Explanation of each modules used in project.

1. fix_stress_labels

We'll first make sure the labels are ints of either 1 (stressed) or 0 (relaxed).
It takes name of coloumn and dataframe.
And search where `dataframe['stress']>=0.5` and give unique value and return dataframe.

2. missing_values

We'll manually clean up all the rows, and fill missing values with the mean.
Note that the heart rate columns are further cleaned with a median filter, this is to make sure errors in readings are smoothed out.
It takes dataframe.
Find not a number value and replace with mean.

3. do_tpot

Machine learning model is created with TPOT, an automated machine learning tool that optimizes machine learning pipelines with genetic programmings. TPOT makes a bunch of models with different variations of algorithms and data processing, and only the "strongest" ones survive each generations. This works similar to natural selection where only the strongest species (pipeline) survives and in the end (in theory) the best possible evolution left. TPOT has to run for a good while to successfully have enough time to evolve, so at this stage.

TPOT is initialized to train with five-fold cross validation based on 80% of the data, and then validates the trained model on 20% of the data.

TPOT was initialized with a population size of 100 and the number of generations was set to 400. Essentially, this means that TPOT will train 100 models for each generation it iterates through.