To approach the problem of predicting age ranges from the provided sensor data, we first manually examined the data and determined that there were no obvious sensor artifacts and missing data was typically brief compared to the full data set; thus, we determined that a linear interpolation to impute missing data regions would be appropriate for this task; regions of missingness which began at the start of the time series were replaced with zeros.

To develop the predictive model, we chose to use a convolutional neural network (CNN) linked to an XGBoost classifier. In addition to the raw signal data, a channel of derived data was added by calculating the ratio of aortic pressure to brachial pressure (as this is well recognized to be correlated with age) at each time step. Data was split into training and validation sets using an 80/20 split. The structure of the CNN is 6 convolutional layers each with kernel size of 65 followed by 3 fully connected layers. The CNN was trained using the training set, and at the end of each trained epoch, the network was tested on the validation set. Training was stopped when the validation score did not improve for 10 epochs. In order to improve performance, the output of the CNN after the convolution layers and before the fully connected layers was fed to an XGBoost classifier. The XGBoost classifier was able to classify with better performance than the fully connected layers of the initial CNN.