Computational Approaches to Mental Health Disorders Detection from Social Media Texts, Images and Videos

Ana-Maria Bucur

Ph.D. Advisors: Prof. Dr. Liviu P. Dinu, University of Bucharest, Romania
Prof. Dr. Paolo Rosso, Universitat Politècnica de València, Spain

Abstract

Mental disorders affect 1 in 8 people worldwide, with nearly half of the population expected to experience a mental disorder by age 75. Stigma and limited access to care often prevent individuals from seeking the help they need. However, integrating digital technologies can improve accessibility to mental health care. This thesis investigates the use of computational approaches to detect depression, self-harm, and pathological gambling through social media data.

Depression affects approximately 4% of the population and is characterized by various symptoms that can impair a person's ability to function. Self-harm, which mainly impacts young people, involves non-suicidal self-injury practices, while pathological gambling refers to problematic gambling behaviors. Research has shown that specific language patterns in social media data correlate with mental disorders, revealing the potential for using online content in digital screening tools.

This thesis aims to leverage computational models to identify linguistic patterns associated with mental health problems, contributing to early detection and intervention efforts. First, we examine the differences in part-of-speech usage, focusing on self-focused language, pronouns, and verbs, which previous studies have linked to depression. Second, we investigate the prevalence of offensive language, analyzing whether individuals with depression use higher rates of profanity or hate speech. Third, we assess the feasibility of using recent advancements in natural language processing for the early detection of mental disorders, including pathological gambling, self-harm, and depression, by analyzing social media data. In addition, we evaluate the effectiveness of multimodal models that integrate text, images, and video content in detecting depression. Finally, we build upon existing psychological literature, which reveals that despite the challenges of depression, individuals may still experience happiness. Through a large-scale analysis of social media posts, we use a

comprehensive deep learning framework to extract, annotate, and classify happy moments based on semantic topics, gender, agency, and sociality.

This thesis highlights the important role of computational approaches in identifying mental health disorders through social media data. By using natural language processing and multimodal models, we can improve early detection and intervention efforts for mental disorders, which have been worsened by the COVID-19 pandemic. Integrating these tools into mental health care can enhance accessibility and support for those who need it.