Stress Detection with Smartwatches: A Comparative Analysis of Machine Learning Approaches

Degree programme: Master of Science in Engineering
Specialisation: Data Science
Thesis advisor: Prof. Dr. Souhir Ben Souissi
Expert: Dr. Gabriella Casalino

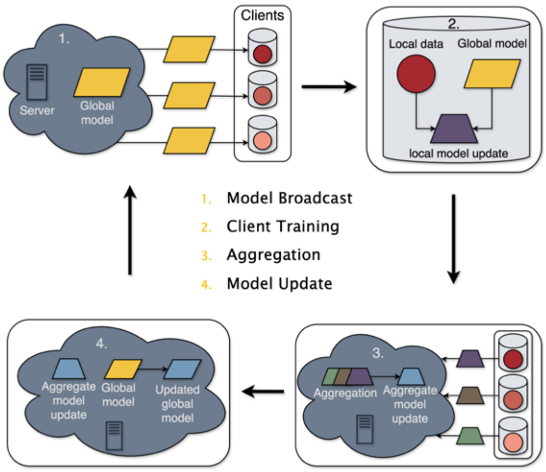
Recent advances in wearable technology have introduced innovative solutions for stress management, offering new ways to monitor and mitigate the negative effects of stress. But the necessity to collect and aggregate data from various sources into a central server can lead to concerns about privacy. This thesis therefore explored and compared different machine learning approaches to leverage sensor data for stress detection while focusing on data-privacy.

Introduction

In contemporary society, stress is a pervasive element of daily life, affecting individuals across diverse demographics. Stress detection is particularly crucial within healthcare, where professionals such as nurses often encounter elevated occupational stress due to demanding work conditions. Wearable sensor-equipped devices can monitor physiological indicators, including heart rate variability (HRV) and electrodermal activity (EDA), offering real-time insights into an individual's stress levels. A primary obstacle in developing robust machine learning models is the necessity for extensive, high-quality datasets. Conventionally, this involves aggregating data from numerous individuals onto a central server, which introduces substantial privacy considerations.

Methods

This thesis investigated and compared three distinct learning methodologies, two of which offer privacy-preserving machine learning capabilities. **Individual learning**, which involves training a dedicated model for each person on their local device. A novel approach called **federated learning**, where multiple participants contribute to the training of a shared model. Models are trained locally, and only model updates are shared with a central server, preserving data privacy. And **centralized learning**, where data from all participants is uploaded to a central server, thereby compromising data privacy. This thesis utilized two publicly available datasets for stress detection, incorporating sensor data from smartwatches, including a dataset focused on stress detection in nurses. XGBoost, neural networks, and logistic regression models were employed for training, alongside various pre-processing and feature engineering techniques.



Federated Learning Process

Cedric Aebi
cedric.aebi@me.com

Results

The experimental results indicated that highly personalized models developed within the individual learning framework achieved exceptional performance, with F_1 scores reaching up to 0.99 on both datasets. However, the outcomes from centralized and federated learning suggest that stress is still a subjective experience, limiting the generalizability of models trained in these scenarios. Centralized learning demonstrated strong performance on training data but exhibited reduced efficacy when applied to data from new, unseen participants. Federated learning encountered additional challenges, with data heterogeneity impeding the training process.

Conclusion

While highly personalized models excel at stress detection and user privacy, generalizing to diverse populations remains a hurdle. Leveraging a foundation model trained on physiological data, followed by fine-tuning with the wearable sensor data used in this thesis, offers a promising solution, potentially leveraging existing knowledge to achieve superior performance and broader applicability.