Burnout Analysis using NLP and Medical Biomarkers

Abstract

The increasing number of burnout cases have determined people to share their experience through social platforms, so that they could find support and understanding. The clinical indicators that can determine the burnout risk are usually conducted in psychological tests or by medical biomarkers, but these methods cannot hold the complex spectrum of emotions. This investigation employs a multimodal approach to burnout prediction, incorporating natural language processing of social media content alongside physiological biomarker analysis. The goal is to determine better results in the prediction phase, keeping in mind the possibility of a detection in the early phases of burnout.

The database for the NLP analysis was created by scraping subreddit comments related to: burnout, anxiety, toxic workplaces and academic stress to identify burnout and completely opposed sections to track the non-burnout content. To keep our direction, we generated the score for the main psychological indicators: PANAS, STAI, SSSQ by applying Natural Language Processing and custom dictionaries for the specific spectrum of emotions. The biomarkers and surveys are provided by the WESAD dataset and our results were combined with the processed biomarkers, by identifying common PANAS scores, so that the study can be relevant.

The analysis revealed a strong correlation between: the negative emotions, the social media reactions, the contrast of the speech and the burnout. It also indicated multiple phases of burnout from low, medium, to high risk, giving us the possibility of early prediction.

Authors:
Mădălina-Andreea Călin, Politehnica University of Bucharest, madalina.calin3008@upb.ro
Mădălin Frunzete, Politehnica University of Bucharest, madalin.frunzete@upb.ro
Vasile Chetroi, Technical University of Moldova, vasile.chetroi@adm.utm.md