Machine learning module

The main function of this module is to provide prediction on optimal heater setting when mobile application provides real time data during sleeping period. This module is composed of two elements: model training and prediction block. The model training block persist the incoming data and apply training algorithm to adapt the underlying model when certain amount of data is collected to achieve online learning. Meanwhile, the prediction block handle the incoming data as prediction input and respond to the mobile application by providing prediction result back. Specifically, the response will be either binary information on heater setting or sleeping quality. The mobile application receiving this outcome should be further interpreted and produce regarding control signal to the smart heating system.

Implementation plan

Our primal solution is to build a sleep quality classifier and used its result with the current body temperature to adjust the heater setting. The classification is defined as binary model based on the heart rate, skin impedance, accelerometer and skin temperature readings collected from wearable sensors around wrist.

The development plan for this classifier can be split into three stages: Feature generation, Model Selection and Model Deployment. The primary aim for feature generation is to discover the optimal set of parameters to accomplish the classification between “good” and ‘bad’ sleeping quality. Since sleeping quality is highly subjective to everyone, we decided to obtain data for ourselves by using Microsoft band 2 with a made mobile application to collect data and label the sleeping quality. Next, we need to investigate on parameters that can represent the raw data in terms of sleeping quality. From previous studies, there exists some parameters which are universally used to interpret their raw format. <add literature quote on HR, skin Temp etc>. Since all the data collected are time series format, the time interval used to calculate these suggested parameters can be varied. Hence, we need to leverage the MATLAB machine learning tool box to perform preliminary analysis on these features. Upon a new set of features, multiple models will be trained and the optimal set will be one with highest average accuracy across model. Furthermore, it is necessary to obtain the statistical properties on each feature, for example, their distribution function and interclass correlations. Most importantly, we need to investigate on correlation between sleeping quality and the body temperature.

Model selection stage focus on determining optimal model for this application. The choice of each model is based on the required accuracy and performance to handle real time prediction. In this application, prediction speed is essential as it requires to provide response when mobile application send a request. Furthermore, the model should be selected based on the feature properties and the preliminary benchmarking result from MATLAB. In addition to accuracy, it is important that there are few parameters to be concerned and they are shown below:

After a specific model is chosen, advance techniques can be applied to improve model computational and behavioral performance. For example, principal component analysis can reduce the feature dimension and allow the model to perform prediction based on smaller feature dimension and hence fewer computational iterations. Furthermore, feature rescaling may allow models to fairly learn the importance of each feature and hence prevent overfitting.

Finally, model deployment stage implement the designed machine learning model into the whole system. Our initial design platform is to apply python library Scikit-learn, Django and Celery to link machine learning with backend end server maintenance. Scikit-learn will provide the machine learning model and training framework for our design while Django and Celery can act as mobile application interface and model online training respectively.