PAR: Political Actor Representation Learning with Social Context and Expert Knowledge

Shangbin Feng, Zhaoxuan Tan, Zilong Chen, Ningnan Wang, Peisheng Yu, Qinghua Zheng, Xiaojun Chang, Minnan Luo

EMNLP 2022

Why political actor representation learning?

Analyzing political text with language models and NLP models has become a thriving research field at the intersection of natural language processing and computational political science.

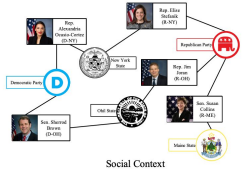
These **political texts**, such as political op-eds, legislator statements, and political news articles frequently mention various political actors.

While the names of political actors are just tokens to a language model, **real-world political analysis often goes beyond tokens and requires much more domain knowledge**, such as the referenced legislator's party affiliation, elected office, voting records, ideological position, and more.

However, language models often could not achieve such fine-grained understanding of external knowledge on their own, hence the need for political actor representation learning to improve political text analysis.

Why social context and expert knowledge?

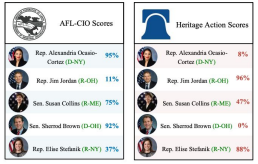
While existing works on political actor analysis achieved relative success, they fall short of incorporating the social context and expert knowledge that are integral to the understanding of political actors.



Social Context

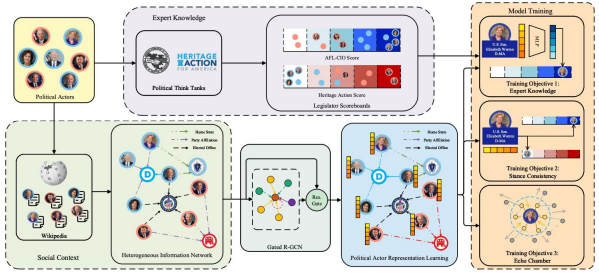
Social context information such as home state and party affiliation helps connect, compare, and contrast different political actors.

Expert knowledge from political think tanks helps to guide and ground the political actor representation process in the political reality.



Expert Knowledge

PAR Methodology



PAR follows a three-step approach to learn and use legislator representations:

1. Construct **heterogeneous information networks** to jointly represent political actors and the affiliated social context
2. Learn legislator representation with **three training objectives**: expert knowledge, stance consistency, and echo chamber.
3. Apply PAR representations to political text analysis tasks as **features**

Analyzing Political Text with PAR representations

Instead of knowledge base embeddings and ideal point vectors, we use PAR as features to represent political actors in two political text analysis tasks.

Method	Setting	SemEval Acc	SemEval MaF	AllSides Acc	AllSides MaF
CNN	GloVe	79.63	*N/A*	*N/A*	*N/A*
	ELMo	84.04	*N/A*	*N/A*	*N/A*
HLSTM	GloVe	81.58	*N/A*	*N/A*	*N/A*
	ELMo	83.28	*N/A*	*N/A*	*N/A*
	Embed	81.71	*N/A*	76.45	74.95
	Output	81.25	*N/A*	76.66	75.39
BERT	base	84.03	82.60	81.55	80.13
MAN	GloVe	81.58	79.29	78.29	76.96
	ELMo	84.66	83.09	81.41	80.44
	Ensemble	86.21	84.33	85.00	84.25
KGAP	TransE	89.56	84.94	86.02	85.52
	TransR	88.54	83.45	85.15	84.61
	DistMult	88.51	83.63	84.47	83.90
	HolE	88.85	83.68	84.78	84.24
	RotatE	88.84	84.04	85.61	85.11
KGAP	PAR	**91.30**	**87.78**	**86.81**	**86.33**

Table 1: Political perspective detection performance on two benchmark datasets. Acc and MaF denote accuracy and macro-averaged F1-score. N/A indicates that the result is not reported in previous works.

Method	Setting random	Setting time-based
majority	77.48	77.40
ideal-point-wf	85.37	*N/A*
ideal-point-tfidf	86.48	*N/A*
ideal-vector	87.35	*N/A*
CNN	87.28	81.97
CNN+meta	88.02	84.30
LSTM+GCN	88.41	85.82
Vote	90.22	89.76
RoBERTa	87.59	87.56
TransE	82.70	80.06
PAR	**90.33**	**89.92**

Table 2: Roll call vote prediction performance (accuracy) with two experiment settings. N/A indicates that the result is not reported in previous works.

PAR and Blue/Red/Swing States

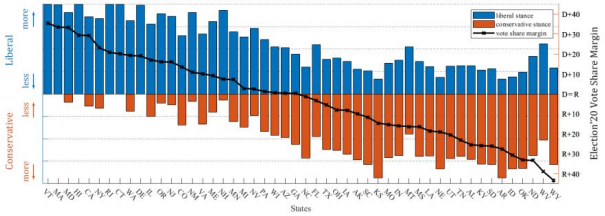
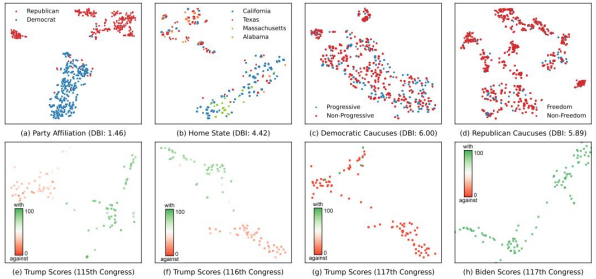


Figure 4: State stances predicted by PAR compared to vote shares in the 2020 U.S. presidential election.

PAR learns state ideological preferences that correlate well with the 2020 election.
Finding 1: certain traditional swing states are no longer as competitive? (PA, NC)
Finding 2: Georgia is the most electorally competitive state in the United States?

PAR Representation Learning



PAR learns political actor representations that correlate well with various **socio-political factors**, such as home state, congressional caucuses, and voting records.

Paper: https://arxiv.org/abs/2210.08362

Code: https://github.com/BunsenFeng/PAR