

Mining Arguments in US Presidential Campaign Debates

Mariia Poiaganova (poiaganova@uni-potsdam.de), Chen Peng and Milena Voskanyan
Project Module “Mining Opinions and Arguments”, taught by Prof. Dr. Manfred Stede

Introduction

Study focus 1: argumentation mining (AM) based on data from political discourse

For AM, political debates = a natural collection of controversial statements and supporting judgements

Study focus 2: replication potential of the paper*, given the dataset with original train, validation and test splits, but no source code

***Paper** “Yes, we can! Mining Arguments in 50 Years of US Presidential Campaign Debates” by Shohreh Haddadan, Elena Cabrio, Serena Villata, 2019.

Data

Dataset [USElecDeb60To16 v.01](https://github.com/ElecDeb60To16/Dataset) (https://github.com/ElecDeb60To16/Dataset)

- 39 debate transcripts from years **1960 to 2016**

Table 1. Corpus statistics (per number of sentences)

Total	Non-Arguments	Arguments	Premises	Claims
36.201	13.089	23.112	10.526	12.528

Example of premises and claims

R. Nixon: “Senator Kennedy’s position and mine completely different on this. I favor the present depletion allowance. I favor it not because I want to make a lot of oil men rich, but because I want to make America rich”.

Experimental Setting

Problem Setting: binary classification on a sentence level

Task 1: Detect if a sentence contains an argument component (either a claim or a premise)



Task 2: Detect if a sentence with an argument component contains a claim or a premise

For each task, we perform:

1. Tf-idf vectorization + SVM classifier
2. **Features*** + SVM classifier
3. FastText Embeddings + Long Short-Term Memory Neural Network
4. **Features*** + Feedforward Neural Network

*** Features**

- Tf-idf unigrams, bi- and trigrams
- NER labels
- Part-Of-Speech for adverbs and adjectives
- Tenses for verbs
- Syntactic features
- Discourse connectives
- Sentence sentiments
- **+** First person pronouns

Results

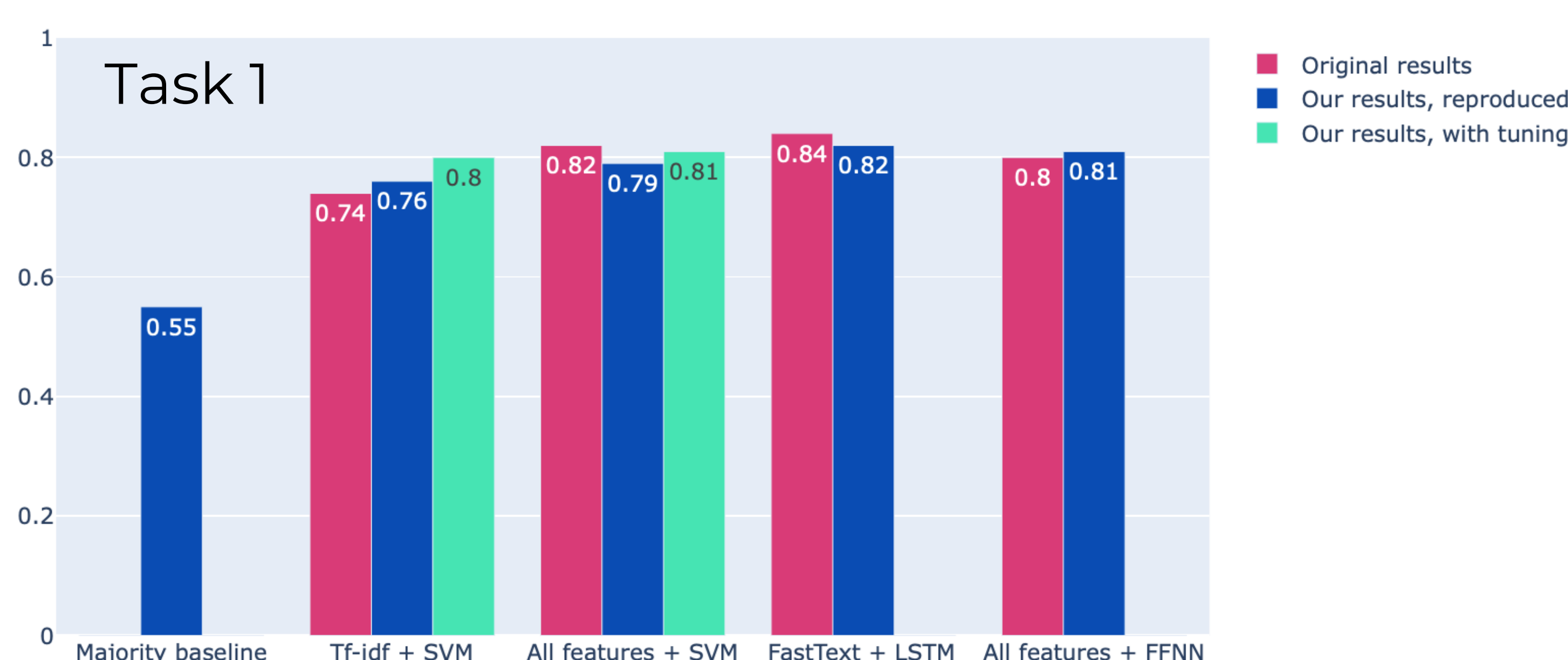


Figure 1. F1-scores for task 1: classification of argumentative vs. non-argumentative sentences

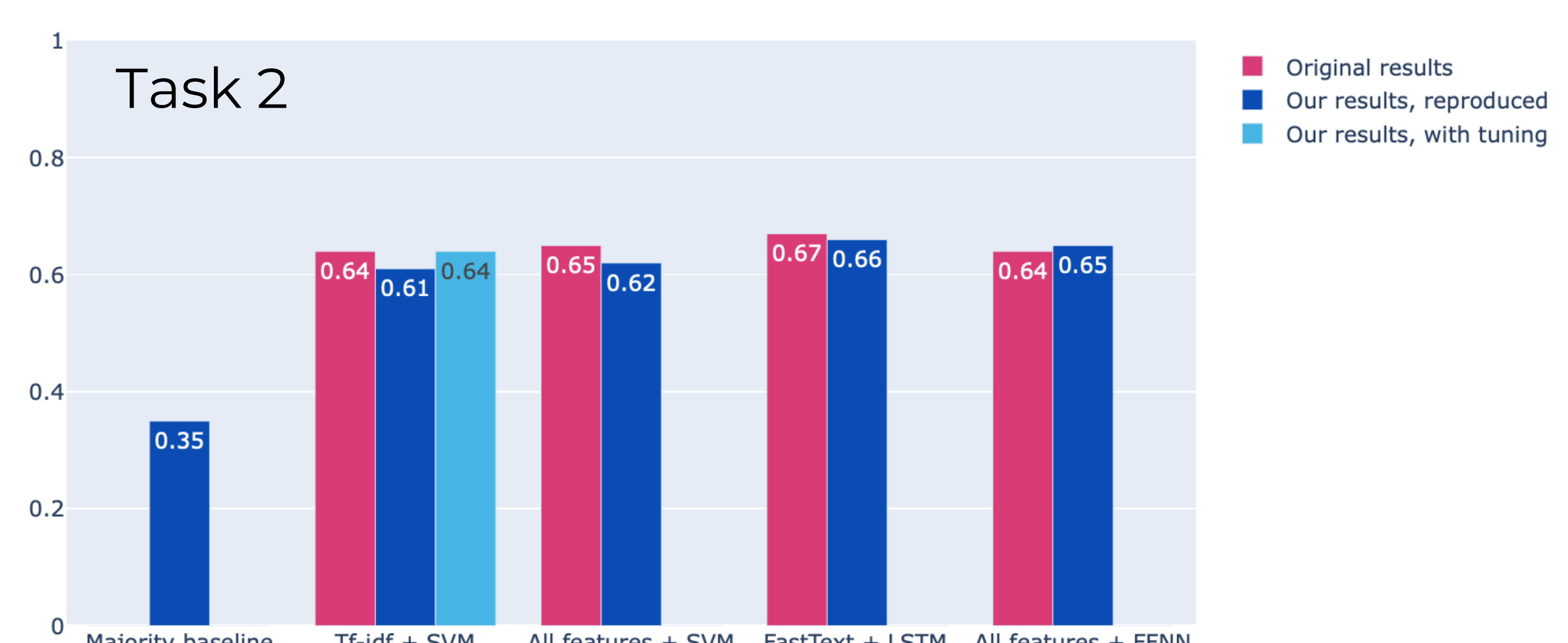


Figure 2. F1-scores for task 2: classification of claims vs. premises

- We replicated the authors’ methodology from scratch and for three out of eight proposed experiments, we exceeded the reported results. For the rest, we achieved comparable performance.
- Result variations can be due to randomization events, differences in the used kinds or versions of software frameworks (we use SpaCy, vader against the authors’ CoreNLP), hyperparameter inconsistencies.

Linguistic Error Analysis for the task 1: Sentences formulated like an argument and verbalized in non-argumentative context were often incorrectly classified as an argument (*I think we need to go back and take a look*). As for the false negatives, the errors were found in incomplete or conditional sentences.

References

Project’s open-source repository:
<https://github.com/a-moi/Argument-Mining-Presidential-Debates>

- Haddadan, S., Cabrio, E., & Villata, S. (2019). Yes, we can! Mining Arguments in 50 Years of US Presidential Campaign Debates. *ACL*.
Honnibal, M., & Montani, I. (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*.
Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.