

An Analysis of Cross-Genre and In-Genre Performance for Author Profiling in Social Media



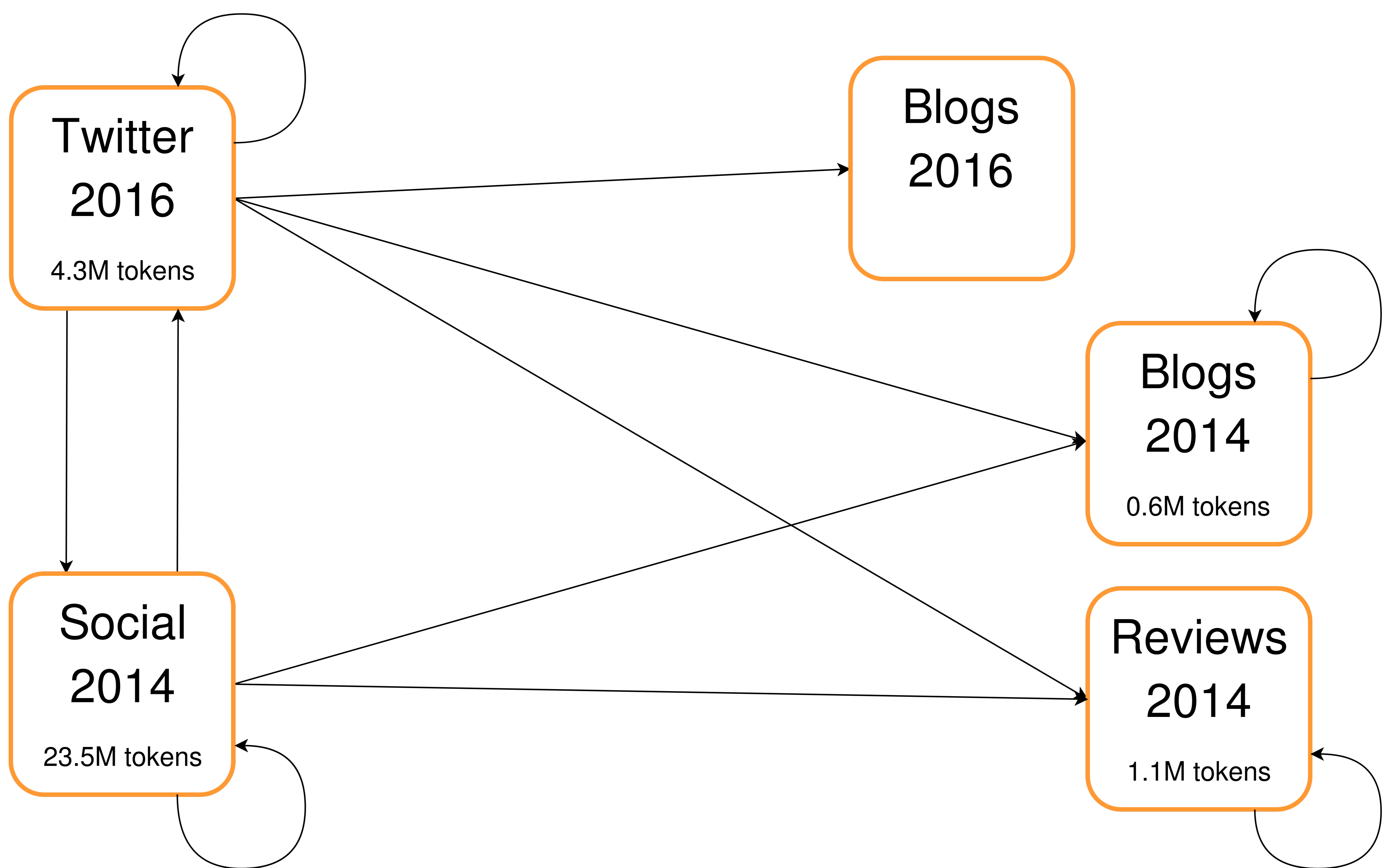
Maria Medvedeva, Hessel Haagsma & Malvina Nissim

Q1 Is the Twitter-trained model that won the PAN 2016 cross-genre task on author profiling truly cross-genre so that good results can be observed on other datasets?

A1 Maybe, confounding factors like *training data size* and *genre gap* provide an unclear picture.

Q2 If the features truly capture some general aspects of demographics, can the model be trained on datasets other than Twitter and still yield a good performance?

A2 Not on Social Media. A more structured evaluation on a wider range of genres is needed to give a definitive answer.



#	Training Set	Test Set	Age (en)	Gender (en)	Age (es)	Gender (es)	Average
1	Twitter 2016	(cross-val)	0.4573	0.7067	0.4899	0.7085	0.5906
2	Blogs 2014	(cross-val)	0.3810	0.7143	0.4091	0.6590	0.5409
3	Reviews 2014	(cross-val)	0.3236	0.6526	–	–	0.4881
4	Social 2014	(cross-val)	0.3277	0.4946	0.3855	0.5951	0.4507
5	Twitter 2016	Blogs 2016 (test)	0.5897	0.6410	0.5179	0.7143	0.6157
6	Twitter 2016	Blogs 2014	0.5374	0.7347	0.4205	0.6818	0.5936
7	Twitter 2016	Reviews 2014	0.2377	0.5000	–	–	0.3689
8	Twitter 2016	Social 2014	0.3273	0.4960	0.3097	0.5628	0.4240
9	Social 2014	Blogs 2014	0.4354	0.5714	0.4318	0.5795	0.5045
10	Social 2014	Reviews 2014	0.2413	0.5115	–	–	0.3764
11	Social 2014	Twitter 2016	0.3601	0.5367	0.3985	0.5188	0.4535
12	Twitter 2014	Twitter 2014 (test)	0.5065	0.7338	0.6111	0.6556	0.6268
13	Blogs 2014	Blogs 2014 (test)	0.4615	0.6795	0.4821	0.5893	0.5531
14	Reviews 2014	Reviews 2014 (test)	0.3502	0.7259	–	–	0.5381
15	Social 2014	Social 2014 (test)	0.3652	0.5421	0.4894	0.6837	0.5201