# Project Report: Sensor Data for Behavioural Activity Recognition

# Introduction:

The aim of this project was to utilize machine learning techniques to recognize behavioral activities using sensor data. The project progressed through several stages, including the initial implementation of a random forest model using the iris dataset. Subsequently, the provided data was imported, and seven handcrafted features were extracted, including minimum, maximum, average, standard deviation, and mean values from a single sensor, namely the accelerometer. The initial attempt resulted in a score of 36%.

Day 1: **Random Forest with Iris Dataset**

On the first day, we programmed a random forest model using the well-known iris dataset. This initial step helped us gain familiarity with the implementation and understand the basic concepts of the random forest algorithm.

Day 2: **Handcrafted Features with Single Sensor**

The next day, we imported the provided data and extracted seven handcrafted features from a single sensor, the accelerometer. These features included minimum, maximum, average, standard deviation, and mean values. However, upon testing, we encountered NaN (not a number) values. To address this issue, we implemented additional checks and used the HistGradientBoostingClassifier algorithm. Despite the improvements, the score only increased to 45%.

Day 3: Handcrafted Features with All Sensors

On the following day, we decided to utilize all 18 handcrafted features extracted from a single sensor. We observed a significant improvement, with the score reaching 68%. This step allowed us to leverage more comprehensive information and achieve better results.

Day 4: Combination of Sensor Feature Matrices

Continuing our exploration, we concatenated the feature matrices of all available sensors. This approach enabled us to capture data from multiple sources and led to a substantial improvement in performance. With the combined feature matrix, we obtained a score of 68%.

Day 5: Evaluation of Sensor Combinations

To further refine our results, we investigated different combinations of sensor feature matrices. We discovered that by using only seven sensors, we achieved a score of 71%. This finding indicates that a subset of sensors provides optimal information for behavioral activity recognition.

Conclusion:

In conclusion, our machine learning project focused on utilizing sensor data for behavioral activity recognition. We progressed through various stages, starting with a random forest implementation using the iris dataset. Subsequently, we extracted handcrafted features from a single sensor, resolving NaN value issues along the way. By incorporating all sensors' feature matrices, we achieved a score of 68%. Further experimentation revealed that using a specific combination of seven sensors improved the score to 71%. These findings showcase the potential of machine learning techniques in recognizing behavioral activities based on sensor data.