Trojan Horses at Touché: Logistic Regression for Classification of Political Debates

Notebook for the Touché Lab at CLEF 2024

Deepak Chandar S, Diya Seshan, Avaneesh Koushik and P Mirunalini

Department of Computer Science and Engineering, Sri Sivasubramaniya Nadar College of Engineering, Chennai, Tamil Nadu, India

CLEF 2024, Grenoble, France

Introduction

Our research focuses on multi-lingual parliamentary speech analysis, specifically:

1. Identification of political ideology of speakers, and
2. Classification of speaker's party as governing or opposition

The model employed is:

- Logistic Regression model (trained with embeddings from CountVectorizer)

An F1-score of 0.59 and 0.69 for ideology and power classification respectively was achieved, with the effectiveness of the model being demonstrated in the context of evaluating parliamentary speeches from multiple countries.

Motivation

Understanding political environment is crucial in parliamentary discussions, with two crucial factors being:

- Philosophy of the speaker's party, and
- Whether the party is in power or not.

Determining above components accurately improves our comprehension of the speaker's viewpoint and the larger political forces at work.
By leveraging Logistic Regression framework, we can deepen our understanding of the intricate interplay between political ideology, governing status, and parliamentary discourse, thereby contributing to the broader discourse on computational approaches to political analysis.

Literature Survey

Key references that shaped our research and methodology:

- **K. Kato et al. (2024)**: LLM-based Political Ideology Nowcasting.
- **G. Glavaš et al. (2019)**: Computational Analysis of Political Texts: Bridging Research efforts across communities.
- **K. Miok et al. (2023)**: Multi-aspect Multilingual and Cross-lingual Parliamentary Speech Analysis.
- **S.V.K.Rohit (2018)**: Analysis of Speeches in Indian Parliamentary Debates.

Proposed System - Overview

Two models were explored:

- The initial model experimented with was BERT uncased classifier - trained with embeddings from CountVectorizer for each language. While some languages achieved F1-scores around 0.60, others had much lower scores, possibly due to varying text lengths and sub-optimal batch size. Hence, this model was ruled out for further implementation.
- **Logistic Regression** was chosen due to its superior accuracy and proven effectiveness across various domains. Initially, the model was trained on a combined dataset of all languages, which resulted in a notably low average F1-score. However, training the model separately for each language improved the average F1-scores, leading to the adoption of this approach for further analysis and evaluation.

Experimental Setup

Dataset: ParlaMint - Multilingual parliamentary debates corpus.
Data Preprocessing:

- Extracting the relevant fields from dataset
- Convert text to lowercase
- Vectorize text using CountVectorizer

Model Training:

- BERT-base uncased trained on embeddings from CountVectorizer
- Logistic Regression (i) for aggregated dataset and (ii) separately for each parliament dataset

Results

Table: Top F1-scores - Power Task

Parliament	**F1-score**	**Baseline F1-score**
Greece	0.83	0.79
Austria	0.72	0.67
Italy	0.71	0.65
Bosnia and Herzegovina	0.57	0.41

Table: Top F1-scores - Orientation Task

Parliament	**F1-score**	**Baseline F1-score**
Spain	0.72	0.72
Italy	0.66	0.65
Denmark	0.60	0.56
Netherlands	0.59	0.58

Results

- The proposed Logistic Regression model's F1-score surpassed that of the baseline for several languages. On average, the metrics measured were 3 to 5 percent higher than those of the baseline model.
- Reason for the improved performance as compared to the baseline model: Inclusion of the parameter C and fine-tuning the other hyper-parameters.
- It was found that the model performed better for the power sub-task which could be attributed to the presence of well-defined discriminating features.

Table: Average Results for each Sub-Task

	Task	Precision	Recall	F1-score
	Orientation	0.62	0.60	0.59
	Power	0.67	0.70	0.69

Conclusion

This research demonstrates the efficacy of logistic regression as a reliable technique for binary classification of multilingual parliamentary speech achieving good results for both the sub-tasks of this lab.

- Ideology classification (F1 Score: 0.59)
- Power classification (F1 Score: 0.69)

Looking ahead, further refinements and extensions of the approach hold promises in enhancing the predictive capabilities and applicability across a broader spectrum of parliamentary contexts. This includes exploring the integration of advanced large language models (LLMs).