Developing a Lithuanian Textual Corpus for Identifying Propaganda and Detecting Propaganda Techniques

In the age of information war, explicit and implicit propaganda spreads globally through social media and networks. Propaganda aims to sway public opinion using biased or misleading information, while organized propaganda employs troll factories and bots as information warfare tools [1]. Propaganda constantly evolves, adapts, and transforms, making detection a challenging field. With the rapid advancement of artificial intelligence, more sophisticated natural language processing models are being developed. While machine learning algorithms can be powerful for detecting potential propaganda content, it is essential to consider potential limitations. The model can effectively detect and classify propaganda in news articles that target high-resource languages. Low-resource languages, such as the Lithuanian language, are challenging. As recent research shows, one of the most effective ways to build propaganda detection models for low-resource languages is by creating high-quality propaganda textual corpora in the target languages and then fine-tuning the advanced large language models for this specific task. Research efforts have shown that the use of deep learning models, such as CNN, LSTM, and transformer-based models like BERT, can be effective in detecting propaganda in low-resource languages [2, 3].

The main goal of my doctoral research project is to develop a hybrid deep learning model for potential propaganda identification tailored explicitly for the Lithuanian language. The proposed research project aims to address the challenges of detecting propaganda in the Lithuanian language by leveraging advanced deep-learning models and creating high-quality propaganda textual corpora, which will be the first specialized Lithuanian dataset that facilitates the identification and analysis of propaganda. The annotation process is currently underway, and a method of cross-annotation has been selected to maintain a high level of agreement among annotators. The texts will be marked for the identification of propaganda, with a specific focus on highlighting propaganda techniques. This will yield a comprehensive dataset that can serve multiple purposes, including identifying propaganda, detecting propaganda techniques, and analyzing the linguistic and semantic aspects of such techniques in the Lithuanian language. However, as this is the first research of its kind in Lithuania, there are no experts available to discuss annotation corpus methodology, making the task quite challenging.

Explanation how the participation in the training school will be useful for the project.

Attending the UNIDIVE Summer School will provide me with a deeper understanding of corpus annotation infrastructure and multiword expression analysis from top experts. This knowledge will directly benefit my PhD thesis, particularly in enhancing the process of annotating propaganda textual data. The brainstorming hackathon will also offer an opportunity to exchange ideas on challenges and solutions in collocation detection with fellow researchers.

Open questions related to the project which could be addressed during the brainstorming hackathon.

- What methodologies can be employed to mitigate annotator bias in the manual labelling of textual datasets, especially in the context of sensitive subjects like propaganda?

- What are effective methods for automatically extracting MWE that are highly indicative of propaganda?
- How can machine learning be used to improve the detection and analysis of MWE in low-resource languages such as Lithuanian?
- What are the challenges in ensuring that MWE detection algorithms remain context-sensitive?

A short statement of the project phase (planning, started, in the process of creation)

The project is currently in an active development phase, with some of the data successfully annotated. Initial machine learning tests have been carried out on this annotated data, resulting in findings that have been reported in a paper recently accepted for publication (I. Rizgelienė, G. Korvel „Comparative Analysis of Various Data Balancing Techniques for Propaganda Detection in Lithuanian News Articles“, the Proceedings of 16th International Baltic Conference on Digital Business and Intelligent Systems, Baltic DB&IS 2024). A full linguistic analysis is planned as the next critical step in the project. The insights from the summer school will be crucial in refining both the machine learning models and the linguistic analysis framework.

References

1. Prier, J. (2017). Commanding the Trend: Social Media as Information Warfare. Strategic Studies Quarterly, 11(4), 50–85. http://www.jstor.org/stable/26271634
2. Chaudhari, D., & Pawar, A. (2023, November 15). Empowering Propaganda Detection in Resource-Restraint Languages: A Transformer-Based Framework for Classifying Hindi News Articles. https://doi.org/10.3390/bdcc7040175
3. Das, M., Banerjee, S., & Mukherjee, A. (2022, June 28). Data Bootstrapping Approaches to Improve Low Resource Abusive Language Detection for Indic Languages. ACM. https://doi.org/10.1145/3511095.3531277