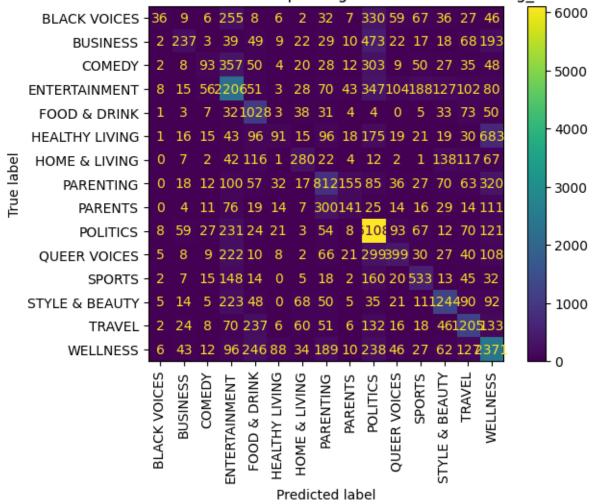
1 AdaBoostClassifier orig avg w2v

Fit time: 122.84558200836182 Test time: 0.5259435176849365

Sanity check: avg f1: 0.5426, avg precision: 0.5416, avg recall: 0.5742

No params.

Confusion matrix for sample orig and vectorization avg w2v



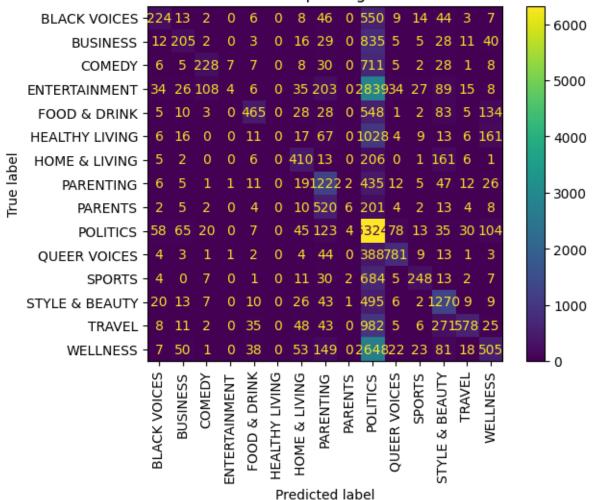
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.3846	0.0486	0.9975	0.0863	0.2202	0.0439	926.0000
BUSINESS	0.4145	0.2729	0.9837	0.3291	0.5181	0.2493	1191.0000
COMEDY	0.2949	0.1004	0.9911	0.1498	0.3154	0.0906	1046.0000
ENTERTAINMENT	0.5235	0.6412	0.9228	0.5764	0.7692	0.5750	3428.0000
FOOD and DRINK	0.5415	0.7508	0.9702	0.6292	0.8535	0.7124	1312.0000
HEALTHY LIVING	0.2600	0.0777	0.9894	0.1197	0.2773	0.0699	1338.0000
HOME and LIVING	0.5843	0.3120	0.9937	0.4068	0.5568	0.2889	811.0000
PARENTING	0.4543	0.4157	0.9673	0.4342	0.6341	0.3800	1804.0000
PARENTS	0.3097	0.1511	0.9908	0.2031	0.3869	0.1371	781.0000
POLITICS	0.7207	0.8697	0.8962	0.7882	0.8828	0.7773	6906.0000
QUEER VOICES	0.5032	0.3732	0.9835	0.4286	0.6059	0.3447	1254.0000
SPORTS	0.4937	0.4606	0.9831	0.4765	0.6729	0.4291	1014.0000
STYLE and BEAUTY	0.6359	0.6818	0.9728	0.6581	0.8144	0.6440	1911.0000
TRAVEL	0.6328	0.6246	0.9733	0.6287	0.7797	0.5867	2014.0000
WELLNESS	0.5053	0.6949	0.9050	0.5851	0.7930	0.6156	3595.0000
avg precision	0.5442	0.5442	0.5442	0.5442	0.5442	0.5442	0.0000
avg recall	0.5756	0.5756	0.5756	0.5756	0.5756	0.5756	0.0000
avg specificity	0.9447	0.9447	0.9447	0.9447	0.9447	0.9447	0.0000
avg f1	0.5439	0.5439	0.5439	0.5439	0.5439	0.5439	0.0000
avg geom. mean	0.7055	0.7055	0.7055	0.7055	0.7055	0.7055	0.0000
avg ind. bal. acc.	0.5223	0.5223	0.5223	0.5223	0.5223	0.5223	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

2 AdaBoostClassifier orig tfidf

Fit time: 378.5199770927429 Test time: 10.508871078491211

Sanity check: avg f1: 0.359, avg precision: 0.4825, avg recall: 0.4271

No params.



	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
DI ACIZ MOTORO							
BLACK VOICES	0.5586	0.2419	0.9938	0.3376	0.4903	0.2223	926.0000
BUSINESS	0.4779	0.1721	0.9920	0.2531	0.4132	0.1568	1191.0000
COMEDY	0.5938	0.2180	0.9945	0.3189	0.4656	0.1999	1046.0000
ENTERTAINMENT	0.3077	0.0012	0.9997	0.0023	0.0342	0.0010	3428.0000
FOOD and DRINK	0.7598	0.3544	0.9948	0.4834	0.5938	0.3300	1312.0000
HEALTHY LIVING	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1338.0000
HOME and LIVING	0.5556	0.5055	0.9885	0.5294	0.7069	0.4756	811.0000
PARENTING	0.4718	0.6774	0.9503	0.5562	0.8023	0.6262	1804.0000
PARENTS	0.4000	0.0077	0.9997	0.0151	0.0876	0.0069	781.0000
POLITICS	0.3351	0.9157	0.4404	0.4906	0.6350	0.4224	6906.0000
QUEER VOICES	0.8043	0.6228	0.9932	0.7020	0.7865	0.5957	1254.0000
SPORTS	0.6739	0.2446	0.9958	0.3589	0.4935	0.2252	1014.0000
STYLE and BEAUTY	0.5802	0.6646	0.9665	0.6195	0.8014	0.6229	1911.0000
TRAVEL	0.8245	0.2870	0.9955	0.4258	0.5345	0.2655	2014.0000
WELLNESS	0.4828	0.1405	0.9790	0.2176	0.3708	0.1260	3595.0000
avg precision	0.4734	0.4734	0.4734	0.4734	0.4734	0.4734	0.0000
avg recall	0.4251	0.4251	0.4251	0.4251	0.4251	0.4251	0.0000
avg specificity	0.8584	0.8584	0.8584	0.8584	0.8584	0.8584	0.0000
avg f1	0.3577	0.3577	0.3577	0.3577	0.3577	0.3577	0.0000
avg geom. mean	0.4852	0.4852	0.4852	0.4852	0.4852	0.4852	0.0000
avg ind. bal. acc.	0.2942	0.2942	0.2942	0.2942	0.2942	0.2942	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

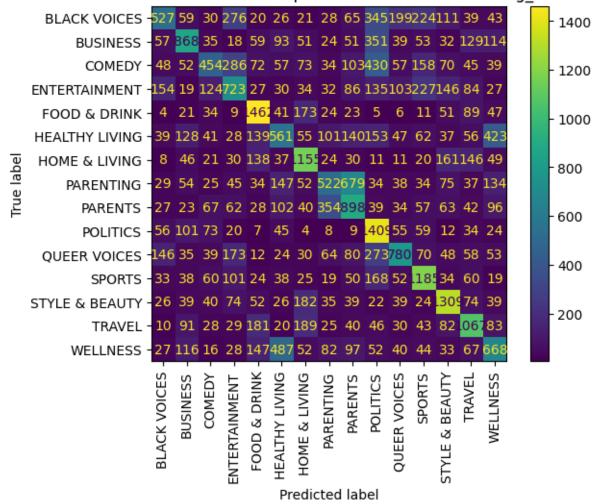
3 AdaBoostClassifier ros avg w2v

Fit time: 109.74340319633484 Test time: 0.4973893165588379

Sanity check: avg f1: 0.507, avg precision: 0.5364, avg recall: 0.5006

No params.

Confusion matrix for sample ros and vectorization avg w2v



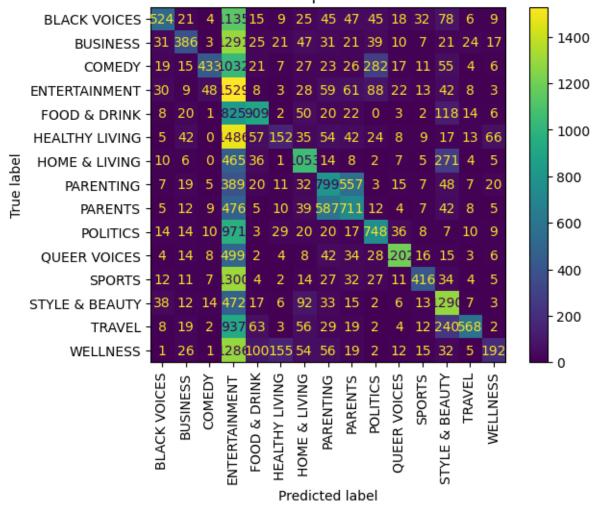
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.3895	0.2082	0.9763	0.2714	0.4509	0.1877	1988.0000
BUSINESS	0.4602	0.4395	0.9631	0.4496	0.6506	0.4011	1959.0000
COMEDY	0.4260	0.2308	0.9778	0.2994	0.4751	0.2089	1958.0000
ENTERTAINMENT	0.3533	0.3955	0.9480	0.3732	0.6123	0.3542	1967.0000
FOOD and DRINK	0.5769	0.7276	0.9620	0.6435	0.8366	0.6835	1949.0000
HEALTHY LIVING	0.3136	0.2748	0.9577	0.2929	0.5130	0.2452	1929.0000
HOME and LIVING	0.5286	0.5890	0.9627	0.5572	0.7530	0.5458	1944.0000
PARENTING	0.3455	0.2386	0.9673	0.2823	0.4804	0.2140	1978.0000
PARENTS	0.3825	0.4318	0.9485	0.4057	0.6400	0.3884	2017.0000
POLITICS	0.4116	0.7275	0.9279	0.5258	0.8216	0.6615	1901.0000
QUEER VOICES	0.4973	0.3737	0.9728	0.4267	0.6029	0.3417	1967.0000
SPORTS	0.5477	0.6809	0.9607	0.6071	0.8088	0.6359	1915.0000
STYLE and BEAUTY	0.6147	0.5745	0.9749	0.5939	0.7484	0.5376	1913.0000
TRAVEL	0.5528	0.5604	0.9676	0.5565	0.7364	0.5202	1954.0000
WELLNESS	0.3792	0.4167	0.9503	0.3970	0.6292	0.3748	1992.0000
avg precision	0.4513	0.4513	0.4513	0.4513	0.4513	0.4513	0.0000
avg recall	0.4564	0.4564	0.4564	0.4564	0.4564	0.4564	0.0000
avg specificity	0.9612	0.9612	0.9612	0.9612	0.9612	0.9612	0.0000
avg f1	0.4444	0.4444	0.4444	0.4444	0.4444	0.4444	0.0000
avg geom. mean	0.6495	0.6495	0.6495	0.6495	0.6495	0.6495	0.0000
avg ind. bal. acc.	0.4186	0.4186	0.4186	0.4186	0.4186	0.4186	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

4 AdaBoostClassifier ros tfidf

Fit time: 377.17439317703247 Test time: 10.9009850025177

Sanity check: avg f1: 0.4075, avg precision: 0.6062, avg recall: 0.3931

No params.



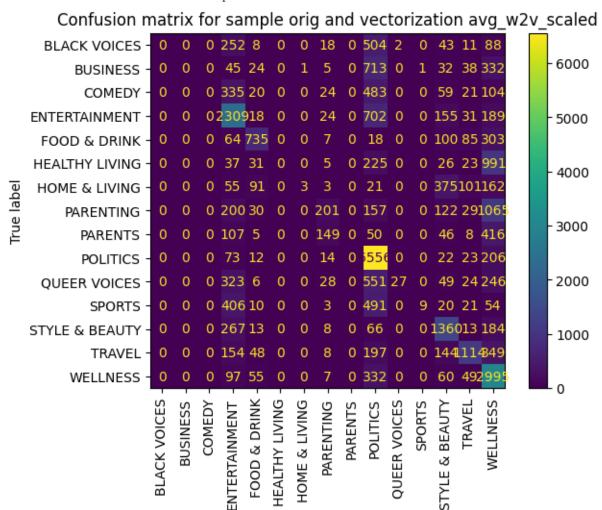
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.7349	0.2565	0.9933	0.3803	0.5048	0.2360	1988.0000
BUSINESS	0.5833	0.1787	0.9909	0.2735	0.4208	0.1627	1959.0000
COMEDY	0.7907	0.2181	0.9959	0.3419	0.4660	0.2003	1958.0000
ENTERTAINMENT	0.1087	0.7804	0.5399	0.1908	0.6491	0.4314	1967.0000
FOOD and DRINK	0.7271	0.4854	0.9870	0.5822	0.6922	0.4551	1949.0000
HEALTHY LIVING	0.3687	0.0757	0.9909	0.1256	0.2739	0.0681	1929.0000
HOME and LIVING	0.7065	0.5165	0.9848	0.5967	0.7132	0.4848	1944.0000
PARENTING	0.3920	0.3943	0.9558	0.3931	0.6139	0.3557	1978.0000
PARENTS	0.4452	0.3426	0.9685	0.3872	0.5760	0.3110	2017.0000
POLITICS	0.5638	0.4019	0.9785	0.4693	0.6271	0.3706	1901.0000
QUEER VOICES	0.8896	0.6268	0.9944	0.7355	0.7895	0.6004	1967.0000
SPORTS	0.7761	0.2642	0.9947	0.3942	0.5127	0.2436	1915.0000
STYLE and BEAUTY	0.5333	0.6273	0.9617	0.5765	0.7767	0.5831	1913.0000
TRAVEL	0.8338	0.2927	0.9958	0.4333	0.5399	0.2710	1954.0000
WELLNESS	0.4712	0.0904	0.9926	0.1516	0.2995	0.0816	1992.0000
avg precision	0.5945	0.5945	0.5945	0.5945	0.5945	0.5945	0.0000
avg recall	0.3697	0.3697	0.3697	0.3697	0.3697	0.3697	0.0000
avg specificity	0.9548	0.9548	0.9548	0.9548	0.9548	0.9548	0.0000
avg f1	0.4015	0.4015	0.4015	0.4015	0.4015	0.4015	0.0000
avg geom. mean	0.5633	0.5633	0.5633	0.5633	0.5633	0.5633	0.0000
avg ind. bal. acc.	0.3232	0.3232	0.3232	0.3232	0.3232	0.3232	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

5 ComplementNB orig avg w2v scaled

Fit time: 33.01250147819519Test time: 0.015288114547729492

Sanity check: avg f1: 0.4267, avg precision: 0.5068, avg recall: 0.5232

'alpha': 0.00013894954943731373



	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	926.0000
BUSINESS	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1191.0000
COMEDY	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1046.0000
ENTERTAINMENT	0.4882	0.6683	0.9073	0.5642	0.7787	0.5919	3428.0000
FOOD and DRINK	0.6833	0.5724	0.9876	0.6230	0.7519	0.5418	1312.0000
HEALTHY LIVING	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1338.0000
HOME and LIVING	0.7500	0.0037	1.0000	0.0074	0.0608	0.0033	811.0000
PARENTING	0.3886	0.1131	0.9883	0.1752	0.3343	0.1020	1804.0000
PARENTS	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	781.0000
POLITICS	0.5946	0.9500	0.8005	0.7314	0.8721	0.7719	6906.0000
QUEER VOICES	0.9583	0.0183	1.0000	0.0360	0.1354	0.0165	1254.0000
SPORTS	0.8571	0.0059	1.0000	0.0118	0.0769	0.0053	1014.0000
STYLE and BEAUTY	0.5327	0.7033	0.9570	0.6062	0.8204	0.6560	1911.0000
TRAVEL	0.6972	0.5660	0.9819	0.6248	0.7455	0.5327	2014.0000
WELLNESS	0.3862	0.8364	0.8143	0.5284	0.8253	0.6826	3595.0000
avg precision	0.4728	0.4728	0.4728	0.4728	0.4728	0.4728	0.0000
avg recall	0.5227	0.5227	0.5227	0.5227	0.5227	0.5227	0.0000
avg specificity	0.9141	0.9141	0.9141	0.9141	0.9141	0.9141	0.0000
avg f1	0.4261	0.4261	0.4261	0.4261	0.4261	0.4261	0.0000
avg geom. mean	0.5665	0.5665	0.5665	0.5665	0.5665	0.5665	0.0000
avg ind. bal. acc.	0.4454	0.4454	0.4454	0.4454	0.4454	0.4454	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

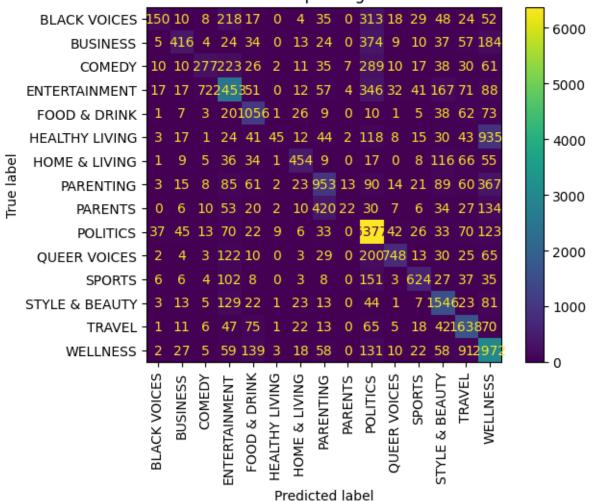
Predicted label

6 ComplementNB orig tfidf

Fit time: 159.33163046836853 Test time: 1.1400470733642578

Sanity check: avg f1: 0.6409, avg precision: 0.6769, avg recall: 0.6769

'alpha': 0.281176869797423



	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.6224	0.1620	0.9968	0.2571	0.4018	0.1480	926.0000
BUSINESS	0.6797	0.3510	0.9930	0.4629	0.5903	0.3261	1191.0000
COMEDY	0.6549	0.2667	0.9948	0.3791	0.5151	0.2460	1046.0000
ENTERTAINMENT	0.6688	0.7153	0.9531	0.6913	0.8257	0.6655	3428.0000
FOOD and DRINK	0.6526	0.8034	0.9800	0.7202	0.8873	0.7734	1312.0000
HEALTHY LIVING	0.6716	0.0336	0.9992	0.0641	0.1833	0.0304	1338.0000
HOME and LIVING	0.7069	0.5561	0.9934	0.6225	0.7433	0.5283	811.0000
PARENTING	0.5475	0.5244	0.9716	0.5357	0.7138	0.4867	1804.0000
PARENTS	0.4783	0.0282	0.9992	0.0532	0.1678	0.0254	781.0000
POLITICS	0.7448	0.9235	0.9026	0.8246	0.9130	0.8353	6906.0000
QUEER VOICES	0.8238	0.5965	0.9943	0.6920	0.7701	0.5695	1254.0000
SPORTS	0.7233	0.6134	0.9916	0.6638	0.7799	0.5853	1014.0000
STYLE and BEAUTY	0.6608	0.8095	0.9710	0.7277	0.8866	0.7734	1911.0000
TRAVEL	0.7036	0.8133	0.9747	0.7545	0.8904	0.7800	2014.0000
WELLNESS	0.5610	0.8253	0.9098	0.6679	0.8665	0.7445	3595.0000
avg precision	0.6703	0.6703	0.6703	0.6703	0.6703	0.6703	0.0000
avg recall	0.6722	0.6722	0.6722	0.6722	0.6722	0.6722	0.0000
avg specificity	0.9529	0.9529	0.9529	0.9529	0.9529	0.9529	0.0000
avg f1	0.6357	0.6357	0.6357	0.6357	0.6357	0.6357	0.0000
avg geom. mean	0.7685	0.7685	0.7685	0.7685	0.7685	0.7685	0.0000
avg ind. bal. acc.	0.6221	0.6221	0.6221	0.6221	0.6221	0.6221	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

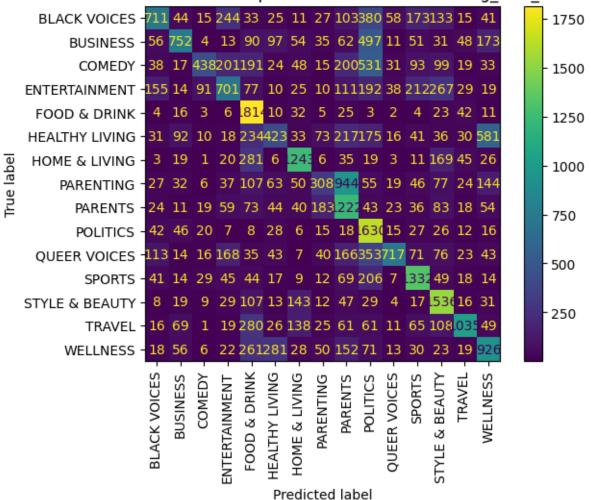
7 ComplementNB ros avg w2v scaled

Fit time: 31.378465175628662Test time: 0.022382259368896484

Sanity check: avg f1: 0.5273, avg precision: 0.5739, avg recall: 0.5479

'alpha': 0.00021209508879201905

Confusion matrix for sample ros and vectorization avg w2v scaled



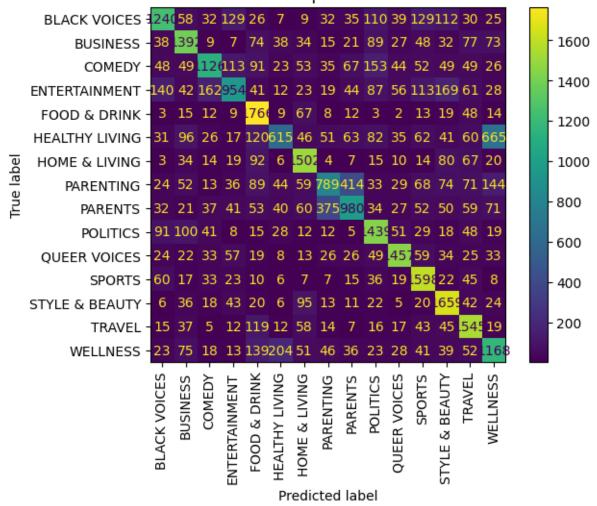
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.6054	0.2470	0.9883	0.3508	0.4941	0.2260	1988.0000
BUSINESS	0.6178	0.3359	0.9851	0.4352	0.5752	0.3094	1959.0000
COMEDY	0.6261	0.1864	0.9920	0.2873	0.4300	0.1700	1958.0000
ENTERTAINMENT	0.4352	0.3823	0.9643	0.4070	0.6072	0.3472	1967.0000
FOOD and DRINK	0.4584	0.8989	0.9244	0.6072	0.9116	0.8288	1949.0000
HEALTHY LIVING	0.3736	0.1716	0.9797	0.2352	0.4100	0.1545	1929.0000
HOME and LIVING	0.6790	0.6322	0.9788	0.6548	0.7866	0.5973	1944.0000
PARENTING	0.3738	0.0809	0.9902	0.1330	0.2830	0.0728	1978.0000
PARENTS	0.3546	0.6371	0.9144	0.4556	0.7632	0.5664	2017.0000
POLITICS	0.3233	0.8832	0.8719	0.4733	0.8775	0.7709	1901.0000
QUEER VOICES	0.7428	0.3274	0.9919	0.4545	0.5699	0.3032	1967.0000
SPORTS	0.6113	0.6799	0.9698	0.6438	0.8120	0.6402	1915.0000
STYLE and BEAUTY	0.5317	0.7852	0.9517	0.6340	0.8644	0.7348	1913.0000
TRAVEL	0.7065	0.5210	0.9845	0.5997	0.7162	0.4892	1954.0000
WELLNESS	0.4072	0.4327	0.9541	0.4196	0.6425	0.3913	1992.0000
avg precision	0.5230	0.5230	0.5230	0.5230	0.5230	0.5230	0.0000
avg recall	0.4783	0.4783	0.4783	0.4783	0.4783	0.4783	0.0000
avg specificity	0.9629	0.9629	0.9629	0.9629	0.9629	0.9629	0.0000
avg f1	0.4518	0.4518	0.4518	0.4518	0.4518	0.4518	0.0000
avg geom. mean	0.6484	0.6484	0.6484	0.6484	0.6484	0.6484	0.0000
avg ind. bal. acc.	0.4385	0.4385	0.4385	0.4385	0.4385	0.4385	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

8 ComplementNB ros tfidf

Fit time: 163.45824480056763 Test time: 1.1326518058776855

Sanity check: avg f1: 0.6595, avg precision: 0.6891, avg recall: 0.6592

'alpha': 0.6551285568595508



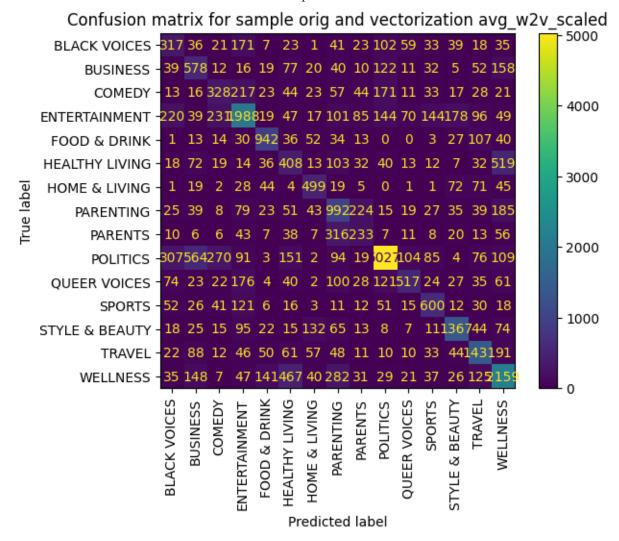
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.6909	0.5971	0.9806	0.6406	0.7652	0.5630	1988.0000
BUSINESS	0.6573	0.6835	0.9745	0.6702	0.8161	0.6467	1959.0000
COMEDY	0.7120	0.5771	0.9833	0.6375	0.7533	0.5444	1958.0000
ENTERTAINMENT	0.6206	0.5165	0.9773	0.5638	0.7105	0.4815	1967.0000
FOOD and DRINK	0.6620	0.8815	0.9680	0.7562	0.9237	0.8459	1949.0000
HEALTHY LIVING	0.5819	0.3002	0.9848	0.3960	0.5437	0.2754	1929.0000
HOME and LIVING	0.7307	0.7901	0.9793	0.7593	0.8797	0.7592	1944.0000
PARENTING	0.5168	0.3504	0.9763	0.4176	0.5849	0.3206	1978.0000
PARENTS	0.5452	0.5112	0.9685	0.5276	0.7036	0.4724	2017.0000
POLITICS	0.6443	0.7564	0.9711	0.6959	0.8571	0.7188	1901.0000
QUEER VOICES	0.8017	0.7707	0.9863	0.7859	0.8719	0.7438	1967.0000
SPORTS	0.7061	0.8507	0.9753	0.7717	0.9108	0.8193	1915.0000
STYLE and BEAUTY	0.6680	0.8129	0.9718	0.7333	0.8888	0.7774	1913.0000
TRAVEL	0.6784	0.8019	0.9729	0.7350	0.8833	0.7668	1954.0000
WELLNESS	0.4996	0.5919	0.9568	0.5418	0.7525	0.5456	1992.0000
avg precision	0.6472	0.6472	0.6472	0.6472	0.6472	0.6472	0.0000
avg recall	0.6516	0.6516	0.6516	0.6516	0.6516	0.6516	0.0000
avg specificity	0.9751	0.9751	0.9751	0.9751	0.9751	0.9751	0.0000
avg f1	0.6414	0.6414	0.6414	0.6414	0.6414	0.6414	0.0000
avg geom. mean	0.7890	0.7890	0.7890	0.7890	0.7890	0.7890	0.0000
avg ind. bal. acc.	0.6175	0.6175	0.6175	0.6175	0.6175	0.6175	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

9 GaussianNB orig avg w2v scaled

Fit time: 0.44372105598449707 Test time: 0.20717692375183105

Sanity check: avg f1: 0.5981, avg precision: 0.6117, avg recall: 0.5911

No params.

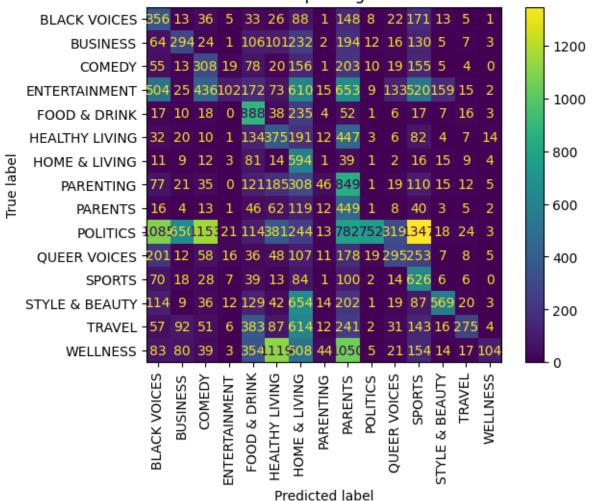


	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.2832	0.3499	0.9711	0.3130	0.5829	0.3187	926.0000
BUSINESS	0.3355	0.4861	0.9592	0.3970	0.6829	0.4443	1191.0000
COMEDY	0.3353	0.3260	0.9761	0.3306	0.5641	0.2975	1046.0000
ENTERTAINMENT	0.6317	0.5770	0.9555	0.6031	0.7425	0.5305	3428.0000
FOOD and DRINK	0.6964	0.7203	0.9853	0.7081	0.8424	0.6909	1312.0000
HEALTHY LIVING	0.2794	0.3057	0.9623	0.2919	0.5424	0.2748	1338.0000
HOME and LIVING	0.5477	0.6153	0.9856	0.5796	0.7787	0.5839	811.0000
PARENTING	0.4362	0.5477	0.9536	0.4856	0.7227	0.5011	1804.0000
PARENTS	0.2989	0.2996	0.9808	0.2992	0.5421	0.2738	781.0000
POLITICS	0.8599	0.7298	0.9634	0.7895	0.8385	0.6867	6906.0000
QUEER VOICES	0.5998	0.4099	0.9878	0.4870	0.6363	0.3815	1254.0000
SPORTS	0.5499	0.5927	0.9826	0.5705	0.7632	0.5597	1014.0000
STYLE and BEAUTY	0.7239	0.7174	0.9809	0.7206	0.8389	0.6852	1911.0000
TRAVEL	0.6570	0.7170	0.9724	0.6857	0.8350	0.6794	2014.0000
WELLNESS	0.5948	0.6006	0.9428	0.5976	0.7525	0.5468	3595.0000
avg precision	0.6145	0.6145	0.6145	0.6145	0.6145	0.6145	0.0000
avg recall	0.5941	0.5941	0.5941	0.5941	0.5941	0.5941	0.0000
avg specificity	0.9654	0.9654	0.9654	0.9654	0.9654	0.9654	0.0000
avg f1	0.6009	0.6009	0.6009	0.6009	0.6009	0.6009	0.0000
avg geom. mean	0.7511	0.7511	0.7511	0.7511	0.7511	0.7511	0.0000
avg ind. bal. acc.	0.5541	0.5541	0.5541	0.5541	0.5541	0.5541	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

10 GaussianNB orig tfidf

Fit time: 9.845949172973633 Test time: 12.109414100646973

Sanity check: avg f1: 0.2036, avg precision: 0.58, avg recall: 0.2302 No params.



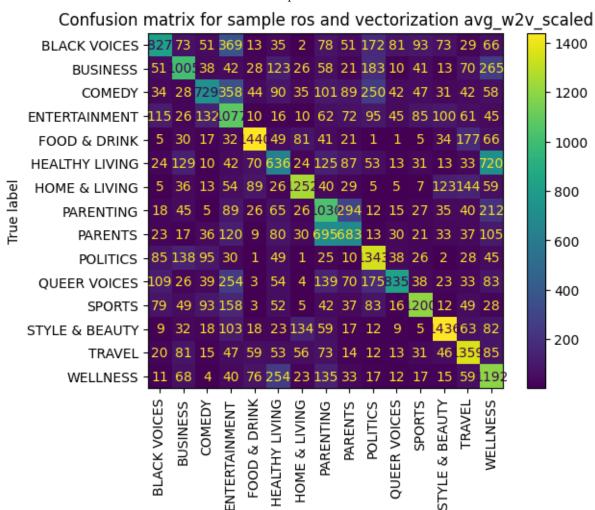
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.1298	0.3844	0.9160	0.1941	0.5934	0.3334	926.0000
BUSINESS	0.2315	0.2469	0.9653	0.2389	0.4881	0.2212	1191.0000
COMEDY	0.1365	0.2945	0.9311	0.1865	0.5236	0.2567	1046.0000
ENTERTAINMENT	0.5178	0.0298	0.9963	0.0563	0.1722	0.0268	3428.0000
FOOD and DRINK	0.3272	0.6768	0.9348	0.4411	0.7954	0.6164	1312.0000
HEALTHY LIVING	0.1451	0.2803	0.9211	0.1912	0.5081	0.2416	1338.0000
HOME and LIVING	0.1252	0.7324	0.8545	0.2139	0.7911	0.6182	811.0000
PARENTING	0.2434	0.0255	0.9948	0.0462	0.1593	0.0229	1804.0000
PARENTS	0.0804	0.5749	0.8200	0.1410	0.6866	0.4599	781.0000
POLITICS	0.9093	0.1089	0.9967	0.1945	0.3294	0.0989	6906.0000
QUEER VOICES	0.3172	0.2352	0.9774	0.2701	0.4795	0.2129	1254.0000
SPORTS	0.1626	0.6174	0.8861	0.2573	0.7396	0.5323	1014.0000
STYLE and BEAUTY	0.6647	0.2977	0.9895	0.4113	0.5428	0.2743	1911.0000
TRAVEL	0.6395	0.1365	0.9943	0.2250	0.3685	0.1241	2014.0000
WELLNESS	0.6797	0.0289	0.9981	0.0555	0.1699	0.0261	3595.0000
avg precision	0.5245	0.5245	0.5245	0.5245	0.5245	0.5245	0.0000
avg recall	0.2057	0.2057	0.2057	0.2057	0.2057	0.2057	0.0000
avg specificity	0.9704	0.9704	0.9704	0.9704	0.9704	0.9704	0.0000
avg f1	0.1853	0.1853	0.1853	0.1853	0.1853	0.1853	0.0000
avg geom. mean	0.3912	0.3912	0.3912	0.3912	0.3912	0.3912	0.0000
avg ind. bal. acc.	0.1815	0.1815	0.1815	0.1815	0.1815	0.1815	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

11 GaussianNB ros avg w2v scaled

Fit time: 0.45205140113830566 Test time: 0.2182445526123047

Sanity check: avg f1: 0.5963, avg precision: 0.6135, avg recall: 0.5877

No params.



	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.5549	0.4014	0.9766	0.4658	0.6261	0.3695	1988.0000
BUSINESS	0.5547	0.5048	0.9710	0.5286	0.7001	0.4674	1959.0000
COMEDY	0.5494	0.3662	0.9785	0.4395	0.5986	0.3364	1958.0000
ENTERTAINMENT	0.3819	0.5506	0.9359	0.4510	0.7179	0.4955	1967.0000
FOOD and DRINK	0.7664	0.7188	0.9844	0.7419	0.8412	0.6888	1949.0000
HEALTHY LIVING	0.4032	0.3142	0.9673	0.3531	0.5512	0.2840	1929.0000
HOME and LIVING	0.7535	0.6636	0.9846	0.7057	0.8083	0.6324	1944.0000
PARENTING	0.3600	0.4985	0.9359	0.4181	0.6830	0.4461	1978.0000
PARENTS	0.4527	0.3510	0.9687	0.3954	0.5831	0.3190	2017.0000
POLITICS	0.5417	0.6933	0.9594	0.6082	0.8156	0.6474	1901.0000
QUEER VOICES	0.7220	0.4331	0.9880	0.5415	0.6542	0.4042	1967.0000
SPORTS	0.7396	0.6334	0.9844	0.6824	0.7897	0.6017	1915.0000
STYLE and BEAUTY	0.6954	0.7350	0.9775	0.7146	0.8476	0.7010	1913.0000
TRAVEL	0.5874	0.6950	0.9652	0.6367	0.8190	0.6526	1954.0000
WELLNESS	0.3831	0.5838	0.9315	0.4626	0.7375	0.5249	1992.0000
avg precision	0.5621	0.5621	0.5621	0.5621	0.5621	0.5621	0.0000
avg recall	0.5417	0.5417	0.5417	0.5417	0.5417	0.5417	0.0000
avg specificity	0.9672	0.9672	0.9672	0.9672	0.9672	0.9672	0.0000
avg f1	0.5419	0.5419	0.5419	0.5419	0.5419	0.5419	0.0000
avg geom. mean	0.7174	0.7174	0.7174	0.7174	0.7174	0.7174	0.0000
avg ind. bal. acc.	0.5035	0.5035	0.5035	0.5035	0.5035	0.5035	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

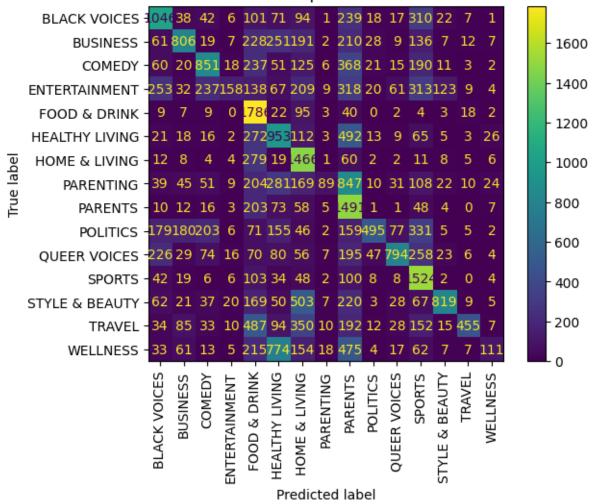
Predicted label

12 GaussianNB ros tfidf

Fit time: 9.718861818313599 Test time: 12.114492416381836

Sanity check: avg f1: 0.3194, avg precision: 0.6132, avg recall: 0.3319

No params.

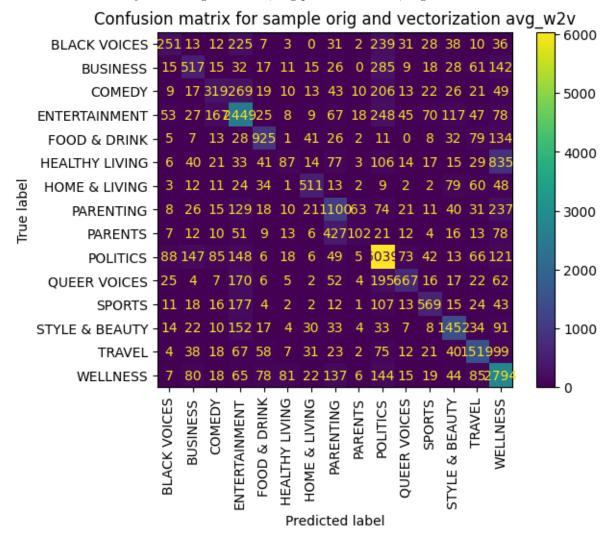


	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.4816	0.5055	0.9604	0.4933	0.6968	0.4634	1988.0000
BUSINESS	0.6269	0.4099	0.9825	0.4957	0.6346	0.3797	1959.0000
COMEDY	0.4959	0.4372	0.9682	0.4647	0.6506	0.4008	1958.0000
ENTERTAINMENT	0.5952	0.0635	0.9969	0.1148	0.2517	0.0574	1967.0000
FOOD and DRINK	0.3983	0.8820	0.9052	0.5488	0.8935	0.7965	1949.0000
HEALTHY LIVING	0.3089	0.4189	0.9340	0.3556	0.6255	0.3711	1929.0000
HOME and LIVING	0.4082	0.7984	0.9178	0.5402	0.8560	0.7240	1944.0000
PARENTING	0.5570	0.0420	0.9976	0.0780	0.2046	0.0379	1978.0000
PARENTS	0.2796	0.8047	0.8469	0.4150	0.8255	0.6786	2017.0000
POLITICS	0.6796	0.2846	0.9907	0.4012	0.5310	0.2620	1901.0000
QUEER VOICES	0.7294	0.3864	0.9897	0.5052	0.6184	0.3593	1967.0000
SPORTS	0.4245	0.8240	0.9220	0.5604	0.8716	0.7523	1915.0000
STYLE and BEAUTY	0.7383	0.4056	0.9900	0.5236	0.6337	0.3781	1913.0000
TRAVEL	0.7904	0.2103	0.9960	0.3323	0.4577	0.1930	1954.0000
WELLNESS	0.5888	0.0633	0.9968	0.1142	0.2511	0.0572	1992.0000
avg precision	0.5396	0.5396	0.5396	0.5396	0.5396	0.5396	0.0000
avg recall	0.4352	0.4352	0.4352	0.4352	0.4352	0.4352	0.0000
avg specificity	0.9595	0.9595	0.9595	0.9595	0.9595	0.9595	0.0000
avg f1	0.3952	0.3952	0.3952	0.3952	0.3952	0.3952	0.0000
avg geom. mean	0.5994	0.5994	0.5994	0.5994	0.5994	0.5994	0.0000
avg ind. bal. acc.	0.3935	0.3935	0.3935	0.3935	0.3935	0.3935	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

13 logregression orig avg w2v

Fit time: 352.5325565338135Test time: 0.06722331047058105

Sanity check: avg f1: 0.6523, avg precision: 0.6614, avg recall: 0.6745



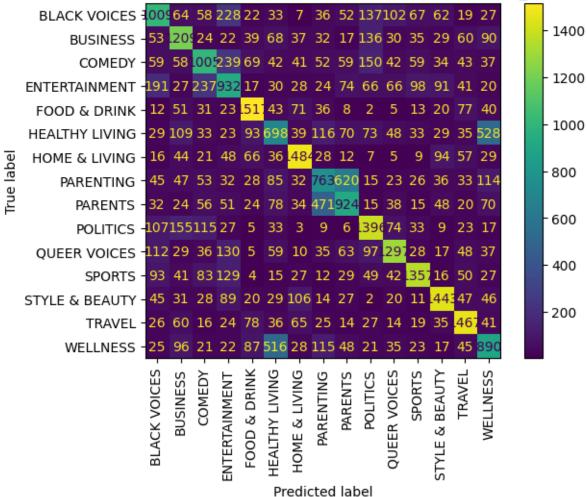
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.5345	0.2592	0.9926	0.3491	0.5072	0.2384	926.0000
BUSINESS	0.5746	0.4593	0.9856	0.5105	0.6728	0.4288	1191.0000
COMEDY	0.5268	0.3002	0.9900	0.3825	0.5452	0.2767	1046.0000
ENTERTAINMENT	0.6215	0.7290	0.9412	0.6710	0.8283	0.6716	3428.0000
FOOD and DRINK	0.7504	0.7264	0.9887	0.7382	0.8474	0.6993	1312.0000
HEALTHY LIVING	0.4292	0.0725	0.9954	0.1240	0.2686	0.0655	1338.0000
HOME and LIVING	0.6973	0.6646	0.9918	0.6806	0.8119	0.6376	811.0000
PARENTING	0.5464	0.6402	0.9652	0.5896	0.7861	0.5979	1804.0000
PARENTS	0.5646	0.1511	0.9968	0.2384	0.3881	0.1379	781.0000
POLITICS	0.7796	0.8804	0.9233	0.8269	0.9016	0.8094	6906.0000
QUEER VOICES	0.7243	0.5781	0.9902	0.6430	0.7566	0.5489	1254.0000
SPORTS	0.7071	0.5809	0.9914	0.6378	0.7589	0.5522	1014.0000
STYLE and BEAUTY	0.7505	0.7698	0.9822	0.7600	0.8695	0.7400	1911.0000
TRAVEL	0.7374	0.7766	0.9796	0.7565	0.8722	0.7453	2014.0000
WELLNESS	0.5762	0.8025	0.9175	0.6708	0.8581	0.7279	3595.0000
avg precision	0.6618	0.6618	0.6618	0.6618	0.6618	0.6618	0.0000
avg recall	0.6742	0.6742	0.6742	0.6742	0.6742	0.6742	0.0000
avg specificity	0.9574	0.9574	0.9574	0.9574	0.9574	0.9574	0.0000
avg f1	0.6518	0.6518	0.6518	0.6518	0.6518	0.6518	0.0000
avg geom. mean	0.7835	0.7835	0.7835	0.7835	0.7835	0.7835	0.0000
avg ind. bal. acc.	0.6274	0.6274	0.6274	0.6274	0.6274	0.6274	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

14 logregression ros avg w2v

Fit time: 324.3756229877472 Test time: 0.07582592964172363

Sanity check: avg f1: 0.6225, avg precision: 0.6591, avg recall: 0.6103

Confusion matrix for sample ros and vectorization avg_w2v



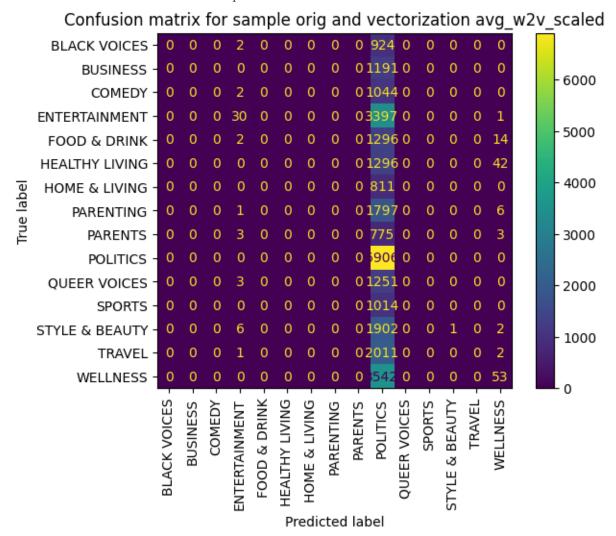
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.5690	0.5060	0.9721	0.5357	0.7014	0.4690	1988.0000
BUSINESS	0.6159	0.6442	0.9712	0.6297	0.7910	0.6052	1959.0000
COMEDY	0.5872	0.5623	0.9717	0.5745	0.7392	0.5240	1958.0000
ENTERTAINMENT	0.4788	0.5043	0.9605	0.4912	0.6960	0.4623	1967.0000
FOOD and DRINK	0.7649	0.7978	0.9825	0.7810	0.8854	0.7694	1949.0000
HEALTHY LIVING	0.4240	0.3831	0.9634	0.4025	0.6075	0.3476	1929.0000
HOME and LIVING	0.7559	0.7598	0.9826	0.7578	0.8640	0.7299	1944.0000
PARENTING	0.4479	0.4368	0.9611	0.4423	0.6479	0.3978	1978.0000
PARENTS	0.5186	0.4829	0.9669	0.5001	0.6833	0.4443	2017.0000
POLITICS	0.6313	0.6802	0.9725	0.6548	0.8133	0.6421	1901.0000
QUEER VOICES	0.7265	0.6563	0.9822	0.6896	0.8029	0.6237	1967.0000
SPORTS	0.7093	0.7478	0.9786	0.7280	0.8554	0.7149	1915.0000
STYLE and BEAUTY	0.7313	0.7569	0.9806	0.7439	0.8615	0.7256	1913.0000
TRAVEL	0.6970	0.7477	0.9768	0.7215	0.8546	0.7136	1954.0000
WELLNESS	0.4682	0.4960	0.9590	0.4817	0.6897	0.4536	1992.0000
avg precision	0.6076	0.6076	0.6076	0.6076	0.6076	0.6076	0.0000
avg recall	0.6097	0.6097	0.6097	0.6097	0.6097	0.6097	0.0000
avg specificity	0.9721	0.9721	0.9721	0.9721	0.9721	0.9721	0.0000
avg f1	0.6080	0.6080	0.6080	0.6080	0.6080	0.6080	0.0000
avg geom. mean	0.7655	0.7655	0.7655	0.7655	0.7655	0.7655	0.0000
avg ind. bal. acc.	0.5737	0.5737	0.5737	0.5737	0.5737	0.5737	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

15 MultinomialNB orig avg w2v scaled

Fit time: 31.032079458236694 Test time: 0.011812210083007812

Sanity check: avg f1: 0.09818, avg precision: 0.303, avg recall: 0.2409

'alpha': 0.0007543120063354615



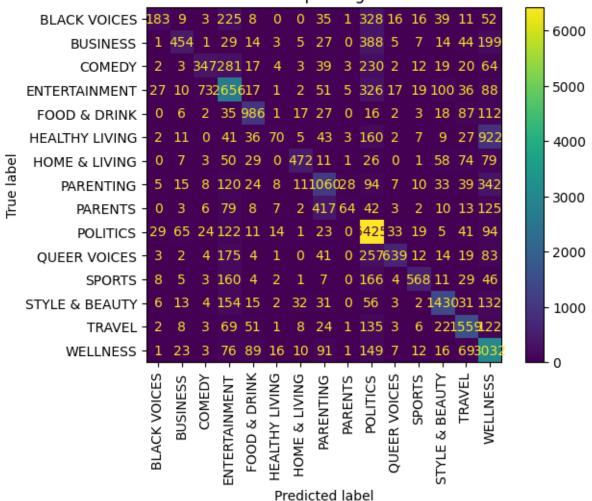
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	926.0000
BUSINESS	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1191.0000
COMEDY	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1046.0000
ENTERTAINMENT	0.6000	0.0088	0.9992	0.0173	0.0935	0.0079	3428.0000
FOOD and DRINK	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1312.0000
HEALTHY LIVING	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1338.0000
HOME and LIVING	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	811.0000
PARENTING	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1804.0000
PARENTS	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	781.0000
POLITICS	0.2369	1.0000	0.0078	0.3830	0.0881	0.0085	6906.0000
QUEER VOICES	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1254.0000
SPORTS	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1014.0000
STYLE and BEAUTY	1.0000	0.0005	1.0000	0.0010	0.0229	0.0005	1911.0000
TRAVEL	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	2014.0000
WELLNESS	0.4309	0.0147	0.9973	0.0285	0.1213	0.0133	3595.0000
avg precision	0.2439	0.2439	0.2439	0.2439	0.2439	0.2439	0.0000
avg recall	0.2383	0.2383	0.2383	0.2383	0.2383	0.2383	0.0000
avg specificity	0.7660	0.7660	0.7660	0.7660	0.7660	0.7660	0.0000
avg f1	0.0958	0.0958	0.0958	0.0958	0.0958	0.0958	0.0000
avg geom. mean	0.0480	0.0480	0.0480	0.0480	0.0480	0.0480	0.0000
avg ind. bal. acc.	0.0046	0.0046	0.0046	0.0046	0.0046	0.0046	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

16 MultinomialNB orig tfidf

Fit time: 155.75294733047485 Test time: 1.0726559162139893

Sanity check: avg f1: 0.6612, avg precision: 0.7019, avg recall: 0.6903

'alpha': 0.281176869797423



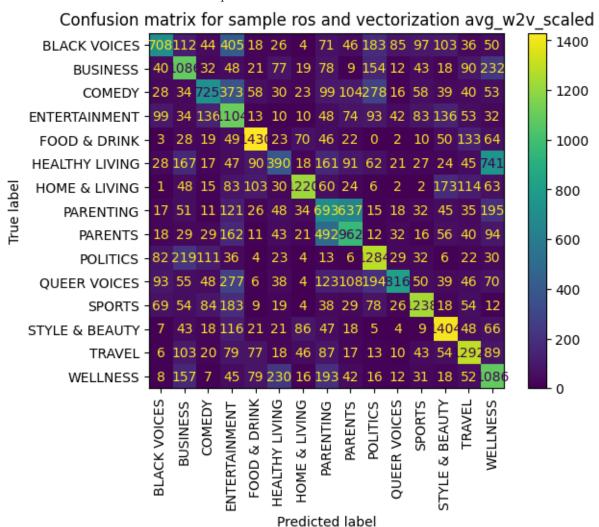
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.6803	0.1976	0.9970	0.3063	0.4439	0.1813	926.0000
BUSINESS	0.7161	0.3812	0.9936	0.4975	0.6154	0.3556	1191.0000
COMEDY	0.7169	0.3317	0.9952	0.4536	0.5746	0.3082	1046.0000
ENTERTAINMENT	0.6217	0.7748	0.9376	0.6899	0.8523	0.7146	3428.0000
FOOD and DRINK	0.7510	0.7515	0.9883	0.7512	0.8618	0.7252	1312.0000
HEALTHY LIVING	0.5385	0.0523	0.9979	0.0954	0.2285	0.0473	1338.0000
HOME and LIVING	0.8295	0.5820	0.9966	0.6841	0.7616	0.5560	811.0000
PARENTING	0.5501	0.5876	0.9685	0.5682	0.7544	0.5474	1804.0000
PARENTS	0.5981	0.0819	0.9985	0.1441	0.2860	0.0743	781.0000
POLITICS	0.7303	0.9304	0.8942	0.8183	0.9121	0.8349	6906.0000
QUEER VOICES	0.8600	0.5096	0.9963	0.6400	0.7125	0.4830	1254.0000
SPORTS	0.8161	0.5602	0.9955	0.6643	0.7467	0.5334	1014.0000
STYLE and BEAUTY	0.7953	0.7483	0.9866	0.7711	0.8592	0.7207	1911.0000
TRAVEL	0.7427	0.7741	0.9802	0.7581	0.8711	0.7431	2014.0000
WELLNESS	0.5521	0.8434	0.9044	0.6673	0.8734	0.7581	3595.0000
avg precision	0.6870	0.6870	0.6870	0.6870	0.6870	0.6870	0.0000
avg recall	0.6800	0.6800	0.6800	0.6800	0.6800	0.6800	0.0000
avg specificity	0.9503	0.9503	0.9503	0.9503	0.9503	0.9503	0.0000
avg f1	0.6494	0.6494	0.6494	0.6494	0.6494	0.6494	0.0000
avg geom. mean	0.7770	0.7770	0.7770	0.7770	0.7770	0.7770	0.0000
avg ind. bal. acc.	0.6269	0.6269	0.6269	0.6269	0.6269	0.6269	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

17 MultinomialNB ros avg w2v scaled

Fit time: 29.743656873703003 Test time: 0.01397395133972168

Sanity check: avg f1: 0.5749, avg precision: 0.5967, avg recall: 0.566

'alpha': 1.0481131341546853e-07



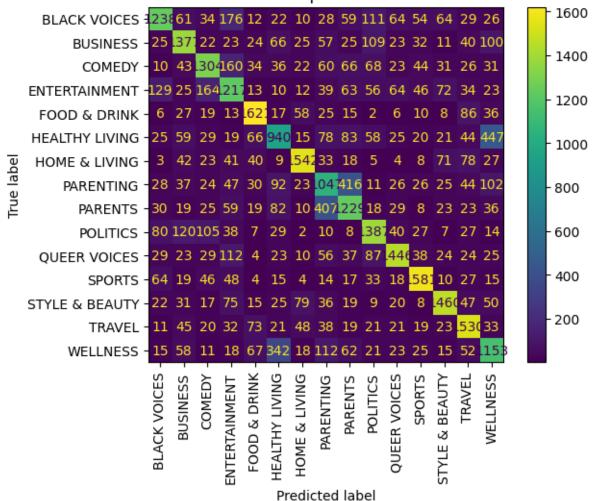
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.5866	0.3561	0.9818	0.4432	0.5913	0.3278	1988.0000
BUSINESS	0.4892	0.5544	0.9586	0.5197	0.7290	0.5099	1959.0000
COMEDY	0.5509	0.3703	0.9784	0.4429	0.6019	0.3402	1958.0000
ENTERTAINMENT	0.3529	0.5613	0.9260	0.4334	0.7209	0.5008	1967.0000
FOOD and DRINK	0.7274	0.7337	0.9804	0.7305	0.8481	0.7016	1949.0000
HEALTHY LIVING	0.3801	0.2022	0.9768	0.2640	0.4444	0.1822	1929.0000
HOME and LIVING	0.7726	0.6276	0.9869	0.6926	0.7870	0.5971	1944.0000
PARENTING	0.3081	0.3504	0.9431	0.3279	0.5748	0.3108	1978.0000
PARENTS	0.4395	0.4769	0.9551	0.4574	0.6749	0.4337	2017.0000
POLITICS	0.5366	0.6754	0.9596	0.5980	0.8051	0.6297	1901.0000
QUEER VOICES	0.7240	0.4148	0.9886	0.5275	0.6404	0.3866	1967.0000
SPORTS	0.6990	0.6465	0.9806	0.6717	0.7962	0.6127	1915.0000
STYLE and BEAUTY	0.6432	0.7339	0.9716	0.6855	0.8444	0.6961	1913.0000
TRAVEL	0.6152	0.6612	0.9705	0.6374	0.8011	0.6218	1954.0000
WELLNESS	0.3775	0.5452	0.9345	0.4461	0.7138	0.4896	1992.0000
avg precision	0.5460	0.5460	0.5460	0.5460	0.5460	0.5460	0.0000
avg recall	0.5263	0.5263	0.5263	0.5263	0.5263	0.5263	0.0000
avg specificity	0.9661	0.9661	0.9661	0.9661	0.9661	0.9661	0.0000
avg f1	0.5242	0.5242	0.5242	0.5242	0.5242	0.5242	0.0000
avg geom. mean	0.7042	0.7042	0.7042	0.7042	0.7042	0.7042	0.0000
avg ind. bal. acc.	0.4884	0.4884	0.4884	0.4884	0.4884	0.4884	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

18 MultinomialNB ros tfidf

Fit time: 148.98659086227417 Test time: 1.1193251609802246

Sanity check: avg f1: 0.6945, avg precision: 0.7187, avg recall: 0.6863

'alpha': 0.009540954763499934



	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.7219	0.6227	0.9826	0.6686	0.7822	0.5899	1988.0000
BUSINESS	0.6934	0.7029	0.9778	0.6981	0.8290	0.6684	1959.0000
COMEDY	0.6966	0.6660	0.9792	0.6809	0.8076	0.6317	1958.0000
ENTERTAINMENT	0.5857	0.6187	0.9685	0.6017	0.7741	0.5783	1967.0000
FOOD and DRINK	0.7989	0.8317	0.9851	0.8150	0.9052	0.8067	1949.0000
HEALTHY LIVING	0.5437	0.4873	0.9712	0.5139	0.6879	0.4504	1929.0000
HOME and LIVING	0.8211	0.7932	0.9877	0.8069	0.8851	0.7682	1944.0000
PARENTING	0.5132	0.5293	0.9637	0.5212	0.7142	0.4879	1978.0000
PARENTS	0.5754	0.6093	0.9668	0.5919	0.7675	0.5680	2017.0000
POLITICS	0.6949	0.7296	0.9778	0.7118	0.8446	0.6957	1901.0000
QUEER VOICES	0.7893		0.9859	0.7613	0.8513	0.7066	1967.0000
SPORTS	0.8124	0.8256	0.9867	0.8190	0.9025	0.8015	1915.0000
STYLE and BEAUTY	0.7828	0.7632	0.9852	0.7729	0.8671	0.7352	1913.0000
TRAVEL	0.7248	0.7830	0.9788	0.7528	0.8754	0.7514	1954.0000
WELLNESS	0.5444	0.5788	0.9647	0.5611	0.7473	0.5368	1992.0000
avg precision	0.6858	0.6858	0.6858	0.6858	0.6858	0.6858	0.0000
avg recall	0.6843	0.6843	0.6843	0.6843	0.6843	0.6843	0.0000
avg specificity	0.9774	0.9774	0.9774	0.9774	0.9774	0.9774	0.0000
avg f1	0.6844	0.6844	0.6844	0.6844	0.6844	0.6844	0.0000
avg geom. mean	0.8156	0.8156	0.8156	0.8156	0.8156	0.8156	0.0000
avg ind. bal. acc.	0.6510	0.6510	0.6510	0.6510	0.6510	0.6510	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

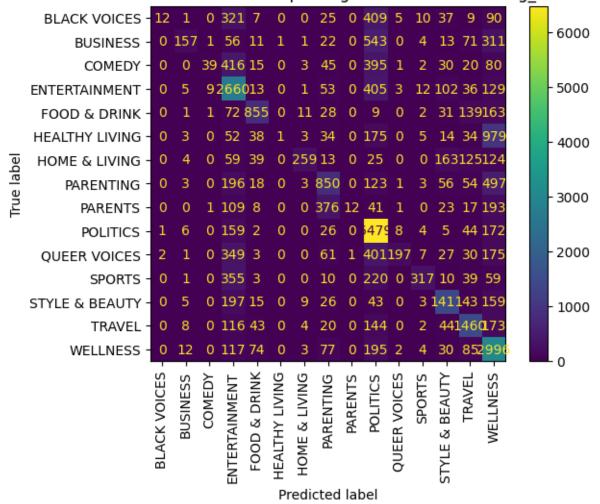
19 RandomForestClassifier orig avg w2v

Fit time: 115.16363644599915 Test time: 0.6198108196258545

Sanity check: avg f1: 0.556, avg precision: 0.6948, avg recall: 0.6166

No params.

Confusion matrix for sample orig and vectorization avg w2v



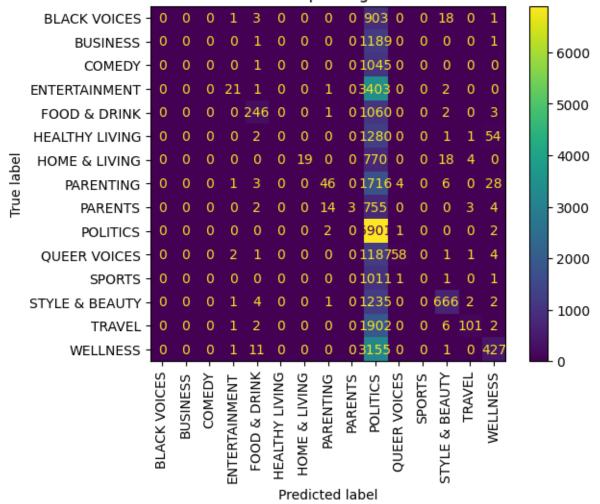
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.7647	0.0140	0.9999	0.0276	0.1185	0.0127	926.0000
BUSINESS	0.7765	0.1167	0.9986	0.2029	0.3414	0.1063	1191.0000
COMEDY	0.7794	0.0507	0.9995	0.0952	0.2250	0.0458	1046.0000
ENTERTAINMENT	0.5032	0.7646	0.9001	0.6069	0.8296	0.6789	3428.0000
FOOD and DRINK	0.7423	0.6654	0.9892	0.7018	0.8113	0.6369	1312.0000
HEALTHY LIVING	1.0000	0.0030	1.0000	0.0060	0.0547	0.0027	1338.0000
HOME and LIVING	0.8809	0.3465	0.9987	0.4973	0.5882	0.3235	811.0000
PARENTING	0.5081	0.4873	0.9691	0.4975	0.6872	0.4494	1804.0000
PARENTS	1.0000	0.0166	1.0000	0.0327	0.1290	0.0150	781.0000
POLITICS	0.6757	0.9380	0.8614	0.7855	0.8989	0.8142	6906.0000
QUEER VOICES	0.8858	0.1547	0.9991	0.2634	0.3932	0.1415	1254.0000
SPORTS	0.8295	0.2880	0.9979	0.4275	0.5361	0.2670	1014.0000
STYLE and BEAUTY	0.7152	0.7268	0.9798	0.7210	0.8439	0.6942	1911.0000
TRAVEL	0.6409	0.7339	0.9697	0.6843	0.8436	0.6948	2014.0000
WELLNESS	0.4783	0.8261	0.8741	0.6058	0.8498	0.7187	3595.0000
avg precision	0.6782	0.6782	0.6782	0.6782	0.6782	0.6782	0.0000
avg recall	0.6027	0.6027	0.6027	0.6027	0.6027	0.6027	0.0000
avg specificity	0.9342	0.9342	0.9342	0.9342	0.9342	0.9342	0.0000
avg f1	0.5395	0.5395	0.5395	0.5395	0.5395	0.5395	0.0000
avg geom. mean	0.6874	0.6874	0.6874	0.6874	0.6874	0.6874	0.0000
avg ind. bal. acc.	0.5393	0.5393	0.5393	0.5393	0.5393	0.5393	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

20 RandomForestClassifier orig tfidf

Fit time: 53.965874910354614 Test time: 1.2370004653930664

Sanity check: avg f1: 0.1852, avg precision: 0.5607, avg recall: 0.2921

No params.



	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	926.0000
BUSINESS	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1191.0000
COMEDY	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1046.0000
ENTERTAINMENT	0.7500	0.0061	0.9997	0.0122	0.0783	0.0055	3428.0000
FOOD and DRINK	0.8881	0.1875	0.9989	0.3096	0.4328	0.1721	1312.0000
HEALTHY LIVING	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1338.0000
HOME and LIVING	1.0000	0.0234	1.0000	0.0458	0.1531	0.0211	811.0000
PARENTING	0.7077	0.0255	0.9993	0.0492	0.1596	0.0230	1804.0000
PARENTS	1.0000	0.0038	1.0000	0.0077	0.0620	0.0035	781.0000
POLITICS	0.2508	0.9993	0.0809	0.4010	0.2843	0.0883	6906.0000
QUEER VOICES	0.9062	0.0463	0.9998	0.0880	0.2150	0.0418	1254.0000
SPORTS	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1014.0000
STYLE and BEAUTY	0.9224	0.3485	0.9980	0.5059	0.5897	0.3252	1911.0000
TRAVEL	0.9018	0.0501	0.9996	0.0950	0.2239	0.0454	2014.0000
WELLNESS	0.8072	0.1188	0.9960	0.2071	0.3440	0.1079	3595.0000
avg precision	0.5439	0.5439	0.5439	0.5439	0.5439	0.5439	0.0000
avg recall	0.2894	0.2894	0.2894	0.2894	0.2894	0.2894	0.0000
avg specificity	0.7828	0.7828	0.7828	0.7828	0.7828	0.7828	0.0000
avg f1	0.1828	0.1828	0.1828	0.1828	0.1828	0.1828	0.0000
avg geom. mean	0.2163	0.2163	0.2163	0.2163	0.2163	0.2163	0.0000
avg ind. bal. acc.	0.0705	0.0705	0.0705	0.0705	0.0705	0.0705	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

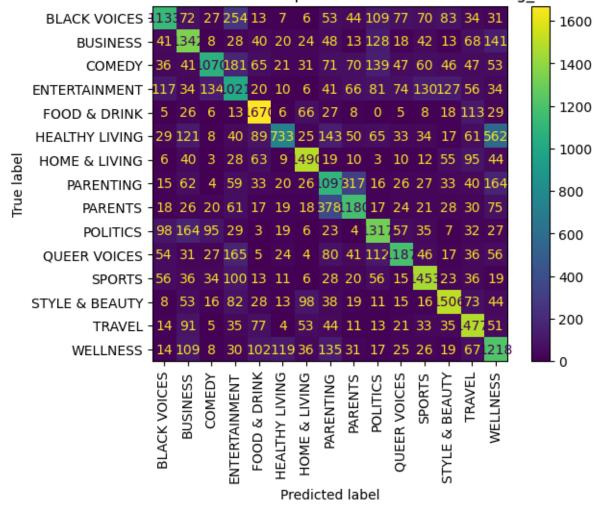
21 RandomForestClassifier ros avg w2v

Fit time: 103.47948789596558 Test time: 0.6246218681335449

Sanity check: avg f1: 0.6595, avg precision: 0.6777, avg recall: 0.6558

No params.

Confusion matrix for sample ros and vectorization avg w2v

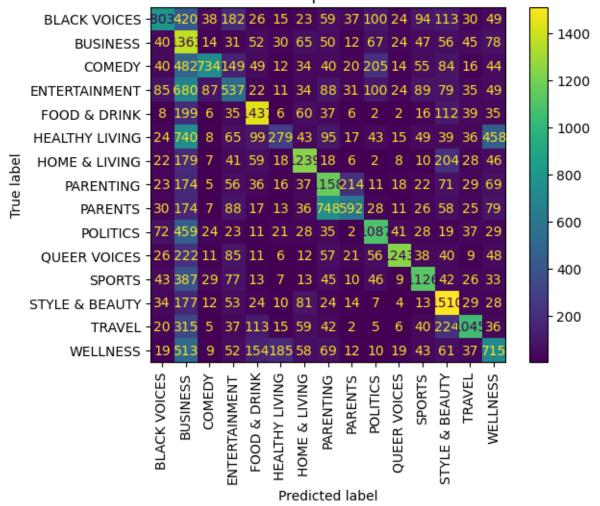


	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.7249	0.5156	0.9858	0.6026	0.7129	0.4844	1988.0000
BUSINESS	0.6563	0.6483	0.9757	0.6523	0.7953	0.6118	1959.0000
COMEDY	0.7286	0.5306	0.9859	0.6141	0.7233	0.4993	1958.0000
ENTERTAINMENT	0.4554	0.5679	0.9512	0.5054	0.7349	0.5194	1967.0000
FOOD and DRINK	0.7577	0.8102	0.9816	0.7830	0.8917	0.7816	1949.0000
HEALTHY LIVING	0.7290	0.3919	0.9897	0.5098	0.6228	0.3647	1929.0000
HOME and LIVING	0.7484	0.7958	0.9810	0.7714	0.8836	0.7662	1944.0000
PARENTING	0.4888	0.5086	0.9615	0.4985	0.6993	0.4669	1978.0000
PARENTS	0.6271	0.6386	0.9720	0.6328	0.7878	0.6000	2017.0000
POLITICS	0.6081	0.7044	0.9685	0.6527	0.8260	0.6642	1901.0000
QUEER VOICES	0.7437	0.6121	0.9848	0.6715	0.7764	0.5803	1967.0000
SPORTS	0.7396	0.7995	0.9803	0.7684	0.8853	0.7696	1915.0000
STYLE and BEAUTY	0.7119	0.7595	0.9786	0.7350	0.8621	0.7270	1913.0000
TRAVEL	0.6654	0.7308	0.9738	0.6966	0.8436	0.6944	1954.0000
WELLNESS	0.4642	0.6350	0.9466	0.5364	0.7753	0.5824	1992.0000
avg precision	0.6560	0.6560	0.6560	0.6560	0.6560	0.6560	0.0000
avg recall	0.6426	0.6426	0.6426	0.6426	0.6426	0.6426	0.0000
avg specificity	0.9744	0.9744	0.9744	0.9744	0.9744	0.9744	0.0000
avg f1	0.6414	0.6414	0.6414	0.6414	0.6414	0.6414	0.0000
avg geom. mean	0.7876	0.7876	0.7876	0.7876	0.7876	0.7876	0.0000
avg ind. bal. acc.	0.6067	0.6067	0.6067	0.6067	0.6067	0.6067	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

22 RandomForestClassifier ros tfidf

Fit time: 53.62057900428772 Test time: 1.2318823337554932

Sanity check: avg f1: 0.554, avg precision: 0.597, avg recall: 0.5434 No params.



	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.7124	0.3576	0.9895	0.4762	0.5949	0.3315	1988.0000
BUSINESS	0.5623	0.4308	0.9760	0.4879	0.6485	0.3976	1959.0000
COMEDY	0.7444	0.3585	0.9912	0.4840	0.5961	0.3329	1958.0000
ENTERTAINMENT	0.1800	0.5897	0.8069	0.2758	0.6898	0.4655	1967.0000
FOOD and DRINK	0.7050	0.6706	0.9800	0.6874	0.8107	0.6369	1949.0000
HEALTHY LIVING	0.3803	0.1441	0.9835	0.2090	0.3765	0.1298	1929.0000
HOME and LIVING	0.6693	0.6476	0.9773	0.6583	0.7956	0.6121	1944.0000
PARENTING	0.4607	0.4833	0.9591	0.4717	0.6808	0.4415	1978.0000
PARENTS	0.5432	0.4056	0.9748	0.4644	0.6288	0.3728	2017.0000
POLITICS	0.5925	0.6218	0.9704	0.6068	0.7768	0.5823	1901.0000
QUEER VOICES	0.8713	0.6573	0.9930	0.7493	0.8079	0.6308	1967.0000
SPORTS	0.6796	0.5692	0.9813	0.6195	0.7473	0.5355	1915.0000
STYLE and BEAUTY	0.5423	0.7470	0.9560	0.6284	0.8451	0.6992	1913.0000
TRAVEL	0.7211	0.5399	0.9851	0.6175	0.7293	0.5082	1954.0000
WELLNESS	0.4132	0.4598	0.9524	0.4353	0.6618	0.4164	1992.0000
avg precision	0.5849	0.5849	0.5849	0.5849	0.5849	0.5849	0.0000
avg recall	0.5114	0.5114	0.5114	0.5114	0.5114	0.5114	0.0000
avg specificity	0.9650	0.9650	0.9650	0.9650	0.9650	0.9650	0.0000
avg f1	0.5242	0.5242	0.5242	0.5242	0.5242	0.5242	0.0000
avg geom. mean	0.6922	0.6922	0.6922	0.6922	0.6922	0.6922	0.0000
avg ind. bal. acc.	0.4720	0.4720	0.4720	0.4720	0.4720	0.4720	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

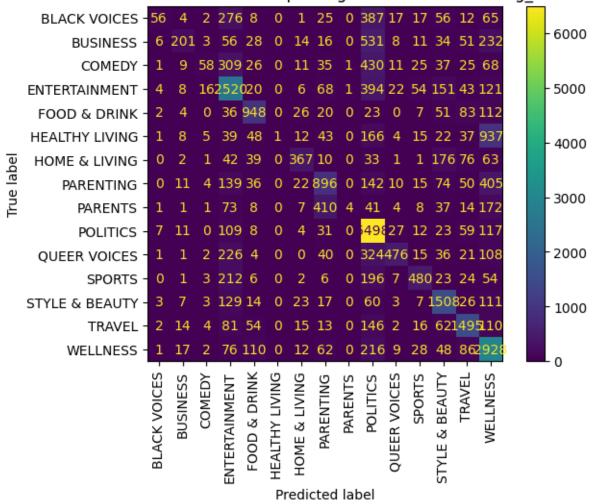
23 ridge orig avg w2v

Fit time: 3.0185704231262207 Test time: 0.08026123046875

Sanity check: avg f1: 0.5798, avg precision: 0.6282, avg recall: 0.6341

No params.

Confusion matrix for sample orig and vectorization avg w2v



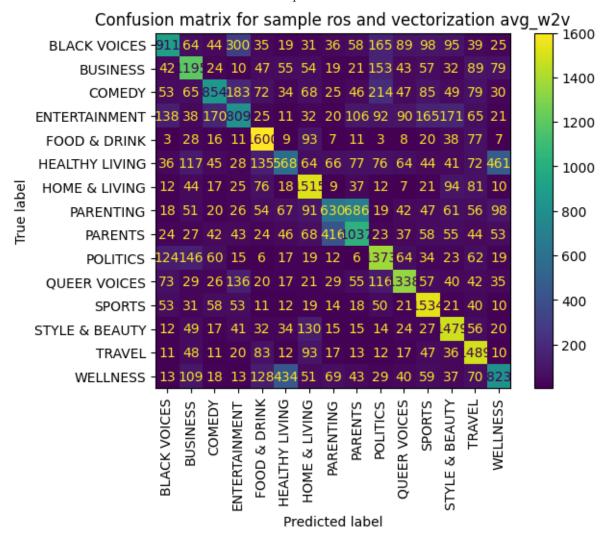
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.7191	0.0691	0.9991	0.1261	0.2628	0.0626	926.0000
BUSINESS	0.6899	0.2074	0.9961	0.3189	0.4545	0.1903	1191.0000
COMEDY	0.6000	0.0373	0.9991	0.0702	0.1930	0.0337	1046.0000
ENTERTAINMENT	0.5751	0.7509	0.9266	0.6513	0.8341	0.6835	3428.0000
FOOD and DRINK	0.6826	0.7424	0.9838	0.7112	0.8546	0.7127	1312.0000
HEALTHY LIVING	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1338.0000
HOME and LIVING	0.7646	0.3884	0.9966	0.5151	0.6222	0.3635	811.0000
PARENTING	0.5239	0.5460	0.9675	0.5347	0.7268	0.5060	1804.0000
PARENTS	1.0000	0.0013	1.0000	0.0026	0.0358	0.0012	781.0000
POLITICS	0.6856	0.9392	0.8674	0.7926	0.9026	0.8205	6906.0000
QUEER VOICES	0.7984	0.3979	0.9955	0.5311	0.6294	0.3725	1254.0000
SPORTS	0.6966	0.4438	0.9931	0.5422	0.6639	0.4165	1014.0000
STYLE and BEAUTY	0.6631	0.7797	0.9724	0.7167	0.8707	0.7436	1911.0000
TRAVEL	0.6955	0.7666	0.9753	0.7293	0.8647	0.7321	2014.0000
WELLNESS	0.5349	0.8072	0.9020	0.6435	0.8533	0.7212	3595.0000
avg precision	0.6260	0.6260	0.6260	0.6260	0.6260	0.6260	0.0000
avg recall	0.6331	0.6331	0.6331	0.6331	0.6331	0.6331	0.0000
avg specificity	0.9412	0.9412	0.9412	0.9412	0.9412	0.9412	0.0000
avg f1	0.5783	0.5783	0.5783	0.5783	0.5783	0.5783	0.0000
avg geom. mean	0.7153	0.7153	0.7153	0.7153	0.7153	0.7153	0.0000
avg ind. bal. acc.	0.5745	0.5745	0.5745	0.5745	0.5745	0.5745	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

24 ridge ros avg w2v

Fit time: 3.0338809490203857 Test time: 0.09703421592712402

Sanity check: avg f1: 0.5834, avg precision: 0.6284, avg recall: 0.5814

No params.



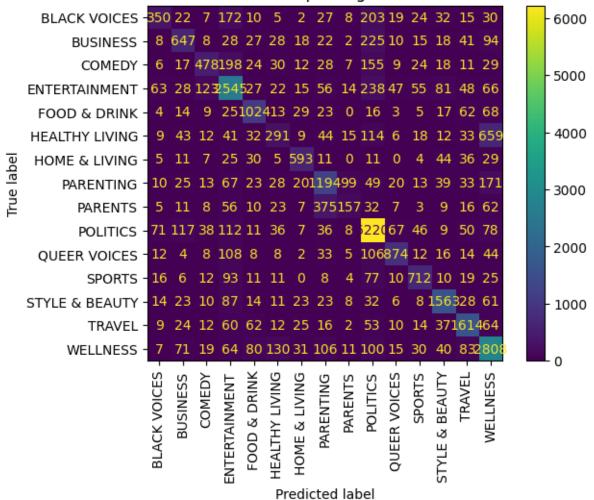
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.5970	0.4396	0.9784	0.5064	0.6559	0.4070	1988.0000
BUSINESS	0.5804	0.6488	0.9664	0.6127	0.7918	0.6071	1959.0000
COMEDY	0.6230	0.4785	0.9793	0.5413	0.6846	0.4452	1958.0000
ENTERTAINMENT	0.5006	0.4103	0.9706	0.4510	0.6310	0.3759	1967.0000
FOOD and DRINK	0.6659	0.8456	0.9698	0.7450	0.9056	0.8098	1949.0000
HEALTHY LIVING	0.4438	0.2701	0.9762	0.3358	0.5135	0.2450	1929.0000
HOME and LIVING	0.6625	0.7593	0.9725	0.7076	0.8593	0.7227	1944.0000
PARENTING	0.4531	0.2391	0.9791	0.3130	0.4839	0.2168	1978.0000
PARENTS	0.4759	0.5493	0.9553	0.5100		0.5035	2017.0000
POLITICS	0.5900	0.6912	0.9667	0.6366	0.8174	0.6498	1901.0000
QUEER VOICES	0.6651	0.6512	0.9764	0.6581	0.7974	0.6152	1967.0000
SPORTS	0.6349	0.7828	0.9686	0.7011	0.8707	0.7441	1915.0000
STYLE and BEAUTY	0.6408	0.7982	0.9688	0.7109	0.8794	0.7601	1913.0000
TRAVEL	0.6159	0.7682	0.9658	0.6837	0.8613	0.7272	1954.0000
WELLNESS	0.4697	0.4664	0.9616	0.4680	0.6697	0.4263	1992.0000
avg precision	0.5740	0.5740	0.5740	0.5740	0.5740	0.5740	0.0000
avg recall	0.5853	0.5853	0.5853	0.5853	0.5853	0.5853	0.0000
avg specificity	0.9704	0.9704	0.9704	0.9704	0.9704	0.9704	0.0000
avg f1	0.5712	0.5712	0.5712	0.5712	0.5712	0.5712	0.0000
avg geom. mean	0.7422	0.7422	0.7422	0.7422	0.7422	0.7422	0.0000
avg ind. bal. acc.	0.5491	0.5491	0.5491	0.5491	0.5491	0.5491	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

25 svc orig tfidf

Fit time: 270.16023111343384 Test time: 0.974750280380249

Sanity check: avg f1: 0.7178, avg precision: 0.7349, avg recall: 0.7374

'C': 0.1



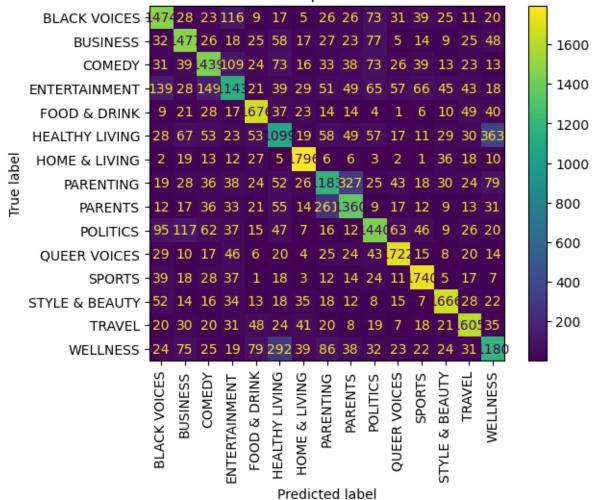
	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.6593	0.3218	0.9946	0.4325	0.5657	0.2985	926.0000
BUSINESS	0.6426	0.5071	0.9881	0.5669	0.7079	0.4770	1191.0000
COMEDY	0.6762	0.4273	0.9924	0.5237	0.6512	0.4001	1046.0000
ENTERTAINMENT	0.6830	0.7713	0.9526	0.7245	0.8572	0.7214	3428.0000
FOOD and DRINK	0.7408	0.7866	0.9871	0.7630	0.8812	0.7609	1312.0000
HEALTHY LIVING	0.6040	0.1345	0.9958	0.2200	0.3660	0.1224	1338.0000
HOME and LIVING	0.7753	0.7361	0.9939	0.7552	0.8554	0.7128	811.0000
PARENTING	0.5870	0.6846	0.9684	0.6320	0.8142	0.6442	1804.0000
PARENTS	0.5779	0.1140	0.9977	0.1904	0.3372	0.1036	781.0000
POLITICS	0.7949	0.9170	0.9271	0.8516	0.9221	0.8494	6906.0000
QUEER VOICES	0.8353	0.6834	0.9940	0.7518	0.8242	0.6582	1254.0000
SPORTS	0.7823	0.6982	0.9930	0.7379	0.8327	0.6729	1014.0000
STYLE and BEAUTY	0.7906	0.8257	0.9848	0.8078	0.9018	0.8002	1911.0000
TRAVEL	0.7778	0.8083	0.9830	0.7928	0.8914	0.7807	2014.0000
WELLNESS	0.6320	0.8264	0.9328	0.7162	0.8780	0.7627	3595.0000
avg precision	0.7168	0.7168	0.7168	0.7168	0.7168	0.7168	0.0000
avg recall	0.7228	0.7228	0.7228	0.7228	0.7228	0.7228	0.0000
avg specificity	0.9625	0.9625	0.9625	0.9625	0.9625	0.9625	0.0000
avg f1	0.7020	0.7020	0.7020	0.7020	0.7020	0.7020	0.0000
avg geom. mean	0.8175	0.8175	0.8175	0.8175	0.8175	0.8175	0.0000
avg ind. bal. acc.	0.6797	0.6797	0.6797	0.6797	0.6797	0.6797	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000

26 svc ros tfidf

Fit time: 305.685373544693 Test time: 0.9919228553771973

Sanity check: avg f1: 0.7344, avg precision: 0.7603, avg recall: 0.7292

'C': 10



	precision	recall	specificity	f1	geom. mean	ind. bal. acc.	support
BLACK VOICES	0.7576	0.7626	0.9823	0.7601	0.8655	0.7326	1988.0000
BUSINESS	0.7378	0.7928	0.9798	0.7643	0.8813	0.7622	1959.0000
COMEDY	0.7340	0.7457	0.9807	0.7398	0.8551	0.7141	1958.0000
ENTERTAINMENT	0.6714	0.6035	0.9788	0.6356	0.7685	0.5685	1967.0000
FOOD and DRINK	0.8431	0.8794	0.9884	0.8609	0.9323	0.8597	1949.0000
HEALTHY LIVING	0.5927	0.5702	0.9724	0.5812	0.7447	0.5322	1929.0000
HOME and LIVING	0.8792	0.9208	0.9910	0.8995	0.9553	0.9061	1944.0000
PARENTING	0.6561	0.5940	0.9775	0.6235	0.7620	0.5584	1978.0000
PARENTS	0.6781	0.7050	0.9753	0.6913	0.8292	0.6690	2017.0000
POLITICS	0.7251	0.7133	0.9813	0.7192	0.8366	0.6812	1901.0000
QUEER VOICES	0.8585	0.8485	0.9900	0.8535	0.9165	0.8281	1967.0000
SPORTS	0.8266	0.8935	0.9869	0.8587	0.9390	0.8735	1915.0000
STYLE and BEAUTY	0.8494	0.8610	0.9894	0.8551	0.9229	0.8408	1913.0000
TRAVEL	0.8103	0.8153	0.9864	0.8128	0.8967	0.7904	1954.0000
WELLNESS	0.6158	0.5768	0.9738	0.5956	0.7495	0.5394	1992.0000
avg precision	0.7485	0.7485	0.7485	0.7485	0.7485	0.7485	0.0000
avg recall	0.7515	0.7515	0.7515	0.7515	0.7515	0.7515	0.0000
avg specificity	0.9822	0.9822	0.9822	0.9822	0.9822	0.9822	0.0000
avg f1	0.7495	0.7495	0.7495	0.7495	0.7495	0.7495	0.0000
avg geom. mean	0.8566	0.8566	0.8566	0.8566	0.8566	0.8566	0.0000
avg ind. bal. acc.	0.7230	0.7230	0.7230	0.7230	0.7230	0.7230	0.0000
total support	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	29331.0000