

# Kaggle Challenge - 1 (30 marks)

## Unsupervised Learning

---

### Instructions.

- This is a kaggle competition so you have to follow these instructions carefully.
- Create a single notebook in the competition. Multiple notebooks of same submission will lead to **0 marks**
- You are free to use any available classical unsupervised machine learning techniques, method, data pre-processing techniques available to perform this task.
- In *train.csv* labels are given for your reference and help in data pre-processing. Do not use them in model training, doing so will reward you **0 marks**.
- You have to submit code and your code will be checked for plagiarism, IIITD plagiarism policy will be followed strictly.

### Dataset Description

Each of the 19 activities is performed by eight subjects (4 female, 4 male, between the ages 20 and 30) for 5 minutes. Total signal duration is 5 minutes for each activity of each subject. The subjects are asked to perform the activities in their own style and were not restricted on how the activities should be performed. For this reason, there are inter-subject variations in the speeds and amplitudes of some activities. The activities are performed at the Bilkent University Sports Hall, in the Electrical and Electronics Engineering Building, and in a flat outdoor area on campus. Sensor units are calibrated to acquire data at 25 Hz sampling frequency.

The 19 activities are:

- Sitting (a01)
- Standing (a02)
- Lying on back (a03)
- Lying on right side (a04)
- Ascending Stairs (a05)
- Descending stairs (a06)
- Standing in an elevator still (a07)
- Moving around in an elevator (a08)
- Walking in a parking lot (a09)
- Walking on a treadmill with a speed of 4 km/h in flat position) (a10)
- Walking on a treadmill with a speed of 4 km/h in 15 deg position) (a11)
- Running on a treadmill with a speed of 8 km/h (a12)
- Exercising on a stepper (a13)
- Exercising on a cross trainer (a14)
- Cycling on an exercise bike in horizontal position (a15)
- Cycling on an exercise bike in horizontal position (a16)
- Rowing (a17)
- Jumping (a18)
- Playing basketball (a19)

There are 5 sensory units on torso (T), right arm (RA), left arm (LA), right leg (RL), left leg (LL) and 9 sensors on each unit (x,y,z accelerometers, x,y,z gyroscopes, x,y,z magnetometers). Each row of CSV is the data recorded by the sensor while subject is performing the activity. *Sensor* column tell the type of sensor used to record the data and its position, *Action\_Person* column tells the activity and person seperated by \_, (a01\_p1)where a01 is activity and p1 is person. All unique entries in Action\_Person column are respective classes and indexing starts from 0.

Sensors:

*T-xacc, T-yacc, T-zacc, T-xgyro, T-ygyro, T-zgyro, T-xmag, T-ymag, T-zmag, RA-xacc, RA-yacc, RA-zacc, RA-xgyro, RA-ygyro, RA-zgyro, RA-xmag, RA-ymag, RA-zmag, LA-xacc, LA-yacc, LA-zacc, LA-xgyro, LA-ygyro, LA-zgyro, LA-xmag, LA-ymag, LA-zmag, RL-xacc, RL-yacc, RL-zacc, RL-xgyro, RL-ygyro, RL-zgyro, RL-xmag, RL-ymag, RL-zmag, LL-xacc, LL-yacc, LL-zacc, LL-xgyro, LL-ygyro, LL-zgyro, LL-xmag, LL-ymag, LL-zmag.*

## Competition Link

## Task Description

Use unsupervised learning to train a model of your choice on training data and perform classification task on test data and submit *submission.csv* as your output. Labels given in *train.csv* are for your reference and for data pre-processing. Do not use them in your model training. your code will be verified for the same and will be rewarded 0 marks for using them in model training.

ID,	TARGET
1234,	34
..	.

Table 1: Prescribed format for uploading your submission.csv

## Marking Scheme

This is a competition so markings will be according to leader-board

- 1<sup>st</sup> - 100%, 2<sup>nd</sup> - 98%, 3<sup>rd</sup> - 96%
- 4<sup>th</sup> - 10<sup>th</sup> position - 95%
- 11<sup>th</sup> - 20<sup>th</sup> position - 90%
- 21<sup>st</sup> - 30<sup>th</sup> position - 80%
- 31<sup>st</sup> - 40<sup>th</sup> position - 70%
- 41<sup>st</sup> - 50<sup>th</sup> position - 60%
- 51<sup>st</sup> - 60<sup>th</sup> position - 50%
- 61<sup>st</sup> - Last position - 40%