HALE Lab NITK at Touché 2024: A Hybrid Approach for Identifying Political Ideology and Power in Multilingual Parliamentary Speeches

Notebook for the Touché Lab at CLEF 2024

Sevitha Simhadri[1,*,†], Mauli Mehulkumar Patel[1,†] and Sowmya Kamath S.[1,†]

[1]*Department of Information Technology, National Institute of Technology Karnataka, Surathkal, Mangalore 575025, India*

Abstract

In this article, an approach to determine the political views and stances of speakers for identifying whether they support or oppose the government in parliamentary discussions is presented. The work was carried out as part of the Touché 2024 Task 2, "Ideology and Power Identification in Parliamentary Debates". Towards this, two systems were developed, the first employs traditional machine learning methods with TF-IDF embeddings, while the second utilizes advanced NLP techniques with the LASER encoder for multilingual embeddings. Both systems incorporate standard preprocessing techniques and also integrates a variety of models, after which a voting classifier is used to combine the predictions from both approaches. Experiments revealed that this comprehensive framework effectively addresses the complexities and nuances of political discourse, providing valuable insights into speakers' ideologies and governing statuses within parliamentary debates.

Keywords

Parliamentary Debates, Governing Status, Natural Language Processing, Multilingual Embeddings

1. Introduction

In recent years, analyzing parliamentary debates has become a crucial area of study in political science and natural language processing (NLP). Understanding speakers' political ideologies and governing statuses in these debates can offer profound insights into legislative processes and power dynamics. The Touché 2024 Task 2, "*Ideology and Power Identification in Parliamentary Debates*", addresses these analytical challenges, inviting participants to develop systems that accurately identify parliamentary speakers' political stances and leadership roles [1]. This task is significant because it advances the field of NLP and provides practical tools for political analysts and researchers. The Hale Lab team participated in this task to contribute to developing these analytical tools. By identifying the underlying ideologies and power structures in parliamentary debates, it is possible to understand better how political narratives are constructed and conveyed, thus offering a deeper understanding of the legislative process and political communication.

In this paper, we cover both tasks; first, we provide a high-level overview of the first task and then detail our strategy. The same strategy is used for the second task, which is described subsequently. Lastly, we will highlight our primary contributions and conclusions, along with suggestions for future work, to wrap up the paper.

The paper is organized as follows: Section 2 describes the related work in this area; Section 3 outlines the competition details; Section 4 provides an in-depth explanation of our approach; Section 5 discusses our main findings; and finally, Section 6 draws conclusions and suggests directions for future research.

CLEF 2024: Conference and Labs of the Evaluation Forum, September 09–12, 2024, Grenoble, France

*Corresponding author.

†These authors contributed equally.

munnysimhadri5544@gmail.com (S. Simhadri); maulipatel03@gmail.com (M. M. Patel); sowmyakamath@nitk.edu.in (S. K. S.)

0000-0002-0888-7238 (S. K. S.)

2. Literature Survey

Political viewpoint identification study encompasses a wide range of approaches and perspectives that are intended to help interpret the intricacies of political debate. In order to gain a clear understanding of who is powerful and influential in political systems without conducting any practical experiments, Abercrombie and Batista-Navarro [2] worked on the sentiment and position analysis of the parliamentary structure. In order to reduce ambiguity and make biases in textual data across the political spectrum more understandable, Doan and Gulla (2022) [3] developed bias learning techniques. In a similar vein, they developed a language model for Scandinavian languages[4] and tested it using political datasets. Preoţiuc Pietro et al. (2017)[5], in contrast, provided a multimodal approach that incorporates fine-grained ideology labels from surveys together with linguistic features including unigrams, LIWC, Word2Vec themes, and sentiment analysis.

Automatic political orientation prediction using social media postings has been studied by Pietro et al. (2017), and it has shown to be rather successful in differentiating between openly avowed liberals and conservatives in the US. Through the usage of language on Twitter, they sought to identify user groups that were politically involved and to develop an improved model that could predict the political ideology of users who are not visible.

Multi-task learning(MTL) was investigated by Barnes et al. (2019)[6]as a means of integrating external knowledge into neural networks for sentiment analysis. A straightforward method for identifying ideological learnings in documents based on sentiment expressions toward various topics was presented by Bhatia and P (2018)[7].

The study conducted by Ahmadalinezhad and Makrehchi (2018)[8] centered on identifying points of agreement and disagreement in political discourse. The importance of identifying agreement and disagreement in political speech is emphasized in their abstract, which also presents their work as a contribution to the social and cultural modeling field.

3. Overview of Tasks and Dataset

3.1. Task Definition

The task consists of two sub-tasks on identifying two important aspects of a speaker in parliamentary debates (a) *Sub-Task 1*: Given a parliamentary speech in one of several languages, identify the ideology of the speaker's party, and (b) *Sub-Task 2*: Given a parliamentary speech in one of several languages, identify whether the speaker's party is currently governing or in opposition.

3.2. Dataset specifics

The dataset for this task is provided from ParlaMint [9], a multilingual corpus of parliamentary debates. The data is curated to minimize potential confounding variables, such as speaker identity, to ensure a balanced and unbiased dataset. The dataset is provided as tab-separated text files with the fields like, *id* (a unique ID for each text), *speaker* (a unique ID for each speaker, multiple speeches from the same speaker may be included), *sex* (the binary/biological sex of the speaker, which can be Female, Male, or Unspecified/Unknown), *text* (the transcribed text of the parliamentary speech, which may include line breaks and special sequences), *text_en* (automatic translation of the text to English. This field may be empty for English speeches or for some non-English speeches where the translation is unavailable), *label* (a binary/numeric label indicating political orientation [0 for left, 1 for right] or power identification [1 for opposition, 0 for coalition/governing party]). The training data encompasses parliamentary speeches from 28 countries for the political orientation task and 25 countries for the power identification task. The test files will have the same fields except for the label. A sample dataset for a single country (e.g., Latvia) for both sub-task 1 *(political orientation)* and sub-task 2 *(power identification)* is illustrated in Fig. 1a and 1b.

id	speaker	sex	text	text_en	label
lv00000	12e4d29ea4c1af0c8	M	Godājamais Prezidij	Honorable Bureau! D	1
lv00001	382e15bb9fedd997c	F	Labrīt, cienījamā Sa	Good morning, hono	1
lv00002	985899224d296176	M	Labdien, kolēģi! Vēl	Hello, colleagues! I'c	1
lv00003	e587d55787bf7611f	M	Cienītā Saeimas prie	President of the Sae	1
lv00004	6660d123d6afa6545	M	Labdien, cienījamā	Good afternoon, Ma	1
lv00005	f1bb423f1c0e74ebe	M	Augsti godājamā Sa	Highly honored Spea	1
lv00006	382e15bb9fedd997c	F	Cienījamā Saeimas	Dear Chairman of the	1
lv00007	508faeedcac361491	M	Cienījamie kolēģi, m	Ladies and gentleme	0
lv00008	12e4d29ea4c1af0c8	M	Cienījamā priekšsēd	Dear President, Colle	1
lv00009	25e789d80c9636e5	M	Godātais sēdes vadīt	Honorary Head of the	1
lv00010	857ecf8abad953155	M	Cienījamie kolēģi, tā	Dear colleagues, this	1

(a) Sub-Task 1 (Political Orientation)

id	speaker	sex	text	text_en	label
lv02806	7af40836f6973e473	M	Labdien, dāmas un k	Hello, ladies and ge	1
lv02807	8444d76c9a547a92	F	Labrīt, godātie kolēģ	Good morning, dear	0
lv02808	68b92d492fe282a16	M	Godātais Prezidij! Kc	Dear Bureau! Collea	0
lv02809	084456182fd5359da	F	Cienījamā Saeimas	Dear President of the	0
lv02810	0a97cdc65b4cc276	F	Labdien, kolēģi! Es n	Hello, colleagues! I	1
lv02811	85e3463cd1a85669	M	Labdien, cienījamā	Good afternoon, Ma	1
lv02812	e98e7c657fc992eee	M	Kolēģi! Šā gada 3.jū	Colleagues! On 3 Jul	0
lv02813	49c6c3fa99b3e38ef	M	Labrīt, cienījamie ko	Good morning, hono	0
lv02814	84ec9b59fa385d60e	F	Labdien, godātie kol	Good afternoon, dea	1
lv02815	0a97cdc65b4cc276	F	Labdien, kolēģi! Izgl	Hello, colleagues! T	1
lv02816	e587d55787bf7611f	M	Cienītā Saeimas prie	President of the Sae	0

(b) Sub-Task 2 (Power Identification)

Figure 1: Data samples from the data provided for the Sub-tasks

4. System Overview

4.1. Data Preprocessing

Before feeding text data into either system, the following preprocessing steps are performed to improve the quality of the analysis. First, the text is broken down into individual words or meaningful units, a process known as tokenization. Next, common words that don't contribute to sentiment analysis, such as "the", "a", and "an" are eliminated using language-specific stopword lists. Finally, words may be reduced to their base form (lemma) to improve consistency, a process called lemmatization, although this step is optional and not necessary for all languages.

Various language-specific libraries and tools cater to a wide range of languages, facilitating text analysis and processing tasks. SpaCy [10] supports languages such as Catalan, Croatian, Danish, Dutch, English, Finnish, French, German, Greek, Italian, Polish, Portuguese, Romanian, Russian, Slovenian, Spanish, Swedish, and Ukrainian. NLTK [11] provides robust support for languages including Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Italian, Polish, Portuguese, Russian, Slovenian, Spanish, Swedish, and Turkish. Additionally, StanfordNLP [12] specializes in Bulgarian, Croatian, Serbian, and Slovenian. These tools are essential for preprocessing, analyzing, and understanding text across diverse linguistic contexts, enhancing the capability of NLP applications worldwide.

4.2. Approach 1: Text Vectorization using TF-IDF

After preprocessing, the textual data is transformed using the Term Frequency-Inverse Document Frequency (TF-IDF) method [13]. TF-IDF converts text into numerical features by assessing word frequency and importance across documents. This transformation is crucial for identifying ideological leanings within parliamentary speeches and distinguishing whether a speaker represents an opposition or governing party. The implementation of TF-IDF in this project is pivotal for analyzing parliamentary

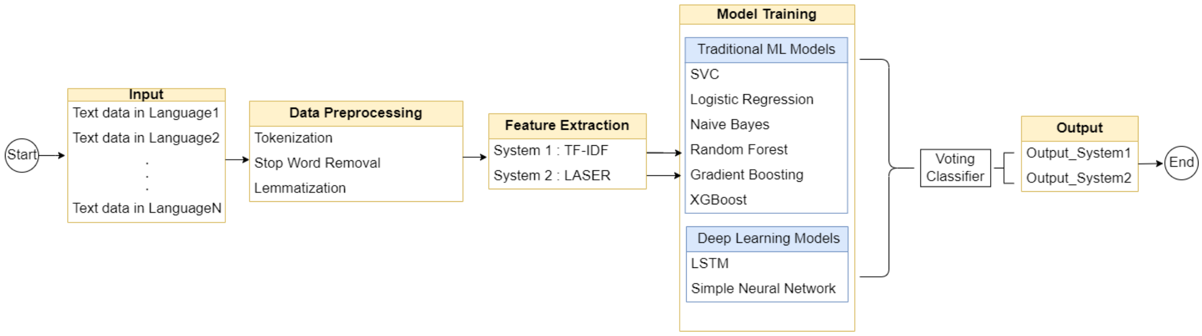

Figure 2: Proposed Approach

speeches, as it emphasizes words that are distinctive within specific documents yet less common across the entire corpus. This approach enhances the understanding of the content and context of individual speeches, facilitating nuanced analysis of political discourse.

TF-IDF plays a crucial role in ideology detection by identifying key terms and phrases indicative of left-leaning or right-leaning ideologies. Terms frequently associated with specific ideological stances receive higher TF-IDF scores, enabling machine learning models to effectively differentiate between speeches with contrasting ideological orientations. Additionally, TF-IDF assists in distinguishing speeches from opposition members versus those representing governing parties. By analyzing the prevalence and importance of specific terms, TF-IDF reveals linguistic patterns characteristic of opposition or governing party discourse.

4.3. Approach 2: Multilingual Sentence Embeddings

Instead of employing the TF-IDF method for text embedding, the LASER (Language-Agnostic SEntence Representations) encoder was utilized to transform the textual data. Developed by Facebook, LASER [14] is designed to enhance performance by providing highly effective multilingual sentence representations. This toolkit supports over 90 languages written in 28 different alphabets, embedding all languages jointly in a unified space rather than requiring separate models for each language. This capability makes LASER particularly advantageous for zero-shot transfer learning (Fig. 3), where a model trained on one language can generalize to others, including low-resource languages. The LASER encoder employs a five-layer bidirectional Long Short-Term Memory (BiLSTM) network (Fig. 4) to generate a fixed-size vector representation of input sentences in 1,024 dimensions. This high-dimensional vector is derived by max-pooling over the final states of the BiLSTM, ensuring that the embeddings encapsulate the semantic essence of sentences, irrespective of their written language. This universal, language-agnostic sentence embedding simplifies the comparison of sentence representations and supports their direct application in diverse classifiers.

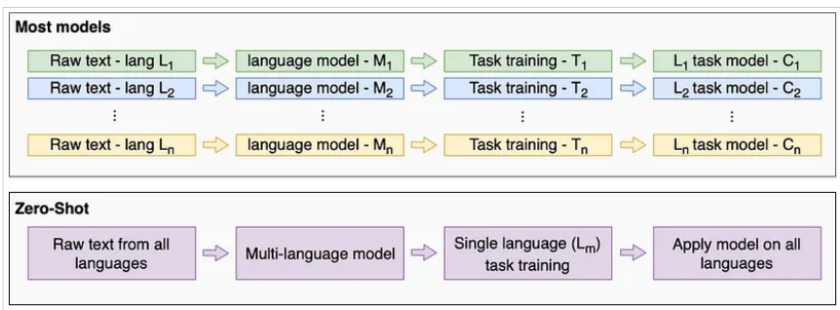

Figure 3: Zero-Shot approach vs other models

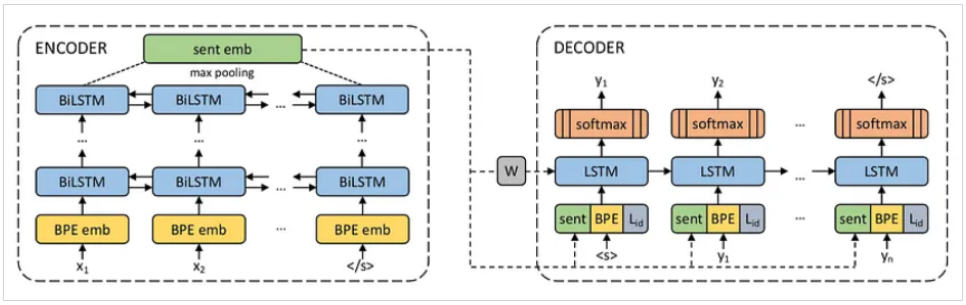

Figure 4: LASER Model Architecture [14]

4.4. Prediction Models

Features extracted from both TF-IDF and LASER encodings are utilized in various traditional machine learning models to classify speeches based on ideological orientation and political affiliation. The models employed include Logistic Regression, Support Vector Classifier (SVC), Naive Bayes, Random Forest Classifier, Gradient Boosting Classifier, and XGBoost Classifier. These models leverage the numerical features derived from the embeddings to effectively categorize the speeches, distinguishing between different ideological leanings and political affiliations.

To address the limitations encountered with traditional machine learning models, advanced deep learning architectures were incorporated into the classification process. A multi-layered LSTM [15] architecture with an embedding layer was utilized to convert inputs into denser representations. Regular dropout and recurrent dropouts were integrated to ensure the model's ability to generalize well. Additionally, a Simple Neural Network with two hidden layers was employed, featuring input layers, multiple hidden layers with ReLU activations, and dropout layers. Furthermore, a Voting Classifier was employed, combining the predictions of all the above classifiers—including the ML models, LSTM, and Simple Neural Network—to enhance classification accuracy.

4.5. Integrating BERT

In an attempt to further enhance performance, we considered integrating BERT [16], a transformer-based model known for its contextual word embeddings. However, due to the demanding computational requirements and the unsatisfactory results obtained during the training phase, we decided to terminate the integration of BERT into the classification pipeline. While BERT holds promise for improving classification accuracy by capturing important contextual information, our preliminary experimentation indicated that computational infrastructure constraints and performance limitations made it impractical for deployment in our project's context. Additionally, as BERT is language-specific and multilingual versions of BERT were not readily available, we halted further testing and deployment of BERT.

5. Experimental Results

As the leaderboard results are yet to be released, we are currently comparing our outcomes solely with the baseline. In the Touché 2024 Task 2, our team, Hale Lab, explored two distinct approaches. Both strategies demonstrated remarkable performance enhancements compared to the baseline across a range of metrics. In System 1, we achieved an F1 score of 0.6055 for political orientation and 0.6724 for power identification. Similarly, System 2 yielded promising results with an F1 score of 0.6154 for political orientation and 0.6983 for power identification, as shown in Table 1. When compared to the baseline, our methodologies consistently showcased improved performance metrics, including precision, recall, and F1 scores, across various countries. These outcomes underscore the effectiveness of our approaches in accurately deciphering political ideologies and power dynamics within parliamentary debates.

Table 1
Touché Task2 2024 Preliminary Results

Model	F1_orientation	F1_power
Baseline	0.569	0.640
System 1	0.615	0.672
System 2	0.605	0.698

6. Conclusion and Future Work

In this paper, the various approaches designed for addressing the Touché 2024 Task 2 requirements, focusing on the identification of political ideologies and power structures within parliamentary debates, were presented. Our methodology involved leveraging diverse feature sets, including linguistic, contextual, and speaker-related features, and applying advanced classification models to accurately detect the political orientation and power status of speakers. Despite the complexities introduced by the multilingual and heterogeneous nature of the dataset, our experiments yielded significant insights into the ideological and power dynamics of parliamentary discourse. These findings underscore the importance of robust preprocessing and the integration of various linguistic and contextual features to enhance model performance.

As part of extended work, we plan to further optimize our model to specifically identify the relationships between speeches, to determine which speeches are replies to others. This relational context is currently missing in the dataset but is crucial for a comprehensive understanding of parliamentary debates. Techniques such as dialogue act recognition and sequential modeling to map the conversational flow between speeches will also be explored. Including more languages and legislative contexts, and expert feedback, to enhance the generalizability of our models.

Acknowledgments

We extend our sincere gratitude to the Touché Lab for providing us with the opportunity to participate in this challenging task and for their support throughout the process. Special thanks to Çağrı Çöltekin for his invaluable assistance.

References

[1] J. Kiesel, Ç. Çöltekin, M. Heinrich, M. Fröbe, M. Alshomary, B. De Longueville, T. Erjavec, N. Handke, M. Kopp, N. Ljubešić, et al., Overview of touché 2024: Argumentation systems, in: European Conference on Information Retrieval, Springer, 2024, pp. 466–473.

[2] G. Abercrombie, R. Batista-Navarro, Sentiment and position-taking analysis of parliamentary debates: a systematic literature review, Journal of Computational Social Science 3 (2020) 245–270.

[3] T. M. Doan, J. A. Gulla, A survey on political viewpoints identification, Online Social Networks and Media 30 (2022) 100208. URL: https://www.sciencedirect.com/science/article/pii/S246869642200012X. doi:https://doi.org/10.1016/j.osnem.2022.100208.

[4] T. M. Doan, B. Kille, J. A. Gulla, SP-BERT: A language model for political text in scandinavian languages, in: Natural Language Processing and Information Systems, Lecture notes in computer science, Springer Nature Switzerland, Cham, 2023, pp. 467–477.

[5] D. Preoţiuc-Pietro, Y. Liu, D. Hopkins, L. Ungar, Beyond binary labels: Political ideology prediction of Twitter users, in: R. Barzilay, M.-Y. Kan (Eds.), Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 729–740. URL: https://aclanthology.org/P17-1068. doi:10.18653/v1/P17-1068.

[6] J. Barnes, S. Touileb, L. Øvrelid, E. Velldal, Lexicon information in neural sentiment analysis: a multi-task learning approach, in: M. Hartmann, B. Plank (Eds.), Proceedings of the 22nd Nordic Conference on Computational Linguistics, Linköping University Electronic Press, Turku, Finland, 2019, pp. 175–186. URL: https://aclanthology.org/W19-6119.

[7] S. Bhatia, D. P, Topic-specific sentiment analysis can help identify political ideology, in: A. Balahur, S. M. Mohammad, V. Hoste, R. Klinger (Eds.), Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Association for Computational Linguistics, Brussels, Belgium, 2018, pp. 79–84. URL: https://aclanthology.org/W18-6212. doi:10.18653/v1/W18-6212.

[8] M. Ahmadalinezhad, M. Makrehchi, Detecting agreement and disagreement in political debates, in: Social, Cultural, and Behavioral Modeling: 11th International Conference, SBP-BRiMS 2018, Washington, DC, USA, July 10-13, 2018, Proceedings 11, Springer, 2018, pp. 54–60.

[9] T. Erjavec, M. Ogrodniczuk, P. Osenova, N. Ljubešić, K. Simov, A. Pančur, M. Rudolf, M. Kopp, S. Barkarson, S. Steingrímsson, Çöltekin, J. de Does, K. Depuydt, T. Agnoloni, G. Venturi, M. Calzada Pérez, L. D. de Macedo, C. Navarretta, G. Luxardo, M. Coole, P. Rayson, V. Morkevičius, T. Krilavičius, R. Darǵis, O. Ring, R. van Heusden, M. Marx, D. Fišer, The parlamint corpora of parliamentary proceedings, Language resources and evaluation 57 (2022) 415–448. doi:10.1007/s10579-021-09574-0.

[10] M. Honnibal, I. Montani, spacy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing, To appear 7 (2017) 411–420.

[11] S. Bird, E. Klein, E. Loper, Natural language processing with Python: analyzing text with the natural language toolkit, " O'Reilly Media, Inc.", 2009.

[12] P. Qi, Y. Zhang, Y. Zhang, J. Bolton, C. D. Manning, Stanza: A python natural language processing toolkit for many human languages, arXiv preprint arXiv:2003.07082 (2020).

[13] J. Ramos, et al., Using tf-idf to determine word relevance in document queries, in: Proceedings of the first instructional conference on machine learning, volume 242, Citeseer, 2003, pp. 29–48.

[14] M. Artetxe, H. Schwenk, Massively multilingual sentence embeddings for zero-shot cross-lingual transfer and beyond, Transactions of the association for computational linguistics 7 (2019) 597–610.

[15] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9 (1997) 1735–1780.

[16] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).