

\*\*\*\*\*

**\* Description \***

\*\*\*\*\*

Write a sensor data analytics program to build the relationship model between the sensor data and the parameters (S, D), and predict S and D from future sensor data. Python language is expected for the sensor data analytics program. The goal is to achieve a low MAE (less than 3 ideally) for two parameters (S, D) on the test data set and show a matching trend comparison plot of labels and prediction results (see main\_plot.py as the example]).

In the data file, each row includes sensor data (10 seconds \* 100Hz) + ID + Time + H + R + S + D, which may be visualized with the command `./view_data.py xxx.npy 6`. H, R, S and D are heartrate, respiratory rate, systolic and diastolic blood pressure. Here 10 seconds \* 100Hz means that each row contains 10 seconds data and the data sampling rate is 100Hz (e.g., 100 data points per second).

The data analytics program may be based on advanced signal processing and machine learning (DSPML), or deep learning (DL), or their combinations. The data set is a simulated data set with ground truth, and a DSPML method can solve the problem well and is preferred in this interview. DL method works too - if you choose to use DL method and demonstrate it works, we wish to hear your insights on why this model and how to fine tune it.

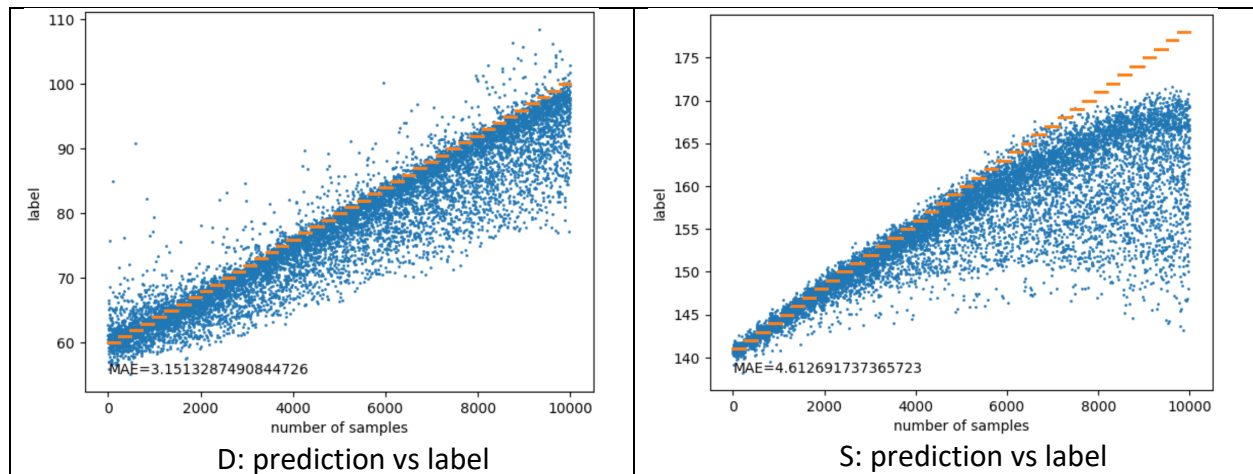
If you choose to perform the supervised training, please use the train data set (\*\_train.npy) for training, and do not include the test data set (\*\_test.npy) in training, which shall be used for test only.

\*\*\*\*\*

**\* Delivery \***

\*\*\*\*\*

1. Please save your trained model in a file and write a prediction program that reads the saved model, takes the test data set as the input, outputs the prediction result (S, D), and prints their MAE and use plot\_2vectors() function to save the plots of the label and prediction result (see example plots below and example code in main\_plot.py), so that we may verify your result by calling your prediction function only, without having to run your training program.



2. The goal is to achieve  $MAE \leq 3$  for the parameters (S, D) on the test data set. Write a one or more pages report to summarize your algorithms, results and findings.

3. Please structure your source code in a professional data science coding style and framework like <https://github.com/jc-audet/WOODS> or <https://github.com/timeseriesAI/tsai/tree/main/tsai> or <https://github.com/fraunhoferportugal/tsfel>