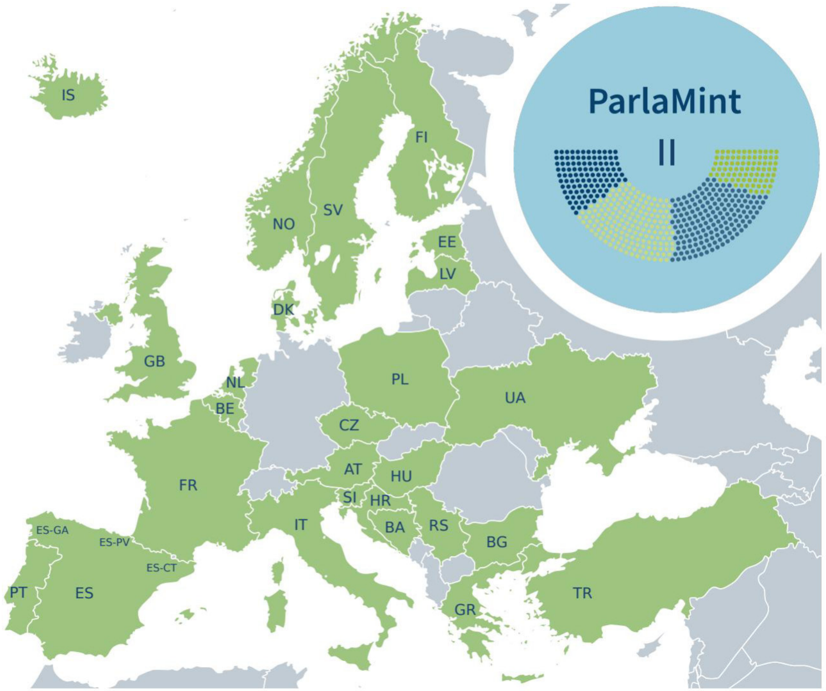
Multilingual Ideology and Power Identification in Parliamentary Debates

Touché’25 Task 2



Çağrı Çöltekin

Katja Meden

Nikola Ljubešić

Tomaž Erjavec

Vaidas Morkevičius

Matyáš Kopp

Multilingual Ideology and Power Identification in Parliamentary Debates

Introduction

- Parliamentary debates result in decisions with high societal impact

- Political/parliamentary language is difficult to analyze
  - highly conventionalized
  - strategies like evasion, circumlocution or the use of metaphors are common

- This task is about identifying three fundamental aspects in political discourse
  - *Political orientation*: the ‘classic’ left–right spectrum
  - *Populism index*: another ‘popular’ dimension of recent political discourse
  - *Power role*: central in discourse analysis, virtually no computational studies

Multilingual Ideology and Power Identification in Parliamentary Debates

Task Description

Scenario: Identify the political orientation and the power role of the speaker from their speeches in parliamentary debates.

Task: Given a transcribed speech delivered in a parliament

Subtask 1: identify political orientation of the speaker (left–right)

Subtask 2: identify the position of the speaker's party in populsit–pluralist scale (4 values)

Subtask 3: identify power role of the speaker (coalition–opposition)

Data:
- A subset of the ParlaMint version 4.1
- 29 national and regional parliaments (some available only for one of the tasks)
- 30 languages (also automatic translation to English)
- Date range varies by parliament, but includes at least from 2015 to 2022
- Typically long texts (approx. 600 words on average)

Multilingual Ideology and Power Identification in Parliamentary Debates

Results - orientation

Rank	Team	Approach	Precision	Recall	F_1-score
1	Munibuc	SVM + NV-Embed-v2	0.680	0.665	0.660
2	GIL_UNAM_Iztacala	SVM/RF/LR/NB + n-grams	0.664	0.655	0.652
3	TüNLP	XLM-RoBERTa	0.684	0.660	0.648
	Baseline	Logistic Regression + Char n-grams	0.661	0.597	0.570
Only on GB					
1	Munibuc	SVM + NV-Embed-v2	0.826	0.828	0.827
2	GIL_UNAM_Iztacala	SVM/RF/LR/NB + n-grams	0.801	0.802	0.801
3	TüNLP	XLM-RoBERTa	0.805	0.802	0.797
	Baseline	Logistic Regression + Char n-grams	0.770	0.771	0.770
4	DEMA^{2}IN	Event Extraction + Logistic Regression	0.727	0.724	0.719

Multilingual Ideology and Power Identification in Parliamentary Debates

Results - populsim

Rank	Team	Approach	Precision	Recall	F_1-score
1	GIL_UNAM_Iztacala	SVM/RF/LR/NB + n-grams	0.533	0.522	0.512
2	Munibuc	SVM + NV-Embed-v2	0.559	0.496	0.497
	Baseline	Logistic Regression + Char n-grams	0.571	0.442	0.419
Only on GB					
1	Munibuc	SVM + NV-Embed-v2	0.710	0.573	0.593
2	GIL_UNAM_Iztacala	SVM/RF/LR/NB + n-grams	0.570	0.565	0.565
3	DEMA^{2}IN	Event Extraction + Logistic Regression	0.560	0.556	0.558
	Baseline	Logistic Regression + Char n-grams	0.717	0.517	0.501

Multilingual Ideology and Power Identification in Parliamentary Debates

Results - populsim

Rank	Team	Approach	Precision	Recall	F_1-score
1	GIL_UNAM_Iztacala	SVM/RF/LR/NB + n-grams	0.709	0.707	0.703
	Baseline	Logistic Regression + Char n-grams	0.708	0.637	0.626
Only on GB					
1	GIL_UNAM_Iztacala	SVM/RF/LR/NB + n-grams	0.801	0.788	0.729
	Baseline	Logistic Regression + Char n-grams	0.784	0.762	0.765
2	DEMA^{2}IN	Event Extraction + Logistic Regression	0.737	0.727	0.729

Multilingual Ideology and Power Identification in Parliamentary Debates

Results: observations

- Similar approaches to last year (with slightly reduced participant nunbers)
- Many teams used ‘traditional’ ML methods and (large) language models to extract features
  - likely the due to cost of processing long texts
- Finetuning a single multilingual model also seems promising
- Focused participation based on event extraction from one of the teams (DEMA^{2}IN)
- Populism identification proves to be most difficult
- Scores on English are much better than the average performance