PoliticIT at EVALITA 2023: Overview of the Political Ideology Detection in Italian Texts Task

Daniel Russo[1,2], Salud María Jiménez-Zafra[3], José Antonio García-Díaz[4], Tommaso Caselli[5], Marco Guerini[1], L. Alfonso Ureña-López[3] and Rafael Valencia-García[4]

[1]LanD, Fondazione Bruno Kessler, Via Sommarive 18, Povo, Trento, Italy
[2]University of Trento, Italy
[3]SINAI, Universidad de Jaén, Spain
[4]UMUTeam, Universidad de Murcia, Spain
[5]CLCG, University of Groningen, Netherlands

Abstract

This paper presents the PoliticIT 2023 shared task, organised at EVALITA 2023 workshop. The task aims to extract politicians' ideology information from a set of tweets in Italian framed as a binary and a multiclass classification. The task is designed to be privacy-preserving and it is accompanied by a subtask targeting the identification of self-assigned gender as a demographic trait. The PoliticIT task attracted 7 teams that registered for the task, submitted results and presented working notes describing their systems. Most of the teams proposed transformer-based approaches, while some of them also used traditional machine learning algorithms or even a combination of both.

Keywords

Author profiling, Political ideology, Author analysis, Demographic and psychographic traits

1. Introduction and Motivations

The study of the political discourse on Social Media Platforms is of paramount importance in order to understand where society is heading. Political discourse is by definition ideologically based and political ideologies are spread with discourse [1]: for this reason, the analysis of the latter cannot go without the understanding of the former.

Political ideology is defined as a psychographic trait that can help comprehend both individual and social behaviour, including moral and ethical values, attitudes, biases, and prejudices [2]. In fact, this trait helps understanding how individuals think that society should be organised and has a strong relationship with personality traits as demonstrated in [3]. For instance, they found that conscientiousness is strongly correlated with the right wing, whereas openness to experience and agreeability were notably more correlated to the left wing. Moreover, political ideology has a great influence in the daily lives of each citizen. For example, [4] found a correlation between political ideology and the attitude of citizens to vaccination campaigns. Still, citizens react to the political messages they are exposed to. Therefore studying how politicians spread their ideology using social media discourses is useful to better analyse the policies and perspectives that are proposed on how society should be organized and work.

In this scenario, the PoliticIT shared task organized at EVALITA 2023 [5] aims to extract political ideology information from texts. To this end, an author profiling task is proposed. In particular, the PoliticIT task focuses on the identification of the political leaning and the self-assigned gender of the author from a binary and multiclass perspective.

In recent years, several shared tasks for author analysis have been organized under the PAN workshop series [6], targeting author attribution, bot detection, gender detection, and author obfuscation, among others. Other initiatives, such as the PoliticES shared task [7], have focused on capturing other traits such as the political ideology expressed in a message. PoliticIT is a twin-task of PoliticES and aims at analysing political ideology while being privacy-preserving. For this reason, a novel methodology of text clustering, that creates "virtual users" has been used on top of traditional anonymisation procedures.

The rest of the paper is organized as follows. Section

EVALITA 2023: 8th Evaluation Campaign of Natural Language Processing and Speech Tools for Italian, Sep 7 – 8, Parma, IT

drusso@fbk.eu (D. Russo); sjzafra@ujaen.es (S. M. Jiménez-Zafra); joseantonio.garcia8@um.es (J. A. García-Díaz); t.caselli@rug.nl (T. Caselli); guerini@fbk.eu (M. Guerini); laurena@ujaen.es (L. A. Ureña-López); valencia@um.es (R. Valencia-García)
https://www.rug.nl/staff/t.caselli (T. Caselli); https://www.marcoguerini.eu (M. Guerini)
0009-0006-9123-5316 (D. Russo); 0000-0003-3274-8825 (S. M. Jiménez-Zafra); 0000-0002-3651-2660 (J. A. García-Díaz); 0000-0003-2936-0256 (T. Caselli); 0000-0003-1582-6617 (M. Guerini); 0000-0001-7540-4059 (L. A. Ureña-López); 0000-0003-2457-1791 (R. Valencia-García)

CEUR Workshop Proceedings (CEUR-WS.org)

2 describes the PoliticIT shared task. Section 3 presents the dataset provided in the competition. Section 4 summarises the participant approaches. Section 5 shows the results and a discussion thereof. Finally, Section 6 concludes the paper with a discussion of the most interesting outcomes of this task and possible future works.

2. Task Description

The PoliticIT task is structured along three subtasks:

- **Subtask A** - *Self-assigned Gender*: Given a message, the system must assign a value for the gender of the author. The set of labels has been determined according to the personal web pages of the politicians of the Italian Parliament. The task has been framed as a binary classification task with **M** for men and **F** for women.
- **Subtask B** - *Political Ideology (binary)*: systems are required to determine the political orientation of a message; the binary version of the task presents two macro-categories: **Left** and **Right**.
- **Subtask C** - *Political Ideology (multi-class)*: this subtask presents a more fine-grained set of labels for the political orientation expressed by a given message. In this case, we employed four labels: **Left**, **Moderate-Left**, **Moderate-Right**, and **Right**.

PoliticIT was organized through CodaLab.[1] The run of the task is divided into three phases: (i) Practice, (ii) Evaluation, and (iii) Post-evaluation. In the Practice phase, the participants were initially provided with a subset of the training data in order to familiarise themselves with the training data format. During this phase, we also provided a notebook comprising the code for our Logistic Regression baseline, as a starting point for the development of more efficient systems. The full training set was released in February 2023. Currently, the task is in its post-evaluation phase, where participation is publicly open to other teams and research groups from the community.

3. Datasets and Format

This section provides the reader with an overview of the dataset proposed for the PoliticIT 2023 shared task along with a comprehensive description of the modalities employed for creating it.

[1]https://codalab.lisn.upsaclay.fr/competitions/8507

3.1. Data Collection

The dataset was collected from the Twitter timelines[2] of Italian politicians using the UMUCorpusClassifier [8], following a strategy similar to the one adopted in PoliticES 2022 [7] and in [9]. In particular, the data refer to the politicians from the legislature XIX of the Italian Republic. The list of deputies, senators, and ministers was taken from the institutional websites of the Italian Parliament[3] and Government.[4] All the politicians' Twitter accounts were manually retrieved, as they are not reported on the institutional websites. We discarded politicians that did not have a Twitter account or that were highly inactive on this social media (i.e. whose accounts present very few or old tweets). The time window for the corpus compilation was December 2022 as the oldest date, but no start date was set. In the first iteration, we compiled 371,822 tweets from 468 politicians between November 2010 and December 2022. The average number of tweets per politician is 794.49 but with a large standard deviation of 847.12, which suggest that not all politicians are equally active on Twitter. Thus, we decided to remove from the dataset those politicians with less than 25 tweets, leaving a total of 408 politicians.

To balance the number of tweets per politician, we first removed those tweets that are not written in Italian. To detect the language, we employed FastText language identification model [10]. Secondly, we removed the documents that shared content from news websites without retweeting. To do this, we discarded tweets that contained mentions of news websites, by detecting linguistic clues within the text, such as the pipe symbol, which is commonly employed by news websites to categorise their content. Thirdly, we selected tweets based on topics. An initial list of topics was extracted with BERTopic [11], a topic modelling technique for the creation of interpretable clusters based on Transformers and c-TF-IDF. In particular, we leveraged the Italian BERT model from [12]. We obtained a list of topics organised into 21 categories. This list was manually checked to introduce additional keywords for categories such as *European Union*, *immigration*, *energy*, *feminism*, *sports*, *mafia* or *religion*. Next, we identified which topics appeared in each tweet and prioritised those tweets that contained at least one topic. We then selected the tweets according to their topic in order to avoid any possible bias in the dataset.

[2]https://developer.twitter.com/en/docs/twitter-api/v1/tweets/timelines/api-reference/get-statuses-home_timeline

[3]Chamber of Deputies: https://www.camera.it/leg19/28
Senate of the Republic: https://www.senato.it/leg/19/BGT/Schede/Attsen/Sena.html

[4]https://www.governo.it/it/ministri-e-sottosegretari

3.2. Data Annotation and Anonymization

We enriched the dataset by assigning to each politician a label indicating their political ideology. Political ideologies have been directly derived from the politicians' affiliation party. In particular, the mapping from the politician to the political ideology was obtained through a two-step procedure:

1. Automatic labelling of politicians with their current political party affiliation. The party affiliation has been inferred from the parliamentary group to which the parliament party belongs. The data were extracted from the Italian institutional websites on October 31, 2023, thus they do not reflect changes in parliamentary groups following this date.
2. Mapping of the political parties to specific political ideology labels. The set of labels has been identified using Wikipedia.[5] We used four political ideology labels, i.e. *left*, *moderate left*, *right*, *moderate right*. Parties that are mapped in the centre, or cross-party, were nevertheless assigned one of the four aforementioned labels on the basis of their political alliances and the programme they presented during the 2022 Italian election campaign. The decision to "force" this classification was made to avoid excessive imbalance within each class. Therefore, we labelled "Movimento 5 Stelle" as *left*, whereas "Azione" and "Italia Viva" as *moderate left*.

Gender labels were assigned through three different approaches, depending on the source of the data. for the Italian deputies, gender was directly extracted from the institutional website, which allows the filtering of members according to this trait. The website of the Senate of the Republic does not clearly states the gender of the members. In this case, employed linguistic cues present on the personal page of each senator to infer the gender. Specifically, we looked at the Italian verb "nascere" in its past participle form as "nato" for the male label and "nata" for the female label. Finally, for the ministers, gender was manually assigned as the official Government website does not comprise this information and does not present helpful linguistic cues. In this case, ministers were labelled according to their biological sex.

Subsequently, to build a privacy-compliant approach we took a two-step procedure including *anonymization* and *clustering*:

- **Anonymized references in text** - References to politicians within Twitter mentions were anonymized by replacing them with the token `@user`. The rest of the in-text mentions were also replaced with the `@user` token. Consequently, the text traits cannot be guessed trivially by reading a politician's name and searching for personal information on the Internet. We also replaced the name of the political parties and of their Twitter accounts with the `POLITICAL_PARTY` token.
- **Clustering procedure** - Subsequently, we created clusters of texts by mixing some of the extracted tweets in order to prevent ethical and privacy issues related to author profiling on Twitter. All the clusters are composed of tweets written by different politicians that share the same traits under evaluation, i.e. political ideology and self-assigned gender. For this, we divide the politicians into training and test in order to prevent that tweets from the same politician from appearing both in training and validation. To generate a cluster, we first set their demographic and psychographic traits, and then randomly pick tweets from users that share these traits. Thereby, each cluster represents "virtual users", with their self-assigned gender (male, female) and political spectrum. For the latter, we labelled the data according to two axes: binary (left, right) and multiclass (left, moderate left, moderate right and right). At the end of this process, we obtained 1751 clusters with 80 tweets per cluster. It should be noted that the clusters from the training and test sets are independent to prevent machine learning approaches from identifying the authors rather than the demographic and psychographic traits.

3.3. Data Formats

The training and test sets are produced in a ratio of nearly 75%-25%. Table 1 presents a summary of the distribution of labels per subtask. In no case, the labels are evenly distributed. Male politicians are almost double the number of female politicians, and more than 200 politicians are from the left wing than from the right. As regards the multi-class ideology, moderate left and right are the most represented labels.

Ultimately, each entry of the PoliticIT dataset comprises four elements: a cluster id, the self-assigned gender label and the political ideology labels for binary and multiclass classification. The dataset is organised at tweet level. This means that each row represents one tweet. Each line also contains a `cluster id` to identify the cluster to which the tweet belongs, as well as the demographic and psychographic traits of the cluster. Examples are provided in Table 2. The full dataset, including the gold labels, is available on Codalab [6].

[5] https://it.wikipedia.org/wiki/Partiti_politici_italiani

[6] https://codalab.lisn.upsaclay.fr/competitions/8507

Table 1
PoliticIT: Data distribution per subclass for Train and Test splits.

Subtask	Label	Train	Test	Total
Subtask A - Gender	Male	810	318	1128
	Female	488	135	623
Subtask B - Ideology (binary)	Left	720	248	968
	Right	578	205	783
Subtask C - Ideology (multi-class)	Left	162	100	262
	Moderate left	558	148	706
	Moderate right	447	154	601
	Right	131	51	182

Table 2
Examples from our dataset. Each dataset entry comprises: the cluster id, the self-assigned gender label, and the political ideology labels, both binary (pib) and multiclass (pim).

cluster id	gender	pib	pim	tweet
72b...	female	left	left	Non possiamo permettere che gli imprenditori si ritrovino dal giorno alla notte non più competitivi e che le imprese vengano lasciate sole. E per questo come @user stiamo lavorando senza sosta.
8f5...	male	right	right	Se spacci, vai in carcere e ci resti!. Con [POLITICIAN] proposta #[POLITICAL_PARTY], presentata in Parlamento, città più sicure e più rispetto per le Forze dell'Ordine, che rischiano per arrestare i pusher e poi se li ritrovano per strada il giorno dopo.

4. Systems Overview

A total of seven teams participated in the PoliticIT task, with all teams involved in each of the subtasks. The majority of the participants represented academic institutions. An overview of the system approaches can be found in Table 3, while Section 4.1 provides further details on the systems proposed.

4.1. System Architectures

ExtremITA [13]. The team proposed two systems. The first system is based on Camoscio [14], the Italian version of the Standford Alpaca model [15], pre-trained to generate text as a response to users' instructions passed as input. The team performed further fine-tuning on triples <`task`, `input`, `output`>. More precisely, the model used the phrasal forms derived from the training data of all EVALITA 2023 challenges: the `task` is a linguistic description of the task to be solved, whereas the `input-output` pairs are task-specific. On the other hand, the second system is based on IT5 Transformer [16], for which fine-tuning was done on `input-output` pairs. More precisely, the model used the phrasal forms derived from the training data of all EVALITA 2023 challenges, where the `input-output` pair is task-specific.

INFOTEC-LaBD [17]. The team employed SVM classifiers with linear and nonlinear kernels. Specific attention was given to the data representation. Indeed, the authors employed low-dimensional projections to concisely represent the dataset and its associated labels. These vectors were used for training an SVM classifier.

INGEOTEC. The team implemented different configurations of Bag of Words (BoW) classifier in all the subtasks. In particular, for gender and political ideology multiclass classification, INGEOTEC employed a stack generalization approach leveraging three BoW classifiers: two BoW classifiers pre-trained on 5M Italian tweets, and a BoW classifier trained on PoliticIT training set. Instead, for the political ideology binary subtask, the team proposed a BoW classifier trained on the training set.

Teeeech. The team proposed three Transformer-based classifiers trained independently of each other. The authors did not specify which Transformer models were used.

Tübingen [18]. The team proposed two main approaches: an SVM-based approach and a Transformer-based approach. The former was the best-performing, i.e.,.e simple linear SVMs with sparse word/character n-gram features, trained separately for each task, only

Table 3
Overview of the participants' systems. For each team, we summarize the pre-trained language models employed, and the method(s) used for their runs.

	Language Model					Method					
Team	**dbmdz BERT Italian**	**AlBERTO**	**IT5**	**Camoscio**	**XLM-RoBERTA**	**SVM**	**BoW classifier**	**Fine-tuning**	**Low-dimensional projection**	**Prompting**	**Data augmentation**
extremITA			✓	✓						✓	
INFOTEC-LaBD						✓			✓		
INGEOTEC							✓				
Teeeech								✓			
Tübingen	✓				✓	✓		✓			✓
UMUTeam	✓							✓			
URJC_TEAM	✓	✓						✓			

using the data provided for the shared task. For the Transformer-based approach, the team experimented with Italian BERT monolingual model [12] and XLM-ROBERTA multilingual model [19] for the classification task, multi-task/multi-label learning, and a number of transfer learning approaches. As additional data they employed the Italian section of the ParlaMint corpus [20] and PoliticES data [7].

UMUTeam [21]. The team proposed a system based on the Italian BERT [12] fine-tuned on PoliticIT data. Demographic and psychographic traits were classified according to two strategies: (i) *mode*, which selects the most repeated label after classifying each user's tweet individually, and (2) *highest probability*, which selects the label with the highest probability.

URJC_TEAM [22]. The system developed by the team leverages three BERT models. Each model was pre-trained on different data according to the subtask it was employed. In particular, the model for gender classification model was trained using text data sourced from Wikipedia dump and the OPUS corpora collection [23]. The binary classification model used to identify political polarity was trained both on the OPUS corpora [24] and on the OSCAR corpus. Lastly, for the multilabel classification subtask, the AlBERTo language model [25] was employed.

5. Results and Discussion

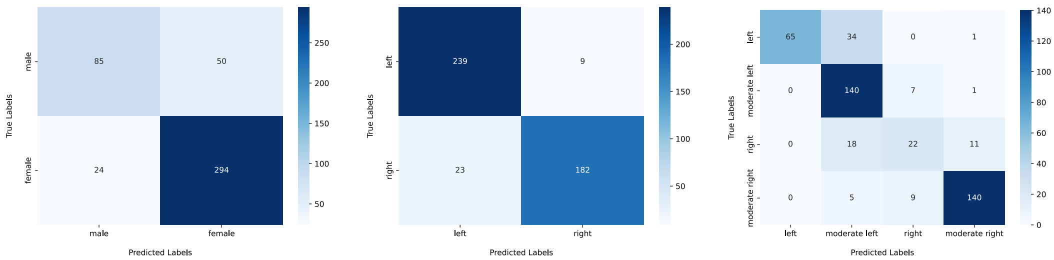
For each subtask, participants were allowed to make a maximum of 10 submissions in CodaLab during the Evaluation phase. For the ranking, participants were asked to select their best model. Systems were evaluated using standard evaluation metrics for text classification, namely Precision, Recall, and F1-score per class. The ranking is based on the arithmetic mean of the macro F1-scores of each subtask.

For all subtasks, we implemented the same baseline using TF–IDF based on word-unigrams, 50,000 as maximum features, minimum feature ratio of 10%, and applying lowercase. These features were trained using a Logistic Regression classifier. Table 4 summarises the leaderboard results and rankings.

Overall, all systems, with the sole exclusion of **URCJ_TEAM** on binary ideology classification, managed to outperform the baseline approach on all subtasks. We can notice that differences in performances across the top five systems are relatively large, ranging from a minimum of 0.024 points to a maximum of 0.071. This highlights a different trend when compared to other shared tasks where usually the top systems are all very close. On the other hand, a similar behaviour can be observed in PoliticES [7], where the range across the top five systems is 0.0.11 and 0.0827, suggesting that the task could present inherent challenges, especially for Subtask A (gender detection) and Subtask C (fine-grained political ideology classification).

Table 4
PoliticIT official leaderboard.

Team	Macro F1 (average)	F1-gender	F1-ideology (binary)	F1-ideology (multiclass)	Rank
Tübingen	0.824	0.792	0.928	0.751	1
INFOTEC-LaBD	0.800	0.824	0.860	0.717	2
extremITA	0.771	0.769	0.925	0.621	3
INGEOTEC	0.762	0.732	0.848	0.705	4
Teeeech	0.751	0.752	0.887	0.613	5
UMUTeam	0.704	0.712	0.866	0.533	6
URJC_TEAM	0.684	0.661	0.770	0.621	7
Baseline	0.569	0.536	0.809	0.362	–



(a) Gender (b) Ideology (binary) (c) Ideology (multiclass)

Figure 1: Confusion matrices for each subtask – self-assigned gender detection (1a), binary political ideology (1b) and multi-class political ideology classification (1c) – from the best performing system (team **Tübingen**).

Focusing on each subtask, it appears evident that binary classification of political ideologies is relatively easier when compared to the other two subtasks, with the best results being 0.928 obtained by **Tübingen**. The more fine-grained the distinction of political ideologies is the more challenging the task. This is not just an effect of having multiple classes and the distribution of the data, but it involves also the subtleties and nuances in policies across the four groups. In Figure 1 we display the confusion matrices obtained by the Tübingen team for the three classification subtasks. Focusing on the errors of Subtask C, multi-class ideology classification, we can notice that most of the errors concern misclassification of the "extremes" (i.e., **Left** and **Right**) into the **Moderate Left** category) rather than of the moderate positions (**Moderate Left** *vs.* **Moderate Right**). This indicates that - at least in their communication on Twitter - differences in the moderate political areas are stronger. A further noticeable result is the fact that messages from politicians on the **Right** spectrum tend to be assimilated mostly with the **Moderate Left** and **Moderate Right**, suggesting that the narratives of moderate groups tend to assimilate issues and expressions of the **Right**.

The identification of the gender of the author from a tweet is also quite challenging even if it is framed as a binary task. In this case, as illustrated by Figure 1a most of the errors affect the classification of male politicians into females. The best system, **INFOTEC-LaBD**, obtains 0.824 macro F1, with a positive Δ from the second best system (**Tübingen**) of 0.032 points.

6. Conclusions

In this paper, we have summarized the outcomes of the first edition of the PoliticIT task at EVALITA 2023. PoliticIT targets the identification of the political ideology and gender of the author of a tweet. Political ideology is a psychographic trait that can be used to understand individual and social behaviour, and thus contribute to a better understanding of the society. The task introduces an innovative method concerning the anonymization of users to preserve privacy, allowing the investigation of these sensitive topics in a fair and ethical way.

PoliticIT has seen the participation of seven teams, five of whom submitted a full report describing their approach. The results indicate that fine-grained political ideology distinction is more challenging than binary classification between two extreme values. This appears to be due to the absorption of the narratives of the extremes into the moderate positions, especially from the **Moderate Left**. This is a datum that could provide political scientists and sociologists additional insights into the evolution of the Italian political systems. A further aspect that raises interest concerns the errors in Subtask A, gender classification. In this case, most of the errors are male politicians classified as females. A deep exploration of the communication styles of the two genres and their correlation with political affiliations is a promising path to better understanding this behaviour.

As expected, approaches based on Transformers are the trend solutions presented by participating teams, but some of them also used feature-based linear machine learning systems. It is quite impressive that the best performing team, **Tübingen**, uses linear SVMs with word and character n-grams as features weighting each feature using TF-IDF. This indicates that simple methods are still competitive and far from being fully overcome by neural approaches.

PoliticIT has been fully run on CodaLab. The task is now in the Post-evaluation phase and it is still open to submissions from other teams, making this resource freely available to the NLP community for research purposes.

Future work will investigate the addition of an extra subtask related to stance detection, to determine which authors are in favor of certain topics and which users are against. We can use this information to define clusters of users and to observe whether there is a relationship between the topics and the political ideology.

Acknowledgments

This work is part of the research projects LaTe4PoliticES (PID2022-138099OB-I00) funded by MCIN/AEI/10.13039/501100011033 and the European Fund for Regional Development (FEDER)-a way to make Europe and LaTe4PSP (PID2019-107652RB-I00/AEI/ 10.13039/501100011033) funded by MCIN/AEI/10.13039/501100011033. This work is also part of the research projects AIInFunds (PDC2021-121112-I00) and LT-SWM (TED2021-131167B-I00) funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR. It also has been partially supported by Project CONSENSO (PID2021-122263OB-C21), Project MODERATES (TED2021-130145B-I00) and Project SocialTox (PDC2022-133146-C21) funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR, Project PRECOM (SUBV-00016) funded by the Ministry of Consumer Affairs of the Spanish Government, Project FedDAP (PID2020-116118GA-I00) supported by MICINN/AEI/10.13039/501100011033 and WeLee project (1380939, FEDER Andalucía 2014-2020) funded by the Andalusian Regional Government. Salud María Jiménez-Zafra has been partially supported by a grant from Fondo Social Europeo and the Administration of the Junta de Andalucía (DOC_01073).

References

[1] N. Fairclough, Critical discourse analysis: The critical study of language, Routledge, 2013.

[2] B. Verhulst, L. J. Eaves, P. K. Hatemi, Correlation not causation: The relationship between personality traits and political ideologies, American journal of political science 56 (2012) 34–51.

[3] M. Fatke, Personality traits and political ideology: A first global assessment, Political Psychology 38 (2017) 881–899.

[4] B. Baumgaertner, J. E. Carlisle, F. Justwan, The influence of political ideology and trust on willingness to vaccinate, PloS one 13 (2018) e0191728.

[5] M. Lai, S. Menini, M. Polignano, V. Russo, R. Sprugnoli, G. Venturi, EVALITA 2023: Overview of the 8th Evaluation Campaign of Natural Language Processing and Speech Tools for Italian, in: M. Lai, S. Menini, M. Polignano, V. Russo, R. Sprugnoli, G. Venturi (Eds.), Proceedings of the Eighth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2023), CEUR.org, Parma, Italy, 2023.

[6] J. Bevendorff, B. Chulvi, G. L. D. L. Peña Sarracén, M. Kestemont, E. Manjavacas, I. Markov, M. Mayerl, M. Potthast, F. Rangel, P. Rosso, et al., Overview of PAN 2021: authorship verification, profiling hate speech spreaders on twitter, and style change detection, in: International Conference of the Cross-Language Evaluation Forum for European Languages, Springer, 2021, pp. 419–431.

[7] J. A. García-Díaz, S. M. Jiménez-Zafra, M.-T. M. Valdivia, F. García-Sánchez, L. A. Ureña-López, R. Valencia-García, Overview of PoliticEs 2022: Spanish Author Profiling for Political Ideology, Procesamiento del Lenguaje Natural 69 (2022) 265–272.

[8] J. A. García-Díaz, Á. Almela, G. Alcaraz-Mármol, R. Valencia-García, Umucorpusclassifier: Compilation and evaluation of linguistic corpus for natural language processing tasks, Procesamiento del Lenguaje Natural 65 (2020) 139–142.

[9] J. A. García-Díaz, R. Colomo-Palacios, R. Valencia-García, Psychographic traits identification based on political ideology: An author analysis study on spanish politicians' tweets posted in 2020, Future Generation Computer Systems 130 (2022) 59–74.

[10] A. Joulin, É. Grave, P. Bojanowski, T. Mikolov, Bag of tricks for efficient text classification, in: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers, 2017, pp. 427–431.

[11] M. Grootendorst, Bertopic: Neural topic modeling with a class-based tf-idf procedure, arXiv preprint arXiv:2203.05794 (2022).
[12] S. Schweter, Italian bert and electra models, 2020. URL: https://doi.org/10.5281/zenodo.4263142. doi:10.5281/zenodo.4263142.
[13] C. D. Hromei, D. Croce, V. Basile, R. Basili, ExtremITA at EVALITA 2023: Multi-Task Sustainable Scaling to Large Language Models at its Extreme, EVALITA 2023 Eigth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (2023) –.
[14] A. Santilli, Camoscio: An italian instruction-tuned llama, https://github.com/teelinsan/camoscio, 2023.
[15] R. Taori, I. Gulrajani, T. Zhang, Y. Dubois, X. Li, C. Guestrin, P. Liang, T. B. Hashimoto, Stanford alpaca: An instruction-following llama model, https://github.com/tatsu-lab/stanford_alpaca, 2023.
[16] G. Sarti, M. Nissim, It5: Large-scale text-to-text pretraining for italian language understanding and generation, ArXiv preprint 2203.03759 (2022). URL: https://arxiv.org/abs/2203.03759.
[17] H. Cabrera-Pineda, E. S. Téllez, S. Miranda, INFOTEC-LaBD at PoliticIT: Political Ideology Detection in Italian Texts, EVALITA 2023 Eigth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (2023) –.
[18] C. Çöltekin, M. Brivio, F. Can, Tübingen at PoliticIT: Exploring SVMs, Pretrained Language Models, and Linguistic Transfer for Ideology Detection in Social Media, EVALITA 2023 Eigth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (2023) –.
[19] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, É. Grave, M. Ott, L. Zettlemoyer, V. Stoyanov, Unsupervised cross-lingual representation learning at scale, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 8440–8451.
[20] T. Erjavec, M. Ogrodniczuk, P. Osenova, N. Ljubešić, K. Simov, A. Pančur, M. Rudolf, M. Kopp, S. Barkarson, S. Steingrímsson, et al., The parlamint corpora of parliamentary proceedings, Language resources and evaluation 57 (2023) 415–448.
[21] R. Pan, Á. Almela, F. García-Sánchez, UMUTeam at PoliticIT-EVALITA2023: Evaluating Transformer Model for Detecting Political Ideology in Italian Texts, EVALITA 2023 Eigth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (2023) –.
[22] M. Á. Rodríguez-García, URJC-Team at EVALITA 2023: Political Ideology Detection in Italian Texts Using Transformers Architectures, EVALITA 2023 Eigth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (2023) –.
[23] J. Tiedemann, Parallel data, tools and interfaces in OPUS, in: Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12), European Language Resources Association (ELRA), Istanbul, Turkey, 2012, pp. 2214–2218. URL: http://www.lrec-conf.org/proceedings/lrec2012/pdf/463_Paper.pdf.
[24] P. J. Ortiz Suárez, B. Sagot, L. Romary, Asynchronous pipelines for processing huge corpora on medium to low resource infrastructures, Proceedings of the Workshop on Challenges in the Management of Large Corpora (CMLC-7) 2019. Cardiff, 22nd July 2019, Leibniz-Institut für Deutsche Sprache, Mannheim, 2019, pp. 9 – 16. URL: http://nbn-resolving.de/urn:nbn:de:bsz:mh39-90215. doi:10.14618/ids-pub-9021.
[25] M. Polignano, P. Basile, M. de Gemmis, G. Semeraro, V. Basile, AlBERTo: Italian BERT Language Understanding Model for NLP Challenging Tasks Based on Tweets, in: Proceedings of the Sixth Italian Conference on Computational Linguistics (CLiC-it 2019), volume 2481, CEUR, 2019. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85074851349&partnerID=40&md5=7abed946e06f76b3825ae5e294ffac14.