

Marco Di Giovanni

PhD candidate, 3rd year. Adv: Marco Brambilla
Thesis submission: October 2020



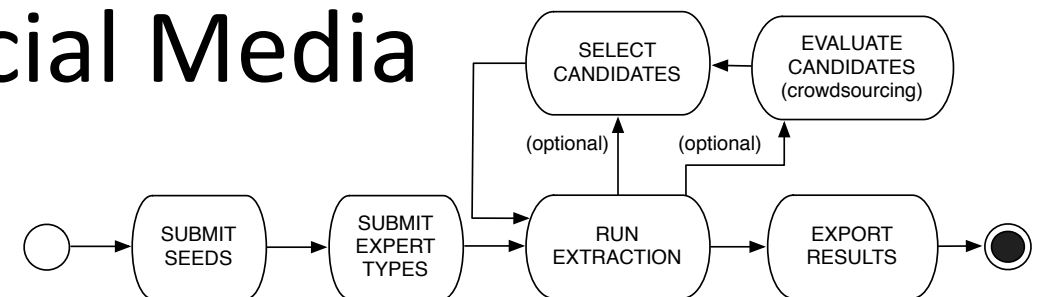
Syntactic and Semantic Similarity on Social Networks

Relevant publications/submissions

- Marco Brambilla, Stefano Ceri, Florian Daniel, Marco Di Giovanni, Andrea Mauri, Giorgia Ramponi: **Iterative Knowledge Extraction from Social Networks**. WWW (Companion Volume) 2018: 1359-1364
- Marco Di Giovanni, Marco Brambilla, Stefano Ceri, Florian Daniel, Giorgia Ramponi: **Content-based Classification of Political Inclinations of Twitter Users**. BigData 2018: 4321-4327
- Giorgia Ramponi*, Marco Brambilla, Stefano Ceri, Florian Daniel, Marco Di Giovanni*: **Content-based Characterization of Online Social Communities**. Information Processing and Management

Initial investigation

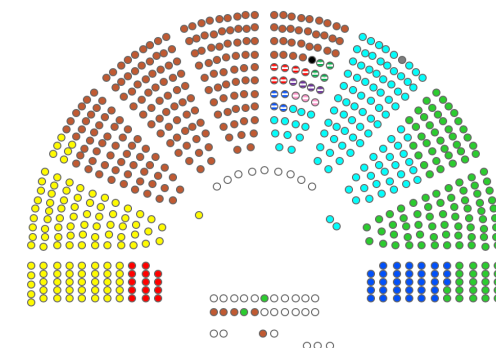
- Knowledge Extraction from Social Media



- Content-based Communities Detection and Characterization

Domain	Feature	$cd_{precision}$	cd_{recall}	$KLD_{precision}$	KLD_{recall}	$l1_{precision}$	$l1_{recall}$	$l2_{precision}$	$l2_{recall}$
Chess	NNP	0.800	0.905	0.770	0.870	0.800	0.885	0.140	0.270
	Noun	0.270	0.335	0.690	0.825	0.660	0.795	0.165	0.215
	Verb	0.155	0.235	0.130	0.330	0.200	0.350	0.135	0.200
	Instance	0.835	0.875	0.775	0.860	0.750	0.810	0.320	0.385
	Type	0.385	0.430	0.700	0.785	0.420	0.560	0.360	0.410
	Topic	0.726	0.824	0.702	0.834	0.734	0.868	0.732	0.822
Fashion	NNP	0.510	0.695	0.560	0.745	0.625	0.690	0.001	0.040
	Noun	0.180	0.345	0.485	0.610	0.710	0.770	0.075	0.150
	Verb	0.010	0.030	0.100	0.105	0.070	0.105	0.010	0.015
	Instance	0.695	0.765	0.595	0.765	0.705	0.750	0.001	0.015
	Type	0.120	0.250	0.165	0.195	0.235	0.315	0.125	0.240
	Topic	0.780	0.870	0.736	0.816	0.654	0.764	0.656	0.748
AW	NNP	0.245	0.435	0.265	0.385	0.310	0.450	0.030	0.030
	Noun	0.095	0.130	0.075	0.220	0.200	0.415	0.110	0.170
	Verb	0.120	0.190	0.005	0.155	0.085	0.190	0.115	0.165
	Instance	0.390	0.515	0.335	0.560	0.245	0.415	0.075	0.115
	Type	0.110	0.245	0.095	0.190	0.165	0.250	0.110	0.230
	Topic	0.522	0.642	0.444	0.570	0.406	0.532	0.378	0.484

- Political Orientation Classification

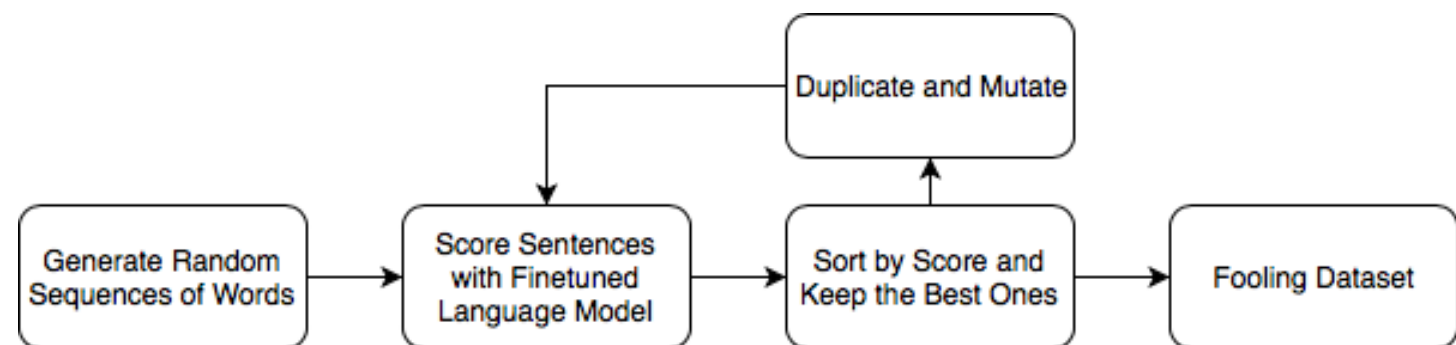


Pre-trained Language Models are Easily Fooled

Language Models: BERT and RoBERTa

Task: Language Acceptability (CoLA)

Evolutionary algorithm



Example of fooling sentences:

Sentence	Score
block characters rude cedar plum pronoun climbers	0.993
Pfizer Dracula articles lousy scissors firemen Genie Letters	0.990
umps teacher Abbey Gillespie brave scones answer exercises	0.980

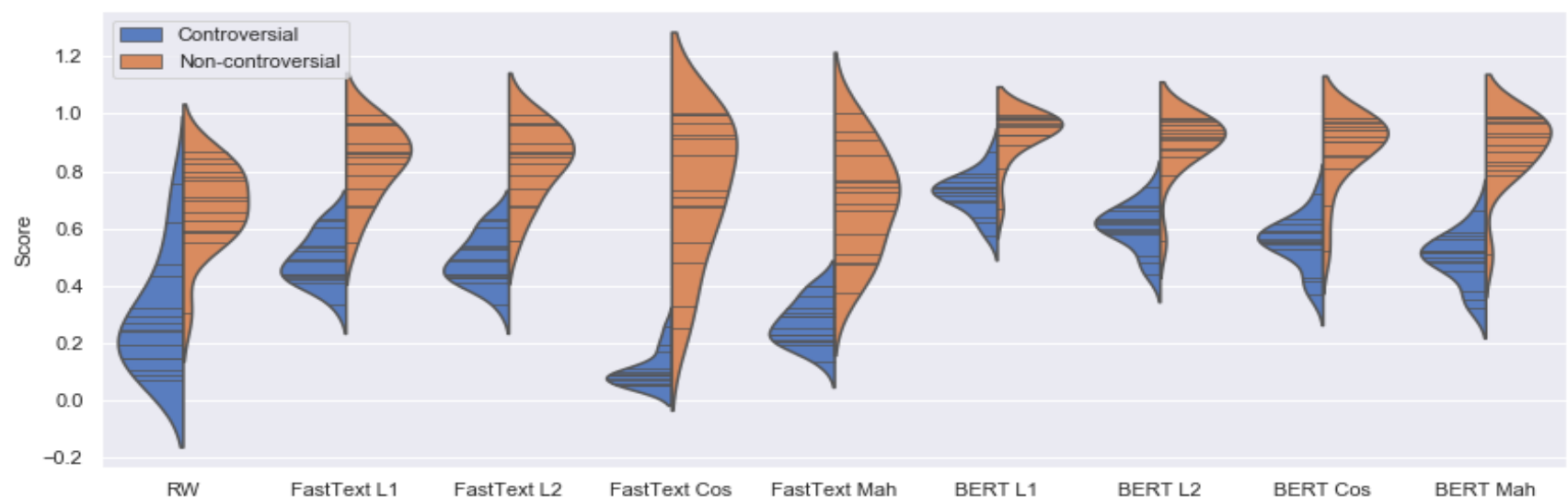
Submitted at ACL2020

Measuring Controversy in Social Networks through NLP

30 Twitter topics

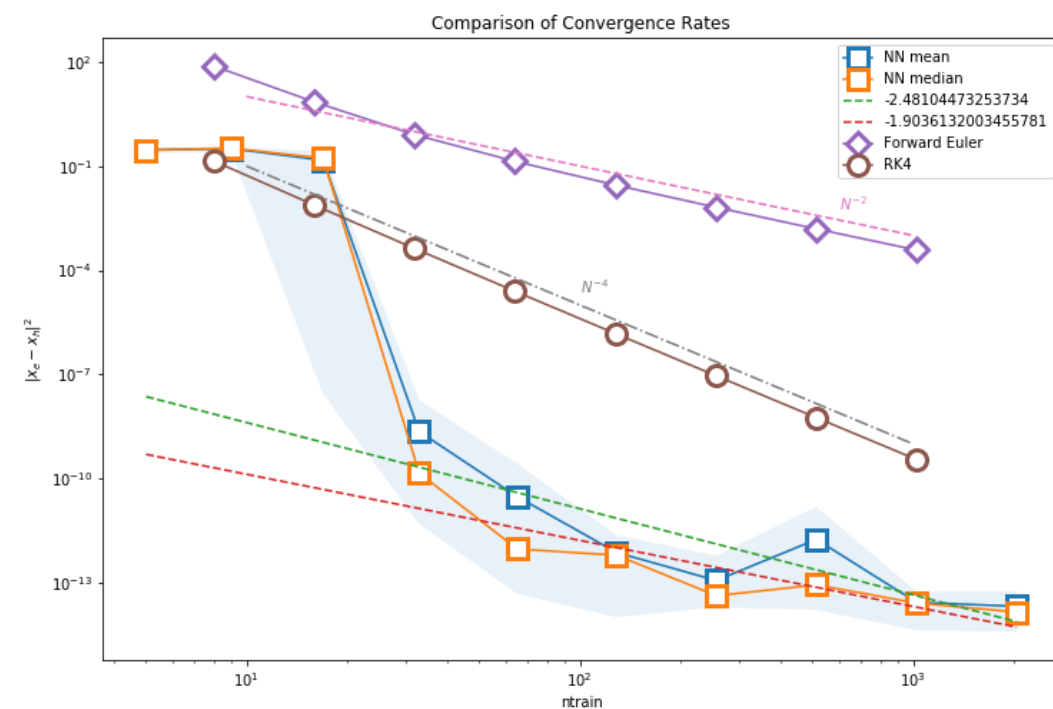
- Graph Building
- Community Identification
- Sanitization and Preprocessing
- Embedding
- Controversy Score

Submitted at SIGIR2020

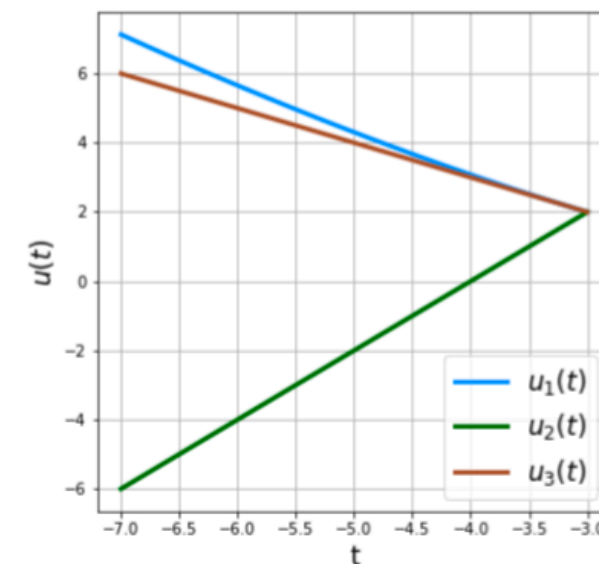


Side Project: Differential Equations solved with Neural Networks

Analysis of convergence rates:



Multiple solutions:



About to be submitted

Accepted at AAAI Spring Symposia
23-25 March 2020