CG3002 Embedded System Design Project
AY2017/18 Semester I
Software Progress Checkpoint (Deadline: Sept. 21/22, Week 6)

Overview

This assessment (worth 5%) is to make sure you are on the right track and can test your software implementation on a dataset of your choice, either taken from an existing dataset or captured by yourself. To meet the criteria for this checkpoint, you will learn how to use libraries and software to perform multi-class classification on a dataset and get familiar with the basics and applications of Machine Learning. Note that we are going to test your software on any laptop / desktop you choose. We will **not** test on Raspberry Pi yet ;)

Your goal is to classify human activities into three classes using Machine Learning:

1. Choose the human activities you would like to classify

Choose three different human activities from the following: sitting, walking, jumping, or the provided project dance moves.

2. Acquire or produce a dataset

The dataset can be acquired from any online sources or collected on your own.

The data can be from any modality, such as inertial signals (accelerometers) or images from cameras. Note that the modality of this data should be same as your design report.

3. Apply machine learning approaches

You should use at least **two** machine learning algorithms/models to do classification. The choice of which one to use is up to you, there are many options such as Neural Networks (NN), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), or any others you can think of! You will need to extract features from your dataset as discussed in the lecture. Tell us which ones did you pick and why? You should use some form of model validation to test your approach, for example k-fold cross-validation.

You should indicate the libraries and software you used for classification. Again, there are many options for you to use, scikit-learn, tensorflow, keras, and weka. Remember that you will have to eventually run your software on the Raspberry Pi, so make sure it is compatible with the Linux OS on the Pi 3.

4. Evaluation and expected outputs

Your software should output the following two metrics of your proposed models:

- Classification accuracy
- Confusion matrix

Try and interpret them (as you might get asked some questions!). Which model performed better and why? What do you think makes the activities different and separable in your chosen modalities?

Then, pick one sample from your dataset and show us which activity it is classified into. This is critical as you will have to be able to run your trained model on new and unseen data.