Comparison of Objective and Subjective Measures Using Wearables: A Machine Learning Approach

University of Applied Sciences and Arts of Southern Switzerland
SUPSI

Università della Svizzera italiana

Alvise Dei Rossi[1,2], Radoslava Švihrová[1,3], Davide Marzorati[1], Max Grossenbacher[4], Francesca Dalia Faraci[1]

[1]University of Applied Sciences and Arts of Southern Switzerland, Department of Innovative Technologies - Institute of Digital Technologies for Personalized Healthcare, Lugano, Switzerland
[2]Università della Svizzera italiana, Faculty of Informatics, Lugano, Switzerland
[3]University of Bern, Institute of Computer Science - Cognitive Computational Neuroscience Group, Bern, Switzerland
[4]Resilient Ag, Lausanne, Switzerland

UNIVERSITÄT BERN

resilient

Background & Objectives

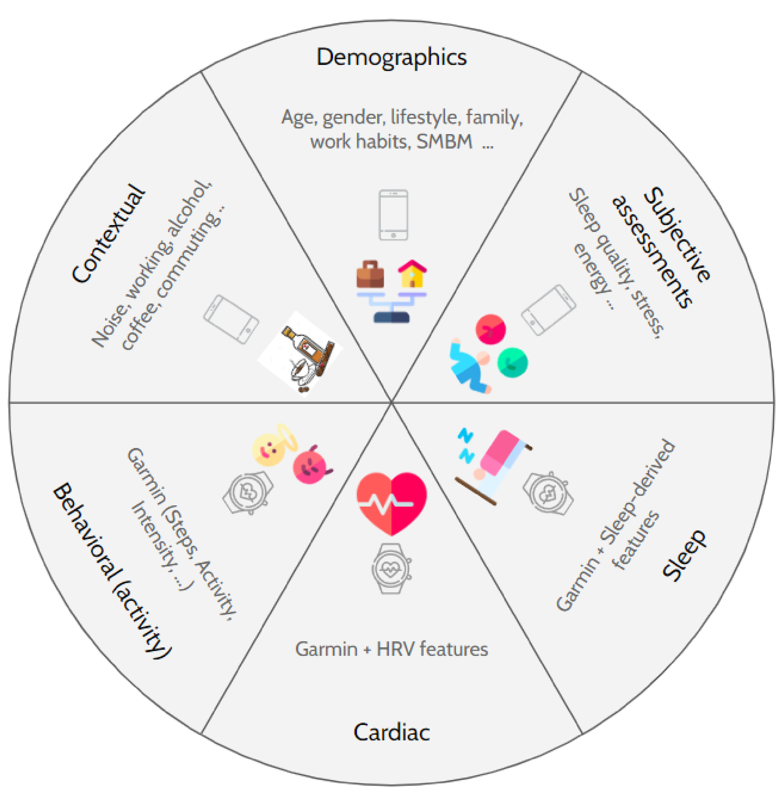
Wearable health devices have significantly impacted how we monitor health and fitness, showing potential for promoting active lifestyles, managing diseases, and enhancing well-being [1]. In this preliminary study, we explore the relationship between physiological features from smartwatches and subjective measures, as well as contextual variables, collected through daily EMA.

Research Design

N=27 individuals were recruited from two separate cohorts. The collected data included:

- **One-time questionnaires** capturing information on demographics, lifestyle, sleep habits, work-related factors, and burnout levels, evaluated through SMBM questionnaire.
- **Daily surveys** encompassing participants' subjective perceptions (e.g., sleep quality, energy levels, and stress) measured using 5-point Likert scales, along with contextual factors (e.g., working time, social interactions, and alcohol/coffee intake). Perceived sleep quality and daily stress levels, were gathered respectively through morning assessments immediately upon waking, and evening assessments at day's end.
- **Objective measures** gauging sleep quality (sleep duration, composition, efficiency, circadian disruption, and recovery), physical activity (daily steps, distance, and intensity), cardiac features (RHR, MHR, HRV), and stress levels. These metrics were derived at daily level from sensor-data acquired via Garmin Venu Sq 2 smartwatches, worn consistently throughout both daytime and nighttime hours over the span of one month.

This study was carried out in compliance with the Declaration of Helsinki.

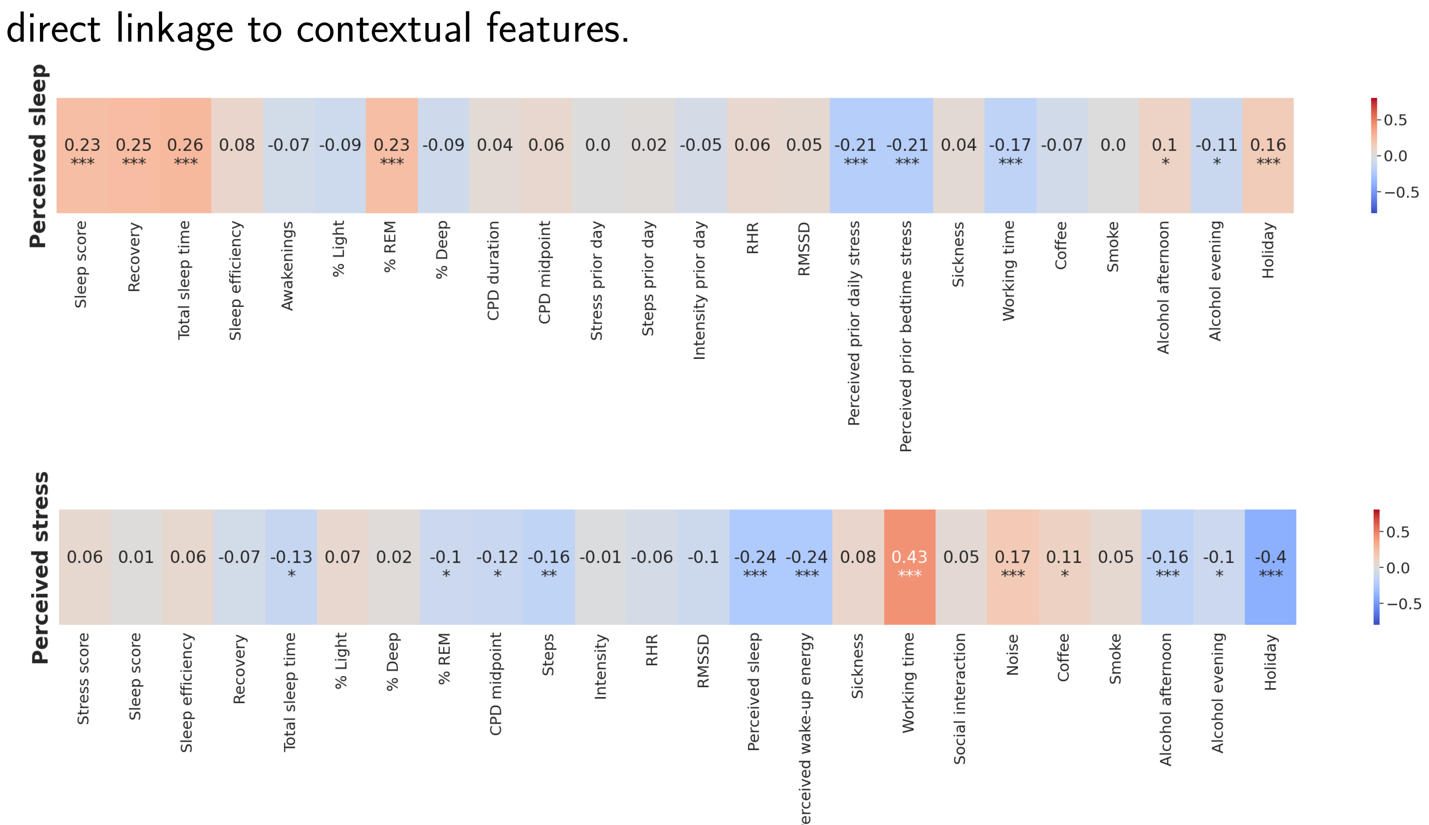


Methodology

Patterns and associations between subjective and objective measures were investigated using **Spearman's rank correlation** coefficients along with visual representations. Subsequently, **xgboost models were optimized for the prediction of perceived sleep quality and stress**. These models incorporated physiological attributes and contextual information. The task was framed as multi-class classification, under standard machine learning assumptions.

Results

Revealed trends and correlations show that perceived sleep is significantly, although weakly, associated with the objective summary sleep score metric derived by the wearable device and some sleep components, in accordance with was observed in previous studies [2]. In contrast, daily stress perception exhibits no significant correlation with the proposed daily stress metric from the smartwatch and only very weak correlations with physiological attributes, while manifesting a more direct linkage to contextual features.



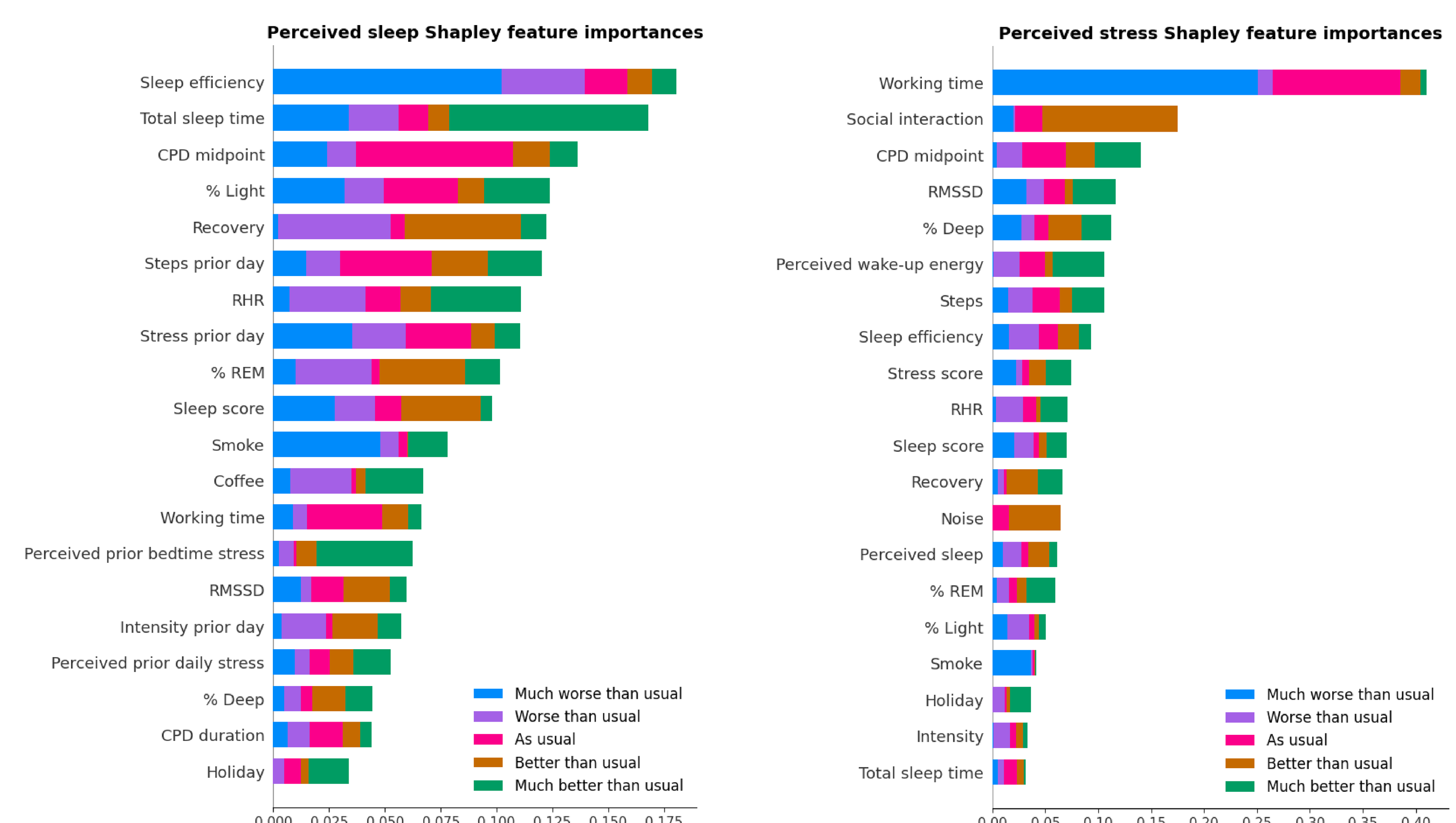
XGBoost models, incorporating sensor-derived and contextual features, only slightly improve predicting perceived sleep quality compared to those using exclusively sensor-derived attributes. While their accuracy marginally improve over random guessing, a significantly lower MAE is achieved.

Perceived sleep model	Balanced accuracy	Macro ROC AUC	MAE
XGBoost (sensor-derived and contextual)	25.2	0.593	0.93
XGBoost (only sensor-derived features)	23.3	0.585	0.98

For perceived stress, the XGBoost model that incorporates contextual data showed substantial additional predictive capabilities with respect to the one relying solely on sensor-based variables.

Perceived stress model	Balanced accuracy	Macro ROC AUC	MAE
XGBoost (sensor-derived and contextual)	46.8	0.718	0.76
XGBoost (only sensor-derived features)	32.1	0.614	0.87

Feature importance through Shapley values [3] can offer valuable insights into which features are the most relevant at single-class level for prediction of perceived sleep and stress.



Conclusions & Future works

A notable discrepancy between perceived subjective and objective measures was observed. Both perceived sleep quality and stress represent intricate constructs entwining both objective, subjective, and contextual facets.

In this initial analysis, observations were treated as independent, adhering to the traditional machine learning paradigm. Our future plans involve a more profound exploration, where we aim to provide personalized insights at the individual user level. This will be achieved by employing longitudinal methods that depart from the assumption of data being independently and identically distributed.

References

[1] Sophie Huhn, Miriam Axt, Hanns-Christian Gunga, et al.
The impact of wearable technologies in health research: scoping review.
JMIR mHealth and uHealth, 10(1):e34384, 2022.

[2] Lidia Alecci, Nouran Abdalazim, Leonardo Alchieri, Shkurta Gashi, and Silvia Santini.
On the mismatch between measured and perceived sleep quality.
In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*, pages 148–152, 2022.

[3] Erik Štrumbelj and Igor Kononenko.
Explaining prediction models and individual predictions with feature contributions.
Knowledge and information systems, 41:647–665, 2014.

Funding

The project has been funded by Swiss Secretariat for Education, Research and Innovation (SERI), INNO-256 55968.1