

Goal

The goal of the project is to create a unified library of continuous glucose monitor (CGM) data sets that allows for rapid experimentation and testing of various time-series models.

At high level, we want to design an algorithm that given the data and the pre-defined configurations is going to do the pre-processing and produce either a table (Pandas DataFrame) or a expose an iterator (PyTorch DataLoader) object. An example desired use case would be:

```
config = load_yaml('path_to_config')
data_formatter = DataFormatter(config)

table = data_formatter.get_table()
iterator = data_formatter.get_iter()
```

Data formatter

Each data set usually comes as `.csv` file.