Multi-task learning beyond GLUE

How emotions or metaphors can help in socially relevant tasks

Pere-Lluis Huguet Cabot
huguetcabot@babelscape.com

SAPIENZA
UNIVERSITÀ DI ROMA

Multi-task Learning

Multi-task Learning (MTL) Aspects

- **Data amplification effect**: MTL provides more data, as the information from other tasks will not be present in Single-task Learning (STL).
- **Eavesdropping on other tasks:** The features learned for one of the tasks may be useful for another related one.
- **Information bias:** By jointly training more than one task, the combination may provide a beneficial information bias.

 Caruana - Multitask Learning: A Knowledge-Based Source of Inductive Bias - ICML (1993)

MTL types

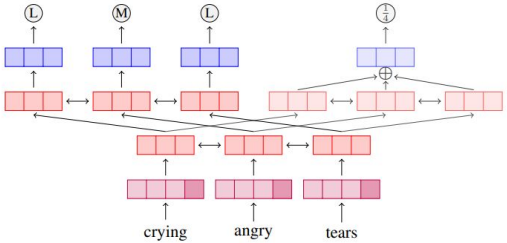
Same Level Vs Auxiliary Tasks

In auxiliary MTL, we want to improve performance on a main task.

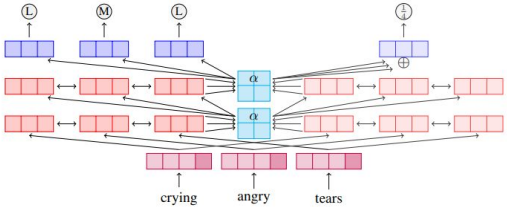
Hard Vs Soft parameter sharing

In hard-parameter sharing, each task shares some layers of the network, while in soft-parameter sharing, there is information shared between the tasks, but not layers.

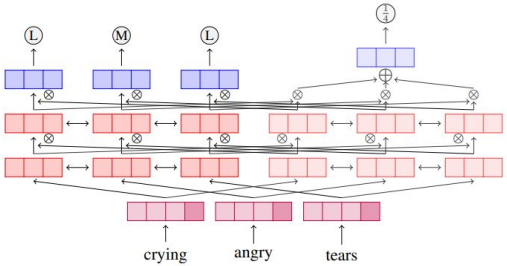
Examples



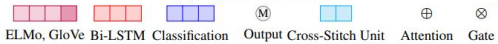
(a) Hard parameter sharing



(b) Cross-stitch network



(c) Gated network



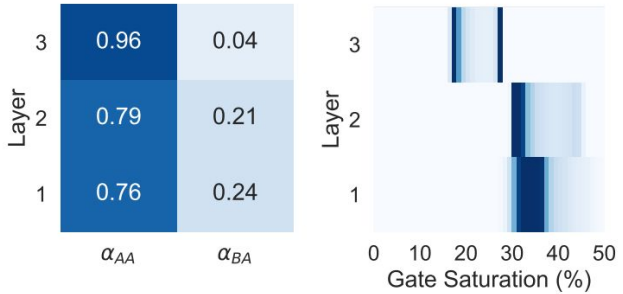
$$\widetilde{\mathbf{h}}_A = \alpha_{AA}\mathbf{h}_A + \alpha_{BA}\mathbf{h}_B$$
$$\widetilde{\mathbf{h}}_B = \alpha_{BB}\mathbf{h}_B + \alpha_{AB}\mathbf{h}_A$$

$$\mathbf{g}_A = \sigma(\mathbf{W}_A[\mathbf{h}_A; \mathbf{h}_B] + \mathbf{b}_A)$$
$$\widetilde{\mathbf{h}}_A = (1 - \mathbf{g}_A) \odot \mathbf{h}_A + \mathbf{g}_A \odot \mathbf{h}_B$$
$$\mathbf{g}_B = \sigma(\mathbf{W}_B[\mathbf{h}_A; \mathbf{h}_B] + \mathbf{b}_B)$$
$$\widetilde{\mathbf{h}}_B = (1 - \mathbf{g}_B) \odot \mathbf{h}_B + \mathbf{g}_B \odot \mathbf{h}_A$$

 Dankers, Verna et al. - Modelling the interplay of metaphor and emotion through multitask learning - EMNLP (2019)

Examples

Approach	Metaphor Task	
	Word (F_1)	Sent. (r)
Gao et al. (2018)	.726	-
LSTM (single task)	.737	.544
Hard Sharing		
+ Valence	.740	**.559**
+ Arousal	.740	**.558**
+ Dominance	**.743**	**.560**
Cross-Stitch Network		
+ Valence	.741	**.556**
+ Arousal	.740	**.558**
+ Dominance	**.743**	**.563**
Gated Network		
+ Valence	**.742**	**.561**
+ Arousal	.741	**.558**
+ Dominance	**.745**	**.560**



(a) Cross-Stitch Units (b) Gates

Figure 2: Illustration of the information flow in between the Bi-LSTM layers, for the dominance regression (B) and metaphor identification (A) tasks. Gate saturation % is calculated by averaging across the hidden dimensionality for every word in the test set.

Modelling Metaphor, Framing and Emotion in Political Discourse

Huguet Cabot et al. - The Pragmatics behind Politics - Findings of EMNLP (2020)

Tasks

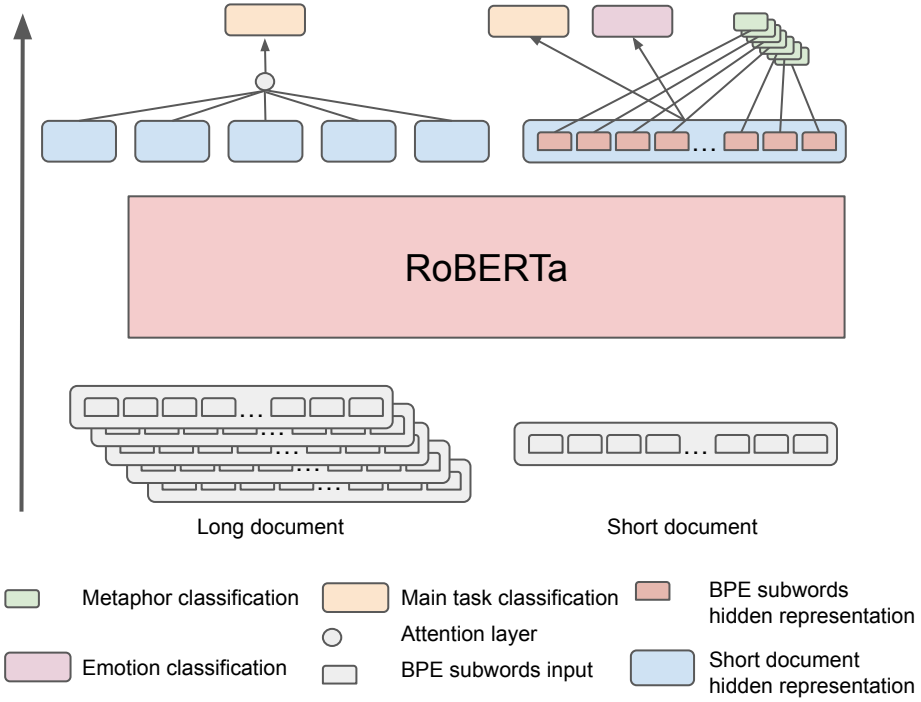
Main Tasks:

- Political Perspective in News. Allsides (Li and Goldwasser, 2019)
- Political Affiliation. RtGender (Voigt et al. ,2018)
- Framing. Media Frames Corpus (Card et al., 2015)

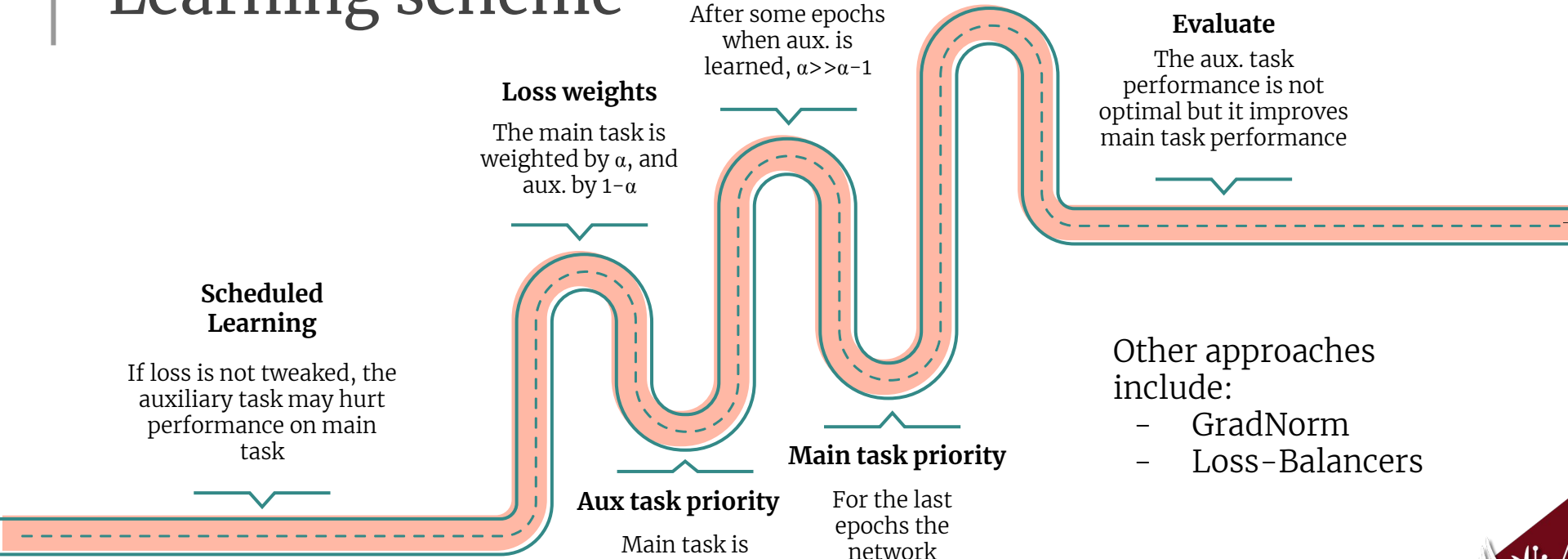
Auxiliary Tasks:

- Metaphor. VU Amsterdam (Steen et al., 2010)
- Emotions. SemEval-2018 Task 1 (Mohammad et al., 2018)

Model



Learning scheme



Other approaches include:

- GradNorm
- Loss-Balancers

Results

	Framing	Affiliation	Perspective
Li and Goldwasser (2019)			
- HLSTM (text-based)	-	-	.746
- GCN-HLSTM (using social information)			.917
STL	.707	.794	.848
MTL, Metaphor	**.716**	**.805**	**.854**
MTL, Emotion	.708	**.802**	**.860**

TABLE 3.2: Accuracy scores for the main political tasks. Significance compared to STL is bolded ($p < 0.05$).

	Anger	Anticipation	Disgust	Fear	Joy	Love	Optimism	Pessimism	Sadness	Surprise	Trust
Democrat	34.0%	42.9%	42.2%	23.1%	61.9%	73.6%	54.0%	82.5%	76.4%	75.4%	41.6%
Republican	66.0%	57.1%	57.8%	76.9%	38.1%	26.4%	46.0%	17.5%	23.6%	24.6%	58.4%

Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions

Modeling Populist attitude

Challenges:

- Lack of training data. Existing datasets are shallow or very small.
- Populism is not a well defined ideology, rather a “thin” ideology.

Solutions:

- Create our own dataset.
- Focus on a particular aspect of populist attitude.

 Huguet Cabot et al. - Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions - EACL (2021)

The Us Vs Them rhetoric

Here's what's gonna happen. One of these days, these guys aren't just going to ice a few European civilians. They're going to blow up the Notre Dame, or destroy a priceless work of European art at the Louvre, and then the shit will really hit the fan. I want to be careful how I word this but ask yourself a question. If the Nazis had won WWII, do you think Islam would be threatening European civilians in the hearts of London, Paris, and Berlin today?

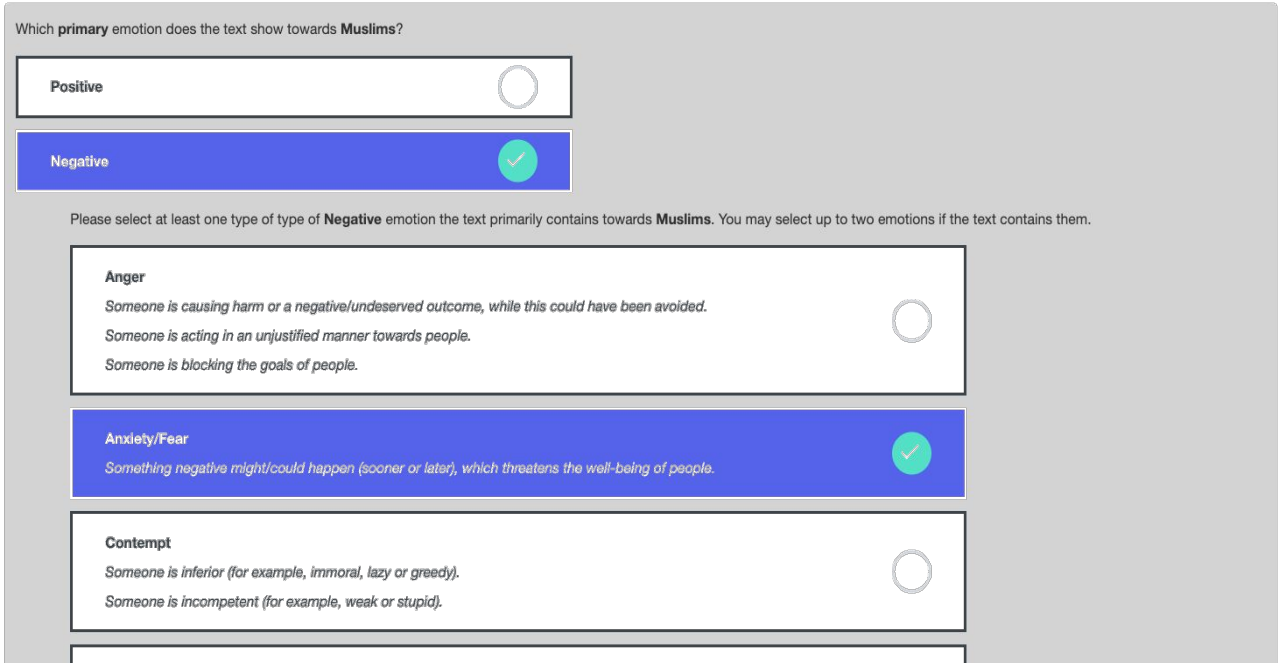
What kind of language does this comment contain towards **Muslims**?

- **Discriminatory or Alienating** ✓
- **Critical but not Discriminatory**
- **Supportive or Favorable**
- **Neutral**

The previous comment implies that all who follow **Islam** threaten European civilization, even suggesting that Nazis would have prevented that. Therefore, the answer is **Discriminatory or Alienating**

 Huguet Cabot et al. - Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions - EACL (2021)

Include Emotions



 Huguet Cabot et al. - Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions - EACL (2021)

Crowdsource annotation

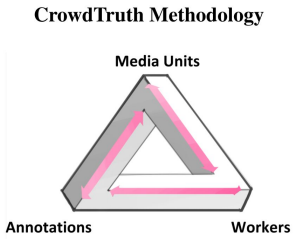


Figure 1: Triangle of Disagreement

We used **CrowdTruth 2.0**, a set of metrics computed recursively to account for disagreement:

Media Unit Quality Score (UQS): expresses the overall worker agreement over one media unit.

Worker Quality Score (WQS): the overall agreement of one crowd worker with the other workers.

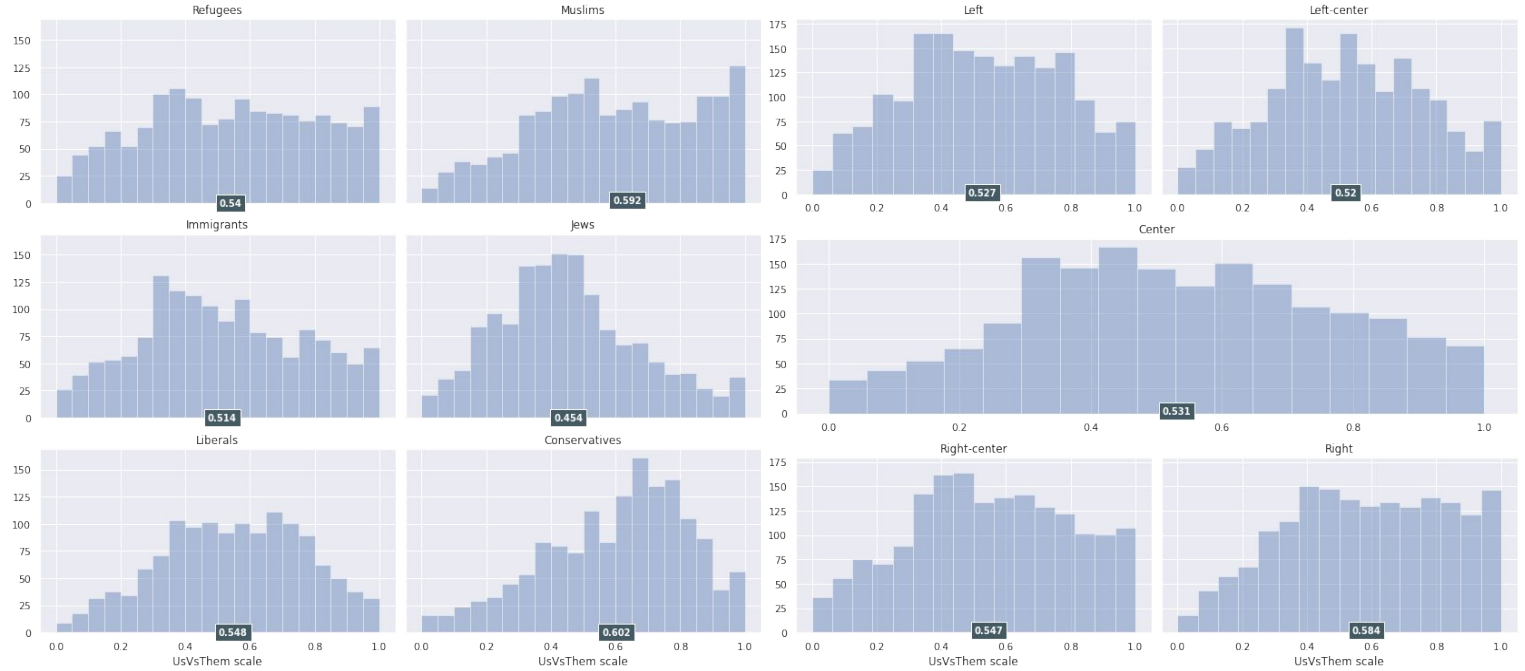
Annotation Quality Score (AQS): the agreement over an annotation in all media units that it appears

$$UQS(u) = \frac{\sum_{i,j} WorkVecWcos(i,j,u)\ WQS(i)\ WQS(j)}{\sum_{i,j} WQS(i)\ WQS(j)}, \qquad WQS(i) = WUA(i)\ WWA(i). \qquad AQS(a) = \frac{\sum_{i,j} WQS(i)\ WQS(j)\ P_a(i|j)}{\sum_{i,j} WQS(i)\ WQS(j)},$$

$$\forall i, j \text{ workers}, i \neq j.$$

 https://github.com/CrowdTruth/CrowdTruth-core

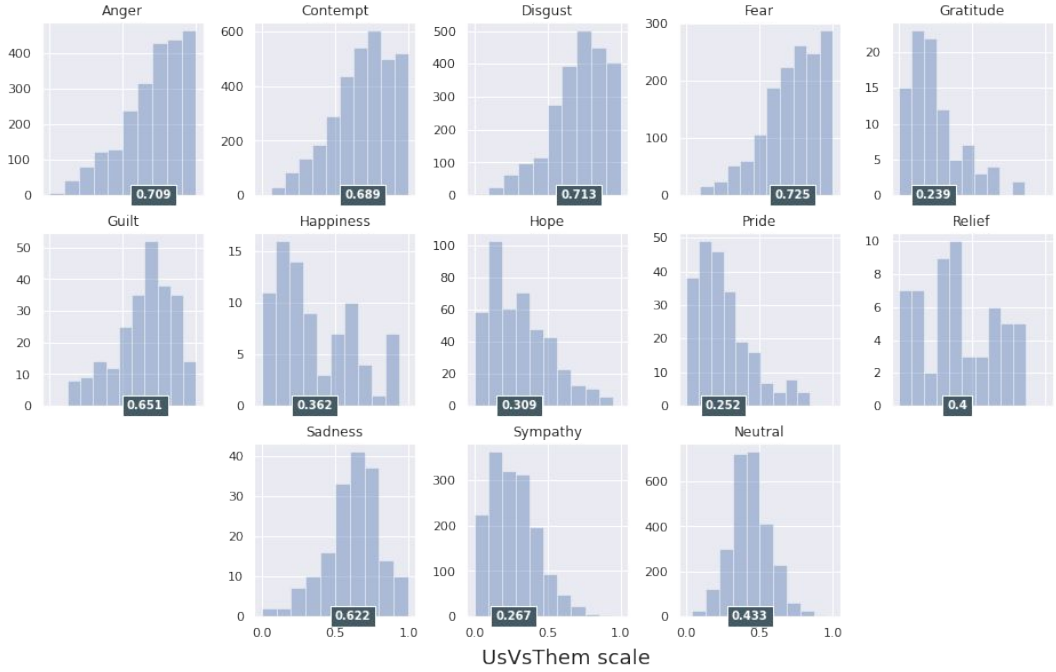
UsVsThem scale

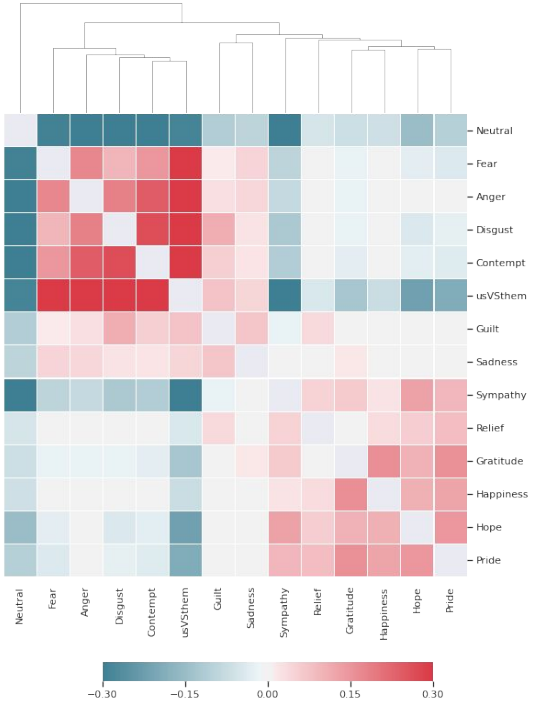


 Huguet Cabot et al. - Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions - EACL (2021)

SAPIENZA NLP

Emotions





 Huguet Cabot et al. - Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions - EACL (2021)

SAPIENZA NLP

Tasks

- **Regression** on the UsVsThem scale: Mean Squared Error loss.
- **Classification** between Negative (Discriminatory and Critical) or not (Neutral and Supportive): Binary Cross-entropy loss.

Auxiliary tasks:

- Emotion detection: Binary Cross-entropy loss (multiclass).
- Group identification: Cross-entropy loss.

Model:

Hard-shared RoBERTa with last layer (12) specific for each task.

Results

	STL	MTL, Emotion	MTL, Group	MTL, Emotion & Group
Pearson R	0.545 ± 0.005	$\mathbf{0.553 \pm 0.009}$	$\mathbf{0.557 \pm 0.012}$	$\underline{\mathbf{0.570 \pm 0.009}}$
Accuracy	0.705 ± 0.006	0.710 ± 0.009	0.711 ± 0.007	$\mathbf{0.717 \pm 0.004}$

Table 1: Results for the *Us vs. Them* rhetoric as regression and classification tasks. Significance compared to STL is bolded ($p < 0.05$). Significance compared to two-task MTL is underlined ($p < 0.05$). Average over 10 seeds.

Analysis

Improvement:

- Comments with high emotion valence
- Group targeted language

Error:

- Comments with emotionally charged language
- Reference to multiple groups
- Annotation error

 Huguet Cabot et al. - Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions - EACL (2021)

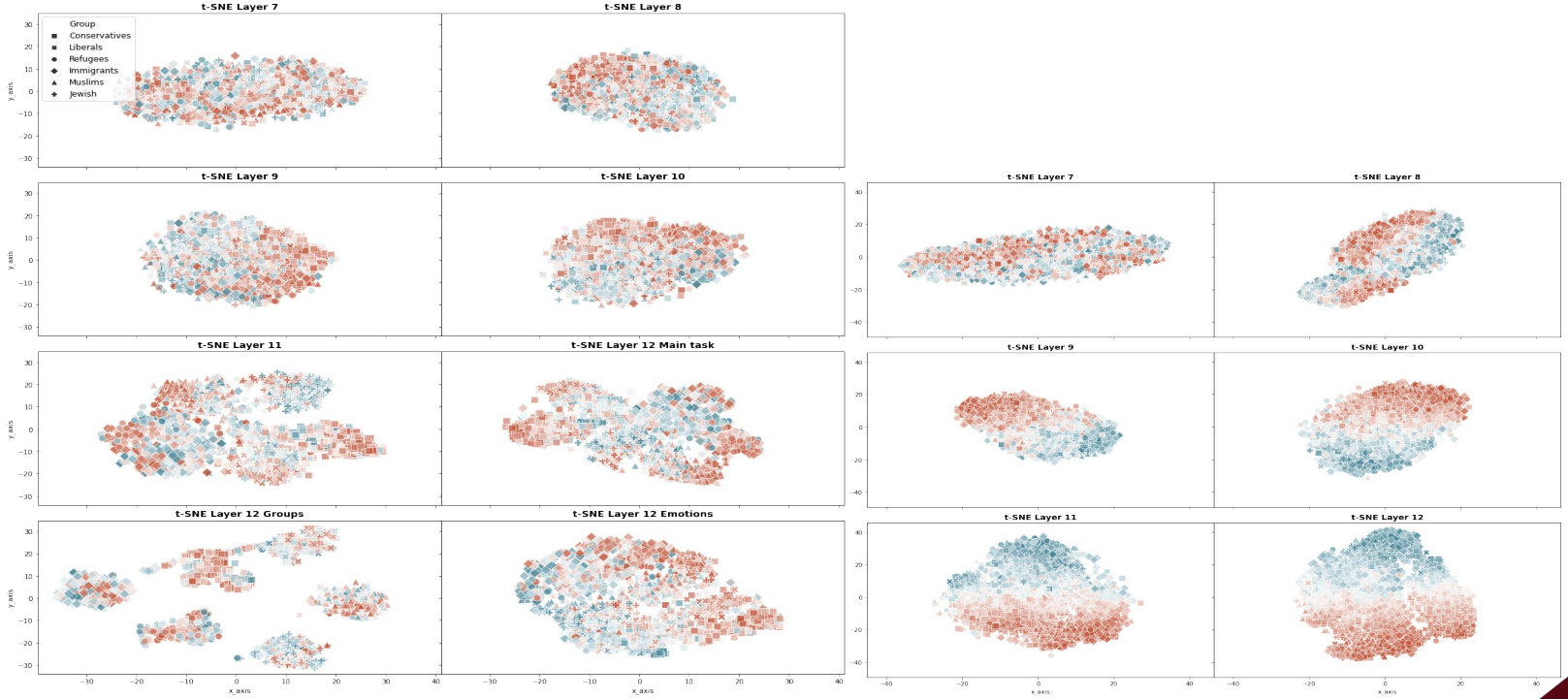
Examples

n	**Reddit Comment**
1	Does anyone else think it's absurd that we have Muslims serving in our congress? Literal foreign agents acting as if they hold the interests of the country and its people at heart? They never talk about the will of the people. It's always some bullshit about how white men (who founded and built this country) are evil and we need to let in more 3rd worlders who want to bomb and kill us. This is literal clown world tier nonsense.
2	You proud of yourselves, making 3 year olds represent themselves in immigration court? You fucking proud of that insanity? All for the sake of keeping out a gang that has already been in America for a long time, meanwhile regular home grown white kids are murdering dozens of their own classmates but goddam, at least they we're legal, amirite
3	Conservatives have every right to revolt. If we don't get our way we will destroy the country. I hope the left keeps pushing us to provoke a civil war. Or maybe Commiefornia should secede. Maybe that's the best thing that can happen, a complete break up. That way we can have our ethnostate, and the left can have their degenerate cesspool without us paying taxes for it. The US is dead anyway. It's time to burn this diverse shithole to the ground. It will be the ultimate proof that diversity doesn't work.

	Label	**MTL, E. & G.**	**MTL, Emo.**	**MTL, Groups**	**STL**	**Group**	**Emotions**
1	1.000	0.872	0.870	0.847	0.759	Muslims	Anger, Contempt, Disgust & Fear
2	0.02	0.774	0.874	0.740	0.834	Immigrants	Sympathy
3	0.071	0.729	0.773	0.747	0.8	Conservatives	Hope & Pride

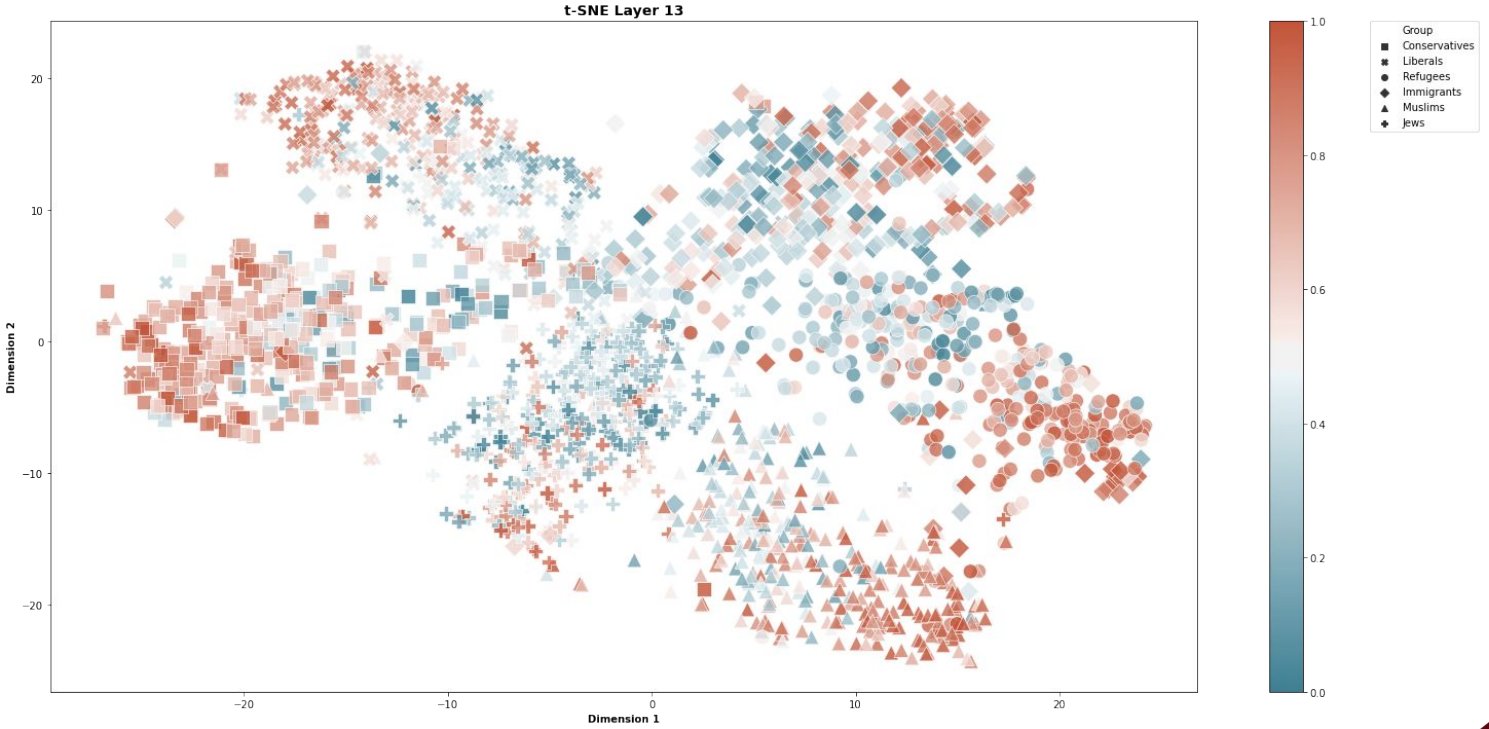
Table 2: Examples of predictions for comments. Predictions are averages over 10 seeds for each model.

Analysis of model representations



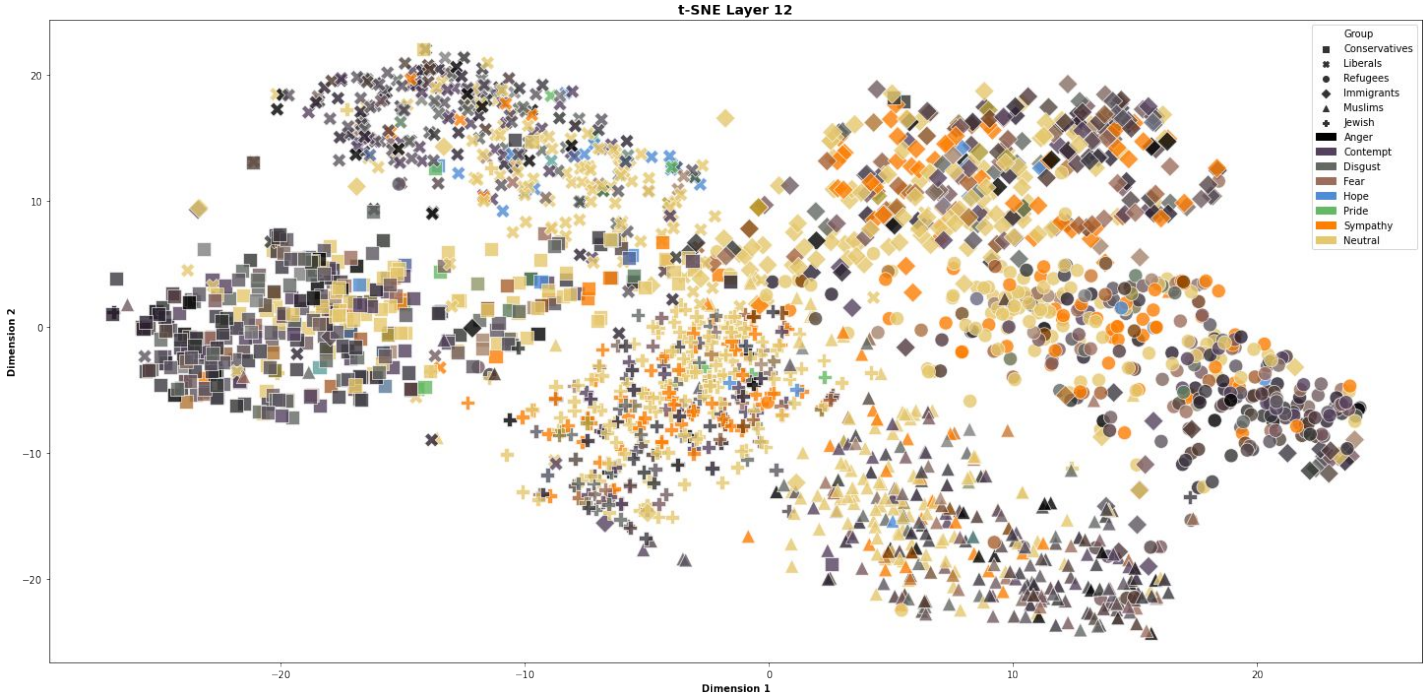
 Huguet Cabot et al. - Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions - EACL (2021)

Analysis of model representations



 Huguet Cabot et al. - Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions - EACL (2021)

Analysis of model representations



 Huguet Cabot et al. - Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions - EACL (2021)

Thank you for your attention!

SAPIENZA
UNIVERSITÀ DI ROMA