

# Activity Detection using Machine Learning

## Description

Human Activity Detection is the problem of predicting what a person is doing based on a trace of their movement using sensors.

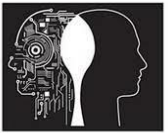
The "Human Activity Detection Dataset" includes data collected from 34 subjects, each of whom were asked to perform 18 tasks for 3 minutes each. Each subject had a smartwatch placed on his/her dominant hand and a smartphone in their pocket. The data collection was controlled by a custom-made app that ran on the smartphone and smartwatch. The sensor data that was collected was from the accelerometer and gyroscope on both the smartphone and smartwatch, yielding four total sensors. The sensor data was collected at a rate of 20 Hz (i.e., every 50ms). The smartphone was either the Google Nexus 5/5X or Samsung Galaxy S5 running Android 6.0 (Marshmallow). The smartwatch was the LG G Watch running Android Wear 1.5. The general characteristics of the data and data collection process are summarized in Table 1. More detailed information is presented later in this document.

Summary of Information for Dataset	
Number of Subjects	34
Number of Activities	18
Minutes collected per activity	3
Sensor polling rate	20 Hz
Smartphones used	Google Nexus 5/5x or Samsung Galaxy S5
Smartwatch used	LG G Watch
Number of raw measurements	15,630,426

**Table 1**

Table 2 lists the 18 activities that were performed. The actual data files specify the activities using the code from Table 2. Similar activities are not necessarily grouped together (e.g., eating activities are not all together).

Activity	Code
Walking	A
Jogging	B
Stairs	C
Sitting	D
Standing	E
Typing	F
Brushing teeth	G
Eating Soup	H



prithvi.AI

Eating Chips	I
Eating Pasta	J
Drinking from Cup	K
Eating Sandwich	L
Kicking	M
Playing Catch	O
Dribbling	P
Writing	Q
Clapping	R
Folding cloth	S

**Table 2**

## About Data

The raw sensor data is located in the directory. Each user has its own data file which is tagged with their subject id, the sensor, and the device. Within the data file, each line is:

**Subject-id, Activity Label, Timestamp, x, y, z**

The features are defined as follows:

**subject-id:** Identifies the subject and is an integer value between 1600 and 1650.

**activity-label:** see Table 2 for a mapping from 18 characters to the activity name

**timestamp:** time that the reading was taken (Unix Time)

**x:** x sensor value (real valued)

**y:** y sensor value (real valued)

**z:** z sensor value (real valued)

## Evaluation

Applicants are provided with training dataset. Training dataset consists of data of 20 subjects. Test data is also provided of 14 subjects to test the model. Evaluation will be done by us using validation dataset. We will evaluate the prediction using a metric. In this case we will use classification accuracy that will capture the performance (or error) of a model given the balance observations across the activities (or classes). Scoring mechanism is given below,

S.No.	Criteria	Marks
1.	<30% accuracy	0
2.	X% accuracy	0.4*X (maximum 40)
3.	Smart Pagination	5

## Submission

Students are requested to submit the Code, files in .csv format and screenshots of the accuracy achieved. The final evaluation will be done on our side using the validation dataset. Submission can be in the form of .zip file uploaded in cloud or a Git repository.