



Debating Europe: A Multilingual Multi-Target Stance Classification Dataset of Online Debates

Valentin Barriere, Alexandra Balahur, Brian Ravenet

06/24/22

PoliticalNLP – LREC



Debates
Browse all debates



In Focus
Themed debate series



Schools
Student-led debates



Leaders
Debating YOUR ideas

Debates

Discuss YOUR ideas with Europe's leaders



Should we give every homeless person a home?

In 2007, Finland adopted a "Housing First" policy. This approach, building on efforts Finland has adopted

22 comments



Would a 'European FBI' help prevent terror attacks?

In the 1930s, Bonnie and Clyde seemed unstoppable. Their gang was infamous for a series of brazen bank

17 comments



How can we end food waste?

Roughly 30-40% of total food production in the world is lost before it reaches the market. In the EU alone,

8 comments

[Join the debate](#)

Search debates...

Channels

[Asia-Europe](#)

[Future](#)

[Global](#)

[Greener](#)

[Quality](#)

[Security](#)

[Smarter](#)

Initial problem

Lack of multilingual multi-target stance-annotated debate dataset.

- We created a corpus of annotated stance in English online debates
- We assess the annotation quality by using our data to improve stance recognition in non-English
- We took advantage of the interactional structure of the debate using a context-aware transformer
- We show self-training can be used with the unlabeled data to improve the performances

Dataset collection

Does Europe need natural gas and nuclear to transition to green energy?

Started 08/02/2022 Greener



Should nuclear and natural gas be counted as "green energy"? This is the question that European policymakers are grappling with at the moment, as the EU puts together a set of technical rules known as the "EU taxonomy", classifying precisely what counts as a "sustainable" investment.

Germany's environment minister has called the plans "absolutely wrong". His country is in the process of phasing out the last of its nuclear power plants, though critics argue this has left Germany dependent on polluting coal, as well as natural gas imports from Russia.

What do our readers think? We had a comment sent in from Jessica on our German-language platform, saying:

"As long as [Germany's] traffic light coalition finally gets serious about climate protection, I'm happy. The plans look better than before, but I have my doubts whether the coal phase-out will really happen by 2030. Germany should definitely implement the Green Deal laws consistently!"



24
comments

JOIN the debate!

24 comments

Post a comment

What do YOU think?



Alfredo • February 11th, 2022 •

Let it freeze, let it freeze...

Write a reply



José • February 11th, 2022 •

The transition should be done but not that fast. If Russia shuts down gas how will we produce energy? the solar doesn't work all year and neither does the wind. it will be beautiful.

Write a reply



George • February 11th, 2022 •

Gas and nuclear IS green energy. Deal with it.

Write a reply



Jevgeni • February 15th, 2022 •

We are just cutting the branch of the tree we sit on by closing nuke stations, fighting cheap gas.

"green", "sustainable" and other marketing bs is not bringing us anywhere. Solar panels are dirty to produce and recycle and their lifespan is super short, same for windmills - we already have an issue in Europe where to store and how to recycle the one that we have (a huge amount).

We should do things gradually, not just cutting off everything in the name of clean energy. ... there is no such thing like clean energy.

Like Putin told in Berlin years ago: "what Europe will do, heating the houses with wood but for wood we would need to go to Siberia..."

Write a reply

Does Europe need natural gas and nuclear to transition to green energy?

24
comments

JOIN the debate!

<https://www.debatingeurope.eu/debates/>

Dataset collection and annotation

- The different debate questions, debates texts, debates tags, comments and user ids were collected: 6.5M tokens
- We selected **18 debates** related to the topic of the '*European Green Deal*': *Should we consume less energy?, Should we make the cities greener?, Can renewables ever replace fossil fuels 100?, ...*
- One expert annotator tagged at the comment level regarding the text was *yes/no/neutral/not answering* toward the close question's debate, using the INCEpTION platform [1]

Label	% DE	Unit	μ_{com}	μ_{deb}	Σ
✗	100%	Comments	\emptyset	89.5	125,798
		Words	51.7	4,623	6,499,625
✓	2.0%	Comments	\emptyset	140	2,523
		Words	33.4	4,683	84,289

Table 1: Low-level statistics on the DE dataset.

Experiment I: Multilinguality

Our labels come from one annotator, how to validate them using another dataset? Is our dataset useful for multilingual transfer learning? We used the multilingual X-Stance dataset [2] and studied transfer learning

Experiment II: Interactional Context

The data is from online debates. Is it possible to simply integrate the interactional context inside the transformer?

Experiment III: Semi-supervised learning

We annotated only 2% of the website. How to take advantage of the unbalanced and unlabeled data from the whole website?

Experiment I: Multilinguality

- X-stance: a dataset of multi-lingual multi-target stance recognition [2]
- No interaction between the users
- Multilingual: French, German and Italian
- Pretraining our model over our data helps to improve the results for non-English languages

	Intra-target			X-question			X-Topic			X-lingual
	DE	FR	Mean	DE	FR	Mean	DE	FR	Mean	IT
M-BERT [2]	76.8	76.6	76.6	68.5	68.4	68.4	68.9	70.9	69.9	70.2
XLM-R	76.3	78.0	77.1	71.5	72.9	72.2	71.2	73.7	72.4	73.0
XLM-R _{ft}	77.3	79.0	78.1	71.5	74.8	73.1	72.2	74.7	73.4	73.9

Table 2: Results over X-Stance dataset for a binary classification.

Experiment II: Interactional Context

Can renewables ever replace fossil fuels 100%?

...

That would be the ultimate reality as fossil fuels are a non-renewable and antiquated energy resource.

Carmelo A.

In foreseeable future and without subsidies, not.

Pawel K.

No, well not the current ones, and please get rid of those horrible giant wind turbines

Kevin

...

How to avoid an energy crisis?

...

Hydrogen! Germany is already studying and improving hydrogen as energy for cars. If cars can run...

Vincente S.

Vicente, obtaining hydrogen from water is an endothermic reaction. I believe approximately -8 kCal/mole....

Jovan I.

Well I am not an engineer, but once I saw a documentary about Hydrogen and they were showing a small machine that some...

Vincente S.

...

Should Europeans be encouraged to eat more sustainably?

...

I agree that the key to influence consumers is through education. I think one major problem that...

Alma

What a great idea!

Silvia P.

Where's the freedom to choose whatever you want to eat, to wear, to do?

Borislav V.

...

[CLS] Debate Question [SEP] Sent n [SEP] Sent n-1
[SEP] Sent n-2 [SEP].

Experiment II: Interactional Context

Can renewables ever replace fossil fuels 100%?

How to avoid an energy crisis ?

Should Europeans be encouraged to eat more sustainably?

...

That would be the ultimate reality as fossil fuels are a non-renewable and antiquated energy resource.

Carmelo A.

...

Hydrogen! Germany is already studying and improving hydrogen as energy for cars. If cars can run...

Vincente S.

...

I agree that the key to influence consumers is through education. I think one major problem that...

Alma

In foreseeable future and without subsidies, not.

Pawel K.

Vicente, obtaining hydrogen from water is an endothermic reaction. I believe approximately -8 kCal/mole...

Jovan I.

What a great idea!

Silvia P.

No , well not the current ones ,and please get rid of those horrible giant wind turbines

Kevin

Well I am not an engineer, but once I saw a documentary about Hydrogen and they were showing a small machine that some...

Vincente S.

Where's the freedom to chose whatever you want to eat, to wear, to do?

Borislav V.

...

...

...

[CLS] Debate Question [SEP] Sent n [SEP] Sent n-1
[SEP] Sent n-2 [SEP].

Experiment II: Interactional Context

Can renewables ever replace fossil fuels 100%?

How to avoid an energy crisis ?

Should Europeans be encouraged to eat more sustainably?

...

That would be the ultimate reality as fossil fuels are a non-renewable and antiquated energy resource.

Carmelo A.

In foreseeable future and without subsidies, not.

Pawel K.

No , well not the current ones ,and please get rid of those horrible giant wind turbines

Kevin

...

...

Hydrogen! Germany is already studying and improving hydrogen as energy for cars. If cars can run...

Vincente S.

Vicente, obtaining hydrogen from water is an endothermic reaction. I believe approximately -8 kCal/mole....

Jovan I.

Well I am not an engineer, but once I saw a documentary about Hydrogen and they were showing a small machine that some...

Vincente S.

...

...

I agree that the key to influence consumers is through education. I think one major problem that...

Alma

What a great idea!

Silvia P.

Where's the freedom to choose whatever you want to eat, to wear, to do?

Borislav V.

...

[CLS] Debate Question [SEP] Sent n [SEP] Sent n-1
[SEP] Sent n-2 [SEP].

Experiment II: Interactional Context

Can renewables ever replace fossil fuels 100%?

How to avoid an energy crisis ?

Should Europeans be encouraged to eat more sustainably?

...

That would be the ultimate reality as fossil fuels are a non-renewable and antiquated energy resource.

Carmelo A.

...

Hydrogen! Germany is already studying and improving hydrogen as energy for cars. If cars can run...

Vincente S.

...

I agree that the key to influence consumers is through education. I think one major problem that...

Alma

In foreseeable future and without subsidies, not.

Pawel K.

Vicente, obtaining hydrogen from water is an endothermic reaction. I believe approximately -8 kCal/mole...

Jovan I.

What a great idea!

Silvia P.

No , well not the current ones ,and please get rid of those horrible giant wind turbines

Kevin

Well I am not an engineer, but once I saw a documentary about Hydrogen and they were showing a small machine that some...

Vincente S.

Where's the freedom to choose whatever you want to eat, to wear, to do?

Borislav V.

[CLS] ... Debate Question [SEP] ... Sent n [SEP] Sent n-1
[SEP] Sent n-2 [SEP].

Experience II: Interactional Context

Ctxt	Prec	Rec.	F1	Acc
0	70.7	69.9	70.2	72.1
1	72.1	70.5	71.2	72.7
2	70.7	69.8	70.2	72.7

Table 3: Results over DE for different context windows. All the models were pre-trained over XS (XLM- R_{ft})

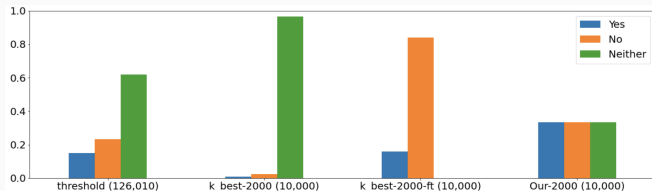
- We tried 3 different size for the context window
- Larger does not necessarily mean better

Experiment III: Semi-supervised learning

Label	% DE	Unit	μ_{com}	μ_{deb}	Σ
✗	100%	Comments	\emptyset	89.5	125,798
		Words	51.7	4,623	6,499,625
✓	2.0%	Comments	\emptyset	140	2,523
		Words	33.4	4,683	84,289

Table 4: Low-level statistics on the DE dataset, regarding there is label annotation or not. μ_{com}/μ_{deb} is the average mean of the respective units (comments or words) at the comment/debate-level.

Figure 1: Distribution of the pseudo-labels



Experience III: Semi-supervised learning

Unsupervised Method	Threshold	k_{max}	Balanced	Model	Prec.	Rec.	F1	Acc
\times	\times	\times	\times	XLM-R	68.6	69.3	68.9	70.1
				XLM-R _{ft}	70.7	69.9	70.2	72.1
thresh-0.99	0.99	\times	\times	XLM-R	68.6	69.8	69.1	70.7
				XLM-R _{ft}	68.9	69.6	69.0	70.9
k-best-2000	\times	2000	\times	XLM-R	67.5	68.3	67.8	69.3
				XLM-R _{ft}	70.4	69.9	69.8	71.9
k-best-600	\times	600	\times	XLM-R	69.4	68.5	68.0	69.5
				XLM-R _{ft}	72.5	70.3	71.1	73.3
our-2000	0.99	2000	✓	XLM-R	69.5	69.4	69.4	71.3
				XLM-R _{ft}	70.5	69.9	69.3	71.7
our-600	0.99	600	✓	XLM-R	70.9	71.6	71.1	72.7
				XLM-R _{ft}	71.5	71.5	71.4	73.5

Table 5: Results over the Debating Europe dataset for a 3-class classification using SSL

Conclusion

- Debating Europe: A new dataset of online political debates
- Annotation of a subpart of the dataset in stance
- We showed our annotation are useful using multi-lingual transfer learning
- We proposed a simple method to integrate context and show its efficiency
- We proposed a method to take advantage of unlabeled and unbalanced data

Questions?



J.-C. Klie, M. Bugert, B. Boullosa, R. E. de Castilho, and I. Gurevych.

The INCEpTION Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation.

Proceedings of the International Conference on Computational Linguistics, pages 5–9, 2018.



J. Vamvas and R. Sennrich.

X-stance: A Multilingual Multi-Target Dataset for Stance Detection.

In SwissText, 2020.