University of Zurich[UZH]

Department of Informatics

University of Zürich
Department of Informatics
Binzmühlestr. 14
CH-8050 Zürich
Phone. +41 44 635 43 11
Fax +41 44 635 68 09
www.ifi.uzh.ch/dbtg

UZH, Dept. of Informatics, Binzmühlestr. 14, CH-8050 Zürich

Prof. Dr. Michael Böhlen
Professor
Phone +41 44 635 43 33
Fax +41 44 635 68 09
boehlen@ifi.uzh.ch

Zürich, October 22, 2019

MSc Project: Timeseries Analysis of Medical Intensive Care Unit Data

Hospitals monitor patients in critical care closely, e.g. those in intensive care units (ICUs), to be able to react quickly when a patient's state starts deteriorating. Sensors collect data about the vital signs represented as time series, which are recorded and analyzed elsewhere. Medical personnel has to make sense of this data, which is far from trivial, as often hundreds of different measurements are taken, some of them several times a second.

One goal is to gain a deeper understanding of medical time series data, so in a first step we want to have a look at the correlation of different signals. Since the human body is a highly interconnected system, we expect to find a considerable number of correlations. However, it is not as simple as comparing the values of time series point for point, as a trend in one time series may be reflected in another one with a time lag.

When monitoring ICU patients, another goal is to cut down on the number of false alarms. In typical ICU scenarios, some monitoring tools raise up to 700 alarms each day, a large percentage of which are false alarms. This is mainly due to the configuration of the monitor tools, which only check whether a signal goes above or below a predefined threshold. The values of these thresholds are determined in a very general way and do not reflect the individual circumstances of a patient. However, patients (and especially ICU patients) show a large variation for certain signals and not all of these fluctuations mean that a patient is in immediate danger. Therefore, it makes sense to build specific models for individual patients as a baseline to determine whether a patient is about to enter a critical state.

For this project, we use the freely and publicly available MIMIC database (`https://physionet.org/content/mimicdb/1.0.0/`). In particular, the project is structured into the following tasks:

- Doing some exploratory data analysis of the data in the database MIMIC to better understand what the data looks like.

- For the following tasks, the dataset might have to be preprocessed to make it suitable for the different algorithms.
- Determining the cross-correlation between two time series is not straightforward, there are various parameters to consider:
  - Due to noise and other stability issues, time series are divided up into smaller windows. There are different approaches for windowing (e.g. tumbling vs. shifting) and various parameters such as window size, time lag, and step size. These parameters need to be calibrated to get good results. The goal is to determine good parameters.
  - A standard way to measure cross-correlation is based on the Pearson correlation coefficient. However, other techniques, such as mutual information, may be better suited. The aim here is to test which cross-correlation measures work best.
- Well-known methods for modeling time series are autoregressive integrated moving average models (ARIMA). The goal is to build ARIMA models for different signals to see if this technique is suitable for modeling ICU data.
- A different approach is to measure how windows evolve over time, i.e., how similar they are to previous windows. Tracking changes over time may give us some indication whether a patient is stable or not. The aim is to find out whether we can track relevant changes in windows over time.
- In a second step the models are tested to see whether they can be used for adapting alarm prediction based on the state of an individual patient to reduce the false alarm rate.

Supervisor: Sven Helmer

Department of Informatics, University of Zurich

Prof. Dr. Michael Böhlen