

Digital ICU : MA - Representation learning and clustering for multivariate time-series patient data

General Info

Contact Person: Kai Wu

Contact Email: k.wu@tum.de

Project Abstract

Digitalization in healthcare has led to the increasing use of digital medical systems in the Intensive Care Unit (ICU). They generate a large amount of data, such as the vital signs of patients, the blood gas analysis results, and the medication that a patient receives. This data can be analyzed using machine learning and data analytics techniques to help clinicians identify clinical deterioration in patients earlier and determine if a patient's treatment is working.

In this project, we want to explore representation learning methods on multivariate time-series patient data. We are interested in comparing data representations learned from both mathematical and deep learning methods, e.g., principal component analysis, minimum covariance determinant, self-organizing map, auto-encoder, graph neural network. In the low dimensional representation space, clusters or boundaries are expected to distinguish between patients who are in stable or risky conditions.

Tasks Description

- Literature reviews on representation learning methods for time-series patient data.
- Clean and preprocess data from public dataset to fit with your task scope
- Implement data representation methods and compare the representation space
- Visualize the representation space and validate the results

Technical Prerequisites

- Good knowledge in probability theory and machine learning
- Python, experiences with libraries like pandas, scikit-learn, pytorch, matplotlib, etc.