Zero-Shot Stance Detection: A Dataset and Model using Generalized Topic Representations

By 李想

2020.11.4

Motivation

- **Topic-specific stance**: requires the existence of numerous, well-labeled training examples in order to build a classifier for a new topic
- **Cross-target stance**: require human knowledge about any new topic and how it is related to the training topics
- Ability to generalize to **a wide variety of topics**

Motivation

- **Zero-shot stance detection :** A classifier is evaluated on a large number of completely new topics
- **Few-shot stance detection:** A classifier is evaluated on a large number of topics for which it has very few training examples

- Existing stance datasets:
  - Small number of topics & described in one way
  - Little linguistic variation in how topics are expressed and limited topics

Contribution

- A new dataset: VAST
  - For zero-shot & few-shot stance detection
- A new model: TGA Net
  - improves performance on a number of challenging linguistic phenomena
  - relies less on sentiment cues

Dataset: VAST

- A large range of topics covering **broad themes**
  - politics (e.g., ‘a Palestinian state’), education (e.g., ‘charter schools’), and public health (e.g., ‘childhood vaccination’)

- Includes a wide range of **similar expressions**
  - ‘guns on campus’ *vs* ‘firearms on campus’

Dataset: VAST

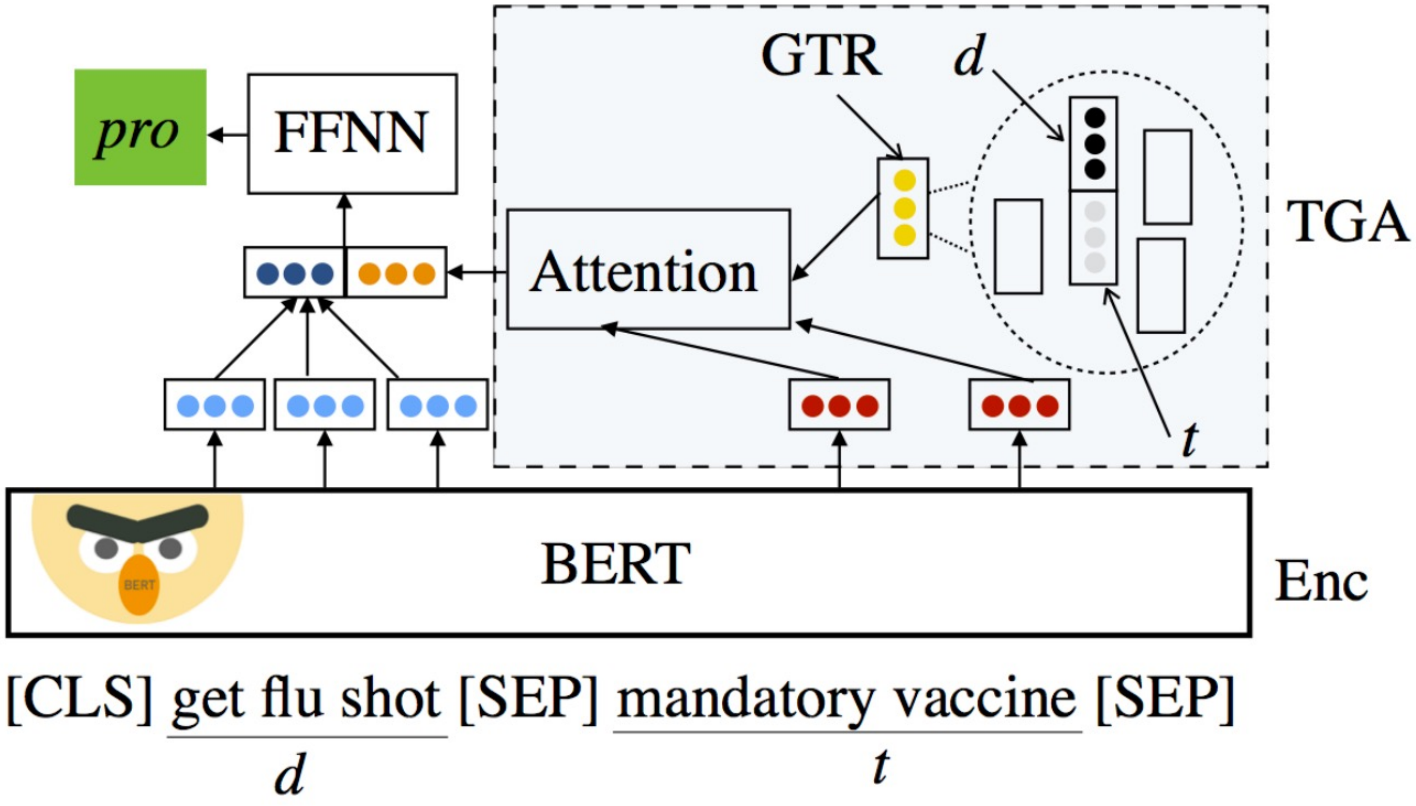
Comment	**ARC Stance**	**Topic**	ℓ	**Type**	
... Instead they have to work jobs (while their tax dollars are going to supporting illegal aliens) in order to put themselves through college [cont]	*Immigration* is really a problem	immigration	Con	Heur	**(1)**
		~~a problem~~ costs to american citizens	Con	List	**(2)**
Why should it be our job to help out the owners of the restaurants and bars? ... If they were paid a living wage ...[cont]	Not to tip	~~workplace~~ living wage	Pro	Corr	**(3)**
		restaurant owners	Con	List	**(4)**
...I like being able to access the internet about my health issues, and find I can talk with my doctors ... [cont]	*Medical websites* are healthful	medical websites	Pro	Heur	**(5)**
		home schoolers	Neu		**(6)**

Dataset: VAST

	#	%P	%C
Type Heur	4416	49	51
Type Corr	3594	44	51
Type List	11531	50	48
Neutral examples	3984	–	–
TOTAL examples	**23525**	40	41
Topics	5634	–	–

- Sematic & stance **complexity**
  - 4 topics/comment
- Great **variety** of topics
  - Few examples each topic (median 1, average 2.4)

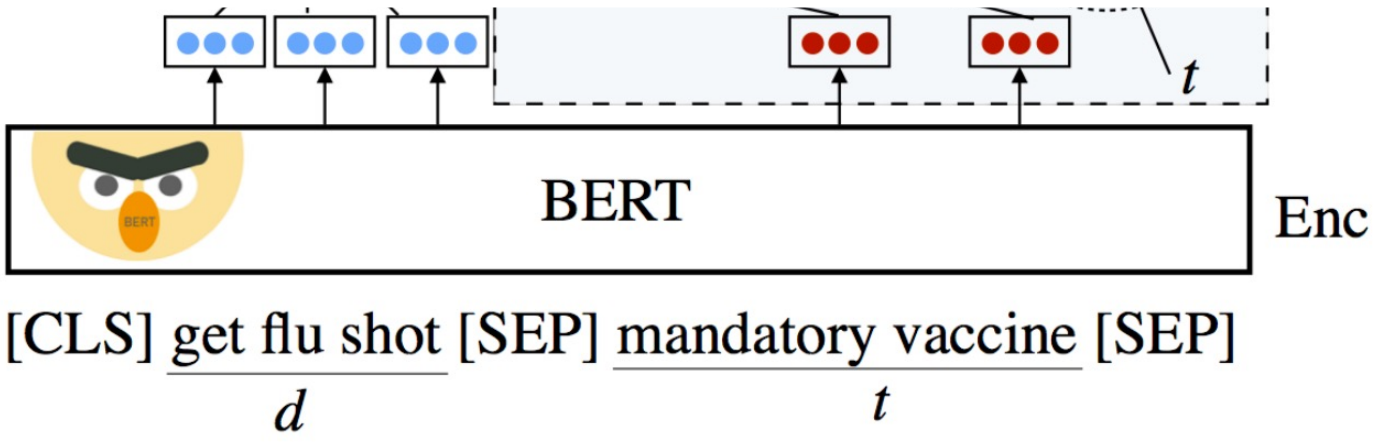
Method



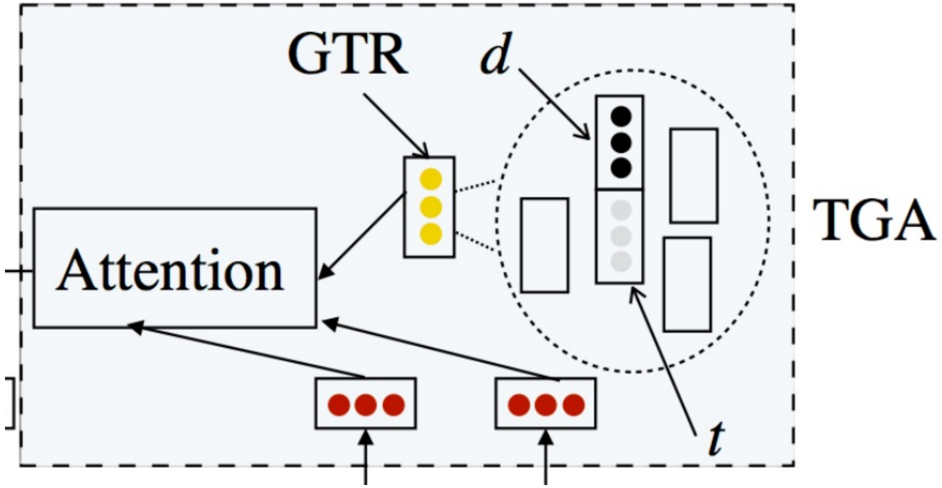
- Enc: Contextual conditional encoding
- GTR: Generalized Topic Representations
- TGA: Topic-Grouped Attention

Method: Contextual conditional encoding

- Embed a document and topic jointly with BERT (treat as a pair)
- Token embedding: $$\bar{t} = t^{(1)}, \ldots, t^{(m)}$$ $$\bar{d} = d^{(1)}, \ldots, d^{(n)}$$

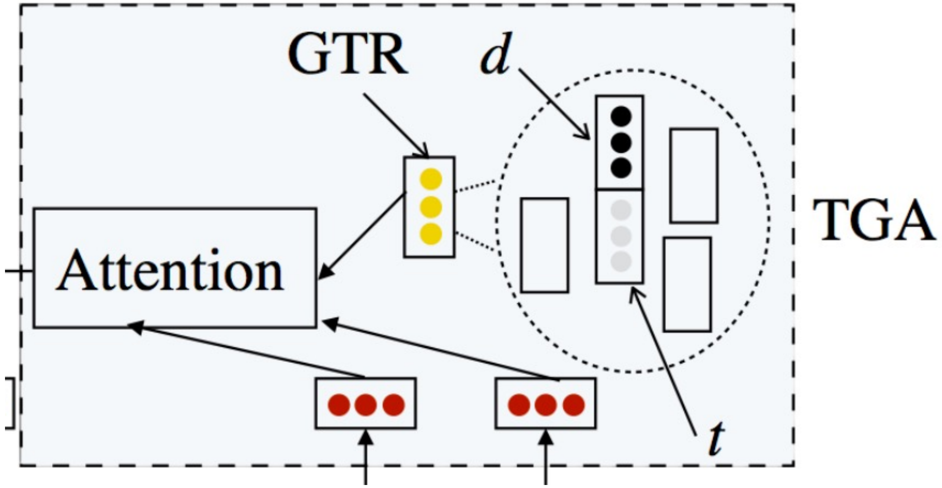


Method: Generalized Topic Representations



- Get v_d and v_t
  - 1) Embed the document and topic separately using BERT
  - 2) Weighted average over $\bar{d}$ ($\bar{t}$)
- $v_{dt} = [v_d; v_t]$ for hierarchical clustering
- r_{dt}: Centroid of the nearest cluster to example $x = (d, t, y)$

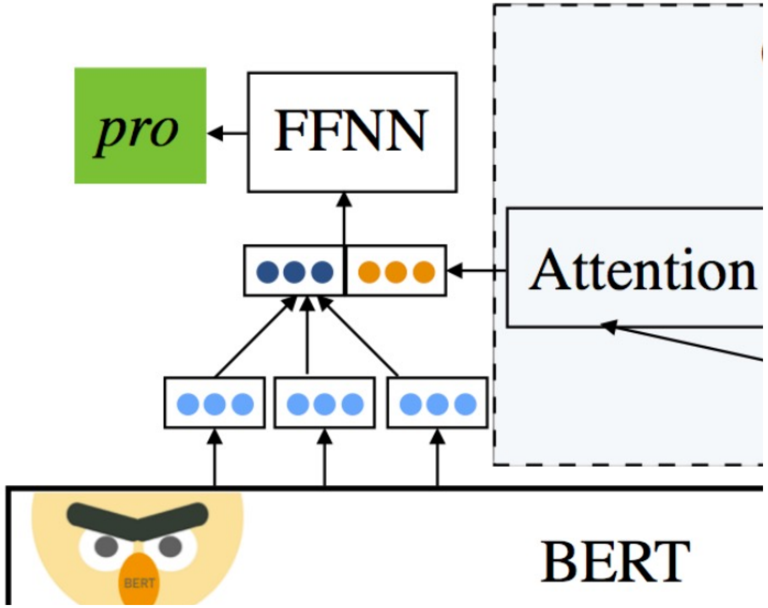
Method: Topic-Grouped Attention



- Scaled dot-product attention
- c_{dt}: captures the relationship between t and related topics and documents

$$c_{dt} = \sum_i s_i t^{(i)}, \; s_i = \mathrm{softmax}\Big(\lambda t^{(i)} \cdot (W_a r_{dt})\Big)$$

Method: Label prediction



- Combine output of TGA with document token embeddings
- Pass through a feed forward neural network (FFNN)

$$p = \mathrm{softmax}(W_2(\tanh(W_1[\tilde{d}; c_{dt}] + b_1) + b_2))$$

Experiments: Results

	F1 All			F1 Zero-Shot			F1 Few-Shot		
	pro	con	all	pro	con	all	pro	con	all
CMaj	.382	.441	.274	.389	.469	.286	.375	.413	.263
BoWV	.457	.402	.372	.429	.409	.349	.486	.395	.393
C-FFNN	.410	.434	.300	.408	.463	.417	.413	.405	.282
BiCond	.469	.470	.415	.446	.474	.428	.489	.466	.400
Cross-Net	.486	.471	.455	.462	.434	.434	.508	.505	.474
BERT-sep	.4734	.522	.5014	.414	.506	.454	.524	.539	.544
BERT-joint	.545	**.591**	.653	.546	.584	.661	.544	**.597**	.646
TGA Net	**.573***	.590	**.665**	**.554**	**.585**	**.666**	**.589***	.595	**.663**

Experiments: Results

Test Topic	**Cluster Topics**
drug addicts	war drug, cannabis, legalization, marijuana popularity, social effect, pot, colorado, american lower class, gateway drug, addiction, smoking marijauana, social drug
oil drilling	natural resource, international cooperation, renewable energy, alternative energy, petroleum age, electric car, solar use, offshore drilling, offshore exploration, planet
free college education	tax break home schooling, public school system, education tax, funding education, public service, school tax, homeschool tax credit, community, home schooling parent

Experiments: Results

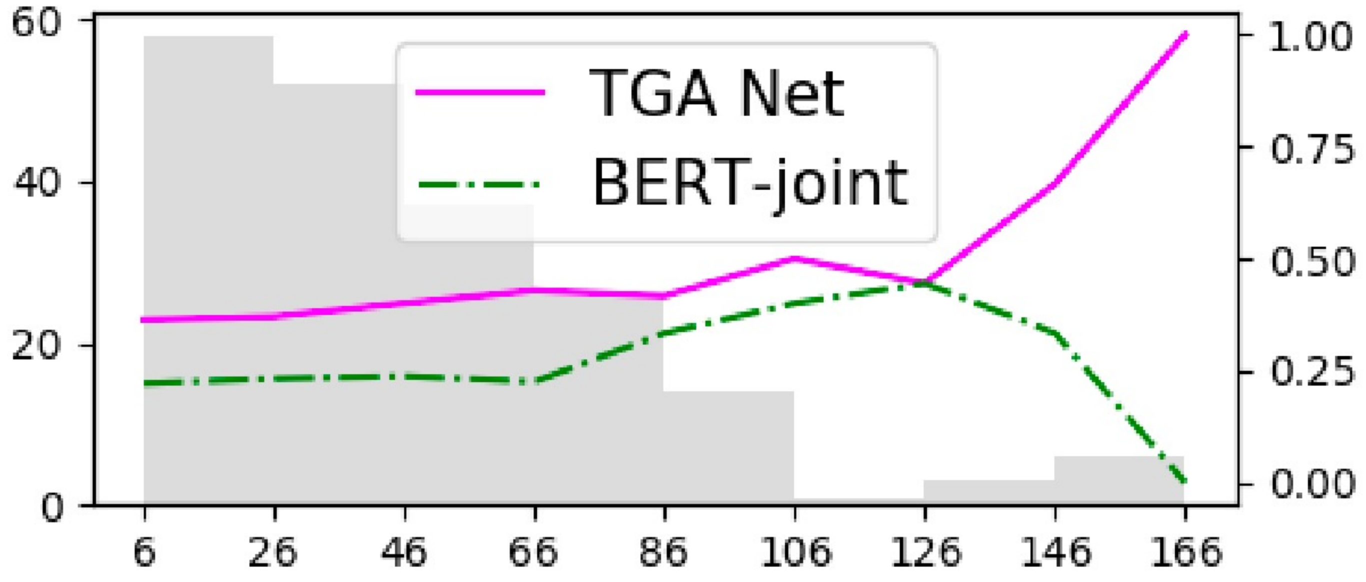


Figure 2: Percentage (right y-axis) each model is best on the test set as a function of the number of *unique topics in each cluster*. Histogram (left y-axis) of unique topics shown in gray.

Experiments: Results

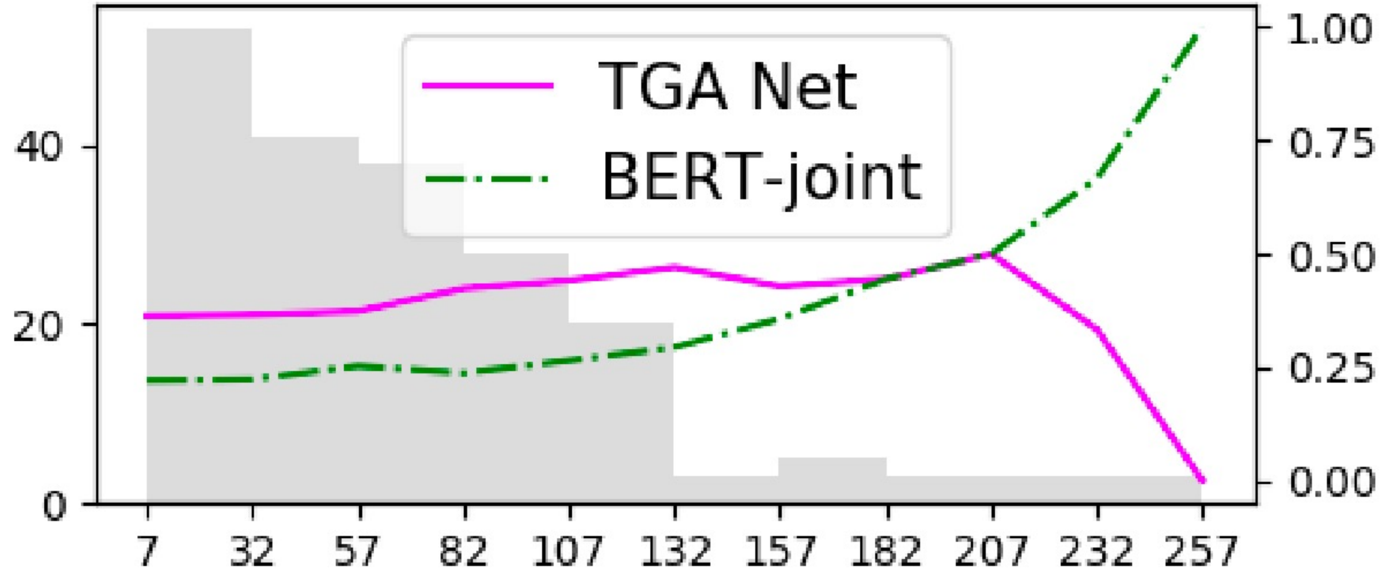


Figure 3: Percentage (right y-axis) each model is best on the test set as a function of the number of *examples per cluster*. Histogram of cluster sizes (left y-axis) shown in gray.

Experiments: Error analysis

		Imp	*mlT*	*mlS*	*Qte*	*Sarc*
BERT	I	.600	.610	.541	.625	.587
joint	O	.710	.748	.713	.657	.662
TGA	I	.623	.624	.547	.661	.637
Net	O	.713	.752	.725	.663	.667

- Imp: topic phrase not in document & not neutral label
- mlT: a document in examples with multiple topics
- mlS: a document in examples with diff & non-neutral label
- Qte: a document with quotations
- Sarc: sarcasm

Experiments: Error analysis

		BERT joint	TGA Net
Pro	$M+$	.73	.77
	$M-$	.65	.68
	$M+ \rightarrow M-$ (↓)	.71→.69	.74→.67
	$M- \rightarrow M+$ (↑)	.71→.74	.71→.70
Con	$M+$	.74	.70
	$M-$	.79	.80
	$M+ \rightarrow M-$ (↑)	.76→.80	.70→.74
	$M- \rightarrow M+$ (↓)	.75→.71	.75→.74

- MPQA: pos/neg sentiment words
- M+: pos > neg
- M-: pos < neg

- Lies less on sentiment cues than other models

Thanks!

Q&A