Ordinal regression for stress levels classification in real-world scenarios

Tiago Bernardes Almeida[1]

Faculdade de Engenharia da Universidade do Porto
up201305665@fe.up.pt

Abstract. Stress and mental health have become a major concern worldwide nowadays. Consequently, stress has become a subject of research in recent years. Some physiological signals have been shown to correlate with stress, among them the electrocardiogram, the electrodermal activity and respiration. Current solutions focus on using predictive models to infer on stress levels using these physiological signals as predictors. Additionally, it has been shown that stress can be experienced in multiple levels, during short (acute) or prolonged (chronic) periods. Thus, multiple stress levels, e.g. low, medium, and high, inherently exhibit a natural order. Performing multi-class prediction ignoring this order information can lead to sub-optimal solutions. Consequently, we present a novel approach for stress detection and prediction. In this work, an ordinal regression approach was applied to infer on multiple stress levels, in this case low, medium and high, using features extracted from physiological signals as predictors provided by the MIT Driver Database. Furthermore, we compare domain-specific features to generic features to better understand their impact on the model's performance and, also, compare the performance with different combinations of physiological signals to investigate which ones are the most relevant. Moreover, we study the impact of the time-window and horizon values on this approach. Different values for the time-window and the horizon were used and compared the performance of the models for every possible combination of those two variables. Results show that ordinal regression presents a competitive predictive performance when compared to a standard multi-class prediction model. Furthermore, domain-specific features proved to be superior to generic features and the features extracted only from the respiration signal presented the best results of all combinations tested. Moreover, the study on horizon impact demonstrated better results for smaller horizons, 0 and 1 minute, whereas the study on time-window impact presented the best results for the larger time-windows, 10 and 15 minutes. Finally, we propose a new data collection protocol for stress level detection in free-living conditions, building upon and adapting currently available datasets. A small data collection was performed for a preliminary analysis of the proposed protocol and preliminary results show that collecting data for stress detection in a free-living environment is a difficult task. Experiments indicate that stressful situations are difficult to encounter when they are not induced intentionally.

Keywords: stress, stress detection, ordinal regression, classification, regression, data collection