# Task

To develop a software system for recognition of activities of daily living (ALDs) by applying computational models. You will develop your recognition system using a dataset with signals from accelerometer and gyroscope collected from six activities (1: Walking, 2: Walking upstairs, 3: Walking downstairs, 4: Sitting, 5: Standing, 6: Laying).

For this task, you will use the Human Activity Recognition datasets (humanADLs.zip file) available in the Moodle page of the EE52036. The datasets are built from the recordings of 30 subjects performing ADLs while carrying a waist-mounted smartphone (Samsung Galaxy S II) with embedded inertial sensors.

# Data Preprocessing

* Inspect data
  + Inspect skewness (histograms)
  + Identify anomalies (box and whiskers plots and scatter plots)
  + Check class balances (bar charts/pie charts)
  + Identify missing data and it’s type
  + Identify other noisy data
* Preprocess data
  + Remove anomalies
  + Balance classes
  + Deal with missing data
  + Deal with noisy data
  + Standardize and normalize
  + Encode actions
  + Data smoothing?
* Feature Engineering
  + Apply transformation
  + Polynomial features?
  + Check for interaction terms and combine if necessary
  + Window techniques for time series features?
  + Build features
  + Dimensionality reduction (PCA/SVD)
  + Feature Selection
  + Correlation heatmaps and Pearsons correlation coefficient

# Training

* Train/test/val split with shuffle
* Can we include Ensemble learning with NN? E.g. boosting

Model 1: Random Forrest

Model 2: Basic NN

Model 3: TBC. Probably something like a RNN/transformer

# Evaluation and visualization

* Confusion matrix
* Residual Plots
* Learning curves
* Cross validation. (Leave one Out?)
* Bias and variance graphs

# Optimization

* Hyperparameter tuning of best model
* Fix any over/under fitting
* Bayesian Optimization backed up by grid search?