Supplementary Materials

Natural Language Processing Tasks

In our analysis we utilize different versions of BERT to perform labelling and classification tasks. BERT models are based on the transformer architecture that employs self-attention mechanisms to assess the importance of various words in a sentence, and they are called foundational models in natural language processing due to their ability to handle text efficiently and capture contextual relationships over long text sequences. We access these pre-trained models through the Hugging Face Transformers library, incorporating them into our dataset. We detail the models used, starting from sentiment analysis, and finally to zero-shot classification concerning immigration, multiculturalism, and Islam. We look at them in this order as they go in ascending complexity.

In the context of sentiment classification, smaller models like DistilBERT often perform equally well as larger models, providing a practical and effective solution. The "uncased" nature of the model indicates that it has been trained on text in lowercase format, treating a "Word" and a "word" as equivalent. This simplification enhances the model's robustness in handling case variations in input text, ensuring consistency in processing. DistilBERT is a streamlined and more efficient version of the BERT architecture. It achieves this efficiency through a process known as knowledge distillation, where the smaller DistilBERT model is trained to mimic the behavior of the larger BERT model. This makes DistilBERT capable of maintaining similar performance levels while using fewer resources, with approximately 66 million parameters and just 6 layers. The existing fine-tuning of the model we use, called distilbert-base-uncased-finetuned-tweets-sentiment, allows it to adapt its general language understanding to the specific linguistic nuances and styles commonly found in tweets, which may include slang, abbreviations, and other informal language constructs.

For classifying texts related to multiculturalism and its related dimensions we use a large variant of the XLM-RoBERTa (XLM-R) model that is fine-tuned for zero-shot classification called xlm-roberta-large-xnli. XLM-R builds upon the robustness of RoBERTa for cross-lingual language tasks, enabling it to handle content in multiple languages, including French. The large variant of XLM-R that we use is among the most substantial configurations in the RoBERTa family, featuring more layers and approximately 550 million parameters. This complexity helps the model capture intricate patterns within the data, which is essential as there is no smaller version available for this specific task.

RoBERTa, which stands for a Robustly Optimized BERT Pretraining Approach, is a refined version of BERT that achieves superior performance by utilizing larger training datasets, extended training times, and adjustments to some of the key hyperparameters of BERT. Unlike BERT, RoBERTa does not incorporate the Next Sentence Prediction (NSP) task and instead focuses solely on masked language modelling (MLM). This adjustment can result in improved performance on language understanding tasks, particularly when dealing with shorter texts.

The model's fine-tuning incorporates the CommonCrawl (CC-100) dataset, which is derived from extensive web crawls and represents one of the largest and most diverse multilingual text datasets available. The CC-100 dataset encompasses text from various sources, including tweets, facilitating the model's ability to learn comprehensive language patterns that apply across different linguistic styles and vocabularies. Despite the proprietary nature and controversies surrounding the web crawl process, the CC-100 dataset comprises a vast collection of internet data since 2008, amounting to multiple petabytes. This comprehensive data allows the model to understand a wide range of texts, regardless of their length or type.

Table A. List of political parties includes in the analyses

Name	Translation	Acronym
Rassemblement National	National Rally	FN|RN (since 2018)
Debout la France	Debout la France	DLR|DLF

Les Verts	The Greens	V
La France Insoumise	France Unbowed	FI
Les Républicains	The Republicans	UMP|LR
Mouvement démocrate	Democratic Movement	UDF|MD
Parti socialiste	Socialist Party	PS
Parti communiste française	French Communist Party	PCF
Renaissance	Renaissance	REM|R
Union des démocrates et indépendants	Union of Democrats and Independents	UDI

43
44 **Table B.** Hypothesis for each category for NLI task.

Category	Hypothesis	Total Count of Occurrences
Secularism	The text contains any reference to secularism or laïcité.	9957
Charlie Hebdo	The text either explicitly or implicitly refers to the Charlie Hebdo terrorist attacks.	2610
Multiculturalism	The text contains any reference to multiculturalism.	1458
Islam	The text contains any reference to Islam or Islam culture.	1286
Jihadism	The text either refers to or mentions jihadism.	1195
Headscarf	The text either refers to or mentions hijab or other forms of head covering.	494
Assimilation	The text contains any negative reference to assimilation.	148
Diversity	The text contains any negative reference to diversity.	68
Immigration	The text contains any reference to immigration.	33
Muslim Culture	The text either refers to or mentions Muslim people or their culture.	25
Bataclan	The text either refers to or mentions the Bataclan terrorist attacks.	9

45
46 **Table C. Relevant events and time frames for the analysis.**

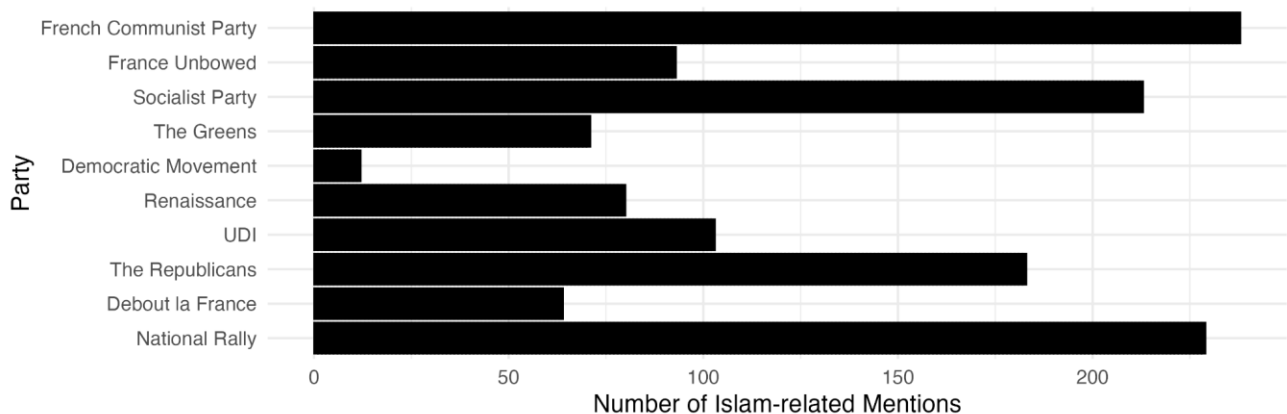| Event	Date	Pre-period Start	Post-period End
Charlie Hebdo Attacks	January 7, 2015	March 7, 2014	November 7, 2015
Bataclan Attacks	November 13, 2015	January 13, 2015	September 13, 2016
Nice Attacks	July 14, 2016	September 14, 2015	May 14, 2017
2017 Presidential Elections	May 7, 2017	July 7, 2016	March 7, 2018

Notre-Dame Cathedral Fire	April 15, 2019	June 15, 2018	February 15, 2020
Murder of Samuel Paty	October 16, 2020	December 16, 2019	August 16, 2021

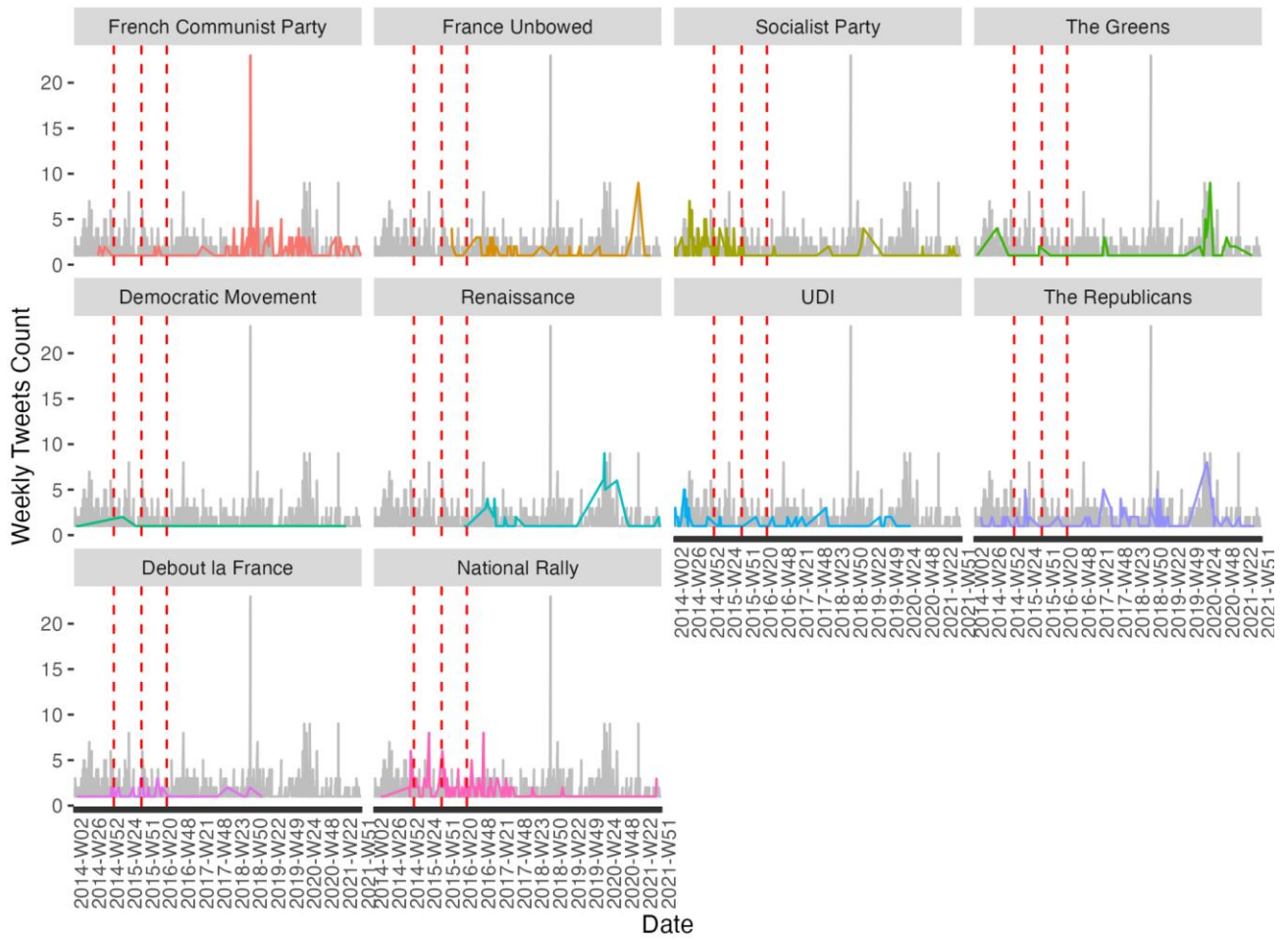**Table D.** Unique number of tweets per party and year coverage for each party.

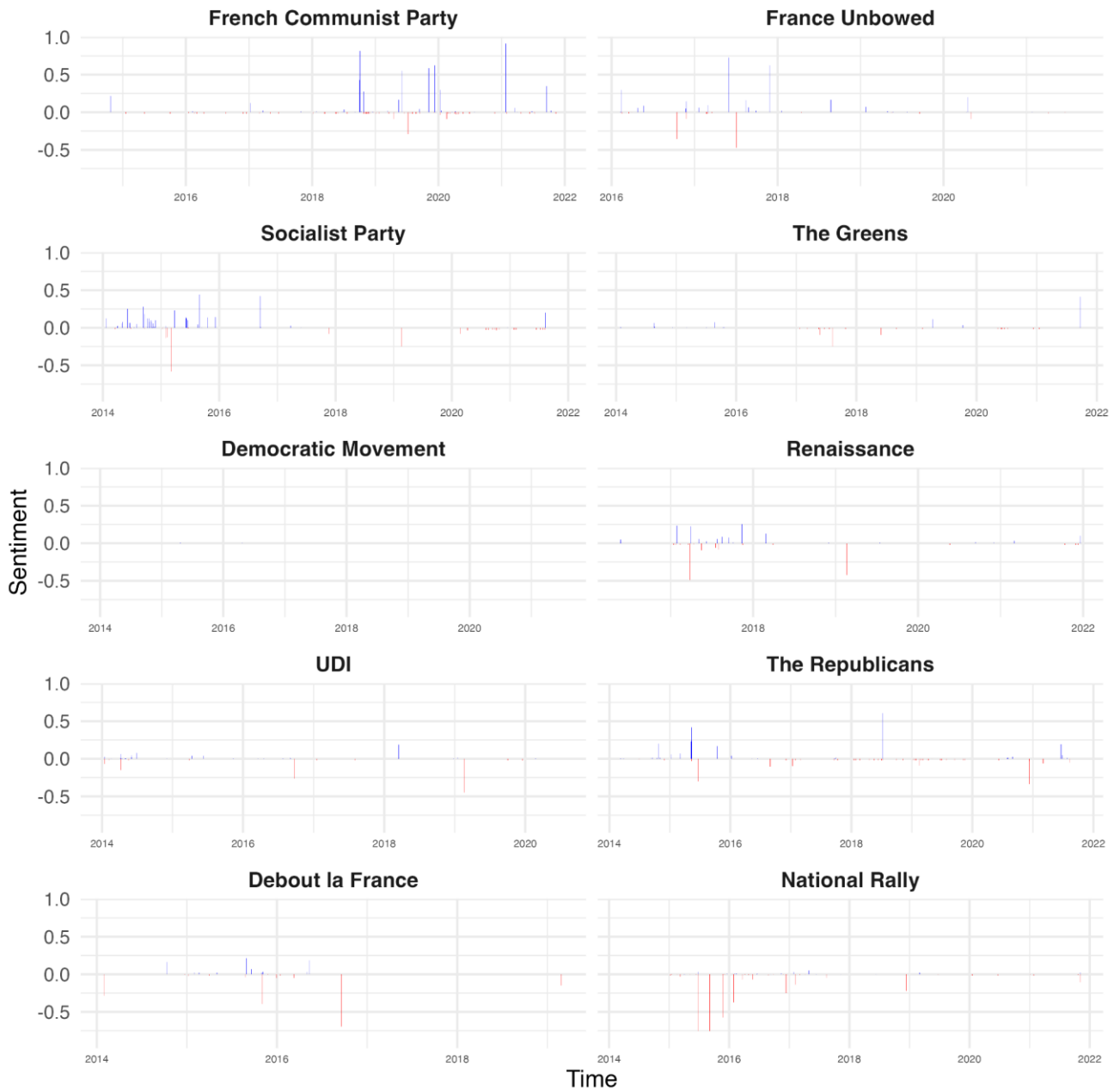| Party	Twitter Handle	French Name	Classification	Year Coverage	Number of Tweets
French Communist Party	pcf	Parti communiste française	Far Left	2014 - 2021	13483
Greens	eelv	Les Verts	Left	2014 - 2021	5358
National Rally	RNational_off	Rassemblement National	Far right	2014 - 2021	27145
Republic Arise	France Arise	dlf_officiel	Debout la république	Debout la France	Far right	2014 - 2021	19554
Socialist Party	partisocialiste	Parti socialiste	Left mainstream	2014 - 2021	31197
The Republic Onwards!	Renaissance	renaissance	La République En Marche!	Renaissance	Centrist mainstream	2016 - 2021	5727
Unbowed France	franceinsoumise	La France Insoumise	Left	2016 - 2021	14045
Union for French Democracy	Democratic Movement	modem	Union pour la démocratie française	Mouvement démocrate	Centrist mainstream	2014 - 2021	5672
Union for a Popular Movement	The Republicans	lesrepublicains	Union pour un mouvement populaire	Les Républicains	Right Wing mainstream	2014 - 2021	11431
Union of Democrats and Independents	udi_off	Union des démocrates et indépendants	Right Wing	2014 - 2021	10075

Figure 1. Total mentions of issues related to Islam by French political parties, 2014-2021. Source: Project ANONYMIZED.

Figure 2. Salience of Islam in tweets in France, by party. Source: Project ANONYMIZED.

55

56   **Figure 3.** Sentiment analysis for the topic *Islam*. Source: Project ANONYMIZED.