**HUMAN ACTIVITY RECOGNITION USING SMARTPHONES**

**INTERIM REPORT**

**Proposed Solution:**

The given problem can be interpreted as multiclass classification problem. It involves predicting the movement of a person based on sensor data and involves deep domain expertise and methods from signal processing to correctly engineer features from raw data in order to fit a machine learning model. The problem can be solved with the use of classical machine learning models and Deep Learning models.

In the given problem, the data set of Human Activity Recognition has been labelled as “Walking”, “Walking Upstairs”, and “Walking”, “Downstairs”, “Sitting “and“ Lying”. The data set is defined into two parts, first is RAW data set and second is pre-engineered data set. The pre-engineered data set is used with classical machine learning to learn from the data, and predict the human activity. Then the RAW dataset is used with Deep Learning models to learn from the data, and predict the human Activity.

**Evaluation metrics:**

The Evaluation metrics, used to find out how good a model is, are given below,

1. Confusion matrix: A confusion matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known.
2. Accuracy score: Accuracy is one metric for evaluating classification models. It is defined as the fraction of predictions our model got right.
3. Multi-class log-loss: it is an important metric for multiclass ML problem.

**Initial Exploratory Data Analysis:**

* The dataset has to be checked to see if there are any null values in any of the features and if any null values are found they have to be removed.
* The dataset has to be balanced. That is, the all observations per subject should be almost same. If it isn’t balance, then it should be balanced using various methods such as SMOOT, etc.
* The activities are to be classified in static and dynamic activities. Static activities include Sitting,Lying,Standing while as Dynamic activities include walking, walking upstairs, walking downstairs.
* The Static and Dynamic Activites of Human has been plotted on a Box Plot for more understanding.

**Summary of Initial Findings:**

* There are no null values in the dataset.
* The dataset is balanced.
* In static activities,motion information will not be very useful
* In the dynamic activities,motion info will be significant.
* From the Box Plot, based on the value of tAccMean the activities of a human are predicted.

**Challenges:**

* Classifying and plotting the activities in Static and Dynamic activities.
* Choosing between 2 different models that have the same level of accuracy.