Digital Health and Connected Healthcare
Prof. Dr. Bert Arnrich
Master's Thesis Proposal

Digital Health Center by Hasso-Plattner-Institut

HPI

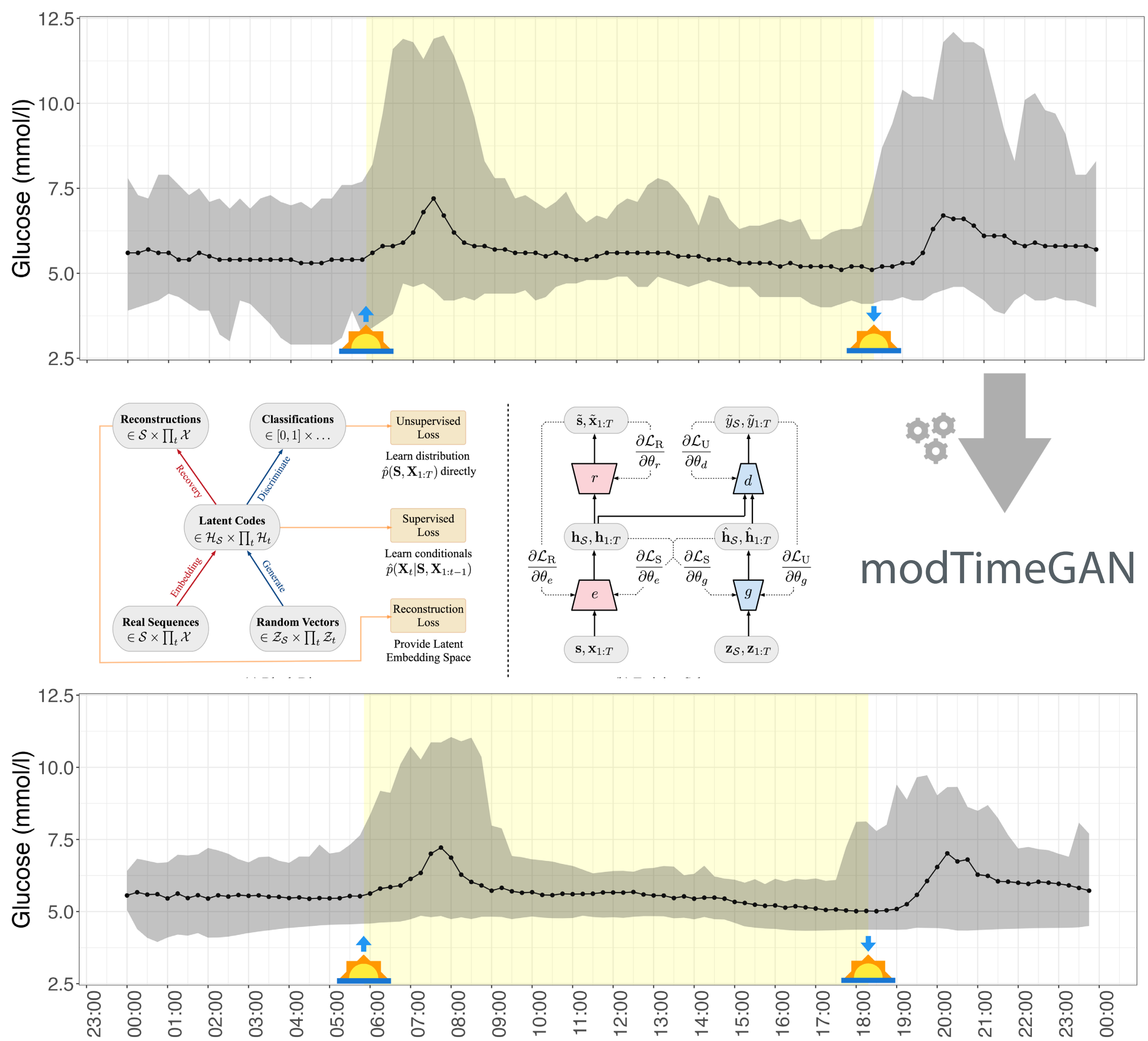
CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN

DIfE Deutsches Institut für Ernährungsforschung Potsdam - Rehbrücke

Synthetic continuous glucose time series

Simulating glucose metabolism using continuous glucose monitoring in different fasting forms

The complex dynamics of glucose metabolism depend on multiple individual factors such as chronobiology and meal patterns and as such are difficult to predict. This project aims to simulate glucose metabolism patterns in different fasting forms, using continuous glucose monitoring (CGM) data to model metabolic reactions to food ingestion and fasting.



Based on CGM data you will modify and implement a generative adversarial network approach, that can synthesize daily glucose time series specific to different fasting forms. The TimeGAN [1] implementation will be modified, and hyperparameters must be adjusted to fit the time series data: balance between performance and runtime. This process will be evaluated by comparing distributions between original and synthetic data, e.g. t-stochastic neighborhood embedding (t-SNE) or principal component analysis (PCA). Discriminative and predictive scores, based on post-hoc recurrent neural networks, can be used to evaluate the performance of the approach.

Our joint project is using clinical trial data from various studies at Charité and DIfE (German Institute for Nutrition). Inside this interdisciplinary project, we use sensors (like continuous glucose monitors, ketone, activity, and vital sign sensors) to digitally phenotype subjects. You will learn to integrate clinical trial data and perform various methodological developments in the field of time series and medical data science (machine learning). You will start synthesizing time series of glucose dynamics using generative adversarial networks to capture circadian patterns based on a variety of fasting trials. You will further explore the increase of accuracy while augmenting trial time series data, to further improve statistical power when data is missing or rare (or expensive to collect). Such an approach not only allows augmenting preliminary trial CGM data but also enables privacy concerned publication of sensible time series data sets for the research community.

Your Responsibilities

- Combining continuous glucose monitoring with biomedical data
- Get familiar with recent developments in generative adversarial networks to generate time series data
- Exploratory time series analysis, visualization, and other interpretation concepts

Your Profile

- Comfortable with data science tools, especially time series analysis (or willing to learn)
- Good programming skills (Python, R)
- Experience with Linux and Windows
- Team player and strong communicator
- Work on own initiative
- Quick learner and willing to share knowledge
- Good English skills
- Some biomedical knowledge is helpful
- Creative thinking is very welcome

If you are interested in interdisciplinary research, time series and generative adversarial networks please contact:

Nico Steckhan
Nico.Steckhan@hpi.de
+49 (0) 331 5509-1374
G-2.1.11, Campus 3

Prof. Dr. Bert Arnrich
Bert.Arnrich@hpi.de
+49 (0) 331 5509-4851
G-2.1.14, Campus 3

[1] Jinsung Yoon, Daniel Jarrett, Mihaela van der Schaar, „Time-series Generative Adversarial Networks," Neural Information Processing Systems (NeurIPS), 2019.