UNSW SYDNEY

COMP6733 - IoT Design Studio
Project Intermediate

Hand washing quality monitoring with sensorTag

Team name: JNY
Team member:
Nan Li, z5182060
Yuhao Qi, z5189309
Jeremy HE, z5142853

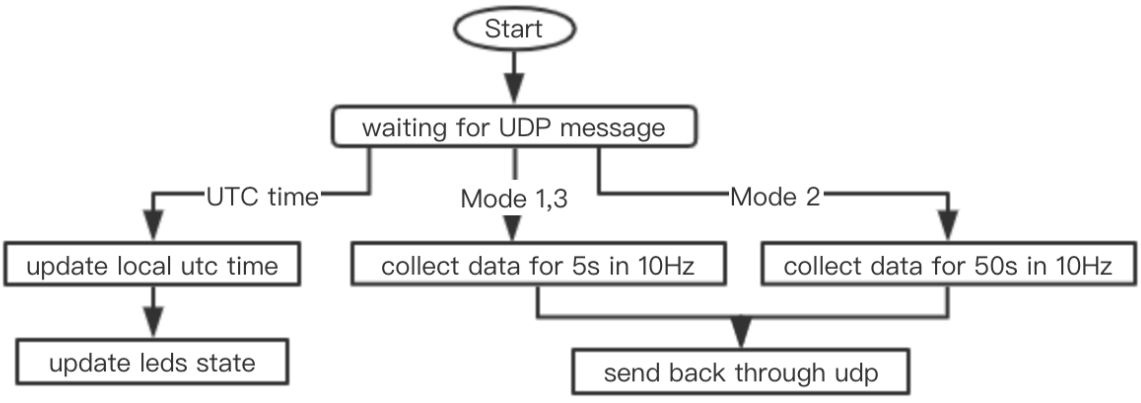
Date:7th, Nov., 2019

Data collection (sensor tag to server with python):

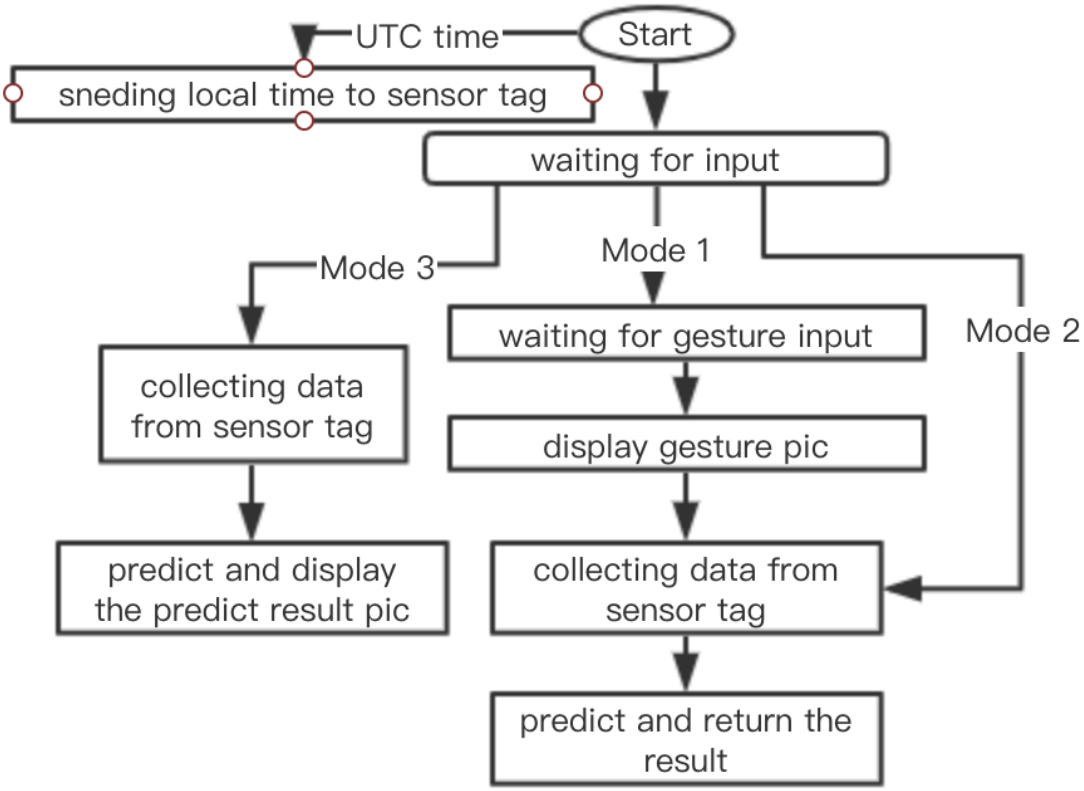
```
mpu_9250_sensor.configure(SENSORS_ACTIVE, MPU_9250_SENSOR_TYPE_ALL);
ctimer_set(&mode1_ctimer,0.1*CLOCK_SECOND,mode1_process,NULL);
void mode1_process() {
    int g_x = mpu_9250_sensor.value(MPU_9250_SENSOR_TYPE_GYRO_X);
    int g_y = mpu_9250_sensor.value(MPU_9250_SENSOR_TYPE_GYRO_Y);
    int g_z = mpu_9250_sensor.value(MPU_9250_SENSOR_TYPE_GYRO_Z);
    int a_x = mpu_9250_sensor.value(MPU_9250_SENSOR_TYPE_ACC_X);
    int a_y = mpu_9250_sensor.value(MPU_9250_SENSOR_TYPE_ACC_Y);
    int a_z = mpu_9250_sensor.value(MPU_9250_SENSOR_TYPE_ACC_Z);
    char result[50];
    sprintf(result,"%d,%d,%d,%d,%d,%d",g_x,g_y,g_z,a_x,a_y,a_z);
    printf("%s:::::%d\r\n",result,strlen(result));
    uip_ipaddr_t ipaddr;
    uip_ip6addr(&ipaddr,0xaaaa, 0, 0, 0, 0, 0, 0, 1);
    uip_udp_packet_sendto(server_conn, &result, strlen(result), &ipaddr,
UIP_HTONS(47371));
    ctimer_reset(&mode1_ctimer);
}
```

We collect the data from sensor tag. Then sending this data to the backend.
The format is [GYRO_X, GYRO_Y, GYRO_Z, ACC_X, ACC_Y, ACC_Z]
The frequency is 10Hz.
At the beginning we also want to get more than 10Hz data, but it will cause some problem, such as, packet lost, network congestion, and so on.

Sensor tag working model:

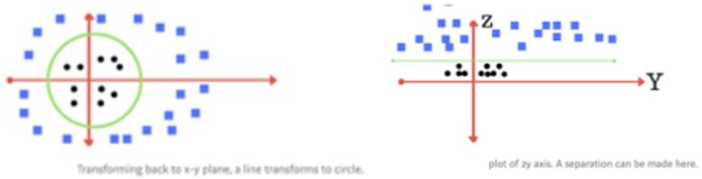


Backend working model:

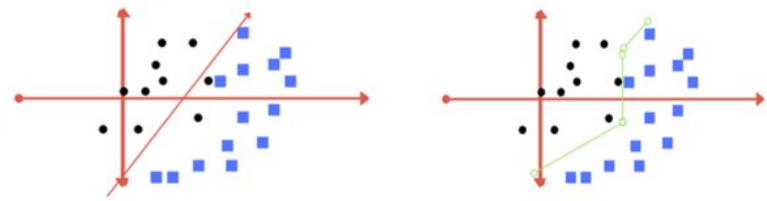


Machine Learning

Currently, we use the SVM algorithm, which also have the best performance for our real test.
SVM is a very common algorithm for supervised learning.
It can split the dataset through different way according to our parameter



When we choose different kernel, then we can separate our data differently



When we set different c value, then we can get higher precision on our separate line.
Image source: https://medium.com/machine-learning-101/

Data handling

The first step we have to handling the data, the original data we get that has six columns, since it is Supervised learning I, so I add the label for gesture.

	gyro_x	gyro_y	gyro_z	acc_x	acc_y	acc_z	class
0	-749	-564	690	1	-86	56	1
1	6950	4362	4733	-59	108	117	1
2	-25013	-5828	-4585	14	-21	49	1
3	25012	7503	1734	-19	-78	42	1
4	-17032	-7091	-2587	-20	44	104	1

Then we add more feature to help machine distinguish them like the length of vector. but the data is still not going to work. since there are many huge data, the kernel will spend more time to get the result, normally it will take a week to learn a model.

ro_y	gyro_z	acc_x	acc_y	acc_z	class	gyro_x_r	gyro_y_r	gyro_z_r	acc_x_r	acc_y_r	acc_z_r	root_gyro	root_acc
-564	690	1	-86	56	1	7699.0	4926.0	4043.0	-60.0	194.0	61.0	1164	102
362	4733	-59	108	117	1	-31963.0	-10190.0	-9318.0	73.0	-129.0	-68.0	9472	169
828	-4585	14	-21	49	1	50025.0	13331.0	6319.0	-33.0	-57.0	-7.0	26089	55
7503	1734	-19	-78	42	1	-42044.0	-14594.0	-4321.0	-1.0	122.0	62.0	26170	90
7091	-2587	-20	44	104	1	6326.0	5307.0	1351.0	22.0	-169.0	-56.0	18629	114

There are two features there, one is the root square number of XYZ value which we also called them length of vector. When the gesture change, the strength and the speed would affect vector length to be different. Then we can distinguish them.

Another one is the vector different between the neighbor. Imagine that two gestures there, one is moving as a circle, another is moving as a straight line, the sensor will use the same time slot to get the vector value, but for the circle one, the vector length could be same at the different time, however for the straight gesture, the difference between two time' s vector length is obvious.

When we done these things, we can train the model. we try the different para and get different accuracy.

We choose the linear one which has the best performance in the real test case.

There is also some risk and challenge for our project, one is overfitting, when we add more gestures, the interference between different gestures become serious. the more gesture we add, the lower accuracy we get, so for the demo, we going to use three gesture. another challenge is Feature extraction, it's actually hard to get the unique feature to distinguish them well. and we have to consider the environment noise, all these factors may affect the final accuracy.

Final Demo link: https://youtu.be/uPGy2R0Q66A --- Nan Li, Yuhao Qi

Contribution for Intermediate:

In this submission, members in charge of below port

Nan Li, z5182060

Coding mainly in charge of Sensor tag and backend python part, try some KNN model but with very low score, Data collection in Sensor tag, help and provide advice in ML

Yuhao Qi, z5189309

Coding mainly in charge of ML part, try several different models and generate it. Final demo work, Data collection in Sensor tag, help and provide advice in backend and sensor tag design, prepare for future app design