8943	0.8969	0.8599	0.8981
	J48	0.6194	0.6565	0.6067	0.7319	0.6006
TwRiloff2013	NB	0.7471	0.7446	0.7478	0.7471	0.7331
1 WKIIOII2015	RF	0.8006	0.8020	0.8125	0.8078	0.8063
	SVM	0.5485	0.7326	0.7339	0.6230	0.7317
	J48	0.7127	0.7376	0.6843	0.7918	0.7234
TwPtáček2014	NB	0.7390	0.7378	0.7386	0.7390	0.7660
1 WF taceK2U14	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

TwImbData08112016 NB NB 0.6620 0.6605 0.66022 0.66020 0.6612 0.6612 0.7748 0.7590 0.7805 0.7805 0.7809 0.7950 0.7805 0.7809 0.7950 0.7805 0.7809 0.7950 0.7805 0.7809 0.7950 0.7805 0.7809 0.7950 0.7805 0.7809 0.7950 0.7805 0.7809 0.7950 0.7805 0.7809 0.7950 0.7805 0.7809 0.7950 0.7805 0.7809 0.7950 0.7748 0.7805 0.7809 0.7815 0.7759 0.7809 0.7815 0.7875 0.7809 0.7815 0.7875 0.7809 0.7815 0.7875 0.7809 0.7815 0.7875 0.7809 0.7815 0.7875 0.7809 0.7815 0.7875 0.7809 0.7815 0.7809 0.8144 0.8235 0.9235 0.9260 0.9203 0.8264 0.9235 0.9260 0.9203 0.8264 0.9231 0.9267 0.9346 0.9330 0.9320 0.9330 0.9320 0.93	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
NumbData08112016 RF		J48	0.5079	0.6385	0.5469	0.7199	0.6562
RF	TImbData08112016	NB	0.6620	0.6605	0.6622	0.6620	0.6416
TwImbData10112016 TwImbData09112016 TwImbData09112016 RF	1 wiiiibData08112010	RF	0.7590	0.7805	0.7809	0.7956	0.7748
TwimbData09112016 NB RF 0.7815 0.7859 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 NB 0.7440 0.8245 0.6766 0.8882 0.7405 RF 0.9231 0.9267 0.9346 0.9350 0.9320 SVM 0.5000 0.8360 0.8593 0.7052 0.8529 MB 0.7937 0.7882 0.6592 0.8701 0.6856 RF 0.9024 0.9081 0.9188 0.9211 0.0856 RF 0.9024 0.9081 0.9188 0.9211 0.9089 SVM 0.5000 0.8134 0.8477 0.6563 0.8310 TwimbData12112016 NB 0.7784 0.8174 0.6549 0.8596 0.7226 TwimbData1212016 NB 0.7895 0.9118 0.9225 0.9103 0.9427		SVM	0.5000	0.6640	0.6759	0.5000	0.6646
NumbData09112016 RF 0.9105 0.9194 0.9235 0.9260 0.9203 NumbData09112016 SVM 0.5000 0.8114 0.8315 0.7090 0.8264 NumbData10112016 RF 0.9231 0.9267 0.9346 0.9350 0.9320 NumbData10112016 RF 0.9231 0.9267 0.9346 0.9350 0.9320 NumbData11112016 NB 0.7937 0.7882 0.6592 0.8701 0.6856 NumbData11112016 RF 0.9024 0.9081 0.9188 0.9111 0.9089 NumbData11112016 RF 0.9024 0.9081 0.9188 0.9211 0.9089 NumbData12112016 NB 0.7784 0.8174 0.6549 0.8526 0.8310 NumbData12112016 RF 0.8937 0.9118 0.9266 0.9225 0.9103 NumbData13112016 NB 0.7784 0.8174 0.6549 0.8526 0.8204 NumbData13112016 RF 0.8937 0.9118 0.9206 0.9225 0.9103 NumbData13112016 NB 0.7734 0.7945 0.6573 0.8628 0.8533 NumbData13112016 NB 0.7734 0.7945 0.6573 0.8628 0.8253 NumbData13112016 NB 0.7785 0.7910 0.7888 0.7885 0.8253 NumbData14112016 NB 0.7885 0.7910 0.7888 0.7885 0.8253 NumbData14112016 NB 0.7875 0.7910 0.7888 0.7885 0.8253 NumbData14112016 NB 0.7875 0.7908 0.6495 0.8675 0.8145 NumbData14112016 NB 0.7875 0.7820 0.7876 0.7875 0.8145 NumbData14112016 NB 0.7875 0.7820 0.7876 0.7875 0.8145 NumbData14112016 NB 0.7884 0.7851 0.7884 0.7851 0.7884 0.8234 NumbData16112016 NB 0.7884 0.7885 0.9059 0.9189 0.9142 0.9165 NB 0.7884 0.7885 0.7860 0.8553 0.8652 0.7175 NB 0.7917 0.7848 0.7894 0.7901 0.8026 NB 0.7917 0.7848 0.7894 0.7901 0.8026 NB 0.7917 0.7848 0.7894 0.7917 0.7917 0.8372 NUmbData17112016 NB 0.7917 0.7848 0.7894 0.7917 0.7917 0.8372 0.9036 0.9094 0.9127 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0.9180 0.9152 0		J48	0.7749	0.7863	0.6383	0.8597	0.7226
RF 0.9105 0.9194 0.9235 0.9260 0.89264 SVM 0.5000 0.8114 0.8315 0.7090 0.8264 TwImbData10112016 NB 0.7440 0.8245 0.6766 0.8882 0.7405 RF 0.9231 0.9267 0.9346 0.9350 0.9320 SVM 0.5000 0.8360 0.8593 0.7052 0.8529 MB 0.7937 0.7882 0.6592 0.8701 0.6856 FWImbData11112016 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 0.8214 TwImbData12112016 NB 0.7784 0.8174 0.6549 0.8596 0.7226 TwImbData13112016 NB 0.7961 0.7933 0.7961 0.7960 0.8224 TwImbData13112016 NB 0.7885 0.7910 0.7888 0.7	TwImbData00112016		0.7815	0.7759	0.7809	0.7815	0.8164
TwImbData10112016 TwImbData10112016 RF 0.9231 0.9267 0.9346 0.9350 0.9320 SVM 0.5000 0.8360 0.8593 0.7052 0.8529 TwImbData11112016 RF 0.9231 0.9267 0.9346 0.9350 0.9320 SVM 0.5000 0.8360 0.8593 0.7052 0.8529 J48 0.7937 0.7882 0.6592 0.8701 0.6856 RF 0.9024 0.9081 0.9188 0.9211 0.9089 SVM 0.5000 0.8134 0.8477 0.6563 0.8310 J48 0.7784 0.8174 0.6549 0.8596 0.7226 RF 0.8937 0.9118 0.9206 0.9225 0.9103 SVM 0.5000 0.8227 0.8427 0.7043 0.8313 J48 0.7784 0.9118 0.9206 0.9225 0.9103 SVM 0.5000 0.8227 0.8427 0.7043 0.8313 J48 0.7784 0.9151 0.9239 0.9241 0.9188 SVM 0.5000 0.8150 0.8431 0.6963 0.8313 TwImbData14112016 RF 0.9074 0.9151 0.9239 0.9241 0.9186 SVM 0.5000 0.8150 0.8431 0.6963 0.8313 J48 0.7865 0.7910 0.7888 0.7875 0.8253 RF 0.9014 0.9150 0.9239 0.9241 0.9186 SVM 0.5000 0.8150 0.8431 0.6963 0.8313 J48 0.7865 0.7900 0.8437 0.6966 0.8252 TwImbData14112016 RF 0.9016 0.9129 0.9226 0.9203 0.9183 SVM 0.5000 0.8196 0.8399 0.6996 0.8323 J48 0.7884 0.7851 0.7883 0.7884 0.8234 TwImbData15112016 RF 0.9016 0.9129 0.9226 0.9203 0.9183 SVM 0.5000 0.8196 0.8399 0.6996 0.8323 TwImbData16112016 RF 0.8973 0.9059 0.9189 0.9142 0.9165 SVM 0.5000 0.8069 0.8315 0.6783 0.8272 TwImbData16112016 RF 0.9036 0.9109 0.9212 0.9180 0.9152 SVM 0.5000 0.8140 0.8442 0.7024 0.8332 TwImbData17112016 RF 0.9036 0.9109 0.9212 0.9180 0.9152 SVM 0.5000 0.8146 0.8405 0.6792 0.8342 TwImbData17112016 RF 0.9003 0.9084 0.9127 0.9179 0.9184 SVM 0.5000 0.8146 0.8405 0.6792 0.8342 TwImbData18112016 RF 0.9003 0.9084 0.9127 0.9179 0.9184 SVM 0.5000 0.8146 0.8405 0.6792 0.8342 TwImbData18112016 RF 0.9003 0.9084 0.9127 0.9179 0.9184 SVM 0.5000 0.8146 0.8405 0.6792 0.8342 TwImbData18112016	1 willibData09112010	RF	0.9105	0.9194	0.9235	0.9260	0.9203
TwImbData10112016 NB RF 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 J48 0.7977 0.7982 0.6592 0.8701 0.6856 TwImbData11112016 NB 0.7977 0.7965 0.7977 0.3380 TwImbData12112016 RF 0.9024 0.9081 0.9188 0.9211 0.9089 TwImbData12112016 NB 0.7784 0.8174 0.6563 0.8310 TwImbData12112016 NB 0.7784 0.8174 0.6563 0.8310 TwImbData12112016 RF 0.8937 0.9118 0.9206 0.9225 0.9103 TwImbData13112016 RF 0.8937 0.9118 0.9206 0.9225 0.9103 TwImbData14112016 RF 0.9074 0.9151 0.9380 0.941 0.9168 SVM 0.5000 0.8150		SVM	0.5000	0.8114	0.8315	0.7090	0.8264
NumbData10112016 RF 0.9231 0.9267 0.9346 0.9350 0.9320		J48	0.7440	0.8245	0.6766	0.8882	0.7405
NF	Tw.ImbData10112016	NB	0.7973	0.7948	0.7970	0.7973	0.8492
TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData13112016	1 willibData10112010	RF	0.9231	0.9267	0.9346	0.9350	0.9320
TwImbDatal1112016 NB RF RF RF 0.9024 0.9081 0.9188 0.9211 0.9089 0.5000 0.8134 0.8477 0.6563 0.8310 0.5000 0.8134 0.8477 0.6563 0.8310 0.5000 0.8134 0.8477 0.6563 0.8310 0.7784 0.8174 0.6549 0.8296 0.7226 0.7226 0.7961 0.7933 0.7961 0.7960 0.8224 0.7961 0.7933 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7960 0.8224 0.7961 0.7962 0.926 0.9265 0.9103 0.8260 0.7961 0.7962 0.8628 0.7171 0.7885 0.7910 0.7888 0.7885 0.8253 0.7961 0.7885 0.7910 0.7888 0.7885 0.8253 0.7961 0.9239 0.9241 0.9186 0.9074 0.9151 0.9239 0.9241 0.9186 0.9074 0.9151 0.9239 0.9241 0.9186 0.8074 0.9151 0.9239 0.9241 0.9186 0.8074 0.9151 0.9239 0.9241 0.9186 0.7875 0.7962 0.7876 0.7875 0.7154 0.7862 0.7962 0.7968 0.8313 0.8313 0.8313 0.8313 0.8313 0.8313 0.7884 0.7885 0.7883 0.7884 0.8234 0.7962 0.9262 0.9203 0.9183 0.7884 0.7885 0.7883 0.7884 0.8234 0.7885 0.7883 0.7884 0.8234 0.7885 0.7966 0.8399 0.6996 0.8323 0.9883 0.7884 0.7881 0.7883 0.7884 0.8234 0.8234 0.7885 0.7883 0.7884 0.8234 0.8234 0.7885 0.7883 0.7884 0.8234 0.8234 0.7885 0.7883 0.7884 0.8234 0.8234 0.7885 0.7883 0.8652 0.7175 0.7884 0.7894 0.7991 0.8026 0.7991 0.7884 0.7894 0.7991 0.8026 0.9109 0.9212 0.9180 0.9152 0.9180 0.9180 0.9180 0.9180 0.9180 0.9180 0.9180 0.9180 0.9180 0.9180		SVM	0.5000	0.8360	0.8593	0.7052	0.8529
NumbData11112016 RF 0.9024 0.9081 0.9188 0.9211 0.9089 SVM 0.5000 0.8134 0.8477 0.6563 0.8310 0.8110 0.8184 0.8477 0.6563 0.8310 0.8184 0.8477 0.6563 0.8310 0.8184 0.8477 0.6563 0.8310 0.8184 0.8477 0.6563 0.8310 0.8184 0.8477 0.6563 0.8310 0.9266 0.7226 0.80837 0.9118 0.9206 0.9225 0.9103 0.84827 0.9060 0.9225 0.9103 0.8080 0.8080 0.9225 0.9103 0.9060 0.9225 0.9103 0.9060 0.8227 0.8427 0.7043 0.8313 0.80828 0.7171 0.80826 0.8272 0.8427 0.7043 0.8313 0.80828 0.7171 0.80826 0.8272 0.8427 0.7043 0.8313 0.9080 0.8150 0.8431 0.9663 0.8313 0.9080 0.8150 0.8431 0.9663 0.8313 0.86828 0.8253		J48	0.7937	0.7882	0.6592	0.8701	0.6856
NB 0.7784 0.9081 0.9888 0.9211 0.9089	TwimbData11112016	NB	0.7977	0.7965	0.7978	0.7977	0.8380
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wiiiibData11112010	RF	0.9024	0.9081	0.9188	0.9211	0.9089
TwImbData12112016 NB RF 0.7961 0.8937 0.9118 0.9118 0.9206 0.9225 0.9103 0.9103 SVM 0.5000 0.8227 0.8427 0.7043 0.8313 J48 0.7734 0.7945 0.6573 0.8628 0.7171 TwImbData13112016 NB RF 0.9074 0.9151 0.9239 0.9241 0.9186 SVM 0.5000 0.8150 0.8431 0.6963 0.8313 J48 0.7662 0.7908 0.6495 0.8675 0.7154 WImbData14112016 NB RF 0.9166 0.7908 0.6495 0.8675 0.7154 TwImbData14112016 NB RF 0.9166 0.9129 0.9226 0.9203 0.9183 SVM 0.5000 0.8196 0.8399 0.6996 0.8323 TwImbData15112016 NB RF 0.7884 0.7851 0.7883 0.7884 0.8234 TwImbData16112016 NB RF 0.9003 0.9059 0.9189 0.9142 0.9165 SVM 0.5000		SVM	0.5000	0.8134	0.8477	0.6563	0.8310
NB		J48	0.7784	0.8174	0.6549	0.8596	0.7226
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData12112016	NB	0.7961		0.7961	0.7960	0.8224
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIIDData12112010	RF	0.8937	0.9118	0.9206	0.9225	0.9103
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.5000	0.8227	0.8427	0.7043	0.8313
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.7734	0.7945	0.6573	0.8628	0.7171
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TIh Doto 12112016	NB	0.7885	0.7910	0.7888	0.7885	0.8253
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 willibData15112010	RF	0.9074	0.9151	0.9239	0.9241	0.9186
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.5000	0.8150	0.8431	0.6963	0.8313
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.7662	0.7908	0.6495	0.8675	0.7154
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData14112016	NB	0.7875	0.7820	0.7876	0.7875	0.8145
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wiiiibData14112010	RF	0.9116	0.9129	0.9226	0.9203	0.9183
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.5000	0.8196	0.8399	0.6996	0.8323
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.7785	0.7907	0.6292	0.8608	0.6821
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData15112016		0.7884	0.7851	0.7883	0.7884	0.8234
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIDData15112010	RF	0.8973	0.9059	0.9189	0.9142	0.9165
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.5000	0.8069	0.8315	0.6783	
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$					0.6553		
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData16112016		0.7901	0.7848	0.7894	0.7901	0.8026
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIDData10112010	RF	0.9036	0.9109	0.9212	0.9180	0.9152
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$			0.5000	0.8149	0.8442	0.7024	0.8332
TwlmbData17112016 RF 0.9003 0.9084 0.9127 0.9179 0.9184 SVM 0.5000 0.8146 0.8405 0.6792 0.8342 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		J48	0.7435	0.7936	0.6443	0.8680	0.7236
RF 0.9003 0.9084 0.9127 0.9179 0.9184 SVM 0.5000 0.8146 0.8405 0.6792 0.8342 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	TwImbData17119016		0.7917	0.7878	0.7917	0.7917	0.8372
TwImbData18112016	1 willibData1/112010	RF	0.9003	0.9084	0.9127	0.9179	0.9184
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$			0.5000				
1WIMDData18112016 RF 0.8859 0.9027 0.9109 0.9135 0.9080					0.6393	0.8505	
RF 0.8859 0.9027 0.9109 0.9135 0.9080	TwImbData18119016	NB	0.7874	0.7904	0.7871	0.7874	0.8319
SVM 0.5000 0.8116 0.8351 0.6981 0.8274	1 WIIIDData10112010	RF	0.8859				
		SVM	0.5000	0.8116	0.8351	0.6981	0.8274

Table 6: New Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
	J48	0.8914	0.8917	0.8662	0.9172	0.8780
Tw.Powog9019	NB	0.8198	0.8341	0.8203	0.8198	0.8595
TwReyes2013	RF	0.9724	0.9693	0.9722	0.9711	0.9718
	SVM	0.8601	0.8810	0.8819	0.8511	0.8821
	J48	0.8807	0.8648	0.8490	0.9074	0.8658
Tw.Inany Parkiani 2014	NB	0.8522	0.8490	0.8519	0.8522	0.8502
TwIronyBarbieri2014	RF	0.9598	0.9569	0.9618	0.9603	0.9609
	SVM	0.8250	0.8435	0.8447	0.7401	0.8476
	J48	0.5344	0.5710	0.5168	0.5555	0.5355
TwMohammad2015	NB	0.6364	0.6424	0.6330	0.6364	0.6099
1 wivionammad2015	RF	0.6561	0.6479	0.6397	0.6497	0.6521
	SVM	0.5075	0.6046	0.5984	0.5000	0.5680
	J48	0.9072	0.9020	0.8759	0.9304	0.8936
TwSarcasmBarbieri2014	NB	0.9097	0.9113	0.9099	0.9097	0.8864
1 wsarcasiiibarbieri2014	RF	0.9755	0.9732	0.9766	0.9754	0.9758
	SVM	0.8784	0.8943	0.8969	0.8599	0.8981
	J48	0.6194	0.6565	0.6067	0.7319	0.6006
TwRiloff2013	NB	0.7471	0.7446	0.7478	0.7471	0.7331
1 WKH0H2013	RF	0.8006	0.8020	0.8125	0.8078	0.8063
	SVM	0.5485	0.7326	0.7339	0.6230	0.7317
	J48	0.7127	0.7376	0.6843	0.7918	0.7234
TwPtáček2014	NB	0.7390	0.7378	0.7386	0.7390	0.7660
1 Wr tacek2014	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
	J48	0.5079	0.6385	0.5469	0.7199	0.6562
TwImbData08112016	NB	0.6620	0.6605	0.6622	0.6620	0.6416
1 wiiiibData08112010	RF	0.7590	0.7805	0.7809	0.7956	0.7748
	SVM	0.5000	0.6640	0.6759	0.5000	0.6646
	J48	0.7749	0.7863	0.6383	0.8597	0.7226
TwImbData09112016	NB	0.7815	0.7759	0.7809	0.7815	0.8164
1 wiiiibData09112010	RF	0.9105	0.9194	0.9235	0.9260	0.9203
	SVM	0.5000	0.8114	0.8315	0.7090	0.8264
	J48	0.7440	0.8245	0.6766	0.8882	0.7405
TwImbData10112016	NB	0.7973	0.7948	0.7970	0.7973	0.8492
1 willibData10112010	RF	0.9231	0.9267	0.9346	0.9350	0.9320
	SVM	0.5000	0.8360	0.8593	0.7052	0.8529
	J48	0.7937	0.7882	0.6592	0.8701	0.6856
TwImbData11112016	NB	0.7977	0.7965	0.7978	0.7977	0.8380
1 wimbData11112010	RF	0.9024	0.9081	0.9188	0.9211	0.9089
	SVM	0.5000	0.8134	0.8477	0.6563	0.8310
-	J48	0.7784	0.8174	0.6549	0.8596	0.7226
TwImbData12112016	NB	0.7961	0.7933	0.7961	0.7960	0.8224
1 w1mbData12112010	RF	0.8937	0.9118	0.9206	0.9225	0.9103
	SVM	0.5000	0.8227	0.8427	0.7043	0.8313
	J48	0.7734	0.7945	0.6573	0.8628	0.7171
TIh Do to 19119016	NB	0.7885	0.7910	0.7888	0.7885	0.8253
TwImbData13112016	RF	0.9074	0.9151	0.9239	0.9241	0.9186
	SVM	0.5000	0.8150	0.8431	0.6963	0.8313
-	J48	0.7662	0.7908	0.6495	0.8675	0.7154
TIh Doto 14112016	NB	0.7875	0.7820	0.7876	0.7875	0.8145
TwImbData14112016	RF	0.9116	0.9129	0.9226	0.9203	0.9183
	SVM	0.5000	0.8196	0.8399	0.6996	0.8323
	J48	0.7785	0.7907	0.6292	0.8608	0.6821
TIh Doto 15112016	NB	0.7884	0.7851	0.7883	0.7884	0.8234
TwImbData15112016	RF	0.8973	0.9059	0.9189	0.9142	0.9165
	SVM	0.5000	0.8069	0.8315	0.6783	0.8272
	J48	0.7820	0.7860	0.6553	0.8652	0.7175
TwImbData16112016	NB	0.7901	0.7848	0.7894	0.7901	0.8026
1 w1mbData10112010	RF	0.9036	0.9109	0.9212	0.9180	0.9152
	SVM	0.5000	0.8149	0.8442	0.7024	0.8332
	J48	0.7435	0.7936	0.6443	0.8680	0.7236
TwImbDo4a17110016	NB	0.7917	0.7878	0.7917	0.7917	0.8372
TwImbData17112016	RF	0.9003	0.9084	0.9127	0.9179	0.9184
	SVM	0.5000	0.8146	0.8405	0.6792	0.8342
	J48	0.7484	0.7803	0.6393	0.8505	0.7068
TIh Do to 10110016	NB	0.7874	0.7904	0.7871	0.7874	0.8319
TwImbData18112016	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
	J48	0.6427	0.6002	0.6099	0.7554	0.5288
TD9012	NB	0.5460	0.5596	0.5441	0.5422	0.6335
TwReyes2013	RF	0.9007	0.8538	0.8838	0.8399	0.8713
	SVM	0.1253	0.0587	0.0587	0.0245	0.0668
	J48	0.6251	0.4738	0.5592	0.7519	0.4958
TIncoDombioni2014	NB	0.5276	0.5176	0.5259	0.5262	0.4324
TwIronyBarbieri2014	RF	0.8869	0.8240	0.8684	0.8045	0.8578
	SVM	0.1550	0.0866	0.0883	0.0129	0.0923
	J48	0.2516	0.2143	0.2314	0.2983	0.2234
TwMohammad2015	NB	0.3814	0.3899	0.3762	0.3814	0.3727
1 wivionammad2015	RF	0.4029	0.3960	0.3858	0.4042	0.3947
	SVM	0.2198	0.1449	0.1511	0.0000	0.2189
	J48	0.6734	0.5535	0.6176	0.7938	0.5478
TwSarcasmBarbieri2014	NB	0.6331	0.6311	0.6339	0.6322	0.4461
1 w5arcasiiiDarbieri2014	RF	0.9140	0.8449	0.8951	0.8285	0.8756
	SVM	0.1064	0.0516	0.0513	0.0167	0.0586
	J48	0.2605	0.2308	0.2333	0.3470	0.2243
TwRiloff2013	NB	0.3971	0.3932	0.3968	0.3971	0.4504
I WKHOH2015	RF	0.5245	0.5279	0.5249	0.5251	0.5354
	SVM	0.4096	0.0772	0.0859	0.0133	0.0861
	J48	0.3655	0.3623	0.3444	0.5934	0.3091
TwPtáček2014	NB	0.4273	0.4250	0.4268	0.4273	0.4745
1 WF tacek2014	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
	J48	0.0046	0.0217	0.0501	0.0410	0.0375
TwImbData08112016	NB	0.0308	0.0309	0.0308	0.0308	0.0351
1 wiiiibData08112010	RF	0.1162	0.0847	0.1233	0.0911	0.0991
	SVM	0.0000	0.0108	0.0107	0.0000	0.0108
	J48	0.1329	0.0491	0.1342	0.0878	0.0828
TwImbData09112016	NB	0.0481	0.0466	0.0479	0.0479	0.0667
1 wiiiibData09112010	RF	0.3457	0.2477	0.3533	0.2770	0.3056
	SVM	0.0000	0.0067	0.0066	0.0000	0.0080
	J48	0.1431	0.0585	0.1809	0.1211	0.1016
TwImbData10112016	NB	0.0538	0.0529	0.0534	0.0530	0.0805
1 willibData10112010	RF	0.3880	0.2747	0.4019	0.3114	0.3545
	SVM	0.0000	0.0090	0.0077	0.0000	0.0100
	J48	0.1444	0.0469	0.1649	0.1193	0.0702
TruImbData11119016	NB	0.0535	0.0518	0.0534	0.0524	0.0880
TwImbData11112016	RF	0.3346	0.2111	0.3309	0.2409	0.2663
	SVM	0.0000	0.0094	0.0081	0.0000	0.0130
-	J48	0.1455	0.0512	0.1570	0.0883	0.0805
TwImbData12112016	NB	0.0533	0.0520	0.0524	0.0519	0.0716
1 wimbData12112016	RF	0.3568	0.2211	0.3643	0.2711	0.3207
	SVM	0.0000	0.0093	0.0091	0.0000	0.0126
	J48	0.1363	0.0480	0.1555	0.0949	0.0784
TIh Do to 19119016	NB	0.0509	0.0508	0.0508	0.0502	0.0782
TwImbData13112016	RF	0.3513	0.2493	0.3776	0.2738	0.3024
	SVM	0.0000	0.0097	0.0086	0.0000	0.0113
	J48	0.1201	0.0452	0.1516	0.1061	0.0791
TwImbData14112016	NB	0.0496	0.0484	0.0496	0.0492	0.0728
1 w1mbData14112010	RF	0.3422	0.2182	0.3381	0.2451	0.2890
	SVM	0.0000	0.0078	0.0077	0.0000	0.0103
	J48	0.1264	0.0470	0.1353	0.0857	0.0688
TIh Doto 15112016	NB	0.0513	0.0496	0.0510	0.0506	0.0702
TwImbData15112016	RF	0.3209	0.2050	0.3340	0.2423	0.3031
	SVM	0.0000	0.0089	0.0083	0.0000	0.0090
	J48	0.1380	0.0464	0.1616	0.0888	0.0763
TwImbData16112016	NB	0.0498	0.0484	0.0496	0.0496	0.0735
1 w1mbData16112016	RF	0.3394	0.2232	0.3491	0.2503	0.3013
	SVM	0.0000	0.0085	0.0077	0.0000	0.0107
	J48	0.1152	0.0505	0.1444	0.1088	0.0760
TwImbDoto17119016	NB	0.0513	0.0508	0.0509	0.0507	0.0829
TwImbData17112016	RF	0.3270	0.2061	0.3371	0.2532	0.3002
	SVM	0.0000	0.0094	0.0083	0.0001	0.0095
	J48	0.1052	0.0448	0.1341	0.0863	0.0755
TIh Do to 10110016	NB	0.0502	0.0506	0.0499	0.0497	0.0747
TwImbData18112016	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
	J48	0.6427	0.6002	0.6099	0.7554	0.5288
TvvPovvog2012	NB	0.5452	0.5585	0.5435	0.5414	0.6321
TwReyes2013	RF	0.9007	0.8538	0.8838	0.8398	0.8713
	SVM	0.1253	0.0587	0.0587	0.0245	0.0668
	J48	0.6251	0.4738	0.5592	0.7519	0.4958
Tw.Inony.Parkioni2014	NB	0.5275	0.5173	0.5258	0.5262	0.4324
TwIronyBarbieri2014	RF	0.2854	0.1613	0.6618	0.6573	0.4983
	SVM	0.1550	0.0866	0.0883	0.0129	0.0923
	J48	0.2516	0.2143	0.2314	0.2983	0.2234
TwMohammad2015	NB	0.3814	0.3896	0.3762	0.3811	0.3727
1 wMonammad2015	RF	0.4006	0.3207	0.3816	0.4009	0.3872
	SVM	0.2198	0.1449	0.1511	0.0000	0.2189
	J48	0.6734	0.5535	0.6176	0.7938	0.5478
TwSarcasmBarbieri2014	NB	0.6331	0.6310	0.6337	0.6322	0.4461
1 wsarcasiii barbieri 2014	RF	0.1271	0.0703	0.1429	0.7322	0.1146
	SVM	0.1064	0.0516	0.0513	0.0167	0.0586
	J48	0.2605	0.2308	0.2333	0.3468	0.2243
TwRiloff2013	NB	0.3971	0.3932	0.3968	0.3971	0.4503
1 WKIIOII2015	RF	0.4712	0.2641	0.4620	0.5004	0.4094
	SVM	0.4096	0.0772	0.0859	0.0133	0.0861
	J48	0.1577	0.3623	0.3444	0.1637	0.3091
TwPtáček2014	NB	0.1773	0.4250	0.4268	0.1773	0.4745
1 wr tacekz014	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

$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$.0375 .0351 .0934 .0108 .0828 .0665 .2871 .0080 .1016
TwImbData08112016 RF 0.1074 0.0294 0.1233 0.3053 0 SVM 0.0000 0.0108 0.0107 0.0000 0 J48 0.1329 0.0491 0.1342 0.0877 0 NB 0.0480 0.0466 0.0479 0.0478 0 RF 0.3355 0.0442 0.3542 0.4499 0	.0934 .0108 .0828 .0665 .2871 .0080
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$.0108 .0828 .0665 .2871 .0080
TwImbData09112016	.0828 .0665 .2871 .0080 .1016
TwImbData09112016	.0828 .0665 .2871 .0080 .1016
RF 0.3355 0.0442 0.3542 0.4499 0	.2871 .0080 .1016
$ ext{RF} 0.3355 0.0442 0.3542 0.4499 0$.0080
0.000 0.000 0.000 0.0000	.1016
SVM 0.0000 0.0067 0.0066 0.0000 0	
J48 0.1431 0.0585 0.1809 0.1211 0	0803
Trulmh Data 10112016 NB 0.0537 0.0528 0.0533 0.0528 0	.0000
TwImbData10112016 RF 0.3888 0.0431 0.4015 0.4376 0	.3443
SVM 0.0000 0.0090 0.0077 0.0000 0	.0100
J48 0.1444 0.0469 0.1649 0.1193 0	.0702
Trulmh Detailing O. NB 0.0534 0.0518 0.0534 0.0524 0	.0878
TwImbData11112016 RF 0.3344 0.0542 0.3309 0.3179 0	.2607
SVM 0.0000 0.0094 0.0081 0.0000 0	.0130
J48 0.1455 0.0512 0.1570 0.0883 0	.0805
	.0712
RF 0.3509 0.0572 0.3640 0.4091 0	.3085
SVM 0.0000 0.0093 0.0091 0.0000 0	.0126
J48 0.1363 0.0480 0.1555 0.0949 0	.0784
Trulmh Data 12112016 NB 0.0508 0.0506 0.0507 0.0500 0	.0781
TwImbData13112016 RF 0.3349 0.0705 0.3774 0.3871 0	.2813
SVM 0.0000 0.0097 0.0086 0.0000 0	.0113
J48 0.1201 0.0452 0.1516 0.1061 0	.0791
TwImbData14112016 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
J48 0.1264 0.0470 0.1353 0.0857 0	.0688
TwImbData15112016 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
J48 0.1380 0.0464 0.1616 0.0888 0	.0763
TwImbData16112016 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
	.0760
TwlmbData1/112016	.0827
RF 0.3241 0.0501 0.3370 0.4144 0	.2954
	.0095
	.0755
TwImbData18112016 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
	J48	0.6427	0.6002	0.6099	0.7554	0.5288
TvvPovvog2012	NB	0.5460	0.5596	0.5441	0.5422	0.6335
TwReyes2013	RF	0.9007	0.8538	0.8838	0.8399	0.8713
	SVM	0.1253	0.0587	0.0587	0.0245	0.0668
	J48	0.6251	0.4738	0.5592	0.7519	0.4958
Tw.Inony.Parkioni2014	NB	0.5276	0.5176	0.5259	0.5262	0.4324
TwIronyBarbieri2014	RF	0.8869	0.8240	0.8684	0.8045	0.8578
	SVM	0.1550	0.0866	0.0883	0.0129	0.0923
	J48	0.2516	0.2143	0.2314	0.2983	0.2234
TwMohammad2015	NB	0.3814	0.3899	0.3762	0.3814	0.3727
1 wWonammad2015	RF	0.4029	0.3960	0.3858	0.4042	0.3947
	SVM	0.2198	0.1449	0.1511	0.0000	0.2189
	J48	0.6734	0.5535	0.6176	0.7938	0.5478
TwSarcasmBarbieri2014	NB	0.6331	0.6311	0.6339	0.6322	0.4461
1 wSarcasiiiDarbieri2014	RF	0.9140	0.8449	0.8951	0.8285	0.8756
	SVM	0.1064	0.0516	0.0513	0.0167	0.0586
	J48	0.2605	0.2308	0.2333	0.3470	0.2243
TwRiloff2013	NB	0.3971	0.3932	0.3968	0.3971	0.4504
1 WKIIOII2015	RF	0.5245	0.5279	0.5249	0.5251	0.5354
	SVM	0.4096	0.0772	0.0859	0.0133	0.0861
	J48	0.3655	0.3623	0.3444	0.5934	0.3091
TwPtáček2014	NB	0.4273	0.4250	0.4268	0.4273	0.4745
1 wr tacekz014	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
	J48	0.0046	0.0217	0.0501	0.0410	0.0375
TwImbData08112016	NB	0.0308	0.0309	0.0308	0.0308	0.0351
1 w11110Data08112010	RF	0.1162	0.0847	0.1233	0.0911	0.0991
	SVM	0.0000	0.0108	0.0107	0.0000	0.0108
	J48	0.1329	0.0491	0.1342	0.0878	0.0828
TwImbData09112016	NB	0.0481	0.0466	0.0479	0.0479	0.0667
1 wimbData09112010	RF	0.3457	0.2477	0.3533	0.2770	0.3056
	SVM	0.0000	0.0067	0.0066	0.0000	0.0080
	J48	0.1431	0.0585	0.1809	0.1211	0.1016
TwImbData10112016	NB	0.0538	0.0529	0.0534	0.0530	0.0805
1 willibData10112010	RF	0.3880	0.2747	0.4019	0.3114	0.3545
	SVM	0.0000	0.0090	0.0077	0.0000	0.0100
	J48	0.1444	0.0469	0.1649	0.1193	0.0702
TwImbData11112016	NB	0.0535	0.0518	0.0534	0.0524	0.0880
1 willibData11112010	RF	0.3346	0.2111	0.3309	0.2409	0.2663
	SVM	0.0000	0.0094	0.0081	0.0000	0.0130
	J48	0.1455	0.0512	0.1570	0.0883	0.0805
TIhDoto19119016	NB	0.0533	0.0520	0.0524	0.0519	0.0716
TwImbData12112016	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
	J48	0.1201	0.0452	0.1516	0.1061	0.0791
TIhDoto14112016	NB	0.0496	0.0484	0.0496	0.0492	0.0728
TwImbData14112016	RF	0.3422	0.2182	0.3381	0.2451	0.2890
	SVM	0.0000	0.0078	0.0077	0.0000	0.0103
	J48	0.1264	0.0470	0.1353	0.0857	0.0688
TIhDoto15112016	NB	0.0513	0.0496	0.0510	0.0506	0.0702
TwImbData15112016	RF	0.3209	0.2050	0.3340	0.2423	0.3031
	SVM	0.0000	0.0089	0.0083	0.0000	0.0090
	J48	0.1380	0.0464	0.1616	0.0888	0.0763
TIh Do to 16112016	NB	0.0498	0.0484	0.0496	0.0496	0.0735
TwImbData16112016	RF	0.3394	0.2232	0.3491	0.2503	0.3013
	SVM	0.0000	0.0085	0.0077	0.0000	0.0107
	J48	0.1152	0.0505	0.1444	0.1088	0.0760
TwlmbDo4a17110016	NB	0.0513	0.0508	0.0509	0.0507	0.0829
TwImbData17112016	RF	0.3270	0.2061	0.3371	0.2532	0.3002
	SVM	0.0000	0.0094	0.0083	0.0001	0.0095
	J48	0.1052	0.0448	0.1341	0.0863	0.0755
TInolo Do to 10110016	NB	0.0502	0.0506	0.0499	0.0497	0.0747
TwImbData18112016	RF	0.2959	0.2049	0.3178	0.2262	0.2928
	SVM	0.0000	0.0076	0.0070	0.0001	0.0085
	11 37 0	ъ .	CD1 1	1.11		

Table 14: New Corpora - Prevalence Thresholding

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
	J48	0.6427	0.6002	0.6099	0.7554	0.5288
TvvPovvog2012	NB	0.5460	0.5596	0.5441	0.5422	0.6335
TwReyes2013	RF	0.9007	0.8538	0.8838	0.8399	0.8713
	SVM	0.1253	0.0587	0.0587	0.0245	0.0668
	J48	0.6251	0.4738	0.5592	0.7519	0.4958
Tw.Inony.Parkioni2014	NB	0.5276	0.5176	0.5259	0.5262	0.4324
TwIronyBarbieri2014	RF	0.8869	0.8240	0.8684	0.8045	0.8578
	SVM	0.1550	0.0866	0.0883	0.0129	0.0923
	J48	0.2516	0.2143	0.2314	0.2983	0.2234
TwMohammad2015	NB	0.3814	0.3899	0.3762	0.3814	0.3727
1 wWonammad2015	RF	0.4029	0.3960	0.3858	0.4042	0.3947
	SVM	0.2198	0.1449	0.1511	0.0000	0.2189
	J48	0.6734	0.5535	0.6176	0.7938	0.5478
TwSarcasmBarbieri2014	NB	0.6331	0.6311	0.6339	0.6322	0.4461
1 wsarcasiii barbieri 2014	RF	0.9140	0.8449	0.8951	0.8285	0.8756
	SVM	0.1064	0.0516	0.0513	0.0167	0.0586
	J48	0.2605	0.2308	0.2333	0.3470	0.2243
TwRiloff2013	NB	0.3971	0.3932	0.3968	0.3971	0.4504
1 WKIIOII2015	RF	0.5245	0.5279	0.5249	0.5251	0.5354
	SVM	0.4096	0.0772	0.0859	0.0133	0.0861
	J48	0.3655	0.3623	0.3444	0.5934	0.3091
TwPtáček2014	NB	0.4273	0.4250	0.4268	0.4273	0.4745
1 wr tacekz014	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

TwImbData08112016 RF 0.1162 0.0847 0.1233 0.0911 0.0919 SVM 0.0000 0.0108 0.0107 0.0000 0.0108 RF 0.1162 0.0847 0.1233 0.0911 0.0991 SVM 0.0000 0.0108 0.0107 0.0000 0.0108 RF 0.348 0.1329 0.0491 0.1342 0.0878 0.0828 TwImbData09112016 RF 0.3457 0.2477 0.3533 0.2770 0.3056 SVM 0.0000 0.0067 0.0066 0.0000 0.0080 TwImbData10112016 RF 0.3457 0.2477 0.3533 0.2770 0.3056 SVM 0.0000 0.0067 0.0066 0.0000 0.0080 TwImbData10112016 RF 0.3880 0.2747 0.4019 0.1211 0.1016 RF 0.3880 0.2747 0.4019 0.3114 0.3545 SVM 0.0000 0.00090 0.0077 0.0000 0.1000 J48 0.1444 0.0469 0.1649 0.1193 0.0702 TwImbData11112016 RF 0.3386 0.2511 0.3309 0.2409 0.2663 SVM 0.0000 0.0094 0.0081 0.0000 0.130 TwImbData12112016 RF 0.3568 0.2511 0.3309 0.2409 0.2663 SVM 0.0000 0.0094 0.0081 0.0000 0.0130 TwImbData12112016 RF 0.3568 0.2211 0.3309 0.2409 0.2663 SVM 0.0000 0.0093 0.0051 0.0000 0.0126 TwImbData13112016 RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 TwImbData13112016 RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 TwImbData13112016 RF 0.3513 0.2493 0.3776 0.2738 0.3024 SVM 0.0000 0.0097 0.0086 0.0000 0.0126 TwImbData14112016 RF 0.3513 0.2493 0.3776 0.2738 0.3024 SVM 0.0000 0.0097 0.0086 0.0000 0.0126 TwImbData14112016 RF 0.3513 0.2493 0.3776 0.2738 0.3024 SVM 0.0000 0.0097 0.0086 0.0000 0.0126 TwImbData14112016 RF 0.3510 0.0452 0.1516 0.1661 0.0791 TwImbData15112016 RF 0.3207 0.0088 0.0000 0.0013 TwImbData16112016 RF 0.3209 0.0505 0.3340 0.2423 0.3031 SVM 0.0000 0.0098 0.0083 0.0000 0.0090 TwImbData16112016 RF 0.3209 0.0505 0.340 0.0490 0.0978 TwImbData16112016 RF 0.3209 0.0505 0.3340 0.2430 0.3031 SVM 0.0000 0.0085 0.0077 0.0000 0.0097 TwImbData16112016 RF 0.3209 0.0050 0.0340 0.0375 0.0000 0.0007 TwImbData16112016 RF 0.3209 0.0050 0.0340 0.0000 0.0000 0.00000 0.00000 0.000000	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
No.		J48	0.0046	0.0217	0.0501	0.0410	0.0375
NF	TImbData08112016	NB	0.0308	0.0309	0.0308	0.0308	0.0351
TwImbData10112016 TwImbData09112016 TwImbData09112016 RF RF 0.3457 0.2477 0.3533 0.2770 0.3056 SVM 0.0000 0.0067 0.0066 0.0000 0.0080 J48 0.1431 0.0585 0.1809 0.1211 0.1016 RF 0.3880 0.02747 0.4019 0.3114 0.03545 SVM 0.0000 0.0090 0.0077 0.0006 0.0000 0.0075 0.0006 0.0000 0.0100 TwImbData10112016 RF 0.3880 0.2747 0.4019 0.3114 0.3545 SVM 0.0000 0.0090 0.0077 0.0000 0.0100 TwImbData11112016 RF 0.3346 0.2111 0.3309 0.2409 0.2663 SVM 0.0000 0.0094 0.0081 0.0080 0.0094 0.0081 0.0080 0.0094 0.0081 0.0080 0.0094 0.0081 0.0080 0.0094 0.0081 0.0000 0.0094 0.0081 0.0000 0.0130 TwImbData12112016 RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 TwImbData13112016 RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 TwImbData13112016 RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 TwImbData13112016 RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 TwImbData13112016 RF 0.3513 0.2493 0.3776 0.2738 0.3024 TwImbData14112016 RF 0.3513 0.2493 0.3776 0.2738 0.3024 TwImbData15112016 RF 0.3422 0.2182 0.3381 0.2451 0.2950 0.0000 0.0103 TwImbData16112016 RF 0.3422 0.2182 0.3381 0.2451 0.2950 0.0000 0.0003 0.00077 0.0000 0.0099 TwImbData16112016 RF 0.3422 0.2182 0.3381 0.2451 0.2950 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.00000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.0000 0.0003 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.000000	1 willibData08112010	RF	0.1162	0.0847	0.1233	0.0911	0.0991
TwImbData09112016 NB RF 0.0481 0.0466 0.0479 0.0479 0.0366 SVM 0.0000 0.0667 0.0666 0.0000 0.0066 0.0000 0.0086 J48 0.1431 0.0585 0.1809 0.1211 0.1016 RF 0.3880 0.0529 0.0534 0.0530 0.0805 RF 0.3880 0.2747 0.4019 0.3114 0.3455 SVM 0.0000 0.0090 0.077 0.0000 0.0107 J48 0.1444 0.0469 0.1649 0.1193 0.0702 RF 0.3346 0.2111 0.3309 0.2409 0.2663 SVM 0.0000 0.0094 0.0081 0.0000 0.0130 TwImbData12112016 NB 0.0533 0.0521 0.1519 0.0716 RF 0.3368 0.2211 0.3643 0.2711 0.3207 TwImbData12112016 RF 0.3568 0.2211 0.3643 0.2711 0.3207		SVM	0.0000	0.0108	0.0107	0.0000	0.0108
NumbData09112016 RF		J48	0.1329	0.0491	0.1342	0.0878	0.0828
RF	Tw.ImbData00112016	NB	0.0481	0.0466	0.0479	0.0479	0.0667
TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData13112016 TwImbData131	1 WIIIDData09112010	RF	0.3457	0.2477	0.3533	0.2770	0.3056
TwImbData10112016 NB RF 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 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 NB 0.0533 0.0520 0.0524 0.0619 0.0716 RF 0.3568 0.2211 0.3643 0.2711 0.3207 0.0716 RF 0.3568 0.2211 0.3643 0.2711 0.3207 0.000 0.0126 TwImbData13112016 RF 0.3568 0.2211 0.3643 0.2711 0.320 TwImbData14112016 RF 0.3513 0.2498 0.0508 0.0509 0.0782 RF 0.35		SVM	0.0000	0.0067	0.0066	0.0000	0.0080
TwImbData10112016 RF 0.3880 0.2747 0.4019 0.3114 0.3545 SVM 0.0000 0.0090 0.0077 0.0000 0.0100 TwImbData11112016 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 RF 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 0.091 0.0000 0.0126 TwImbData13112016 RF 0.3568 0.2211 0.3643 0.2711 0.3207 TwImbData13112016 NB 0.0509 0.0508 0.0509 0.0782 TwImbData14112016 NB 0.0509 0.0508 0.0500 0.0713 TwImbData14112016 NB 0.0496 0.0484 0.0496 0.		J48	0.1431	0.0585	0.1809	0.1211	0.1016
RF	Tw.ImbData10112016	NB	0.0538	0.0529	0.0534	0.0530	0.0805
TwImbData11112016 TwImbData11112016 NB 0.0535 0.0518 0.0534 0.0524 0.0880 SVM 0.0000 0.0094 0.0081 0.0000 0.0130 0.0001 0.0082 0.0524 0.0519 0.0716 RF 0.3368 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0094 0.0081 0.0000 0.0130 0.0005 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 RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.05091 0.0000 0.0126 RF 0.3513 0.2493 0.3776 0.2738 0.3024 SVM 0.0000 0.0097 0.0086 0.0000 0.0113 18 0.0509 0.0508 NB 0.04090 0.0496 0.0496 0.0496 0.0492 0.0728 RF 0.3422 0.2182 0.3381 0.2451 0.2890 SVM 0.0000 0.0077 0.0000 0.0103 TwImbData15112016 NB 0.0464 0.0470 0.1353 0.0857 0.0688 NB 0.0509 0.0508 NB 0.0509 0.0508 0.0509 0.0508 NB 0.0509 0.0508 0.0509 0.0508 0.0702 0.0688 NB 0.0513 0.0464 0.0496 0.0496 0.0492 0.0728 RF 0.3209 0.2050 0.3340 0.2423 0.3031 SVM 0.0000 0.0088 0.0000 0.0090 0.0090 0.0088 NB 0.0498 0.0484 0.0496 0.0496 0.0496 0.0705 RF 0.3394 0.2232 0.3491 0.2503 0.3013 SVM 0.0000 0.0085 0.0077 0.0000 0.0107 TwImbData17112016 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 0.0000 0.0090 0.0088 0.0083 0.0000 0.0090 0.0088 0.0083 0.0000 0.0090 0.0088 0.0088 0.0070 0.0090 0.0088 0.0088 0.0070 0.0090 0.0088 0.0088 0.0070 0.0090 0.0088 0.0088 0.0070 0.0090 0.0088 0.0088 0.0070 0.0090 0.0088 0.0088 0.0070 0.0090 0.0088 0.0089 0.0083 0.0000 0.0090 0.0089 0.0083 0.0000 0.0090 0.0083 0.0000 0.0090 0.0083 0.0000 0.0090 0.0083 0.0001 0.0095 0.0088 0.0070 0.0000 0.000000	1 willibData10112010	RF	0.3880	0.2747	0.4019	0.3114	0.3545
TwImbData11112016 NB RF 0.0535 0.0518 0.0534 0.0524 0.0880 SVM 0.0000 0.0094 0.0081 0.0000 0.0130 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 NB 0.0509 0.0508 0.0502 0.0782 RF 0.3513 0.0480 0.1555 0.0949 0.0784 TwImbData13112016 RF 0.3513 0.2493 0.3776 0.2738 0.3024 SVM 0.0000 0.0097 0.0086 0.0000 0.0113 0.0496 0.0444 0.0496 0.0492 0.0728 TwImbData14112016 RF 0.3422 0.2182 0.3381 0.2451 0.2890 TwImbData15112016		SVM	0.0000	0.0090	0.0077	0.0000	0.0100
TwImbData11112016 RF SVM 0.0000 0.0094 0.0081 0.0000 0.0130 148 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 RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 RF 0.3513 0.2493 0.3776 0.2738 0.3024 SVM 0.0000 0.0097 0.0086 0.0000 0.0113 148 0.1363 0.480 0.1555 0.0949 0.0782 RF 0.3513 0.2493 0.3776 0.2738 0.3024 SVM 0.0000 0.0097 0.0086 0.0000 0.0113 HR 0.0496 0.0484 0.0496 0.0492 0.0792 RF 0.3422 0.2182 0.3381 0.2451 0.2890 SVM 0.0000 0.0078 0.0077 0.0000 0.0103 HR 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 0.00990 148 0.0320 0.0484 0.0496 0.0496 0.0496 0.0493 0.0702 RF 0.3209 0.2050 0.3340 0.2423 0.3031 SVM 0.0000 0.0089 0.0083 0.0000 0.00990 148 0.03394 0.2232 0.3491 0.2503 0.3013 SVM 0.0000 0.0085 0.0777 0.0000 0.0107 RF 0.3394 0.2232 0.3491 0.2503 0.3013 SVM 0.0000 0.0085 0.0077 0.0000 0.0107 RF 0.3394 0.2232 0.3491 0.2503 0.3013 SVM 0.0000 0.0085 0.0777 0.0000 0.0107 RF 0.3394 0.2232 0.3491 0.2503 0.3013 0.0000 0.0107 0.0000 0.0107 0.0000 0.0107 0.0000 0.0107 0.0000 0.0000 0.0000 0.0085 0.0077 0.0000 0.0107 0.0000 0.0107 0.0000 0.0000 0.0085 0.0077 0.0000 0.0107 0.0000 0.0000 0.0085 0.0077 0.0000 0.0107 0.0000 0.0000 0.0085 0.0077 0.0000 0.0107 0.0000 0.0000 0.0085 0.0077 0.0000 0.00107 0.0000 0.00107 0.0000 0.00107 0.0000 0.00107 0.0000 0.0000 0.0000 0.00000 0.00000 0.00000 0.00000 0.000000		J48	0.1444	0.0469	0.1649	0.1193	0.0702
SVM 0.0000 0.0094 0.0081 0.0000 0.0130	Trulmah Data 11119016	NB	0.0535	0.0518	0.0534	0.0524	0.0880
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData11112016	RF	0.3346	0.2111	0.3309	0.2409	0.2663
TwImbData12112016 NB RF 0.0533 0.0520 0.0524 0.0519 0.0716 SVM 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 J48 0.1363 0.0480 0.1555 0.0949 0.0782 RF 0.3513 0.02493 0.3776 0.2738 0.3024 SVM 0.0000 0.0097 0.0086 0.0000 0.0113 TwImbData14112016 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 0.0133 TwImbData15112016 NB 0.0513 0.0496 0.0510 0.0506 0.0702 RF 0.3209 0.2050 0.3340 0.2423 0.301 SVM 0.0000 0.0089 0.0083 0.0000 0.0094 T		SVM	0.0000	0.0094	0.0081	0.0000	0.0130
TwimbData12112016 RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 J48 0.1363 0.0480 0.1555 0.0949 0.0784 TwImbData13112016 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 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 RF 0.3209 0.2050 0.3340 0.2423 0.3031 TwImbData16112016 NB 0.0513 0.0496 0.0510 0.0506 0.0735		J48	0.1455	0.0512	0.1570	0.0883	0.0805
RF 0.3568 0.2211 0.3643 0.2711 0.3207 SVM 0.0000 0.0093 0.0091 0.0000 0.0126 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 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 RF 0.3209 0.2050 0.3340 0.2423 0.3031 SVM 0.0000 0.0089 0.0083 0.0000 0.0090 J48 0.1380 0.0464 0.1616 0.0888 0.0763 RF 0.3394	TI	NB	0.0533	0.0520	0.0524	0.0519	0.0716
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData12112016	RF	0.3568	0.2211	0.3643	0.2711	0.3207
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.0093	0.0091	0.0000	0.0126
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	T. I. I.D. (19110016)	J48	0.1363	0.0480	0.1555	0.0949	0.0784
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		NB	0.0509	0.0508	0.0508	0.0502	0.0782
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData13112016	RF	0.3513	0.2493	0.3776	0.2738	0.3024
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.0097	0.0086	0.0000	0.0113
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.1201	0.0452	0.1516	0.1061	0.0791
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.0496	0.0484	0.0496	0.0492	0.0728
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData14112016	RF	0.3422	0.2182	0.3381	0.2451	0.2890
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.0078	0.0077	0.0000	0.0103
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.1264	0.0470	0.1353	0.0857	0.0688
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.0513	0.0496	0.0510	0.0506	0.0702
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData15112010	RF	0.3209	0.2050	0.3340	0.2423	0.3031
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.0089	0.0083	0.0000	0.0090
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.1380	0.0464	0.1616	0.0888	0.0763
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Trulmah Data 16112016	NB	0.0498	0.0484	0.0496	0.0496	0.0735
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData16112016	RF	0.3394	0.2232	0.3491	0.2503	0.3013
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.0085	0.0077	0.0000	0.0107
TwlmbData17112016 RF 0.3270 0.2061 0.3371 0.2532 0.3002 SVM 0.0000 0.0094 0.0083 0.0001 0.0095 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		J48	0.1152	0.0505	0.1444	0.1088	0.0760
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData 17110016	NB	0.0513	0.0508	0.0509	0.0507	0.0829
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	TwImbData17112016	RF	0.3270	0.2061	0.3371	0.2532	0.3002
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.0094	0.0083	0.0001	0.0095
RF 0.2959 0.2049 0.3178 0.2262 0.2928		J48	0.1052	0.0448	0.1341	0.0863	0.0755
RF = 0.2959 - 0.2049 - 0.3178 - 0.2262 - 0.2928	Trulmah Data 10110016	NB	0.0502	0.0506	0.0499	0.0497	0.0747
SVM $0.0000 0.0076 0.0070 0.0001 0.0085$	1 w1mbData18112016	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
	J48	0.8311	0.8091	0.8164	0.8054	0.8160
TwReyes2013	NB	0.5669	0.5981	0.5802	0.5880	0.6673
1 wheyes2013	RF	0.8540	0.8306	0.8600	0.8320	0.8606
	SVM	0.8025	0.7887	0.7903	0.7086	0.7981
	J48	0.7996	0.7289	0.7631	0.7288	0.7707
TIncomPorkioni2014	NB	0.5989	0.5918	0.5957	0.5710	0.6014
TwIronyBarbieri2014	RF	0.8165	0.7982	0.8229	0.8005	0.8256
	SVM	0.7443	0.7137	0.7180	0.4960	0.7289
	J48	0.3238	0.4165	0.3358	0.4162	0.3546
TwMohammad2015	NB	0.4155	0.4393	0.4383	0.4441	0.4240
1 wMohammad2015	RF	0.1294	0.4455	0.2282	0.4667	0.3338
	SVM	0.0384	0.4508	0.4405	0.4323	0.3601
	J48	0.8413	0.7948	0.8140	0.7871	0.8239
TwSarcasmBarbieri2014	NB	0.7028	0.6924	0.6860	0.6534	0.6741
1 w5arcasiiiDarbiei12014	RF	0.8659	0.8360	0.8708	0.8415	0.8709
	SVM	0.8163	0.7798	0.7853	0.6606	0.7993
	J48	0.4118	0.4632	0.3866	0.5011	0.4202
TwRiloff2013	NB	0.4953	0.4979	0.4976	0.4977	0.4792
1 WKIIOII2013	RF	0.3333	0.5318	0.4713	0.5404	0.5372
	SVM	0.1913	0.5411	0.5471	0.4267	0.5438
	J48	0.5904	0.6190	0.5790	0.2193	0.5930
TwPtáček2014	NB	0.5186	0.5337	0.5339	0.4608	0.5632
I WI TACEKZUI4	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
	J48	0.0052	0.0685	0.0979	0.0697	0.0696
TwImbData08112016	NB	0.0625	0.0615	0.0605	0.0562	0.0627
1 w1mbData08112010	RF	0.0375	0.0810	0.0526	0.0915	0.0497
	SVM	0.0000	0.0748	0.0785	0.0385	0.0749
	J48	0.2690	0.1302	0.2556	0.1245	0.2277
TI	NB	0.0948	0.0870	0.0889	0.0761	0.0804
TwImbData09112016	RF	0.0629	0.1356	0.0985	0.1475	0.1497
	SVM	0.0000	0.1287	0.1443	0.0644	0.1489
	J48	0.3064	0.1504	0.3308	0.1534	0.2839
TI	NB	0.1056	0.0990	0.0988	0.0871	0.0941
TwImbData10112016	RF	0.0443	0.1486	0.1033	0.1626	0.2083
	SVM	0.0000	0.1655	0.1821	0.0636	0.1930
	J48	0.2959	0.1310	0.2977	0.1322	0.2160
Trulmah Data 11119016	NB	0.1062	0.1004	0.1008	0.0886	0.0959
TwImbData11112016	RF	0.0570	0.1354	0.0902	0.1513	0.1072
	SVM	0.0000	0.1467	0.1722	0.0552	0.1849
	J48	0.2888	0.1349	0.2930	0.1290	0.2333
TI	NB	0.1045	0.0986	0.0996	0.0882	0.0952
TwImbData12112016	RF	0.0455	0.1401	0.0981	0.1532	0.1481
	SVM	0.0000	0.1546	0.1741	0.0634	0.1844
T. I. I.D. (19110016	J48	0.2828	0.1318	0.2875	0.1266	0.2234
	NB	0.1013	0.0947	0.0944	0.0857	0.0885
TwImbData13112016	RF	0.0716	0.1343	0.0983	0.1522	0.1386
	SVM	0.0000	0.1501	0.1705	0.0618	0.1756
	J48	0.2742	0.1246	0.2840	0.1253	0.2284
TI	NB	0.0974	0.0927	0.0939	0.0808	0.0822
TwImbData14112016	RF	0.0765	0.1307	0.1136	0.1439	0.1448
	SVM	0.0000	0.1421	0.1600	0.0625	0.1703
	J48	0.2825	0.1264	0.2521	0.1237	0.2016
TI	NB	0.0955	0.0950	0.0959	0.0846	0.0807
TwImbData15112016	RF	0.0632	0.1286	0.0811	0.1429	0.1150
	SVM	0.0000	0.1386	0.1561	0.0586	0.1568
	J48	0.2868	0.1274	0.2920	0.1278	0.2178
Trulmah Data 16112016	NB	0.0959	0.0944	0.0949	0.0815	0.0796
TwImbData16112016	RF	0.0516	0.1326	0.0819	0.1462	0.0915
	SVM	0.0000	0.1427	0.1644	0.0630	0.1743
	J48	0.2651	0.1282	0.2735	0.1359	0.2160
TwImbData 17110016	NB	0.1003	0.0946	0.0980	0.0841	0.0861
TwImbData17112016	RF	0.0301	0.1319	0.0610	0.1445	0.1249
	SVM	0.0000	0.1479	0.1654	0.0588	0.1681
	J48	0.2158	0.1203	0.2594	0.1226	0.2116
TII-D-7 10110010	NB	0.1006	0.0953	0.0947	0.0825	0.0848
TwImbData18112016	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
	J48	0.8311	0.8092	0.8164	0.8219	0.8160
Tw.Powos2012	NB	0.5408	0.6008	0.5831	0.5686	0.6693
TwReyes2013	RF	0.8572	0.8249	0.8572	0.8142	0.8599
	SVM	0.8025	0.7887	0.7903	0.7086	0.7981
	J48	0.7996	0.7289	0.7631	0.7793	0.7707
Tw.Inany.Panhiani2014	NB	0.5854	0.5914	0.5953	0.5213	0.6055
TwIronyBarbieri2014	RF	0.8170	0.7998	0.8200	0.6991	0.8264
	SVM	0.7443	0.7137	0.7180	0.0000	0.7289
	J48	0.3238	0.4165	0.3358	0.3652	0.3546
TwMohammad2015	NB	0.0000	0.4459	0.4365	0.0000	0.4222
1 wWollammad2015	RF	0.2875	0.4627	0.2851	0.0941	0.4009
	SVM	0.0384	0.4508	0.4405	0.0000	0.3601
	J48	0.8413	0.7948	0.8140	0.8259	0.8239
TwSarcasmBarbieri2014	NB	0.7038	0.6931	0.6880	0.7012	0.6774
1 wsarcasinbarbieriz014	RF	0.8676	0.8348	0.8704	0.7753	0.8706
	SVM	0.8163	0.7798	0.7853	0.6606	0.7993
	J48	0.4118	0.4632	0.3866	0.5262	0.4202
TwRiloff2013	NB	0.0000	0.4950	0.4972	0.0000	0.4810
1 WKIIOII2015	RF	0.4299	0.5375	0.4659	0.1865	0.5393
	SVM	0.1913	0.5411	0.5471	0.0000	0.5438
	J48	0.2253	0.6190	0.5790	0.3217	0.5930
TwPtáček2014	NB	0.0000	0.5335	0.5334	0.3699	0.5624
1 WF tacek2014	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

$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	696 630 667 749 277
TwImbData08112016 RF 0.0586 0.0785 0.0527 0.0313 0.0 SVM 0.0000 0.0748 0.0785 0.0000 0.0 J48 0.2690 0.1302 0.2556 0.0000 0.2 NB 0.0000 0.0866 0.0885 0.0000 0.0 RF 0.0934 0.1411 0.0822 0.0458 0.2 SVM 0.0000 0.1287 0.1443 0.0000 0.1 J48 0.3064 0.1506 0.3308 0.0000 0.2 NB 0.0000 0.0883 0.0984 0.0000 0.2	667 749
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	749
TwImbData09112016	
TwImbData09112016 NB RF 0.0000 0.0934 0.0885 0.1411 0.0000 0.0822 0.0000 0.458 0.2 0.2 0.0458 SVM 0.0000 0.1287 0.1443 0.0000 0.1 J48 0.3064 0.1506 0.3308 0.0000 0.2 NB 0.0000 0.0983 0.0984 0.0000 0.0	277
RF 0.0934 0.1411 0.0822 0.0458 0.2 SVM 0.0000 0.1287 0.1443 0.0000 0.1 J48 0.3064 0.1506 0.3308 0.0000 0.2 NR 0.0000 0.0983 0.0984 0.0000 0.2	
SVM 0.0000 0.1287 0.1443 0.0000 0.1 J48 0.3064 0.1506 0.3308 0.0000 0.2 NB 0.0000 0.0983 0.0984 0.0000 0.6	813
J48 0.3064 0.1506 0.3308 0.0000 0.2	147
NR 0.0000 0.0983 0.0984 0.0000 0.0	489
NB 0.0000 0.0983 0.0984 0.0000 0.0	839
Therefore 10.119016 0.0000 0.0001 0.0001	972
TwImbData10112016 RF 0.1089 0.1538 0.0726 0.0260 0.2	654
SVM 0.0000 0.1655 0.1821 0.0000 0.1	930
J48 0.2959 0.1310 0.2977 0.0000 0.2	160
Turk Detailing NB 0.0000 0.0993 0.1001 0.0000 0.0	980
TwImbData11112016 RF 0.0927 0.1382 0.0820 0.0375 0.1	595
SVM 0.0000 0.1467 0.1722 0.0000 0.1	849
J48 0.2888 0.1349 0.2930 0.0000 0.2	333
TwImbData12112016 NB 0.0000 0.0977 0.0989 0.0000 0.0	972
RF 0.1052 0.1444 0.0758 0.0260 0.2	093
SVM 0.0000 0.1546 0.1741 0.0000 0.1	844
J48 0.2828 0.1320 0.2875 0.0000 0.2	234
Turk Detailations NB 0.0000 0.0942 0.0936 0.0000 0.0	910
TwImbData13112016 RF 0.1154 0.1429 0.0894 0.0591 0.1	857
SVM 0.0000 0.1501 0.1705 0.0000 0.1	756
J48 0.2742 0.1246 0.2840 0.0000 0.2	284
TwImbData14112016 NB 0.0000 0.0919 0.0945 0.0000 0.0	840
RF 0.1165 0.1376 0.1073 0.0637 0.1	883
	703
J48 0.2825 0.1264 0.2521 0.0000 0.2	016
TwImbData15112016 NB 0.0000 0.0935 0.0953 0.0000 0.0	842
RF 0.0923 0.1369 0.0728 0.0517 0.1	961
	568
J48 0.2868 0.1274 0.2920 0.0000 0.2	178
TwImbData16112016 NB 0.0000 0.0934 0.0942 0.0000 0.0	817
RF 0.0799 0.1359 0.0686 0.0411 0.1	695
SVM 0.0000 0.1427 0.1644 0.0000 0.1	743
	160
TwImbData17112016 NB 0.0000 0.0934 0.0963 0.0000 0.0	902
RF = 0.0769 - 0.1383 - 0.0501 - 0.0227 - 0.1	864
	681
	116
TwImbData18112016 NB 0.0000 0.0944 0.0934 0.0000 0.0	878
RF 0.0696 0.1306 0.0534 0.0310 0.1	832
SVM 0.0000 0.1342 0.1506 0.0000 0.1	535

Table 20: New Corpora - Logistic Calibration

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
	J48	0.8325	0.8111	0.8187	0.8365	0.8165
Tw.Powos2012	NB	0.5600	0.6080	0.5921	0.5600	0.6713
TwReyes2013	RF	0.8594	0.8095	0.8594	0.7762	0.8531
	SVM	0.8025	0.7887	0.7903	0.7086	0.7981
	J48	0.7881	0.7187	0.7676	0.7976	0.7675
Tw.Inany.Panhiani2014	NB	0.5856	0.5901	0.5944	0.5856	0.6116
TwIronyBarbieri2014	RF	0.8191	0.7871	0.8213	0.7396	0.8285
	SVM	0.7443	0.7137	0.7180	0.0000	0.7289
	J48	0.3382	0.4135	0.3278	0.2967	0.3541
TwMohammad2015	NB	0.3852	0.4625	0.4470	0.3852	0.4109
1 wWonammad2015	RF	0.4564	0.4698	0.4564	0.3662	0.4701
	SVM	0.0384	0.4508	0.4405	0.0000	0.3601
	J48	0.8390	0.7870	0.8187	0.8341	0.8238
TwSarcasmBarbieri2014	NB	0.6940	0.6964	0.6941	0.6940	0.6884
1 wsarcasinbarbieriz014	RF	0.8683	0.8249	0.8700	0.8110	0.8699
	SVM	0.8163	0.7798	0.7853	0.6606	0.7993
	J48	0.4497	0.4669	0.3851	0.4681	0.4247
TwRiloff2013	NB	0.4266	0.4915	0.4944	0.4266	0.4866
1 WKIIOII2015	RF	0.5249	0.4926	0.5539	0.3639	0.5426
	SVM	0.1913	0.5411	0.5471	0.0000	0.5438
	J48	0.6007	0.6208	0.5796	0.5001	0.5977
TwPtáček2014	NB	0.4521	0.5288	0.5286	0.4521	0.5566
1 WF tacek2014	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

TwImbData08112016 RF 0.1040 0.0622 0.1053 0.0390 0.1428 SVM 0.0000 0.0748 0.0599 0.0257 0.0637 SVM 0.0000 0.0748 0.0785 0.0000 0.0749 J48 0.2798 0.1344 0.2561 0.0235 0.2294 TwImbData09112016 RF 0.2655 0.1101 0.2917 0.0527 0.3415 SVM 0.0000 0.1287 0.1443 0.0000 0.1489 J48 0.3203 0.1524 0.0320 0.1201 0.2851 TwImbData10112016 RF 0.2565 0.1101 0.2917 0.0527 0.3415 SVM 0.0000 0.1287 0.1443 0.0000 0.1489 J48 0.3203 0.1524 0.0320 0.1201 0.2851 TwImbData10112016 RF 0.3264 0.1148 0.3525 0.0333 0.3723 SVM 0.0000 0.1685 0.1821 0.0000 0.1930 J48 0.3069 0.1324 0.2983 0.1970 0.2177 TwImbData11112016 RF 0.3524 0.0983 0.3343 0.0540 0.2977 SVM 0.0000 0.1467 0.1722 0.0000 0.1849 J48 0.3002 0.1373 0.2945 0.0539 0.2295 TwImbData12112016 RF 0.3043 0.1015 0.3255 0.0333 0.3573 SVM 0.0000 0.1546 0.1741 0.0000 0.1849 J48 0.3002 0.1373 0.2945 0.0539 0.2295 TwImbData12112016 RF 0.3043 0.1015 0.3255 0.0336 0.3575 SVM 0.0000 0.1546 0.1741 0.0000 0.1844 SVM 0.0000 0.1540 0.1701 0.0201 TwImbData13112016 RF 0.3350 0.1014 0.3757 0.0627 0.3444 SVM 0.0000 0.1501 0.1705 0.0000 0.1703 TwImbData14112016 RF 0.3350 0.1014 0.3757 0.0627 0.3444 SVM 0.0000 0.1501 0.1705 0.0000 0.1703 TwImbData15112016 RF 0.3568 0.1019 0.2899 0.0563 0.1047 TwImbData16112016 RF 0.2568 0.1019 0.2897 0.0580 0.3301 TwImbData16112016 RF 0.2568 0.0890 0.0901 0.0512 0.1036 RF 0.2568 0.1019 0.2897 0.0580 0.3301 TwImbData16112016 RF 0.3688 0.0872 0.0890 0.0901 0.0512 0.1036 RF 0.2568 0.1019 0.2897 0.0580 0.3301 TwImbData16112016 RF 0.3044 0.0000 0.1421 0.1600 0.0000 0.1703 J48 0.2806 0.0279 0.2802 0.0563 0.1047 TwImbData16112016 RF 0.3040 0.0000 0.1386 0.1561 0.0000 0.1568 RF 0.3040 0.0000 0.1427 0.1644 0.0000 0.1703 J48 0.2818 0.2817 0.1290 0.0890 0.0901 0.0512 0.3080 SVM 0.0000 0.1479 0.1644 0.0000 0.1681 TwImbData16112016	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
NumbData08112016 RF 0.1040 0.0622 0.1053 0.0390 0.1428 0.0000 0.0748 0.0785 0.0000 0.0748 0.0000 0.0748 0.0785 0.0000 0.0748 0.0285 0.02294 0.0285 0.0235 0.2294 0.0285 0.0		J48	0.0107	0.0688	0.0982	0.0209	0.0828
NF	TIla Data 08112016	NB	0.0257	0.0594	0.0599	0.0257	0.0637
TwImbData10112016 TwImbData09112016 TwImbData09112016 RF RF 0.2565 0.10110 0.2917 0.0527 0.0483 0.0483 0.0483 0.0883 0.0483 0.0483 0.0883 0.0483 0.0483 0.0883 0.0483 0.0577 0.0527 0.3415 SVM 0.0000 0.1287 0.1443 0.0000 0.1287 0.1443 0.0000 0.1287 0.01443 0.0000 0.1287 0.01443 0.0000 0.1287 0.01443 0.0000 0.1287 0.01443 0.0000 0.1287 0.01443 0.0000 0.1287 0.01443 0.0000 0.1287 0.0443 0.0000 0.1524 0.0320 0.0201 0.02851 0.0333 0.3723 0.3723 0.3723 0.3723 0.3723 0.3723 0.3723 0.3723 0.3723 0.3724 0.0000 0.1655 0.1821 0.0000 0.1930 0.0631 0.0950 0.06631 0.0950 0.06631 0.0950 0.06631 0.0950 0.06631 0.0950 0.06631 0.0950 0.06631 0.0950 0.06631 0.0950 0.06631 0.0950 0.06631 0.0540 0.2977 0.02177 0.02177 0.02177 0.02177 0.0227 0.0000 0.1849 0.0000 0.1467 0.1722 0.0000 0.1849 0.0000 0.1467 0.1722 0.0000 0.1849 0.0000 0.1467 0.1722 0.0000 0.1849 0.0000 0.1467 0.1722 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1849 0.0000 0.1546 0.1741 0.0000 0.1680 0.0000 0.1546 0.1741 0.0000 0.1750 0.0000 0.1750 0.0000 0.1750 0.0000 0.1750 0.0000 0.1750 0.0000 0.1750 0.0000 0.1750 0.0000 0.1750 0.0000 0.1750 0.0000 0.1763 0.0000 0.1763 0.0000 0.1763 0.0000 0.1763 0.0000 0.1763 0.0000 0.1763 0.0000 0.1763 0.0000 0.1763 0.0000 0.1764 0.0000 0.1765 0.00000 0.1765 0.00000 0.1765 0.00000 0.1765 0.00000 0.1765 0.00000 0.1765 0.00000 0.1765 0.000000 0.1765 0.000000 0.1764 0.00000 0.1764 0.00000 0.1764 0.00000 0.1764 0.00000 0.1764 0.00000 0.1764	1 wimbData08112010	RF	0.1040	0.0622	0.1053	0.0390	0.1428
TwImbData09112016 NB RF 0.0483 0.2565 0.1010 0.1187 0.0833 0.0527 0.3415 0.03263 SVM 0.0000 0.1287 0.1443 0.0000 0.1489 J48 0.3263 0.1524 0.3320 0.1201 0.2851 TwImbData10112016 NB RF 0.3264 0.1148 0.3525 0.0333 0.3723 SVM 0.0000 0.1655 0.1821 0.0000 0.1930 J48 0.3069 0.1324 0.2983 0.1970 0.2177 TwImbData11112016 NB RF 0.3524 0.0983 0.0343 0.0170 0.2177 TwImbData11112016 NB RF 0.3524 0.0983 0.3343 0.0540 0.0297 SVM 0.0000 0.1467 0.1722 0.0000 0.1647 0.1722 0.0000 0.1849 TwImbData12112016 RF 0.3524 0.0897 0.0756 0.1096 0.0512 0.1096 0.0525 0.0330 0.0252 0.0330 0.0252 0.0330 0.0252		SVM	0.0000	0.0748	0.0785	0.0000	0.0749
NumbData09112016 RF		J48	0.2798	0.1344	0.2561	0.0235	0.2294
RF	TIla Data 00112016	NB	0.0483	0.0857	0.0883	0.0483	0.1030
TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData13112016 TwImbData131	1 wimbData09112010	RF	0.2565	0.1101	0.2917	0.0527	0.3415
TwImbData10112016 RF 0.3264 0.1148 0.3525 0.0333 0.3723 SVM 0.0000 0.1655 0.1821 0.0000 0.1930 J48 0.3609 0.1324 0.2983 0.1970 0.2177 MB 0.0631 0.0950 0.0960 0.06631 0.1161 RF 0.3524 0.0983 0.3343 0.0540 0.2977 SVM 0.0000 0.1467 0.1722 0.0000 0.1849 J48 0.3002 0.1373 0.2945 0.0539 0.2995 TwImbData12112016 RF 0.3002 0.1373 0.2945 0.0539 0.2295 SVM 0.0000 0.1467 0.1722 0.0000 0.1849 J48 0.3002 0.1373 0.2945 0.0539 0.2295 SVM 0.0000 0.1546 0.1741 0.0000 0.1844 J48 0.3052 0.1331 0.2891 0.0337 0.2253 SVM 0.0000 0.1546 0.1741 0.0000 0.1844 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 J48 0.2806 0.1279 0.2852 0.0759 0.2308 NB 0.0512 0.0890 0.0901 0.0512 0.1036 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 J48 0.2877 0.1287 0.2540 0.0170 0.2016 RF 0.2568 0.1019 0.2879 0.0563 0.1047 RF 0.2568 0.1019 0.2879 0.0563 0.3321 TwImbData15112016 RF 0.2568 0.1019 0.2879 0.0563 0.3041 NB 0.0563 0.0872 0.0892 0.0563 0.3041 NB 0.0563 0.0872 0.0892 0.0563 0.3041 NB 0.0563 0.0872 0.0892 0.0563 0.3321 SVM 0.0000 0.1386 0.1561 0.0000 0.1568 SVM 0.0000 0.1421 0.1600 0.0000 0.1568 SVM 0.0000 0.1386 0.1561 0.0000 0.1568 SVM 0.0000 0.1427 0.1644 0.0000 0.1743 TwImbData16112016 RF 0.3046 0.0260 0.3226 0.0496 0.3308 SVM 0.0000 0.1427 0.1644 0.0000 0.1743 TwImbData17112016 RF 0.2940 0.1050 0.3004 0.0282 0.3578 SVM 0.0000 0.1427 0.1644 0.0000 0.1743 TwImbData17112016 RF 0.2940 0.1050 0.3004 0.0282 0.3578 SVM 0.0000 0.1427 0.1644 0.0000 0.1681 RF 0.2940 0.1050 0.3004 0.0282 0.3578 SVM 0.0000 0.1427 0.1644 0.0000 0.1681 RF 0.2940 0.1050 0.3004 0.0282 0.3578 SVM 0.0000 0.1427 0.1644 0.0000 0.1681 RF 0.2940 0.1050 0.3004 0.0282 0.3578 SVM 0.0000 0.1427 0.1644 0.0000 0.1681 RF 0.2940 0.0050 0.3004 0.00504 0.0304 0.0318 SVM 0.0000 0.1427 0.1644 0.0000 0.1681 RF 0.2940 0.1050 0.3004		SVM	0.0000	0.1287	0.1443	0.0000	0.1489
TwimbData10112016 RF 0.3264 0.1148 0.3255 0.0333 0.3723 SVM 0.0000 0.1655 0.1821 0.0000 0.1930 TwImbData11112016 NB 0.03069 0.1324 0.2983 0.1970 0.2177 TwImbData11112016 RF 0.3524 0.0983 0.3343 0.0540 0.2977 SVM 0.0000 0.1467 0.1722 0.0000 0.1849 TwImbData12112016 RF 0.30302 0.1373 0.2945 0.0539 0.2295 SVM 0.0000 0.1546 0.1741 0.0000 0.1849 TwImbData13112016 RF 0.3043 0.1014 0.3757 0.0336 0.3575 SVM 0.0000 0.1546 0.1741 0.0000 0.1844 TwImbData13112016 NB 0.0604 0.0894 0.0886 0.0604 0.1036 RF 0.3350 0.1014 0.3757 0.0627 0.3444 SVM 0.0000 0.1501		J48	0.3263	0.1524	0.3320	0.1201	0.2851
RF	Tw.ImbData10112016	NB	0.0645	0.0924	0.0927	0.0645	0.1156
TwImbData11112016 TwImbData11112016 NB O.0631 O.0950 O.0960 O.0631 O.1161 RF O.3524 O.0983 O.3343 O.0540 O.2977 SVM O.0000 O.1467 O.1722 O.0000 O.1849 J48 O.3002 O.1373 O.2945 O.0896 O.0897 O.0756 O.1096 RF O.3043 O.1015 O.3255 O.0336 O.3575 SVM O.0000 O.1546 O.1741 O.0000 O.1880 TwImbData13112016 NB O.0604 O.0894 O.0896 O.0604 O.1080 RF O.3350 O.1014 O.3757 O.0627 O.3444 SVM O.0000 O.1501 O.1705 O.0890 O.0901 O.1756 RF O.2991 O.1012 O.0890 O.0901 O.0512 O.0890 O.0901 O.0512 O.0890 O.0901 O.0512 O.0890 O.0901 O.0170 O.2016 RF O.2991 O.1012 O.3304 O.0717 O.3172 SVM O.0000 O.1421 O.1600 O.0000 O.1703 TwImbData15112016 NB O.0563 O.8872 O.0892 O.0563 O.1017 O.2016 NB O.0563 O.0872 O.0892 O.0563 O.1047 O.2016 NB O.0563 O.0872 O.0890 O.0901 O.1703 O.2179 TwImbData16112016 NB O.0563 O.0872 O.0890 O.0900 O.1703 O.2179 TwImbData16112016 NB O.0563 O.0872 O.0890 O.0917 O.0441 O.0906 O.1568 TwImbData16112016 NB O.0000 O.1421 O.0000 O.1427 O.0441 O.0906 O.1568 O.0664 O.0894 O.0900 O.1703 O.0000 O.00000 O.00000 O.00000 O.00000 O.00000 O.00000 O.00000 O.00000 O.00000000	1 willibData10112010	RF	0.3264	0.1148	0.3525	0.0333	0.3723
TwImbData11112016 NB RF 0.0631 0.3524 0.0950 0.0983 0.0631 0.3343 0.0540 0.0540 0.2977 0.2977 SVM 0.0000 0.1467 0.1722 0.0000 0.1849 J48 0.3002 0.1373 0.2945 0.0539 0.2295 NB 0.0755 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 MB 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 SVM 0.0000 0.1501 0.1705 0.0000 0.1756 TwImbData14112016 RF 0.2306 0.1279 0.2852 0.0759 0.2308 TwImbData15112016 RF 0.2991 0.1012 0.3304 0.0717 0.3172		SVM	0.0000	0.1655	0.1821	0.0000	0.1930
NumbData11112016 RF 0.3524 0.0983 0.3343 0.0540 0.2977 SVM 0.0000 0.1467 0.1722 0.0000 0.1849 0.0000 0.1467 0.1722 0.0000 0.1849 0.0000 0.1849 0.0000 0.1849 0.0000 0.1849 0.0000 0.1849 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.000000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.0000000 0.000000 0.00000000		J48	0.3069	0.1324	0.2983	0.1970	0.2177
SVM 0.0000 0.1467 0.1722 0.0000 0.1849	TwImbData11119016	NB	0.0631	0.0950	0.0960	0.0631	0.1161
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIDData11112010	RF	0.3524	0.0983	0.3343	0.0540	0.2977
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1467	0.1722	0.0000	0.1849
TwImbData12112016 RF SVM 0.0000 0.1546 0.1741 0.0000 0.1844		J48	0.3002	0.1373	0.2945	0.0539	0.2295
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData12112016	NB	0.0755	0.0896	0.0897	0.0756	0.1096
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wimbData12112010	RF	0.3043	0.1015	0.3255	0.0336	0.3575
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1546	0.1741	0.0000	0.1844
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TIl.D. 4-12112016	J48	0.3052	0.1331	0.2891	0.0337	0.2253
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		NB	0.0604	0.0894	0.0886	0.0604	0.1080
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIDData15112010	RF	0.3350	0.1014	0.3757	0.0627	0.3444
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1501	0.1705	0.0000	0.1756
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.2806	0.1279	0.2852	0.0759	0.2308
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData14119016	NB	0.0512	0.0890	0.0901	0.0512	0.1036
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIDData14112010	RF	0.2991	0.1012	0.3304	0.0717	0.3172
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1421	0.1600	0.0000	0.1703
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48		0.1287	0.2540	0.0170	0.2016
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwimbData15112016	NB	0.0563	0.0872	0.0892	0.0563	0.1047
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wiiiibData15112010	RF	0.2568	0.1019	0.2879	0.0583	0.3321
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000		0.1561	0.0000	
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48		0.1312	0.2927	0.0540	0.2179
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TImbData16112016	NB	0.0441	0.0894	0.0917	0.0441	0.0976
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 willibData10112010	RF	0.3046	0.1026	0.3226	0.0496	0.3308
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1427	0.1644	0.0000	0.1743
TwlmbData17112016 RF 0.2940 0.1050 0.3004 0.0282 0.3578 SVM 0.0000 0.1479 0.1654 0.0000 0.1681 J48 0.2454 0.1220 0.2607 0.0881 0.2140 TwImbData18112016 NB 0.0504 0.0902 0.0894 0.0504 0.1081 RF 0.2818 0.0978 0.2611 0.0384 0.3318		J48	0.2811	0.1314	0.2735	0.0160	0.2149
RF 0.2940 0.1050 0.3004 0.0282 0.3578 SVM 0.0000 0.1479 0.1654 0.0000 0.1681 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	TwImbData17119016	NB	0.0568	0.0885	0.0896	0.0568	0.1114
TwImbData18112016	1 willibbata1/112010	RF	0.2940	0.1050	0.3004	0.0282	0.3578
TwImbData18112016			0.0000	0.1479	0.1654	0.0000	0.1681
RF 0.2818 0.0978 0.2611 0.0384 0.3318		J48			0.2607	0.0881	
RF = 0.2818 - 0.0978 - 0.2611 - 0.0384 - 0.3318	TwImbData19119016	NB	0.0504	0.0902	0.0894	0.0504	0.1081
SVM $0.0000 0.1342 0.1506 0.0000 0.1535$	1 w11110Data10112010		0.2818	0.0978		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
	J48	0.8320	0.8092	0.8164	0.8088	0.8160
Tw.Powog9019	NB	0.5880	0.6033	0.5874	0.5880	0.6625
TwReyes2013	RF	0.8607	0.8283	0.8605	0.8207	0.8580
	SVM	0.8025	0.7887	0.7903	0.7086	0.7981
	J48	0.8003	0.7289	0.7631	0.7311	0.7707
Tw.Inony.Panhioni2014	NB	0.5898	0.5841	0.5882	0.5898	0.6447
TwIronyBarbieri2014	RF	0.8196	0.7879	0.8247	0.7689	0.8282
	SVM	0.7443	0.7137	0.7180	0.4960	0.7289
	J48	0.3325	0.4165	0.3358	0.4168	0.3546
TwMohammad2015	NB	0.4341	0.4409	0.4452	0.4341	0.3879
1 wWonammad2015	RF	0.4564	0.4698	0.4564	0.3662	0.4701
	SVM	0.0384	0.4508	0.4405	0.0000	0.3601
	J48	0.8425	0.7948	0.8140	0.7900	0.8239
TwSarcasmBarbieri2014	NB	0.6902	0.6930	0.6906	0.6902	0.7108
1 wsarcasiii barbieri 2014	RF	0.8683	0.8249	0.8700	0.8110	0.8699
	SVM	0.8163	0.7798	0.7853	0.6606	0.7993
	J48	0.4296	0.4632	0.3866	0.5016	0.4202
TwRiloff2013	NB	0.4961	0.4870	0.4926	0.4961	0.4831
1 WKH0H2013	RF	0.5249	0.4926	0.5539	0.3639	0.5426
	SVM	0.1913	0.5411	0.5471	0.4267	0.5438
	J48	0.5941	0.6189	0.5790	0.6102	0.5930
TwPtáček2014	NB	0.5484	0.5470	0.5483	0.5484	0.5641
I WF tacekZU14	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

TwImbData08112016 RF 0.1040 0.0622 0.1053 0.0390 0.01428 SVM 0.0000 0.0748 0.0755 0.0000 0.0749 J48 0.1046 0.1302 0.2556 0.1274 0.2277 TwImbData09112016 RF 0.2565 0.1101 0.2917 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 J48 0.1801 0.1566 0.3308 0.1553 0.2839 TwImbData10112016 RF 0.2565 0.1101 0.2917 0.0527 0.3415 SVM 0.0000 0.1287 0.1443 0.0644 0.1489 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 J48 0.1552 0.1310 0.2977 0.1359 0.2160 TwImbData11112016 RF 0.3264 0.1148 0.3525 0.0333 0.3723 SVM 0.0000 0.1655 0.1821 0.0636 0.1930 TwImbData11112016 RF 0.3524 0.0983 0.3343 0.0540 0.2977 SVM 0.0000 0.1467 0.1722 0.0552 0.1849 J48 0.1293 0.1349 0.2930 0.1330 0.2333 TwImbData12112016 RF 0.3364 0.0983 0.3343 0.0540 0.2977 SVM 0.0000 0.1467 0.1722 0.0552 0.1849 J48 0.1293 0.1349 0.2930 0.1330 0.2333 TwImbData13112016 RF 0.3043 0.1015 0.3255 0.0336 0.3575 SVM 0.0000 0.1546 0.1741 0.0634 0.1844 TwImbData13112016 RF 0.3350 0.1014 0.3757 0.0627 0.3444 TwImbData14112016 RF 0.3350 0.1014 0.3757 0.0627 0.3444 TwImbData14112016 RF 0.3350 0.1014 0.3757 0.0627 0.3444 TwImbData15112016 RF 0.3568 0.1019 0.0825 0.0618 0.1703 SVM 0.0000 0.1546 0.1741 0.0634 0.1849 TwImbData14112016 RF 0.3560 0.1019 0.2879 0.0618 0.1763 NB 0.0850 0.0807 0.0825 0.0618 0.1703 TwImbData15112016 RF 0.2568 0.1019 0.2879 0.0583 0.3321 TwImbData16112016 RF 0.2568 0.1019 0.2879 0.0583 0.3321 TwImbData16112016 RF 0.3608 0.0865 0.0807 0.0825 0.0800 0.0915 SVM 0.0000 0.1386 0.1561 0.0586 0.1568 NB 0.0850 0.0803 0.0797 0.0805 0.0904 TwImbData16112016 RF 0.3046 0.1026 0.3226 0.0496 0.3308 SVM 0.0000 0.1427 0.1644 0.0630 0.1743 TwImbData16112016 RF 0.3046 0.1026 0.3226 0.0496 0.3308 SVM 0.0000 0.1427 0.1644 0.0630 0.1743 TwImbData16112016 RF 0.3046 0.1026 0.3226 0.0496 0.3308 SVM 0.0000 0.1427 0.1644 0.0630 0.1743 TwImbData16112016	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
NumbData08112016 RF 0.1040 0.0622 0.1053 0.0390 0.1428 0.0000 0.0748 0.0785 0.0000 0.0748 0.0785 0.0000 0.0748 0.0785 0.0000 0.0748 0.0785 0.0000 0.0748 0.0003 0.0763 0.0777 0.0783 0.0933 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.000000 0.00000 0.00000		J48	0.0079	0.0686	0.0979	0.0712	0.0696
NB	Tw.ImbData08112016	NB	0.0569	0.0561	0.0564	0.0569	0.0645
TwImbData09112016 J48 0.1046 0.1302 0.2556 0.1274 0.2277 TwImbData09112016 NB 0.0783 0.0763 0.0777 0.0783 0.0931 SVM 0.0000 0.1287 0.1443 0.0644 0.1489 TwImbData10112016 NB 0.0803 0.0832 0.0834 0.0833 0.1220 RF 0.3264 0.1148 0.5525 0.0333 0.3723 0.3723 SVM 0.0000 0.1655 0.1821 0.0636 0.1930 TwImbData11112016 NB 0.0806 0.0856 0.0819 0.0806 0.1030 RF 0.3524 0.0933 0.3343 0.0540 0.2977 SVM 0.0000 0.1655 0.1819 0.0806 0.0819 0.0806 0.0819 0.0806 0.0819 0.0900 0.1300 0.2977 0.1359 0.2160 0.0819 0.0814 0.0910 0.0812 0.0806 0.0819 0.0801 0.0791 0.0821 0.0813	1 wimbData08112010	RF	0.1040	0.0622	0.1053	0.0390	0.1428
TwImbData09112016 NB RF 0.0783 0.0773 0.0783 0.0933 SVM 0.0000 0.1287 0.1443 0.0444 0.0444 J48 0.1801 0.1506 0.3308 0.1583 0.2839 TwImbData10112016 NB 0.0833 0.0832 0.0834 0.0833 0.1020 RF 0.3264 0.1148 0.3525 0.0333 0.1320 SVM 0.0000 0.1655 0.1810 0.0297 0.1359 0.2160 TwImbData11112016 NB 0.0806 0.0856 0.0819 0.0806 0.1000 RF 0.3524 0.0933 0.3343 0.0500 0.1000 0.1000 0.0000 0.10467 0.1329 0.2160 0.0000 0.1467 0.1722 0.0552 0.1849 0.0000 0.1467 0.1722 0.0552 0.1849 0.0000 0.1467 0.1722 0.0552 0.1849 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 <td></td> <td>SVM</td> <td>0.0000</td> <td>0.0748</td> <td>0.0785</td> <td>0.0000</td> <td>0.0749</td>		SVM	0.0000	0.0748	0.0785	0.0000	0.0749
Numbright Numb		J48	0.1046	0.1302	0.2556	0.1274	0.2277
RF 0.2965 0.1101 0.2917 0.0341 SVM 0.0000 0.1287 0.1443 0.0644 0.1489 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 0.1810 0.0636 0.1000 RF 0.3524 0.0831 0.2977 0.1359 0.2160 SVM 0.0000 0.1665 0.0819 0.0806 0.0907 TwImbData12112016 RF 0.3524 0.0831 0.3343 0.0540 0.2977 TwImbData12112016 RF 0.3524 0.0831 0.3343 0.0540 0.2977 TwImbData12112016 RF 0.3443 0.1017 0.0712 0.0520 0.0336 0.3575 TwImbData13112016 RF 0.3360 0.1501 0.0703 0.0764 0.0939	Tw.ImbData00112016	NB	0.0783	0.0763	0.0777	0.0783	0.0933
TwImbData10112016 J48 0.1801 0.1506 0.3308 0.1583 0.2839 TwImbData10112016 NB 0.0833 0.0832 0.0834 0.0833 0.0323 SVM 0.0000 0.1655 0.1821 0.0636 0.1930 TwImbData11112016 NB 0.0806 0.0856 0.0819 0.0806 0.1000 RF 0.3524 0.0983 0.3343 0.0500 0.2977 SVM 0.0000 0.1467 0.1722 0.0552 0.1849 MB 0.0806 0.0856 0.0819 0.0297 SVM 0.0000 0.1467 0.1722 0.0552 0.1849 TwImbData12112016 RF 0.3403 0.1015 0.0255 0.0819 0.0233 TwImbData12112016 NB 0.0819 0.0791 0.0821 0.0819 0.0233 TwImbData13112016 NB 0.0819 0.0791 0.0821 0.0819 0.0234 TwImbData13112016 NB 0.0764 0.080	1 wiiiiDData09112010	RF	0.2565	0.1101	0.2917	0.0527	0.3415
TwimbData10112016 NB RF 0.0834 0.0834 0.0833 0.0723 TwimbData10112016 RF 0.3264 0.1148 0.3525 0.0333 0.3723 TwimbData11112016 NB 0.0800 0.1655 0.1821 0.0636 0.1930 TwimbData11112016 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 NB 0.0819 0.0791 0.0819 0.0297 RF 0.3524 0.0983 0.3343 0.0540 0.2977 TwimbData12112016 NB 0.0819 0.0793 0.0730 0.0233 TwimbData13112016 NB 0.0819 0.0791 0.0819 0.0819 0.0923 TwimbData14112016 RF 0.3350 0.1714 0.0634 0.0744 0.0933 TwimbData14112016 NB 0.0564<		SVM	0.0000	0.1287	0.1443	0.0644	0.1489
TwimbData10112016 RF 0.3264 0.1148 0.3255 0.0333 0.3723 SVM 0.0000 0.1655 0.1821 0.0636 0.1930 TwimbData11112016 NB 0.0806 0.0819 0.0806 0.0100 RF 0.3524 0.0983 0.3343 0.0540 0.2977 SVM 0.0000 0.1467 0.1722 0.0552 0.1849 TwimbData12112016 RF 0.3649 0.0791 0.0821 0.0819 0.0233 TwimbData12112016 RF 0.3043 0.1310 0.2333 0.3343 0.0349 RF 0.3040 0.0791 0.0821 0.0819 0.0233 0.0343 0.0336 0.3355 0.0336 0.3355 0.0336 0.3555 0.0336 0.3555 0.0364 0.0819 0.02923 0.0344 0.0000 0.1546 0.1741 0.0634 0.1844 0.1844 0.1240 0.05793 0.0764 0.0893 0.0767 0.0618 0.1750 0.0618 0.		J48	0.1801	0.1506	0.3308	0.1583	0.2839
RF 0.3264 0.1148 0.3255 0.0333 0.3723 TWImbData11112016 NB 0.0000 0.1655 0.1821 0.0636 0.1930 TwImbData11112016 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 NB 0.0819 0.0791 0.0821 0.0819 0.0923 TwImbData12112016 NB 0.0819 0.0791 0.0821 0.0819 0.0923 TwImbData12112016 RF 0.3043 0.1015 0.3255 0.0336 0.3575 SVM 0.0000 0.1546 0.1741 0.0634 0.1844 TwImbData13112016 RF 0.3330 0.0101 0.0793 0.0764 0.0939 TwImbData14112016 RF 0.3350 0.1014 0.3757 0.0627 0.3444 TwImbData15112016	Tw.ImbData10112016		0.0833	0.0832	0.0834	0.0833	0.1020
TwImbData11112016 J48 0.1552 0.1310 0.2977 0.1359 0.2160 TwImbData11112016 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 J48 0.1293 0.1349 0.2930 0.3330 0.2333 RF 0.3043 0.1015 0.3255 0.0336 0.3575 SVM 0.0000 0.1546 0.1741 0.0634 0.1844 TwImbData13112016 NB 0.0764 0.0801 0.0793 0.0764 0.0939 RF 0.3350 0.1310 0.0793 0.0764 0.0939 0.0793 0.0764 0.0939 TwImbData13112016 NB 0.0764 0.0801 0.0793 0.0627 0.3444 TwImbData14112016 RF 0.2901 0.1012 0.1705 0.0618 0.1756 RF 0.2	1 WIIIDData10112010	RF	0.3264	0.1148	0.3525	0.0333	0.3723
TwImbData11112016 NB RF 0.0806 0.0856 0.0819 0.0806 0.1000 SVM 0.0000 0.1467 0.1722 0.0552 0.1849 J48 0.1293 0.1349 0.2930 0.1330 0.2333 TwImbData12112016 RF 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 RF 0.3043 0.1015 0.2875 0.1317 0.2234 TwImbData13112016 RF 0.3350 0.1320 0.2875 0.1317 0.2234 TwImbData13112016 RF 0.3350 0.1014 0.3757 0.0627 0.3444 TwImbData14112016 NB 0.0820 0.0807 0.0825 0.0820 0.0915 TwImbData15112016 RF 0.2991 0.1012 0.3304 0.0717 0.3172 TwImbData16112016		SVM	0.0000	0.1655	0.1821	0.0636	0.1930
NumbData11112016 RF 0.3524 0.0983 0.3343 0.0540 0.2977 0.0000 0.1467 0.1722 0.0552 0.1849 0.0000 0.1467 0.1722 0.0552 0.1849 0.2018 0.2033 0.2333 0.0000 0.0000 0.0000 0.0001 0.0001 0.0001 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.00000 0.0000 0.000000 0.000000 0.00000 0.00000 0.00000 0.00000 0.000		J48	0.1552	0.1310	0.2977	0.1359	0.2160
RF 0.3524 0.0983 0.3343 0.0404 0.2977 J48 0.1293 0.1467 0.1722 0.0552 0.1849 TwImbData12112016 NB 0.0819 0.0791 0.0821 0.0819 0.0923 TwImbData12112016 RF 0.3043 0.1015 0.3255 0.0336 0.3575 SVM 0.0000 0.1546 0.1741 0.0634 0.1844 TwImbData13112016 RF 0.3350 0.1320 0.2875 0.1317 0.2234 RF 0.3350 0.1014 0.0793 0.0627 0.3444 SVM 0.0000 0.1501 0.1705 0.0618 0.1756 RF 0.3350 0.1014 0.3757 0.0627 0.3444 TwImbData14112016 NB 0.0820 0.0807 0.0825 0.0820 0.0915 RF 0.2991 0.1012 0.3304 0.0717 0.3172 TwImbData15112016 NB 0.0820 0.0807 0.0825 0.08	TwImbData11119016	NB	0.0806	0.0856	0.0819	0.0806	0.1000
TwImbData12112016 J48 NB RF 0.1293 0.1349 0.2930 0.1330 0.2333 0.2333 0.0923 TwImbData12112016 NB RF NB NB RF 0.3043 0.1015 0.3255 0.0336 0.3575 0.0375 0.0366 0.3575 SVM 0.0000 0.1546 0.1741 0.0634 0.1844 0.1844 TwImbData13112016 NB NB 0.0764 0.0801 0.0793 0.0764 0.0939 0.0764 0.0801 0.0793 0.0764 0.0939 RF 0.3350 0.1014 0.3757 0.0627 0.3444 0.0000 0.1501 0.1705 0.0618 0.1756 SVM 0.0000 0.1501 0.1705 0.0618 0.1256 0.2284 0.1517 0.1246 0.2840 0.1286 0.2284 TwImbData14112016 NB 0.0820 0.0807 0.0825 0.0820 0.0915 RF 0.2991 0.1012 0.3304 0.0717 0.3172 0.0000 0.1421 0.1600 0.0625 0.1703 TwImbData15112016 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 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 NB 0.0000 0.1427 0.1644 0.0630 0.1743 RF 0.3046 0.1026 0.3226 0.0496 0.3308 SVM 0.0000 0.1427 0.1644 0.0630 0.1743 TwImbData17112016 NB 0.0000 0.1427 0.1644 0.0630 0.1743 NB 0.0000 0.1427 0.1644 0.0630 0.1743 NB 0.0000 0.1427 0.1644	1 WIIIDData11112010	RF	0.3524	0.0983	0.3343	0.0540	0.2977
TwImbData12112016 NB RF 0.0819 0.0791 0.0821 0.0819 0.0923 SVM 0.3043 0.1015 0.3255 0.0336 0.3575 SVM 0.0000 0.1546 0.1741 0.0634 0.1844 TwImbData13112016 NB 0.0764 0.0801 0.0793 0.0764 0.0934 RF 0.3350 0.1014 0.3757 0.0627 0.3444 0.0936 0.0756 0.0801 0.0759 0.0627 0.3444 0.0916 0.0756 0.0801 0.0759 0.0627 0.3444 0.0916 0.0756 0.0801 0.0757 0.0627 0.3444 0.0916 0.0756 0.0818 0.0756 0.0820 0.0915 0.0717 0.0815 0.0820 0.0815 0.0820 0.0820 0.0915 0.0820 0.0825 0.0820 0.0915 0.0820 0.0915 0.0820 0.0915 0.0820 0.0915 0.0820 0.0915 0.0820 0.0915 0.0820 0.0915 0.0820 0.0915		SVM	0.0000	0.1467	0.1722	0.0552	0.1849
TwimbData12112016 RF 0.3043 0.1015 0.3255 0.0336 0.3575 SVM 0.0000 0.1546 0.1741 0.0634 0.1844 TwImbData13112016 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 NB 0.0820 0.0807 0.0825 0.0820 0.0915 RF 0.2991 0.1012 0.3304 0.0717 0.3172 0.0717 0.3172 SVM 0.0000 0.1421 0.1600 0.0625 0.1703 0.0717 0.3172 0.0717 0.3172 0.0717 0.3172 0.0717 0.03172 0.0717 0.03172 0.0717 0.03172 0.0717 0.0717 0.0717 0.0717 0.0717 0.0717 0.0717 0.0717 0.0717 0.0717 0.0717 0.0717 0.0717 0.0717		J48	0.1293	0.1349	0.2930	0.1330	0.2333
RF 0.3043 0.1015 0.3255 0.0336 0.3375 SVM 0.0000 0.1546 0.1741 0.0634 0.1844 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 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 NB 0.0815 0.0798 0.0817 0.0815 0.0886 TwImbData16112016 NB 0.0815 0.0798 0.0817 0.0815 0.0886 TwImbData16112016 NB 0.0805 0.0803 0.0797 0.0805 0.094 TwImbDa	Tw.ImbData12112016	NB	0.0819	0.0791	0.0821	0.0819	0.0923
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwimbData12112016	RF	0.3043	0.1015	0.3255	0.0336	0.3575
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1546	0.1741	0.0634	0.1844
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	T. I. I.D. (19110016	J48	0.1326	0.1320	0.2875	0.1317	0.2234
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		NB	0.0764	0.0801	0.0793	0.0764	0.0939
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wimbData15112010	RF	0.3350	0.1014	0.3757	0.0627	0.3444
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1501	0.1705	0.0618	0.1756
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.1517	0.1246	0.2840	0.1286	0.2284
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Trulock Date 14112016	NB	0.0820	0.0807	0.0825	0.0820	0.0915
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wiiiibData14112010	RF	0.2991	0.1012	0.3304	0.0717	0.3172
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1421	0.1600	0.0625	0.1703
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.1482	0.1264	0.2521	0.1267	0.2016
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwimbData15112016	NB	0.0815	0.0798	0.0817	0.0815	0.0886
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wimbData15112010	RF	0.2568	0.1019	0.2879	0.0583	0.3321
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1386	0.1561	0.0586	0.1568
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.1167	0.1274	0.2920	0.1312	0.2178
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData16112016	NB	0.0805	0.0803	0.0797	0.0805	0.0904
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wimbData10112010	RF	0.3046	0.1026	0.3226	0.0496	0.3308
TwImbData17112016 NB RF 0.0799 0.2940 0.0781 0.1050 0.0801 0.3004 0.0799 0.0282 0.0968 0.3578 SVM 0.0000 0.1479 0.1654 0.0588 0.1681 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	0.1427	0.1644	0.0630	0.1743
TwimbData17112016 RF 0.2940 0.1050 0.3004 0.0282 0.3578 SVM 0.0000 0.1479 0.1654 0.0588 0.1681 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			0.1630	0.1284	0.2735	0.1404	0.2160
RF 0.2940 0.1050 0.3004 0.0282 0.3578 SVM 0.0000 0.1479 0.1654 0.0588 0.1681 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	TwImbData17119016	NB	0.0799	0.0781	0.0801	0.0799	0.0968
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	r willioData1/112010	RF	0.2940	0.1050	0.3004	0.0282	0.3578
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1479	0.1654	0.0588	0.1681
1 WIMDData18112016 RF 0.2818 0.0978 0.2611 0.0384 0.3318		J48	0.1013	0.1203	0.2594	0.1251	0.2116
RF = 0.2818 - 0.0978 - 0.2611 - 0.0384 - 0.3318	TImb Data 10110016	NB	0.0762	0.0770	0.0763	0.0762	0.0977
SVM $0.0000 0.1342 0.1506 0.0622 0.1535$	TwlmbData18112016	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
	J48	0.2409	0.3059	0.2593	0.3137	0.2586
Tw.Powos2012	NB	0.2605	0.2893	0.2902	0.3226	0.2912
TwReyes2013	RF	0.2280	0.2990	0.2422	0.3014	0.2490
	SVM	0.2296	0.3068	0.3059	0.4134	0.2920
	J48	0.1830	0.2700	0.2115	0.2755	0.2061
TwIronyBarbiery2014	NB	0.2599	0.3071	0.3177	0.3901	0.3037
1 whony barbiery 2014	RF	0.1582	0.2257	0.1649	0.2298	0.1678
	SVM	0.1717	0.2459	0.2421	0.5677	0.2332
	J48	0.2431	0.4841	0.2799	0.4624	0.3167
TwMohammad2015	NB	0.3136	0.3773	0.3691	0.4324	0.3764
1 wwonammad2015	RF	0.0425	0.4971	0.1026	0.5521	0.1612
	SVM	0.0098	0.4049	0.3914	1.0000	0.2395
	J48	0.1895	0.2511	0.2106	0.2615	0.2030
TwSarcasmBarbieri2014	NB	0.2614	0.2971	0.3041	0.3516	0.2802
1 wSarcasiiiDarbieri2014	RF	0.1788	0.2326	0.1871	0.2332	0.1903
	SVM	0.1896	0.2539	0.2516	0.3769	0.2377
	J48	0.1822	0.4272	0.2492	0.4355	0.2797
TwRiloff2013	NB	0.3223	0.3833	0.3704	0.4254	0.4212
1 w1(110112013	RF	0.0795	0.3972	0.1387	0.4170	0.2173
	SVM	0.0361	0.3962	0.3731	0.7532	0.3745
	J48	0.2534	0.3888	0.2755	0.0374	0.3088
TwPtáček2014	NB	0.3874	0.4373	0.4388	0.2814	0.4519
I WI tateK2014	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
	J48	0.0005	0.3520	0.0282	0.4154	0.0225
TwImbData08112016	NB	0.2010	0.3210	0.3283	0.5041	0.3287
1 wimbData08112010	RF	0.0004	0.3448	0.0006	0.2924	0.0005
	SVM	0.0000	0.3138	0.3050	1.0000	0.3136
	J48	0.0087	0.2235	0.0259	0.2608	0.0308
TI	NB	0.1492	0.2666	0.2711	0.4327	0.4105
TwImbData09112016	RF	0.0007	0.2438	0.0011	0.2198	0.0023
	SVM	0.0000	0.2402	0.2158	0.5883	0.2023
	J48	0.0109	0.1960	0.0244	0.2109	0.0330
Trulmah Data 10112016	NB	0.1470	0.2284	0.2255	0.3349	0.3416
TwImbData10112016	RF	0.0004	0.2185	0.0012	0.1985	0.0033
	SVM	0.0000	0.1791	0.1684	0.5963	0.1518
	J48	0.0100	0.2212	0.0248	0.2414	0.0356
Trulmah Data 11119016	NB	0.1476	0.2304	0.2254	0.3280	0.3113
TwImbData11112016	RF	0.0006	0.2420	0.0009	0.2123	0.0014
	SVM	0.0000	0.1986	0.1760	0.6921	0.1540
	J48	0.0102	0.2192	0.0242	0.2478	0.0310
TI	NB	0.1446	0.2079	0.2072	0.3019	0.2965
TwImbData12112016	RF	0.0005	0.2295	0.0010	0.2053	0.0020
	SVM	0.0000	0.1891	0.1701	0.5985	0.1539
	J48	0.0101	0.2217	0.0260	0.2550	0.0347
TI	NB	0.1545	0.2241	0.2271	0.3334	0.3470
TwImbData13112016	RF	0.0007	0.2418	0.0011	0.2084	0.0018
	SVM	0.0000	0.1920	0.1754	0.6141	0.1621
	J48	0.0099	0.2311	0.0242	0.2638	0.0312
TI	NB	0.1405	0.2374	0.2356	0.3828	0.3896
TwImbData14112016	RF	0.0008	0.2542	0.0012	0.2262	0.0018
	SVM	0.0000	0.2119	0.1904	0.6078	0.1731
	J48	0.0094	0.2288	0.0233	0.2622	0.0355
TI	NB	0.1501	0.2347	0.2297	0.3482	0.4154
TwImbData15112016	RF	0.0006	0.2551	0.0009	0.2235	0.0014
	SVM	0.0000	0.2099	0.1920	0.6504	0.1878
	J48	0.0104	0.2276	0.0244	0.2582	0.0325
Trulmah Data 16112016	NB	0.1526	0.2443	0.2461	0.3806	0.3848
TwImbData16112016	RF	0.0005	0.2490	0.0009	0.2192	0.0011
	SVM	0.0000	0.2073	0.1857	0.6022	0.1669
	J48	0.0108	0.2242	0.0243	0.2330	0.0350
TwImbData 17110016	NB	0.1377	0.2380	0.2257	0.3526	0.3803
TwImbData17112016	RF	0.0003	0.2437	0.0007	0.2162	0.0018
	SVM	0.0000	0.1960	0.1817	0.6458	0.1738
	J48	0.0089	0.2378	0.0253	0.2643	0.0332
TII-D-7 10110010	NB	0.1521	0.2377	0.2354	0.3569	0.3975
TwImbData18112016	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
	J48	0.2409	0.3058	0.2593	0.2933	0.2586
TD2012	NB	0.2227	0.3010	0.3027	0.2635	0.2957
TwReyes2013	RF	0.2361	0.3072	0.2354	0.1949	0.2575
	SVM	0.2296	0.3068	0.3059	0.4134	0.2920
	J48	0.1830	0.2700	0.2115	0.2248	0.2061
Tw.Inony.Panhioni2014	NB	0.2003	0.3134	0.3198	0.1673	0.2983
TwIronyBarbieri2014	RF	0.1591	0.2221	0.1618	0.1148	0.1709
	SVM	0.1717	0.2459	0.2421	0.0000	0.2332
	J48	0.2431	0.4841	0.2799	0.3623	0.3167
TwMohammad2015	NB	0.0000	0.3955	0.3883	0.0000	0.3614
1 wWollammad2015	RF	0.1254	0.5448	0.1405	0.0228	0.2696
	SVM	0.0098	0.4049	0.3914	0.0000	0.2395
	J48	0.1895	0.2511	0.2106	0.2307	0.2030
TwSarcasmBarbieri2014	NB	0.2254	0.2959	0.3011	0.2717	0.2763
1 wsarcasiiibarbieriz014	RF	0.1803	0.2339	0.1857	0.1351	0.1908
	SVM	0.1896	0.2539	0.2516	0.3769	0.2377
	J48	0.1822	0.4272	0.2492	0.2973	0.2797
TwRiloff2013	NB	0.0000	0.3852	0.3727	0.0000	0.4073
1 WKIIOII2013	RF	0.1170	0.3815	0.1392	0.0365	0.2228
	SVM	0.0361	0.3962	0.3731	0.0000	0.3745
	J48	0.7466	0.3888	0.2755	0.8898	0.3088
TwPtáček2014	NB	1.0000	0.4334	0.4343	0.8758	0.4348
I WI TACEK2U14	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
	J48	0.0005	0.3520	0.0282	0.0000	0.0225
TwImbData08112016	NB	0.0000	0.3323	0.3381	0.0000	0.3254
1 wiiiibData08112010	RF	0.0006	0.3704	0.0005	0.0003	0.0010
	SVM	0.0000	0.3138	0.3050	0.0000	0.3136
	J48	0.0087	0.2235	0.0259	0.0000	0.0308
TwImbData09112016	NB	0.0000	0.2742	0.2770	0.0000	0.3968
1 wiiiibData09112010	RF	0.0010	0.2301	0.0009	0.0005	0.0048
	SVM	0.0000	0.2402	0.2158	0.0000	0.2023
	J48	0.0109	0.1958	0.0244	0.0000	0.0330
TwImbData10112016	NB	0.0000	0.2372	0.2338	0.0000	0.3220
1 willibData10112010	RF	0.0012	0.2088	0.0008	0.0003	0.0050
	SVM	0.0000	0.1791	0.1684	0.0000	0.1518
	J48	0.0100	0.2212	0.0248	0.0000	0.0356
TwImbData11112016	NB	0.0000	0.2378	0.2338	0.0000	0.2923
1 willibData11112010	RF	0.0010	0.2406	0.0008	0.0004	0.0029
	SVM	0.0000	0.1986	0.1760	0.0000	0.1540
-	J48	0.0102	0.2192	0.0242	0.0000	0.0310
TwImbData12112016	NB	0.0000	0.2187	0.2175	0.0000	0.2832
1 wimbData12112010	RF	0.0012	0.2174	0.0008	0.0003	0.0035
	SVM	0.0000	0.1891	0.1701	0.0000	0.1539
	J48	0.0101	0.2215	0.0260	0.0000	0.0347
TIh Do to 19119016	NB	0.0000	0.2335	0.2359	0.0000	0.3235
TwImbData13112016	RF	0.0013	0.2212	0.0010	0.0006	0.0034
	SVM	0.0000	0.1920	0.1754	0.0000	0.1621
	J48	0.0099	0.2311	0.0242	0.0000	0.0312
TIh Doto 14112016	NB	0.0000	0.2478	0.2460	0.0000	0.3742
TwImbData14112016	RF	0.0013	0.2359	0.0012	0.0006	0.0033
	SVM	0.0000	0.2119	0.1904	0.0000	0.1731
	J48	0.0094	0.2288	0.0233	0.0000	0.0355
TIh Doto 15112016	NB	0.0000	0.2443	0.2391	0.0000	0.3881
TwImbData15112016	RF	0.0010	0.2311	0.0008	0.0005	0.0030
	SVM	0.0000	0.2099	0.1920	0.0000	0.1878
	J48	0.0104	0.2276	0.0244	0.0000	0.0325
TIh Do to 16119016	NB	0.0000	0.2526	0.2528	0.0000	0.3658
TwImbData16112016	RF	0.0009	0.2396	0.0008	0.0004	0.0029
	SVM	0.0000	0.2073	0.1857	0.0000	0.1669
	J48	0.0108	0.2237	0.0243	0.0000	0.0350
TI	NB	0.0000	0.2477	0.2363	0.0000	0.3523
TwImbData17112016	RF	0.0008	0.2284	0.0005	0.0002	0.0034
	SVM	0.0000	0.1960	0.1817	0.0000	0.1738
	J48	0.0089	0.2378	0.0253	0.0000	0.0332
TILD-+ 10110012	NB	0.0000	0.2469	0.2440	0.0000	0.3727
TwImbData18112016	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
	J48	0.2490	0.3039	0.2551	0.2516	0.2638
Tw.Powog9019	NB	0.2506	0.3725	0.3780	0.2506	0.3276
TwReyes2013	RF	0.2428	0.3265	0.2559	0.1744	0.2718
	SVM	0.2296	0.3068	0.3059	0.4134	0.2920
	J48	0.2003	0.2818	0.2073	0.1995	0.2154
Tw.Inony.Panhioni2014	NB	0.2001	0.3333	0.3285	0.2001	0.2892
TwIronyBarbieri2014	RF	0.1607	0.2390	0.1630	0.1286	0.1765
	SVM	0.1717	0.2459	0.2421	0.0000	0.2332
	J48	0.2778	0.4826	0.2742	0.2773	0.3297
TwMohammad2015	NB	0.2732	0.4593	0.4469	0.2732	0.3225
1 wWonammad2015	RF	0.4810	0.7911	0.5219	0.2073	0.6008
	SVM	0.0098	0.4049	0.3914	0.0000	0.2395
	J48	0.1998	0.2585	0.2069	0.2006	0.2101
TwSarcasmBarbieri2014	NB	0.2004	0.2891	0.2893	0.2004	0.2604
1 wsarcasiii barbieri 2014	RF	0.1832	0.2433	0.1926	0.1494	0.1964
	SVM	0.1896	0.2539	0.2516	0.3769	0.2377
	J48	0.2164	0.4193	0.2460	0.2270	0.2834
TwRiloff2013	NB	0.2201	0.3935	0.3852	0.2201	0.3306
1 WKH0H2015	RF	0.2289	0.5543	0.2816	0.0855	0.3499
	SVM	0.0361	0.3962	0.3731	0.0000	0.3745
	J48	0.2696	0.3809	0.2692	0.2100	0.3172
TwPtáček2014	NB	0.2680	0.4185	0.4177	0.2680	0.3802
1 Wr tacek2014	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

TwImbData08112016 RF	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
Name		J48	0.0021	0.3539	0.0278	0.0144	0.0330
RF	TI	NB	0.0197	0.3863	0.3825	0.0197	0.3123
TwImbData10112016 TwImbData09112016 TwImbData09112016 TwImbData09112016 TwImbData09112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData11112016 TwImbData111112016 TwImbData11	1 w1mbData08112010	RF	0.0028	0.5525	0.0029	0.0004	0.0322
TwimbData09112016 NB RF 0.0197 0.3053 0.2958 0.0197 0.2237 SVM 0.0000 0.2402 0.2158 0.0006 0.0198 J48 0.0200 0.1894 0.0242 0.0007 0.0354 TwimbData10112016 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 MB 0.0195 0.2868 0.2771 0.0195 0.0225 RF 0.0110 0.3627 0.0131 0.0006 0.0255 SVM 0.0000 0.1986 0.1760 0.0000 0.1541 TwimbData12112016 NB 0.0187 0.2944 0.2920 0.0197 0.2209 TwimbData12112016 RF 0.0074 0.3496 0.0080 0.0003 0.0238 TwimbData13112016 NB 0.0181 0.2184 0.0290		SVM	0.0000	0.3138	0.3050	0.0000	0.3136
NumbData09112016 RF 0.0049 0.3238 0.0068 0.0006 0.0198 SVM 0.0000 0.2402 0.2158 0.0000 0.2023 0.246 0.0097 0.0354 0.0046 0.2866 0.2846 0.0196 0.2216 0.0097 0.0054 0.0006 0.2846 0.0196 0.2210 0.0007 0.3061 0.0081 0.0004 0.0206 0.0007 0.3061 0.0081 0.0004 0.0206 0.0006 0.0000 0.1791 0.1684 0.0000 0.1518 0.0006		J48	0.0184	0.2117	0.0257	0.0054	0.0344
RF	TI	NB	0.0197	0.3053	0.2958	0.0197	0.2237
TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData11112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData13112016 TwImbData131	1 willibData09112010	RF	0.0049	0.3238	0.0068	0.0006	0.0198
TwImbData10112016 NB RF 0.0077 0.3661 0.0246 0.0004 0.0206 SVM 0.0000 0.1791 0.1684 0.0004 0.0381 TwImbData11112016 NB 0.0194 0.2173 0.0247 0.0140 0.0384 TwImbData11112016 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 NB 0.0188 0.2133 0.0238 0.0064 0.0345 RF 0.0074 0.3496 0.080 0.0003 0.0228 SVM 0.0000 0.1891 0.1701 0.0000 0.1539 TwImbData13112016 RF 0.0074 0.3496 0.080 0.0003 0.0228 TwImbData14112016 RF 0.0089 0.3561 0.0106 0.0007 0.0293 TwImbData14112016 RF 0.009		SVM	0.0000	0.2402	0.2158	0.0000	0.2023
Name		J48	0.0200	0.1894	0.0242	0.0097	0.0354
RF	Tw.ImbData10112016	NB	0.0196	0.2869	0.2846	0.0196	0.2210
TwImbData11112016 TwImbData11112016 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 0.0005 0.1986 0.01760 0.0000 0.1540 0.0006 0.1760 0.0000 0.1540 0.0000 0.1986 0.02760 0.0000 0.1540 0.0000 0.1986 0.00197 0.2944 0.2920 0.0197 0.2209 0.0197 0.0209 0.01891 0.1701 0.0000 0.1539 0.2935 0.2934 0.0057 0.0052 0.0382 0.2944 0.0257 0.0052 0.0382 0.2944 0.0195 0.2935 0.2944 0.0196 0.0007 0.0293 0.2944 0.0106 0.0007 0.0293 0.0000 0.1621 0.0007 0.0293 0.0000 0.1621 0.0007 0.0000 0.1920 0.1754 0.0000 0.1621 0.0007 0.02182 0.0007 0.0000 0.1920 0.1754 0.0000 0.1621 0.0007 0.0272 0.0007 0.0000 0.1621 0.0007 0.0000 0.1731 0.0000 0.1878 0.0000 0.2119 0.1904 0.0000 0.1731 0.0350 0.0000 0.1878 0.0000 0.2995 0.2933 0.0196 0.2375 0.0000 0.1878 0.0000 0.2099 0.1920 0.00000 0.1878 0.0000 0.2099 0.1920 0.00000 0.1878 0.0000 0.1878 0.0000 0.2099 0.1920 0.00000 0.1878 0.0000 0.1878 0.0000 0.0000 0.1878 0.0000 0.0000 0.1878 0.0000 0.0000 0.1878 0.0000 0.0000 0.1878 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.00000 0.000000	1 wimbData10112010	RF	0.0077	0.3061	0.0081	0.0004	0.0206
TwImbData11112016 NB RF 0.0195 0.2868 0.2771 0.0195 0.2025 SVM 0.0000 0.1986 0.1760 0.0000 0.1540 J48 0.0188 0.2133 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 NB 0.0195 0.2935 0.2944 0.0957 0.052 0.0382 TwImbData13112016 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 NB 0.0196 0.2978 0.2904 0.0196 0.2182 TwImbData15112016 RF 0.0071 0.3570 0.0102 0		SVM	0.0000	0.1791	0.1684	0.0000	0.1518
NumbData11112016 RF 0.0110 0.3627 0.0131 0.0006 0.0255		J48	0.0194	0.2173	0.0247	0.0140	0.0384
SVM 0.0000 0.1986 0.1760 0.0006 0.0255	TwimbData11112016	NB	0.0195	0.2868	0.2771	0.0195	0.2025
TwImbData12112016	1 wimbData11112010	RF	0.0110	0.3627	0.0131	0.0006	0.0255
TwImbData12112016 NB RF 0.0197 0.2944 0.2920 0.0197 0.2209 SVM 0.0004 0.3496 0.0080 0.0003 0.0228 SVM 0.0000 0.1891 0.1701 0.0000 0.1539 TwImbData13112016 NB 0.0185 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 NB 0.0196 0.2297 0.2904 0.0196 0.2182 RF 0.0071 0.3570 0.0102 0.0077 0.0350 TwImbData14112016 NB 0.0196 0.2978 0.2904 0.0196 0.2182 TwImbData15112016 NB 0.0196 0.2978 0.2904 0.0196 0.2375 RF 0.0071 0.3570 0.0102 0.0007 0.0272 SVM 0.0196 0.2995 0.2933		SVM		0.1986	0.1760	0.0000	0.1540
Invimibility RF 0.0074 0.3496 0.0080 0.0003 0.0228 SVM 0.0000 0.1891 0.1701 0.0000 0.1539 J48 0.0181 0.2184 0.0257 0.0052 0.0382 TwImbData13112016 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 J48 0.0198 0.2211 0.0240 0.0137 0.0350 TwImbData14112016 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 NB 0.0196 0.2995 0.2933 0.0196 0.2375 RF 0.0057 0.3489 0.0072 0.0006 0.0243 </td <td></td> <td>J48</td> <td>0.0188</td> <td>0.2133</td> <td>0.0238</td> <td>0.0064</td> <td>0.0345</td>		J48	0.0188	0.2133	0.0238	0.0064	0.0345
RF 0.0074 0.3496 0.0003 0.0003 0.0228 SVM 0.0000 0.1891 0.1701 0.0000 0.1539 TwImbData13112016 NB 0.0181 0.2184 0.0257 0.0052 0.0382 RF 0.0089 0.3561 0.0106 0.0007 0.0293 SVM 0.0000 0.1920 0.1754 0.0000 0.1621 TwImbData14112016 NB 0.0196 0.2978 0.2904 0.0137 0.0350 TwImbData14112016 NB 0.0196 0.2978 0.2904 0.0196 0.2182 TwImbData15112016 NB 0.0196 0.2978 0.2904 0.0196 0.2182 TwImbData15112016 NB 0.0196 0.2978 0.2904 0.0007 0.0272 SVM 0.0000 0.2119 0.1904 0.0007 0.0272 TwImbData15112016 NB 0.0196 0.2995 0.2933 0.0196 0.2375 TwImbData16112016 NB 0.0196	Tw.ImbData12112016	NB	0.0197	0.2944	0.2920	0.0197	0.2209
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wimbData12112010	RF	0.0074	0.3496	0.0080	0.0003	0.0228
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1891	0.1701	0.0000	0.1539
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.0181	0.2184	0.0257	0.0052	0.0382
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.0195	0.2935	0.2944	0.0195	0.2243
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIDData15112010	RF	0.0089	0.3561	0.0106	0.0007	0.0293
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1920	0.1754	0.0000	0.1621
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.0198	0.2211	0.0240	0.0137	0.0350
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData14119016	NB	0.0196	0.2978	0.2904	0.0196	0.2182
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIDData14112010	RF	0.0071	0.3570	0.0102	0.0007	0.0272
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.2119	0.1904	0.0000	0.1731
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48		0.2197	0.0230	0.0040	0.0384
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwimbData15112016	NB	0.0196	0.2995	0.2933	0.0196	0.2375
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wiiiibData15112010	RF	0.0057	0.3489	0.0072	0.0006	0.0243
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.2099	0.1920	0.0000	0.1878
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48		0.2183	0.0242	0.0099	0.0366
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData16112016	NB	0.0198	0.2947	0.2873	0.0198	0.2186
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 WIIIDData10112010	RF	0.0074	0.3493	0.0080	0.0006	0.0295
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.2073	0.1857	0.0000	0.1669
TwImbData17112016 RF 0.0071 0.3383 0.0077 0.0003 0.0275 SVM 0.0000 0.1960 0.1817 0.0000 0.1738 J48 0.0194 0.2296 0.0251 0.0165 0.0366 TwImbData18112016 NB 0.0196 0.2952 0.2939 0.0196 0.2230 RF 0.0071 0.3661 0.0066 0.0004 0.0328		J48				0.0021	
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData17119016	NB	0.0196	0.2988		0.0196	
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	1 w1111012/ata1/112010	RF	0.0071	0.3383	0.0077	0.0003	0.0275
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$			0.0000	0.1960		0.0000	0.1738
1 WIMDData18112016 RF 0.0071 0.3661 0.0066 0.0004 0.0328		J48			0.0251	0.0165	
RF = 0.0071 - 0.3661 - 0.0066 - 0.0004 - 0.0328	TwImbData18119016	NB	0.0196	0.2952	0.2939	0.0196	0.2230
SVM 0.0000 0.2253 0.2050 0.0000 0.1943	i wimbibata10112010		0.0071			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
	J48	0.2433	0.3058	0.2593	0.3106	0.2586
TvvPovvog2012	NB	0.4252	0.4359	0.4262	0.4252	0.3726
TwReyes2013	RF	0.2444	0.3003	0.2510	0.1988	0.2604
	SVM	0.2296	0.3068	0.3059	0.4134	0.2920
	J48	0.1859	0.2700	0.2115	0.2736	0.2061
Tw.Inony.Parkioni2014	NB	0.3436	0.3543	0.3467	0.3436	0.2338
TwIronyBarbieri2014	RF	0.1615	0.2380	0.1678	0.1373	0.1758
	SVM	0.1717	0.2459	0.2421	0.5677	0.2332
	J48	0.2483	0.4841	0.2799	0.4587	0.3167
TwMohammad2015	NB	0.3868	0.4120	0.4825	0.3868	0.2701
1 wWonammad2015	RF	0.4810	0.7911	0.5219	0.2073	0.6008
	SVM	0.0098	0.4049	0.3914	0.0000	0.2395
	J48	0.1921	0.2511	0.2106	0.2597	0.2030
TwSarcasmBarbieri2014	NB	0.2963	0.2945	0.2963	0.2963	0.2287
1 wsarcasiii barbieri 2014	RF	0.1832	0.2433	0.1926	0.1494	0.1964
	SVM	0.1896	0.2539	0.2516	0.3769	0.2377
	J48	0.1914	0.4272	0.2492	0.4327	0.2797
TwRiloff2013	NB	0.4596	0.4601	0.4610	0.4596	0.3430
1 WKIIOII2015	RF	0.2289	0.5543	0.2816	0.0855	0.3499
	SVM	0.0361	0.3962	0.3731	0.7532	0.3745
	J48	0.2587	0.3888	0.2755	0.4533	0.3088
TwPtáček2014	NB	0.5460	0.5556	0.5448	0.5460	0.4690
1 wr tacekz014	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

TwImbData08112016 RF	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
NumbData08112016 RF 0.0028 0.5525 0.0029 0.0004 0.0322 0.0006 0.3138 0.3050 0.0000 0.3138 0.3050 0.0000 0.3138 0.3050 0.0000 0.3138 0.3050 0.0000 0.3088 0.0068 0.0066 0.1088 0.0066 0.0198 0.0006 0.0198 0.0006 0.0198 0.0006 0.0198 0.0006 0.0198 0.0006 0.0198 0.0006 0.0198 0.0006 0.00		J48	0.0083	0.3516	0.0282	0.4032	0.0225
NF	TII-D-4-09119016	NB	0.4860	0.4822	0.4886	0.4860	0.3037
TwImbData10112016 TwImbData09112016 TwImbData09112016 RF RF 0.0049 0.3238 0.0068 0.0060 0.0198 SVM 0.0000 0.2402 0.2158 0.5883 0.2023 J48 0.0955 0.1958 0.0244 0.22031 0.0330 RF 0.0077 0.3061 0.0081 0.0081 0.0004 0.0206 SVM 0.0000 0.1791 0.1684 0.5963 0.1518 TwImbData11112016 RF 0.0077 0.3061 0.0081 0.0081 0.0004 0.0206 SVM 0.0000 0.1791 0.1684 0.5963 0.1518 TwImbData11112016 RF 0.0110 0.3627 0.0131 0.0066 0.0255 SVM 0.0000 0.1986 0.3779 0.3948 0.2729 TwImbData12112016 RF 0.0010 0.3627 0.0131 0.0066 0.0255 SVM 0.0000 0.1986 0.1760 0.8921 0.1540 0.0001 0.1891 0.1701 0.5985 0.1539 TwImbData13112016 RF 0.0074 0.3496 0.0080 0.0003 0.0228 SVM 0.0000 0.1891 0.1701 0.5985 0.1539 TwImbData14112016 RF 0.0089 0.3561 0.0106 0.0007 0.0223 SVM 0.0000 0.1920 0.1754 0.6141 0.1621 0.488 0.1529 0.2288 0.0330 0.2412 0.2549 0.0312 TwImbData15112016 RF 0.0071 0.3570 0.0102 0.0007 0.0293 SVM 0.0000 0.1910 0.1754 0.6141 0.1621 0.1731 0.0006 0.0235 SVM 0.0000 0.1920 0.1754 0.6141 0.1621 0.1621 0.1731 0.0007 0.0273 0.00000 0.1920 0.1754 0.6141 0.1621 0.1621 0.1731 0.0006 0.0007 0.0275 0.0000 0.1910 0.0000 0.1900 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.6141 0.1621 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.0000 0.1754 0.1731 0.1731 0.00000 0.1754 0.1731 0.00000 0.1754 0.1731 0.1731 0.00000 0.1754 0.1731 0.1731 0.1731 0.1731 0.0	1 w1mbData08112016	RF	0.0028	0.5525	0.0029	0.0004	0.0322
TwImbData10112016 TwImbData09112016 RF RF 0.0049 0.3238 0.0068 0.0060 0.0198 SVM 0.0000 0.2402 0.2158 0.5883 0.2023 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 0.0081 0.0004 0.0206 SVM 0.0000 0.1791 0.1684 0.5963 0.1518 TwImbData11112016 RF 0.0107 0.3661 0.0081 0.0081 0.0004 0.0206 SVM 0.0000 0.1791 0.1684 0.5963 0.1518 RF 0.0110 0.3627 0.0131 0.0006 0.0255 SVM 0.0000 0.1986 0.1760 0.0921 0.1540 NB 0.3812 0.4041 0.3796 0.3812 0.3087 TwImbData12112016 RF 0.0074 0.3496 0.0080 0.0003 0.0228 SVM 0.0000 0.1891 0.1701 0.5985 0.1539 TwImbData13112016 RF 0.0074 0.3496 0.0080 0.0007 0.0293 SVM 0.0000 0.1920 0.1764 0.0007 0.0237 TwImbData14112016 RF 0.0089 0.3561 0.0106 0.0007 0.0293 SVM 0.0000 0.1920 0.1754 0.6141 0.1621 0.4044 0.3498 0.3916 0.4149 0.3038 TwImbData14112016 RF 0.0089 0.3561 0.0106 0.0007 0.0293 SVM 0.0000 0.1920 0.1754 0.6141 0.1621 0.4041 0.3760 0.3812 0.4044 0.3498 0.3916 0.4149 0.3038 TwImbData14112016 RF 0.0071 0.3570 0.0102 0.0007 0.0293 SVM 0.0000 0.1920 0.1754 0.6141 0.1621 0.4041 0.3767 0.3620 0.3643 0.2873 TwImbData15112016 RF 0.0071 0.3570 0.0102 0.0007 0.0275 SVM 0.0000 0.1920 0.3844 0.3851 0.3417 TwImbData16112016 RF 0.0077 0.3888 0.3891 0.3883 0.3851 0.3417 TwImbData16112016 RF 0.0077 0.3489 0.0072 0.0006 0.0244 0.0000 0.2199 0.1920 0.6504 0.1878 TwImbData17112016 RF 0.0077 0.3883 0.3922 0.3844 0.3851 0.3417 TwImbData16112016 RF 0.0077 0.3383 0.0080 0.0006 0.0295 SVM 0.0000 0.2099 0.1920 0.6504 0.1878 TwImbData17112016 RF 0.0071 0.3383 0.0077 0.0006 0.0295 SVM 0.0000 0.1960 0.1817 0.6458 0.1738 TwImbData18112016 RF 0.0071 0.3383 0.0077 0.0003 0.0225 SVM 0.0000 0.1960 0.1817 0.6458 0.1738 TwImbData18112016 RF 0.0071 0.3383 0.0077 0.0003 0.0295 SVM 0.0000 0.1960 0.1817 0.6458 0.1738 TwImbData18112016 RF 0.0071 0.3383 0.0077 0.0003 0.0295		SVM	0.0000	0.3138	0.3050	0.0000	0.3136
NumbData09112016 RF		J48	0.2374	0.2235	0.0259	0.2518	
RF	TII-D-+-00110016	NB	0.4061	0.4197	0.4086	0.4061	0.2887
TwImbData10112016 TwImbData10112016 TwImbData10112016 RF 0.0077 0.3061 0.0081 0.0004 0.0206 SVM 0.0000 0.1791 0.1684 0.5963 0.1518	1 w1mbData09112016	RF	0.0049	0.3238	0.0068	0.0006	0.0198
TwImbData10112016 NB RF 0.3719 0.3711 0.3686 0.3719 0.2948 RF 0.00077 0.3061 0.0081 0.0004 0.0206 SVM 0.0000 0.1791 0.1684 0.5963 0.1518 J48 0.1374 0.2212 0.0248 0.3311 0.0366 RF 0.0110 0.3627 0.0131 0.0006 0.0255 SVM 0.0000 0.1986 0.1760 0.6921 0.1540 TwImbData12112016 NB 0.3812 0.0300 0.0003 0.0222 RF 0.0074 0.3496 0.0080 0.0033 0.0228 SVM 0.0000 0.1891 0.1701 0.5985 0.1539 TwImbData13112016 RF 0.0074 0.3496 0.0080 0.0003 0.0228 SVM 0.0000 0.1891 0.1701 0.5985 0.1539 TwImbData13112016 RF 0.0089 0.3611 0.01449 0.3038 TwImbData1411		SVM	0.0000	0.2402	0.2158	0.5883	0.2023
Name		J48	0.0955	0.1958	0.0244	0.2031	0.0330
RF	TI	NB	0.3719	0.3711	0.3686	0.3719	0.2948
TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData13112016	1 w1mbData10112016	RF	0.0077	0.3061	0.0081	0.0004	0.0206
TwImbData11112016 NB RF 0.3948 0.3465 0.3779 0.3948 0.2729 SVM 0.0000 0.1986 0.1760 0.6921 0.1540 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 RF 0.0000 0.1891 0.1701 0.5985 0.1539 TwImbData13112016 RF 0.0000 0.1891 0.1701 0.5985 0.1539 TwImbData14112016 RF 0.0089 0.3561 0.0106 0.0007 0.0293 TwImbData14112016 NB 0.3643 0.3767 0.3620 0.3643 0.2873 TwImbData15112016 RF 0.0071 0.3570 0.0102 0.0007 0.0272 TwImbData16112016 RF 0.0		SVM	0.0000	0.1791	0.1684	0.5963	0.1518
NumbData11112016 RF 0.0110 0.3627 0.0131 0.0006 0.0255		J48	0.1374	0.2212	0.0248	0.2311	0.0356
SVM 0.0000 0.1986 0.1760 0.6921 0.1540	Trulmah Data 11119016	NB	0.3948	0.3465	0.3779	0.3948	0.2729
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData11112016	RF	0.0110	0.3627	0.0131	0.0006	0.0255
TwImbData12112016 NB RF 0.3812 0.0074 0.4041 0.3496 0.080 0.0080 0.0003 0.0028 0.0228 SVM 0.0000 0.1891 0.1701 0.5985 0.1539 TwImbData13112016 NB RF 0.0037 0.2215 0.0260 0.2426 0.0347 TwImbData13112016 NB RF 0.0089 0.3561 0.0106 0.0007 0.0293 SVM 0.0000 0.1920 0.1754 0.6141 0.1621 TwImbData14112016 NB RF 0.0007 0.0293 0.0016 0.0007 0.0293 TwImbData14112016 NB RF 0.03643 0.3767 0.3620 0.3643 0.2873 TwImbData15112016 NB RF 0.0071 0.3570 0.0102 0.0007 0.0272 SVM 0.0000 0.2119 0.1904 0.6078 0.1731 TwImbData15112016 NB RF 0.0057 0.3489 0.0233 0.2541 0.0355 TwImbData16112016 NB RF 0.0057 0.3489 0.0072 0.0066		SVM	0.0000	0.1986	0.1760	0.6921	0.1540
Invimibility RF 0.0074 0.3496 0.0080 0.0003 0.0228 SVM 0.0000 0.1891 0.1701 0.5985 0.1539 J48 0.1537 0.2215 0.0260 0.2426 0.0347 TwImbData13112016 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 J48 0.1291 0.2311 0.0242 0.2549 0.0312 TwImbData14112016 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 NB 0.3851 0.3992 0.3844 0.3851 0.3417 RF 0.0057 0.3489 0.0072 0.0066 0.0243 </td <td></td> <td>J48</td> <td>0.1825</td> <td>0.2192</td> <td>0.0242</td> <td>0.2369</td> <td>0.0310</td>		J48	0.1825	0.2192	0.0242	0.2369	0.0310
RF 0.0074 0.3496 0.0008 0.0003 0.0228 SVM 0.0000 0.1891 0.1701 0.5985 0.1539 TwImbData13112016 NB 0.1537 0.2215 0.0260 0.2426 0.0347 RF 0.0089 0.3561 0.0106 0.0007 0.0293 SVM 0.0000 0.1920 0.1754 0.6141 0.1621 TwImbData14112016 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 NB 0.3851 0.3992 0.3844 0.3851 0.3417 TwImbData16112016 RF 0.0057 0.3489 0.0072 0.0066 0.0243 SVM 0.0000 0.2099 0.1920 0.6504 0.1878 TwImbData16112016 NB 0.3883 0.3885 0.3921 0.3	TI	NB	0.3812	0.4041	0.3796	0.3812	0.3087
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData12112016	RF	0.0074	0.3496	0.0080	0.0003	0.0228
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1891	0.1701	0.5985	0.1539
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.1537	0.2215	0.0260	0.2426	0.0347
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.4149	0.3908	0.3916	0.4149	0.3038
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData13112016	RF	0.0089	0.3561	0.0106	0.0007	0.0293
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1920	0.1754	0.6141	0.1621
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.1291	0.2311	0.0242	0.2549	0.0312
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.3643	0.3767	0.3620	0.3643	0.2873
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData14112010	RF	0.0071	0.3570	0.0102	0.0007	0.0272
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.2119	0.1904	0.6078	0.1731
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.1529	0.2288	0.0233	0.2541	0.0355
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Trulmah Data 15112016	NB	0.3851	0.3992	0.3844	0.3851	0.3417
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 wimbData15112010	RF	0.0057	0.3489	0.0072	0.0006	0.0243
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.2099	0.1920	0.6504	0.1878
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.2187	0.2276	0.0244	0.2493	0.0325
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.3883	0.3858	0.3921	0.3883	0.2897
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 willibData10112010	RF	0.0074	0.3493	0.0080	0.0006	0.0295
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.2073	0.1857	0.6022	0.1669
TwlmbData17112016 RF 0.0071 0.3383 0.0077 0.0003 0.0275 SVM 0.0000 0.1960 0.1817 0.6458 0.1738 J48 0.2521 0.2378 0.0253 0.2570 0.0332 TwImbData18112016 NB 0.4243 0.4088 0.4258 0.4243 0.2956 RF 0.0071 0.3661 0.0066 0.0004 0.0328		J48	0.1142	0.2237	0.0243	0.2241	0.0350
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData17119016	NB	0.4080	0.4146	0.4075	0.4080	0.2994
TwImbData18112016	1 wiiiiDData1/112010	RF	0.0071	0.3383	0.0077	0.0003	0.0275
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.1960	0.1817	0.6458	0.1738
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		J48	0.2521	0.2378	0.0253	0.2570	0.0332
RF = 0.0071 - 0.3661 - 0.0066 - 0.0004 - 0.0328	Tw.ImbData10110016	NB	0.4243	0.4088	0.4258		0.2956
SVM 0.0000 0.2253 0.2050 0.6095 0.1943	1 willioData16112010	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

5 F-score

Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
	J48	0.8932	0.8496	0.8725	0.8451	0.8726
TD2012	NB	0.7067	0.7164	0.7044	0.7007	0.7612
TwReyes2013	RF	0.9180	0.8656	0.9117	0.8658	0.9077
	SVM	0.8819	0.8359	0.8373	0.7716	0.8466
	J48	0.8898	0.7977	0.8440	0.7961	0.8523
TwIronyBarbieri2014	NB	0.7229	0.7083	0.7092	0.6901	0.7146
1 wirony Barbieri 2014	RF	0.9303	0.8570	0.9254	0.8563	0.9235
	SVM	0.8657	0.7965	0.8006	0.6566	0.8111
	J48	0.5467	0.5513	0.5407	0.5525	0.5420
TwMohammad2015	NB	0.5847	0.5847	0.5853	0.5776	0.5760
1 wMonammad2015	RF	0.6152	0.5713	0.5828	0.5834	0.6090
	SVM	0.3655	0.5871	0.5824	0.0000	0.5748
	J48	0.9101	0.8432	0.8760	0.8348	0.8877
TwSarcasmBarbieri2014	NB	0.7845	0.7699	0.7648	0.7410	0.7623
1 w5arcasiiiDarbiei12014	RF	0.9367	0.8762	0.9311	0.8792	0.9281
	SVM	0.8941	0.8331	0.8372	0.7452	0.8516
	J48	0.6458	0.6167	0.5972	0.6412	0.6094
TwRiloff2013	NB	0.6461	0.6407	0.6411	0.6377	0.6257
1 WKHOH2015	RF	0.7222	0.6620	0.7290	0.6669	0.7063
	SVM	0.7514	0.6678	0.6732	0.6135	0.6711
	J48	0.7274	0.2916	0.2908	0.8266	0.2932
TwPtáček2014	NB	0.6392	0.3558	0.3558	0.6265	0.3353
I WF taceKZU14	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

TwImbData08112016 RF 0.9904 0.5177 0.9655 0.5206 0.9654 SVM 0.0000 0.5147 0.5160 0.0000 0.5147 TwImbData09112016 RF 0.9904 0.5177 0.9655 0.5206 0.9654 SVM 0.0000 0.5147 0.5160 0.0000 0.5147 A48 0.7105 0.5330 0.6052 0.5320 0.5861 TwImbData09112016 RF 0.9760 0.5355 0.9581 0.5389 0.8558 SVM 0.0000 0.5325 0.9581 0.5389 0.8558 SVM 0.0000 0.5329 0.5377 0.5166 0.5389 TwImbData10112016 RF 0.9904 0.5393 0.6435 0.5407 0.50674 RF 0.9904 0.5393 0.6435 0.5407 0.5074 TwImbData11112016 RF 0.9904 0.5392 0.9599 0.5435 0.8568 SVM 0.0000 0.5404 0.5493 0.5164 0.5527 TwImbData11112016 RF 0.9904 0.5392 0.9599 0.5435 0.8569 SVM 0.0000 0.5345 0.6227 0.5227 0.5224 0.5220 TwImbData11112016 RF 0.9904 0.5383 0.6227 0.5227 0.5204 0.5230 SVM 0.0000 0.5381 0.5462 0.5141 0.5505 TwImbData12112016 RF 0.8789 0.5366 0.9836 0.5405 0.9101 SVM 0.0000 0.5405 0.5467 0.5163 0.5501 TwImbData13112016 RF 0.8789 0.5366 0.9836 0.5405 0.9101 SVM 0.0000 0.5405 0.5467 0.5163 0.5501 TwImbData14112016 RF 0.9906 0.5349 0.9664 0.5402 0.9002 SVM 0.0000 0.5391 0.5467 0.5196 0.5223 TwImbData14112016 RF 0.9906 0.5340 0.9604 0.5160 0.8962 TwImbData15112016 RF 0.9906 0.5340 0.9604 0.5160 0.5465 TwImbData15112016 RF 0.9906 0.5340 0.5204 0.5185 0.5194 RF 0.9906 0.5340 0.9000 0.5117 0.5710 TwImbData16112016 RF 0.9906 0.5340 0.9604 0.5160 0.5465 TwImbData17112016 RF 0.9906 0.5340 0.5010 0.5117 0.5710 NB 0.5202 0.5208 0.5217 0.5196 0.5323 TwImbData16112016 RF 0.9905 0.5333 0.9844 0.5462 0.5143 TwImbData16112016 RF 0.9905 0.5333 0.9844 0.5374 0.9212 SVM 0.0000 0.5356 0.5411 0.5151 0.5413 TwImbData16112016 RF 0.9905 0.5330 0.6240 0.5185 0.5159 TwImbData16112016 RF 0.9905 0.5346 0.9348 0.5384 0.9212 SVM 0.0000 0.5358 0.5217 0.5191 0.5207 TwImbData16112016	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
NumbData08112016		J48	0.0559	0.5128	0.5325	0.5144	0.5235
RF	TIh Dot - 00112016	NB	0.5092	0.5096	0.5093	0.5097	0.5101
TwImbData10112016 TwImbData09112016 TwImbData09112016 TwImbData09112016 TwImbData09112016 TwImbData09112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData11112016 TwImbData111	1 w1mbData08112010	RF	0.9904	0.5177	0.9655	0.5206	0.9654
TwimbData09112016 NB RF 0.5200 0.5185 0.5192 0.5178 0.5191 SVM 0.0000 0.5325 0.5387 0.5389 0.5858 J48 0.7076 0.5393 0.6435 0.5407 0.6074 TwimbData10112016 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 J48 0.7147 0.5332 0.6277 0.5342 0.5781 RF 0.7924 0.5333 0.9907 0.5404 0.5236 RF 0.7924 0.5332 0.6277 0.5342 0.5781 TwimbData11112016 NB 0.5238 0.5227 0.5204 0.5236 SVM 0.0000 0.5381 0.5462 0.5141 0.5505 TwimbData12112016 RF 0.8789 0.5366 0.9836 0.5401 RF		SVM	0.0000	0.5147	0.5160	0.0000	0.5147
NumbData(19112016 RF 0.9760 0.5355 0.9581 0.5389 0.8558 SVM 0.0000 0.5329 0.5377 0.5166 0.5389 0.5389 0.5389 0.5389 0.5389 0.5389 0.5389 0.5389 0.5389 0.5328 0.5221 0.5220 0.5199 0.5229 0.5220 0.5199 0.5229 0.5435 0.8569 0.5000 0.5440 0.5493 0.5164 0.5527 0.5260 0.5221 0.5220 0.5199 0.5229 0.5435 0.8569 0.5221 0.5238 0.5247 0.5332 0.6277 0.5342 0.5781 0.5238 0.5227 0.5227 0.5342 0.5238 0.5227 0.5227 0.5204 0.5238 0.5227 0.5227 0.5204 0.5238 0.5227 0.5227 0.5404 0.5238 0.5227 0.52527 0.5204 0.5238 0.5227 0.5227 0.5204 0.5238 0.5227 0.5227 0.5204 0.5238 0.5227 0.5227 0.5204 0.5238 0.5227 0.5258 0.5205 0.520		J48	0.7105	0.5330	0.6052	0.5320	0.5861
RF 0.9760 0.5352 0.5389 0.5389 0.5389 SVM 0.0000 0.5329 0.5377 0.5166 0.5389 J48 0.7076 0.5393 0.6435 0.5407 0.6074 TwImbData10112016 RF 0.9904 0.5392 0.9599 0.5435 0.8569 SVM 0.0000 0.5440 0.5493 0.5164 0.5527 J48 0.7147 0.5332 0.6277 0.5342 0.5781 TwImbData11112016 RF 0.7924 0.5333 0.9907 0.5400 0.8995 SVM 0.0000 0.5381 0.5462 0.5141 0.5505 TwImbData12112016 RF 0.7924 0.5331 0.5462 0.5141 0.5505 TwImbData12112016 RF 0.8789 0.5345 0.6259 0.5332 0.5906 RF 0.8789 0.5366 0.9836 0.5405 0.9101 SVM 0.0000 0.5405 0.5467 0.5163 0.5202	TIh Dot - 00112016	NB	0.5200	0.5185	0.5192	0.5178	0.5191
TwImbData10112016 TwImbData10112016 RF O.9904 O.5326 O.5221 O.5220 O.5199 O.5229 O.5526 RF O.9904 O.5392 O.9599 O.5435 O.8569 SVM O.0000 O.5440 O.5493 O.5493 O.5164 O.5237 TwImbData11112016 RF O.7924 O.5332 O.6277 O.5342 O.5781 TwImbData11112016 RF O.7924 O.5338 O.5227 O.5207 O.5204 O.5906 RF O.7924 O.5353 O.9907 O.5400 O.8995 SVM O.0000 O.5381 O.5462 O.5141 O.5505 AB O.5232 O.5217 O.5220 O.5196 O.5223 RF O.8789 O.5366 O.9836 O.5405 O.9101 RF O.9906 O.5349 O.5647 O.5163 O.5501 TwImbData13112016 RF O.9906 O.5349 O.5040 O.5000 O.5340 O.5040 O.5050 O.5345 O.5467 O.5163 O.5501 TwImbData14112016 RF O.9906 O.5349 O.5040 O.5324 O.5205 O.5193 O.5209 TwImbData14112016 RF O.9906 O.5340 O.5024 O.5155 O.5194 RF O.9906 O.5340 O.9762 O.5204 O.5185 O.5194 TwImbData15112016 RF O.9906 O.5340 O.5000 O.5368 O.5405 O.5193 O.5207 O.5206 O.5193 O.5207 O.5217 O.5210 O.5217 O.5210 O.5185 O.5193 TwImbData16112016 RF O.9906 O.5340 O.5324 O.5682 O.5773 O.5906 TwImbData16112016 RF O.9906 O.5340 O.5324 O.5682 O.5193 O.5207 O.5204 O.5185 O.5193 O.5207 O.5206 O.5312 O.5207 O.5207 O.5217 O.5710 O	1 WIIIDData09112010	RF	0.9760	0.5355	0.9581	0.5389	0.8558
TwImbData10112016 NB RF 0.5236 0.5221 0.5220 0.5199 0.5229 RF 0.9904 0.5332 0.9599 0.5435 0.8569 SVM 0.0000 0.5440 0.5493 0.5164 0.5527 J48 0.7147 0.5332 0.6277 0.5342 0.5781 TwImbData11112016 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 J48 0.7058 0.5345 0.6259 0.5332 0.5906 TwImbData12112016 NB 0.5232 0.5217 0.5220 0.5196 0.5223 TwImbData13112016 NB 0.5221 0.5205 0.5196 0.5825 TwImbData14112016 RF 0.9906 0.5349 0.5467 0.5163 0.5209 TwImbData14112016 RF 0.9906 0.5340 0.9624		SVM	0.0000	0.5329	0.5377	0.5166	0.5389
NumbData10112016 RF 0.9904 0.5392 0.9599 0.5435 0.8569 SVM 0.0000 0.5440 0.5493 0.5164 0.5527 0.5272 0.5272 0.5272 0.5272 0.5272 0.5272 0.5272 0.5204 0.5230 0.5331 0.5462 0.5141 0.5505 0.5230 0.5405 0.9101 0.5230 0.5405 0.9101 0.5230 0.5405 0.5230 0.5230 0.5405 0.5230 0.5330 0.5330 0.5330 0.5330 0.5330 0.5330		J48	0.7076	0.5393	0.6435	0.5407	0.6074
RF	TI	NB	0.5236	0.5221	0.5220	0.5199	0.5229
TwImbData11112016 TwImbData11112016 NB 0.5238 0.5227 0.5227 0.5204 0.5239 0.5227 0.5204 0.5239 0.5227 0.5204 0.5230 0.59907 0.5400 0.8995 SVM 0.0000 0.5381 0.5462 0.5141 0.5505 0.5205 0.5196 0.5223 0.5207 0.5200 0.5223 0.5906 NB 0.5232 0.5217 0.5220 0.5196 0.5223 0.5906 NB 0.0000 0.5405 0.5467 0.5163 0.5501 NB 0.5221 0.5205 0.5205 0.5193 0.5209 SVM 0.0000 0.5349 0.9664 0.5402 0.9002 SVM 0.0000 0.5349 0.9664 0.5402 0.9002 SVM 0.0000 0.5391 0.5456 0.5159 0.5472 0.5205 NB 0.5209 0.5200 0.5204 0.5185 0.5194 RF 0.9906 0.5340 0.9762 0.5375 0.9246 SVM 0.0000 0.5368 0.5424 0.5161 0.5456 SVM 0.0000 0.5368 0.5424 0.5161 0.5456 SVM 0.0000 0.5333 0.9844 0.5374 0.9222 SVM 0.0000 0.5333 0.9844 0.5374 0.9222 SVM 0.0000 0.5335 0.9844 0.5374 0.9222 SVM 0.0000 0.5336 0.5410 0.5151 0.5413 D.5413 D.5413 D.5413 D.5413 D.5413 D.5413 D.5413 D.5511 D.5414 D.5511 D.5413 D.5511 D.5414 D.5511 D.5413 D.551	1 w1mbData10112016	RF	0.9904	0.5392	0.9599	0.5435	0.8569
TwImbData11112016 NB RF 0.5238 0.5227 0.5227 0.5204 0.5239 SVM 0.0000 0.5381 0.5462 0.5141 0.5505 J48 0.7058 0.5341 0.5462 0.5141 0.5505 TwImbData12112016 NB 0.5232 0.5217 0.5229 0.5332 0.5906 RF 0.8789 0.5366 0.9836 0.5405 0.9101 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5201 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5203 0.5204 0.5163 0.5501 0.5204 0.5825 0.5204 0.5163 0.5201 0.5204 0.5163 0.5202 0.5204 0.5193 0.5209 0.5204 0.5193 0.5209 0.5204 0.5193 0.5209 0.5204 0.5185 0.5193 0.5204 0.5825 0.5204 0.5185		SVM	0.0000	0.5440	0.5493	0.5164	0.5527
Invimibility RF 0.7924 0.5353 0.9907 0.5400 0.8995 SVM 0.0000 0.5381 0.5462 0.5141 0.5505 Image: Problem of the pr		J48	0.7147	0.5332	0.6277	0.5342	0.5781
SVM 0.0000 0.5381 0.5462 0.5141 0.5505	TI.m.h.Dt11119016	NB	0.5238	0.5227	0.5227	0.5204	0.5230
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData11112016	RF	0.7924	0.5353	0.9907	0.5400	0.8995
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.5381	0.5462	0.5141	0.5505
TwImbData12112016 RF SVM 0.0000 0.5405 0.5467 0.5163 0.5501 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 J48 0.6967 0.5312 0.6223 0.5324 0.5825 NB 0.5209 0.5200 0.5204 0.5159 0.5472 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 SVM 0.0000 0.5368 0.5424 0.5161 0.5456 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 MB 0.5202 0.5208 0.5211 0.5151 0.5413 NB 0.5204 0.5208 0.5211 0.5151 0.5413 TwImbData16112016 RF 0.9905 0.5331 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 0.5368 0.5424 0.5161 0.5456 0.5468 NB 0.5219 0.5207 0.5217 0.5191 0.5207 TwImbData17112016 RF 0.7923 0.5341 0.9530 0.5376 0.5376 0.5771 0.5771 NB 0.5219 0.5207 0.5217 0.5186 0.5207 0.5370 0.5865		J48	0.7058	0.5345	0.6259	0.5332	0.5906
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.5232	0.5217	0.5220	0.5196	0.5223
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData12112016	RF	0.8789	0.5366	0.9836	0.5405	0.9101
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.5405	0.5467	0.5163	0.5501
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.7036	0.5335	0.6197	0.5326	0.5825
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.5221	0.5205	0.5205	0.5193	0.5209
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData13112016	RF	0.9906	0.5349	0.9664	0.5402	0.9002
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.5391	0.5456	0.5159	0.5472
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.6967	0.5312	0.6223	0.5324	0.5862
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.5209	0.5200	0.5204	0.5185	0.5194
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData14112016	RF	0.9906	0.5340	0.9762	0.5379	0.9246
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.5368	0.5424	0.5161	0.5456
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.7141	0.5318	0.6091	0.5317	0.5710
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.5202	0.5208	0.5211	0.5192	0.5193
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData15112016	RF	0.9905	0.5333	0.9844	0.5374	0.9222
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.5356	0.5411	0.5151	0.5413
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.7009	0.5321	0.6254	0.5331	0.5835
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI.m.h.D. + 0.16119016	NB	0.5204	0.5208	0.5210	0.5187	0.5182
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData16112016	RF	0.9905	0.5346	0.9348	0.5384	0.9212
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.5369	0.5438	0.5162	0.5468
TwimbData17112016 RF 0.7923 0.5341 0.9530 0.5378 0.8742 SVM 0.0000 0.5384 0.5440 0.5151 0.5447 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		J48	0.6792	0.5323	0.6169	0.5352	0.5773
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Tw.ImbData 17119016	NB	0.5219	0.5207	0.5217	0.5191	0.5207
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	i willibDatai (112016	RF	0.7923	0.5341	0.9530	0.5378	0.8742
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		SVM	0.0000	0.5384	0.5440	0.5151	0.5447
RF 0.9904 0.5324 0.9780 0.5370 0.8652		J48	0.6647	0.5298	0.6080		0.5771
RF = 0.9904 0.5324 0.9780 0.5370 0.8652	TI.m.h.D. 4 . 10110016	NB	0.5219	0.5210	0.5207	0.5186	0.5207
SVM $0.0000 0.5344 0.5396 0.5160 0.5403$	1 willibrata18112016	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

	J48	0.8933	0.8497	0.8725	0.8618	0.8726
TI D 0019	NB	0.7085	0.7146	0.7026	0.7065	0.7611
TwReyes2013	RF	0.9140	0.8596	0.9146	0.9221	0.9023
	SVM	0.8819	0.8359	0.8373	0.7716	0.8466
	J48	0.8898	0.7977	0.8440	0.8460	0.8523
Trulmanu Dankiani 2014	NB	0.7407	0.7071	0.7087	0.6929	0.7179
TwIronyBarbieri2014	RF	0.9293	0.8601	0.9277	0.9291	0.9204
	SVM	0.8657	0.7965	0.8006	0.2000	0.8111
	J48	0.5467	0.5513	0.5407	0.5398	0.5420
TruMahamamad2015	NB	0.2758	0.5855	0.5799	0.2758	0.5783
TwMohammad2015	RF	0.6093	0.5820	0.5915	0.6863	0.5928
	SVM	0.4758	0.5871	0.5824	0.2758	0.5748
	J48	0.9101	0.8432	0.8760	0.8711	0.8877
TwSarcasmBarbieri2014	NB	0.7997	0.7705	0.7665	0.7806	0.7652
1 wsarcasin barbieri 2014	RF	0.9362	0.8749	0.9322	0.9404	0.9274
	SVM	0.8941	0.8331	0.8372	0.7452	0.8516
	J48	0.6458	0.6167	0.5972	0.6716	0.6094
TwRiloff2013	NB	0.2191	0.6385	0.6407	0.2191	0.6277
1 WKHOH2015	RF	0.7315	0.6669	0.7270	0.7101	0.7083
	SVM	0.7514	0.6678	0.6732	0.2191	0.6711
	J48	0.7274	0.2916	0.2908	0.7683	0.2932
TwPtáček2014	NB	0.7316	0.3556	0.3558	0.6073	0.3351
1 WF taceK2014	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 TwImbData08112016 NB 0.0196 0.5096 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 J48 0.7105 0.5330 0.6052 0.0196 0.5181 RF 0.9560 0.5370 0.9757 0.9842 0.7704 SVM 0.0196 0.5329 0.5377 0.0196 0.5381 TwImbData10112016 NB 0.0196 0.5220 0.0196 0.5220 RF 0.9605 0.5408 0.9767 0.7962 0.8213 TwImbData10112016 NB 0.0196 0.5420 0.0196 0.5522 TwImbData11112016 NB 0.0196 0.5224 0.5226 0.0196 0.5233
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
TwImbData10112016 RF
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
TwImbData11112016 RF SVM 0.0196 0.5329 0.5377 0.0196 0.5389 0.6435 0.0196 0.6074 0.65200 0.5220 0.6220 0.0196 0.5236 0.6220 0.6220 0.0196 0.5236 0.6220 0.6220 0.0196 0.5236 0.6277 0.0196 0.5236 0.6277 0.0196 0.5236 0.6277 0.0196 0.5226 0.6277 0.0196 0.5236 0.5206 0.6196 0.5226 0.6277 0.0196 0.5236 0.6279 0.0196 0.5236 0.6259 0.0196 0.5206 0.5
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
TwImbData13112016 RF 0.9586 0.5373 0.9682 0.9905 0.8086 0.5373 0.9682 0.9905 0.8086 0.5472 0.6967 0.5312 0.6223 0.0196 0.5862 0.5862 0.5905 0.5906 0.5199 0.5208 0.0196 0.5197 0.9558 0.5359 0.9762 0.9905 0.8278 0.9558 0.5368 0.5424 0.0196 0.5466
TwImbData14112016
TwImbData14112016 NB RF 0.6967 0.5312 0.6223 0.0196 0.5862 0.0196 0.5199 0.5208 0.0196 0.5197 SVM 0.0196 0.5368 0.5359 0.9762 0.9905 0.5424 0.0196 0.5466
TwImbData14112016 NB RF 0.0196 0.9558 0.5199 0.5359 0.5208 0.9762 0.0196 0.9905 0.5197 0.8278 SVM 0.0196 0.5368 0.5424 0.0196 0.5456
RF 0.9558 0.5359 0.9762 0.9905 0.8278 SVM 0.0196 0.5368 0.5424 0.0196 0.5456
SVM 0.0196 0.5368 0.5424 0.0196 0.5466
J48 0.7141 0.5318 0.6091 0.0196 0.5710
TwImbData15112016 NB 0.0196 0.5204 0.5210 0.0196 0.5201
RF 0.9740 0.5355 0.9843 0.9905 0.8749
SVM $0.0196 0.5356 0.5411 0.0196 0.5413$
J48 0.7009 0.5321 0.6254 0.0196 0.5835
TwImbData16112016 NB 0.0196 0.5205 0.5208 0.0196 0.5186
RF 0.9623 0.5354 0.9572 0.9779 0.8293
SVM $0.0196 0.5369 0.5438 0.0196 0.5468$
J48 0.6799 0.5323 0.6169 0.0196 0.5778
TwImbData17112016 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
J48 0.6647 0.5298 0.6080 0.0196 0.5771
TwImbData18112016 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
	J48	0.8676	0.7885	0.8495	0.8740	0.8442
Tw.Inony.Ponbioni2014	NB	0.7410	0.7041	0.7072	0.7410	0.7232
TwIronyBarbieri2014	RF	0.9286	0.8439	0.9272	0.9299	0.9154
	SVM	0.8657	0.7965	0.8006	0.2000	0.8111
	J48	0.5427	0.5489	0.5371	0.5143	0.5380
TwMohammad2015	NB	0.5776	0.5876	0.5775	0.5776	0.5801
1 wwonammad2015	RF	0.5827	0.5970	0.5780	0.6057	0.5854
	SVM	0.4758	0.5871	0.5824	0.2758	0.5748
	J48	0.8996	0.8358	0.8815	0.8959	0.8824
${\bf TwSarcasmBarbieri 2014}$	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
	J48	0.6504	0.6186	0.5972	0.6608	0.6110
TwRiloff2013	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

$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	MOTE 0.5242 0.5104 0.5519 0.5147 0.5831
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$0.5104 \\ 0.5519 \\ 0.5147 \\ \hline 0.5831$
1WImbData08112016 RF 0.7141 0.5141 0.7378 0.9779 0	$0.5519 \\ 0.5147 \\ 0.5831$
	$\frac{0.5147}{0.5831}$
5 1 5 1 5 1 5 1 5 1 5 1 5 1 5 1 5 1 5 1	0.5831
	0 = 00 1
NB 0.5146 0.5187 0.5195 0.5146 (0.5234
TwImbData09112016 RF 0.8282 0.5286 0.7977 0.9749 0	0.6710
	0.5389
J48 0.6554 0.5398 0.6446 0.5649 0	0.6051
T-Jul D-4-10112016 NB 0.5230 0.5209 0.5209 0.5230 0	0.5278
TwImbData10112016 RF 0.7930 0.5299 0.8071 0.9661 0	0.6834
SVM 0.0196 0.5440 0.5493 0.0196 0	0.5527
J48 0.6492 0.5336 0.6282 0.5918 0	0.5759
Trulmh Deta 1111 2016 NB 0.5224 0.5218 0.5220 0.5224 0	0.5278
TwImbData11112016 RF 0.7451 0.5252 0.7138 0.8476 0	0.6310
SVM 0.0196 0.5381 0.5462 0.0196 0	0.5505
J48 0.6468 0.5352 0.6275 0.5383 0	0.5837
Tu-Imb Data 12112016 NB 0.5288 0.5200 0.5200 0.5288 0	0.5258
TwImbData12112016 RF 0.8059 0.5260 0.0250 0.	0.6662
SVM 0.0196 0.5405 0.5467 0.0196 0	0.5501
J48 0.6551 0.5339 0.6210 0.5169 0	0.5792
TwImbData13112016 NB 0.5211 0.5199 0.5196 0.5211 0	0.5253
RF 0.7814 0.5262 0.7716 0.9805 0	0.6424
SVM 0.0196 0.5391 0.5456 0.0196 0	0.5472
J48 0.6327 0.5321 0.6233 0.5530 0	0.5832
TwImbData14112016 NB 0.5159 0.5198 0.5201 0.5159 0	0.5236
RF 0.7990 0.5262 0.7463 0.9781 0	0.6440
SVM 0.0196 0.5368 0.5424 0.0196 0	0.5456
	0.5690
TwImbData15112016 NB 0.5191 0.5192 0.5198 0.5191 0	0.5243
RF 0.8027 0.5262 0.7850 0.9905	0.6518
	0.5413
	0.5776
TWIMBUATAIBITZUIB	0.5215
$ m RF = 0.8020 \ 0.5265 \ 0.7800 \ 0.9429 \ 0.$	0.6432
	0.5468
	0.5750
	0.5263
RF 0.7828 0.5271 0.7802 0.7962 0	0.6539
	0.5447
	0.5752
Twimbijata 1811 2016	0.5252
RF $0.7681 - 0.5252 - 0.7805 - 0.7613 - 0.7613$	0.6304
SVM 0.0196 0.5344 0.5396 0.0196 0	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
TIDk:::9014	J48	0.8874	0.7977	0.8440	0.7980	0.8523
	NB	0.7031	0.6990	0.7020	0.7031	0.7602
TwIronyBarbieri2014	RF	0.9279	0.8450	0.9230	0.9319	0.9159
	SVM	0.8657	0.7965	0.8006	0.6566	0.8111
	J48	0.5507	0.5513	0.5407	0.5536	0.5420
TwMohammad2015	NB	0.5790	0.5810	0.5743	0.5790	0.5856
1 wivionammad2015	RF	0.5827	0.5970	0.5780	0.6057	0.5854
	SVM	0.4758	0.5871	0.5824	0.2758	0.5748
	J48	0.9085	0.8432	0.8760	0.8372	0.8877
TwSarcasmBarbieri2014	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

TwImbData08112016 RF 0.7141 0.5141 0.7378 0.75095 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 J48 0.5243 0.5330 0.6052 0.5328 0.5861 RF 0.8282 0.5286 0.7977 0.9749 0.6710 SVM 0.0196 0.5329 0.5377 0.5166 0.5389 J48 0.5500 0.5393 0.6435 0.5421 0.5074 TwImbData10112016 RF 0.8282 0.5286 0.7977 0.9749 0.6710 SVM 0.0196 0.5329 0.5377 0.5166 0.5389 SVM 0.0196 0.5393 0.6435 0.5421 0.5074 RF 0.7930 0.5299 0.8071 0.9661 0.6834 SVM 0.0196 0.5440 0.5493 0.5164 0.5527 J48 0.5401 0.5332 0.6277 0.5352 0.5781 TwImbData11112016 RF 0.7930 0.5295 0.8071 0.9661 0.6834 SVM 0.0196 0.5349 0.6074 0.5493 0.5164 0.5527 J48 0.5401 0.5332 0.6277 0.5352 0.5781 TwImbData12112016 RF 0.7451 0.5252 0.7138 0.8476 0.6310 SVM 0.0196 0.5381 0.5462 0.5141 0.5505 J48 0.5320 0.5381 0.5462 0.5140 0.5505 TwImbData12112016 RF 0.8059 0.5261 0.7922 0.7963 0.6662 SVM 0.0196 0.5405 0.5467 0.5163 0.5501 TwImbData13112016 RF 0.8059 0.5261 0.7922 0.7963 0.6662 SVM 0.0196 0.5405 0.5467 0.5163 0.5501 TwImbData14112016 RF 0.7814 0.5262 0.7716 0.9805 0.6424 SVM 0.0196 0.5391 0.5467 0.5163 0.5501 TwImbData14112016 RF 0.7814 0.5262 0.7716 0.9805 0.6424 SVM 0.0196 0.5381 0.5190 0.5185 0.5202 TwImbData14112016 RF 0.7814 0.5262 0.7716 0.9805 0.6424 SVM 0.0196 0.5391 0.5465 0.5185 0.5207 TwImbData14112016 RF 0.7814 0.5262 0.7716 0.9805 0.6424 SVM 0.0196 0.5368 0.5445 0.5185 0.5207 TwImbData15112016 RF 0.7814 0.5262 0.7716 0.9805 0.6424 SVM 0.0196 0.5368 0.5424 0.5431 0.5456 SVM 0.0196 0.5368 0.5424 0.5431 0.5456 TwImbData16112016 RF 0.8027 0.5262 0.7800 0.9905 0.6518 SVM 0.0196 0.5368 0.5424 0.5341 0.5835 TwImbData16112016 RF 0.8027 0.5262 0.7800 0.9905 0.6518 SVM 0.0196 0.5368 0.5424 0.5341 0.5835 TwImbData16112016 RF 0.8020 0.5262 0.7800 0.9905 0.6518 SVM 0.0196 0.5368 0.5424 0.5341 0.5835 TwImbData16112016 RF 0.8020 0.5262 0.7800 0.9905 0.6518 SVM 0.0196 0.5364 0.5400 0.5151 0.5407 TwImbData16112016 RF 0.8020 0.5262 0.7800 0.9905 0.56188 0.5205 SVM 0.0196 0.5384 0.5000 0.5361 0.5403	Dataset	Algorithm	ORIGINAL	RUS	ROS	COST	SMOTE
NumbData08112016		J48	0.0693	0.5128	0.5325	0.5148	0.5235
Name		NB	0.5097	0.5093	0.5095		0.5106
TwImbData09112016 TwImbData09112016 RF RF 0.8282 0.5286 0.7977 0.9749 0.6710 SVM 0.0196 0.5329 0.5337 0.5166 0.5389 0.5377 0.5166 0.5389 0.5389 0.5377 0.5166 0.5389 0.5389 0.5387 0.5166 0.5389 0.5389 0.5387 0.5166 0.5389 0.5389 0.5383 0.6435 0.5421 0.6074 0.7930 0.5299 0.8071 0.9661 0.6834 0.5180 0.5187 0.5193 0.5193 0.5193 0.5193 0.5193 0.5193 0.5193 0.5193 0.5193 0.5193 0.5193 0.5193 0.5193 0.5193 0.5247 0.7930 0.5299 0.8071 0.9661 0.6834 0.5401 0.5332 0.6277 0.5352 0.7538 0.5277 0.5352 0.7538 0.5164 0.5527 0.5362 0.5737 0.5362 0.5381 0.5462 0.5318 0.5187 0.5196 0.5188 0.5187 0.5190 0.5190 0.5215 0.5381 0.5462 0.5314 0.5505 0.5381 0.5462 0.5141 0.5505 0.5381 0.5462 0.5314 0.5505 0.5381 0.5462 0.5314 0.5505 0.5381 0.5462 0.5141 0.5505 0.5467 0.5163 0.5504 0.0196 0.5405 0.5467 0.5163 0.5507 0.5382 0.6662 0.7716 0.9805 0.6424 0.5402 0.5402 0.5402 0.5402 0.5402 0.5403 0.5402 0.5403 0.5403 0.5402 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5403 0.5404 0.5403 0.540		RF	0.7141	0.5141			
TwimbData09112016 RF RF 0.8282 0.8282 0.5175 0.5286 0.5179 0.9749 0.6710 0.6710 0.6710 SVM 0.0196 0.5329 0.5339 0.5377 0.5166 0.5389 0.5393 0.6435 0.6435 0.5421 0.6074 TwimbData10112016 NB RF 0.7530 0.7930 0.5193 0.5193 0.5193 0.5193 0.5247 0.6834 TwimbData11112016 NB RF 0.7330 0.5187 0.5193 0.5240 0.5401 0.5440 0.5430 0.5164 0.5430 0.5164 0.5527 0.5781 0.5352 0.5781 0.5232 TwimbData11112016 NB RF 0.7451 0.7451 0.5252 0.7138 0.6177 0.5352 0.5781 0.5238 TwimbData12112016 NB RF 0.7451 0.5230 0.5345 0.5320 0.5345 0.5183 0.5190 0.51462 0.5141 0.5505 0.5467 0.5190 0.5215 0.5020 0.5215 0.5906 TwimbData12112016 NB RF 0.5190 0.5187 0.5183 0.5190 0.5190 0.5190 0.5215 0.5202 0.7963 0.5090 0.5345 0.6259 0.5346 0.5190 0.5190 0.5215 0.5202 0.7963 0.5090 0.5346 0.5190 0.5190 0.5215 0.5466 0.5190 0.5215 0.5202 0.7963 0.5190 0.5202 0.7963 0.5190 0.5202 0.7963 0.5190 0.5202 0.7963 0.5190 0.536		SVM	0.0196	0.5147	0.5160	0.0196	0.5147
RF 0.8282 0.5286 0.7977 0.9749 0.6710		J48	0.5243	0.5330	0.6052	0.5328	0.5861
RF 0.8282 0.5286 0.7977 0.9749 0.6710	TII-D-+-00110016	NB	0.5180	0.5175	0.5178	0.5180	0.5212
TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData10112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData11112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData12112016 TwImbData13112016 TwImbData131	1 w1mbData09112016	RF		0.5286	0.7977	0.9749	0.6710
TwImbData10112016 NB RF 0.5193 0.5192 0.5193 0.5193 0.5247 SVM 0.0196 0.5440 0.5493 0.5164 0.5527 J48 0.5401 0.5332 0.6277 0.5352 0.5781 TwImbData11112016 NB 0.5187 0.5196 0.5188 0.5187 0.5238 RF 0.7451 0.5252 0.7138 0.8476 0.6310 0.5016 0.5188 0.5187 0.5238 TwImbData12112016 RF 0.7451 0.5252 0.7138 0.8476 0.6310 0.5016 0.5183 0.5462 0.5141 0.5505 0.5345 0.6259 0.5343 0.5906 0.5215 0.5016 0.7922 0.7963 0.6662 0.5016 0.7922 0.7963 0.6662 0.5017 0.5020 0.5183 0.5190 0.5215 0.5184 0.5190 0.5215 0.5202 0.7963 0.6662 0.7963 0.6662 0.7963 0.6602 0.7963 0.6602 0.7963 0.6518		SVM	0.0196	0.5329	0.5377	0.5166	0.5389
Name		J48	0.5500	0.5393	0.6435	0.5421	0.6074
NF 0.7930 0.5299 0.8071 0.9661 0.5527	TI.m.h.Dt10112016	NB	0.5193	0.5192	0.5193	0.5193	0.5247
TwImbData11112016 NB O.5187 O.5196 O.5188 O.5187 O.5238 O.5345 O.5252 O.7138 O.8476 O.6331 O.5505 O.5345 O.5252 O.5138 O.5141 O.5505 O.5345 O.5252 O.5345 O.5252 O.5138 O.5141 O.5505 O.5252 O.5345 O.5252 O.5345 O.5252 O.5343 O.5906 O.5215 O.5252 O.5345 O.5290 O.5345 O.5290 O.5183 O.5190 O.5190 O.5190 O.5190 O.5190 O.5215 O.5201 O.5220 O.5345 O.6197 O.5340 O.5825 O.5345 O.5355 O.6197 O.5340 O.5825 O.5345 O.5184 O.5187 O.5184 O.5187 O.5184 O.5187 O.5184 O.5187 O.5187 O.5187 O.5187 O.5187 O.5202 O.5363 O.5907 O.5202 O.5363 O.5303 O.5363 O.5906 O.5472 O.5202 O.5363 O.5303 O.5865 O.5407 O.5202 O.5363 O.5407 O.5202 O.5363 O.5363 O.5407 O.5203 O.5363 O.5363 O.5407 O.5203 O.5407 O.5203 O.5363 O.5207 O.5203 O	1 w1mbData10112016	RF	0.7930	0.5299	0.8071	0.9661	0.6834
TwImbData11112016 NB RF 0.5187 0.7451 0.5185 0.5252 0.5188 0.5187 0.5238 0.6310 TwImbData12112016 J48 NB RF 0.5320 0.5190 0.5341 0.5190 0.5190 0.5183 0.5190 0.5190 0.5215 0.5215 TwImbData12112016 RF RF 0.8059 0.5201 0.5462 0.5467 0.5163 0.5623 0.6662 0.501 TwImbData13112016 NB RF 0.5327 0.5327 0.5335 0.6197 0.5340 0.5825 0.5501 0.5825 TwImbData13112016 NB RF 0.5175 0.5184 0.5181 0.5175 0.5181 0.5185 0.5175 0.5220 0.6424 0.5761 0.5462 0.5467 0.5340 0.5825 TwImbData14112016 RF RF 0.7814 0.5185 0.5181 0.5321 0.6223 0.5333 0.5862 0.5424 0.5161 0.5472 0.5426 TwImbData14112016 NB RF 0.5185 0.5189 0.5185 0.5183 0.5151 0.5413 0.5413 0.5434 0.543		SVM	0.0196	0.5440	0.5493	0.5164	0.5527
NumbData11112016 RF 0.7451 0.5252 0.7138 0.8476 0.6310		J48	0.5401	0.5332	0.6277	0.5352	0.5781
SVM	TI.m.h.Dt11119016	NB	0.5187	0.5196	0.5188	0.5187	0.5238
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData11112016	RF	0.7451	0.5252	0.7138	0.8476	0.6310
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0196	0.5381	0.5462	0.5141	0.5505
TwimbData12112016 RF 0.8059 0.5261 0.7922 0.7963 0.6662 SVM 0.0196 0.5405 0.5467 0.5163 0.5501 TwImbData13112016 NB 0.5327 0.5335 0.6197 0.5340 0.5825 RF 0.7814 0.5262 0.7716 0.9805 0.6424 SVM 0.0196 0.5391 0.5456 0.5159 0.5472 J48 0.5396 0.5312 0.6223 0.5333 0.5862 TwImbData14112016 NB 0.5185 0.5183 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 SVM 0.0196 0.5368 0.5424 0.5161 0.5466 SVM 0.0196 0.5356 0.5189 0.5189 0.5189 TwImbData15112016 NB 0.5189 0.5185 0.5189 0.5181 0.5205		J48	0.5320	0.5345	0.6259	0.5343	0.5906
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI.m.h.Dt19119016	NB	0.5190	0.5183	0.5190	0.5190	0.5215
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData12112016	RF	0.8059	0.5261	0.7922	0.7963	0.6662
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0196	0.5405	0.5467	0.5163	0.5501
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.5327	0.5335	0.6197	0.5340	0.5825
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TI	NB	0.5175	0.5184	0.5181	0.5175	0.5220
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData13112016	RF	0.7814	0.5262	0.7716	0.9805	0.6424
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0196	0.5391	0.5456	0.5159	0.5472
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		J48	0.5396	0.5312	0.6223	0.5333	0.5862
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TII-D-4-14119016	NB	0.5185	0.5183	0.5187	0.5185	0.5207
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	1 w1mbData14112016	RF	0.7990	0.5262	0.7463	0.9781	0.6440
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0196	0.5368	0.5424	0.5161	0.5456
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData15112016		0.5389	0.5318	0.6091	0.5326	0.5710
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		NB	0.5189	0.5185	0.5189	0.5189	0.5207
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		RF	0.8027	0.5262	0.7850	0.9905	0.6518
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0196	0.5356	0.5411	0.5151	0.5413
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData16112016	J48	0.5281	0.5321	0.6254	0.5341	0.5835
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		NB	0.5185	0.5183	0.5183	0.5185	0.5205
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		RF	0.8020	0.5265	0.7800	0.9429	0.6432
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$		SVM	0.0196	0.5369	0.5438	0.5162	0.5468
TwImbData17112016 RF 0.7828 0.5271 0.7802 0.7962 0.6539 SVM 0.0196 0.5384 0.5440 0.5151 0.5447 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		J48	0.5434	0.5323	0.6169	0.5366	0.5773
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	TwImbData17112016	NB	0.5188	0.5182	0.5189	0.5188	0.5228
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		RF					
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		SVM	0.0196	0.5384	0.5440	0.5151	0.5447
TwImbData18112016 RF 0.5176 0.5176 0.5177 0.5176 0.5231 0.7681 0.5252 0.7805 0.7613 0.6304	TwImbData18112016	J48	0.5233	0.5298	0.6080	0.5321	
RF = 0.7681 - 0.5252 - 0.7805 - 0.7613 - 0.6304		NB	0.5176	0.5176	0.5177	0.5176	0.5231
		RF	0.7681	0.5252	0.7805	0.7613	0.6304
		SVM		0.5344	0.5396	0.5160	0.5403

Table 40: New Corpora - Youden Thresholding