A Computational Linguistic Analysis of Political Likeness: Angela Merkel and the SPD

Computational Linguistics Experiment on Political Speech Patterns (1998–2021)

Poster by

Lukas Kubicki
lkubicki@smail.uni-koeln.de
Bachelor Informations-Verarbeitung / English Studies
Course: Anwendungen der Computer Linguistik WS2024/2025

01 Context & Motivation

- In Recent research, Reiter (2014) proposes methods to measure linguistic similarity in narrative texts, focusing on how texts can be classified based on stylistic and structural features. While not political in nature, his approach inspires the application of computational similarity analysis to other domains—such as political discourse.
- Angela Merkel has internally been accused of social-democratizing the CDU, often aligning her ideas with traditionally SPD-held positions in areas like social policy, migration, and climate.
- Although Magazines like [Goethe-Institut, 2016][marx21, 2021] reject these claims, can a position be supported by CL Methods?
- This project investigates such claims by analyzing speech patterns of Merkel compared to SPD speakers between 1998–2021 which includes her time as Chancecllor of Germany (2005-2021).

02 Goal

The objective is to examine claims of "social democratization" of Merkel's CDU by comparing her speeches to those of the SPD.

How to observe Results?

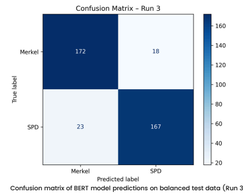
- F1 Accuracy: low = Model cant differentiate between Merkel &SPD's Speeches, high= Clear Discrimination between the Speeches
- Specific Aspects of Speech have high fluctuations in F1/Accuracy

Approach:
Can a machine learning model distinguish between Merkel and SPD speeches based on language alone and proof these claims?

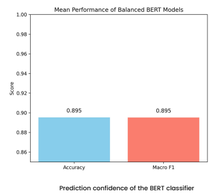
03 Methodology

- Corpus: Filtered All Bundestag speeches by Angela Merkel (1998–2021) and SPD members during the same period, based on the GermaParl corpus.
- Model: Fine-tuned HuggingFace's bert-base-german-cased
- Training Setup in 2 splits:
  Balanced training: 5 runs using equal numbers of Merkel and SPD speeches (stratified random sampling).
  Imbalanced baseline: Additional model trained on the full Merkel set vs. all SPD speeches, reflecting real-world distribution.
- Evaluation: Accuracy and Macro F1 score on a separate test set per run; topic-specific evaluations for interpretability.

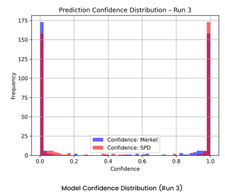
04 Model Performance



Confusion matrix of BERT model predictions on balanced test data (Run 3)



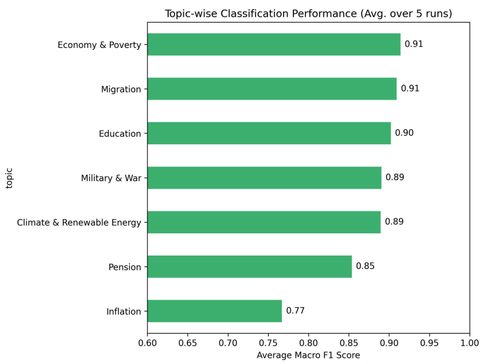
Prediction confidence of the BERT classifier



Model Confidence Distribution (Run 3)

- Balanced BERT model achieves ~90% Accuracy and ~0.90 Macro F1 across 5 randomized training runs.
- Confusion matrix reveals mostly correct predictions for both Merkel and SPD speeches.
- Confidence histogram shows **clear class separation** — the model assigns high certainty to most predictions.
- High per-class precision underlines **reliable linguistic separation**.
- Strong results despite **small dataset (~1,200 speeches)** show BERT's effectiveness on political text.

05 Topic-Based Classification

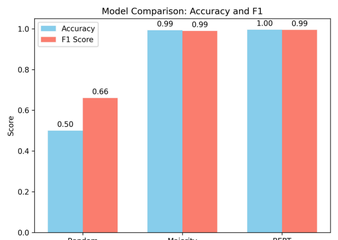


Macro-F1 scores of BERT across topical subsets (e.g., Economy, Climate, Education)

- Aimed to reveal topic-dependent similarities: Helps assess whether Merkel's rhetoric aligns more with SPD positions in certain policy areas, supporting or challenging the thesis of convergence.
- Applied all 5 balanced BERT models to test sets from each run.
- Filtered test speeches by topic-specific keywords (e.g., "Klima", "Wirtschaft")
- Evaluated classification performance (Accuracy & Macro F1) per topic.
- Averaged results across 5 runs to ensure robustness and mitigate outlier bias.
- Inflation had the lowest Macro F1 ≈ 0.77 - likely due to low sample size/less specific vocabulary, but still sugests difficulty for the model to distinguish parties on that topic

06 Model Comparison

Balanced Model on Balanced Test Set (Run 3 Macro F1)



Unbalanced Model on unbalanced Test set (Weighed F1)

Tested on Balanced Test Data

07 Conclusion & Interpretation

- **BERT reliably distinguishes Merkel and SPD speeches** based on linguistic patterns, achieving ~90% accuracy and macro F1 across five randomized training runs.
- **Balanced training significantly improves interpretability**:
  Compared to the imbalanced model, balanced sampling leads to more meaningful macro F1 scores and reduces bias from class distribution.
- **Findings do not support the "social-democratization" thesis:**
  Merkel's speeches remain distinguishable from SPD rhetoric, even during years of supposed ideological convergence.
- **Topic-specific analysis reveals subtle dynamics:**
  High linguistic separation in economic and climate-related topics.
  Lower performance on topics like inflation, likely due to both rhetorical overlap and limited test samples.
- **Confusion matrix and confidence scores underline model reliability**, with most predictions made with high certainty and low class confusion.
- **Political discourse ≠ political reality** While the model reliably distinguishes between Merkel and SPD speeches, it classifies language, not policy. Political likeness in language doesn't always align with legislative decisions or enacted laws. This highlights a gap between discursive similarity and political practice — the model reveals how parties position themselves linguistically, not necessarily how they act in parliament.
- **Outlook:**
  Extend the dataset to other parties for multi-class comparison.
  Include a temporal dimension to observe shifts in rhetoric over time.
  Refine the topic taxonomy for deeper thematic analysis.

References

https://github.com/drippidgeon/SpeechShift
https://github.com/PolMine/GermaParlTEI
https://huggingface.co/google-bert/bert-base-german-cased
https://archiv.ub.uni-heidelberg.de/volltextserver/17042/1/main.pdf
https://www.goethe.de/ins/gb/de/kul/mag/20773671.html
https://www.marx21.de/angela-merkel-16-jahre-bundeskanzlerin-bilanz-schlechter-als-ihr-ruf/