

Doctoral School of Information and Biomedical Technologies

Polish Academy of Sciences (TIB PAN)

SUBJECT:

Explainable Methods for Sentiment Analysis in Social Media with a Focus on Sarcasm and Irony

SUPERVISOR:

Prof. Michał Karpowicz, michal.karpowicz@nask.pl, NASK National Research Institute

Inez Okulska, PhD, inez.okulska@nask.pl, NASK National Research Institute

DESCRIPTION:

Sentiment analysis has become crucial for understanding public opinion on social media platforms. With the increased usage of sarcasm and irony, misinterpreting them may lead to wrong predictions about the actual meaning of an utterance, jeopardizing the overall performance of specific content classification (such as hate speech or cyberbullying, etc.). This research proposal aims to develop explainable methods for sentiment analysis in social media, focusing on detecting and interpreting sarcasm.

Detecting sarcasm requires understanding dynamic contextual cues and entails challenges even among human evaluators due to inherent ambiguity. While large language models (LLMs) offer new insights into this task, concerns arise regarding their explainability. Nevertheless, their demonstrated capacity for reasoning could significantly contribute to constructing datasets that would facilitate the development of less complex yet more interpretable models, such as transformer-based ones.

The outcomes of this research will contribute to the field of sentiment analysis and provide valuable insights into public opinion, especially on social media platforms.

BIBLIOGRAPHY:

Babanejad, Nastaran, et al. "Affective and contextual embedding for sarcasm detection." *Proceedings of the 28th international conference on computational linguistics*. 2020.

Ghosh, Debanjan, Avijit Vajpayee, and Smaranda Muresan. "A report on the 2020 sarcasm detection shared task." *arXiv preprint arXiv:2005.05814* (2020).

Hsieh, Cheng-Yu, et al. "Distilling step-by-step! outperforming larger language models with less training data and smaller model sizes." *arXiv preprint arXiv:2305.02301* (2023).

Lemmens, Jens, et al. "Sarcasm detection using an ensemble approach." *proceedings of the second workshop on figurative language processing*. 2020.

Li, Shiyang, et al. "Explanations from large language models make small reasoners better." *arXiv preprint arXiv:2210.06726* (2022).

Sun, Xiaofei, et al. "Text Classification via Large Language Models." *arXiv preprint arXiv:2305.08377* (2023).