Machine Learning Based Fitness Tracker Platform Using MEMS Accelerometer

Yash Jain
Dept. of Applied Electronics and Instrumentation Engineering,
Heritage Institute of Technology
Kolkata:7000107, India
yashjain9339@gmail.com

Debjyoti Chowdhury
Dept. of Applied Electronics and Instrumentation Engineering,
Heritage Institute of Technology
Kolkata:7000107, India
debjyoti.chowdhury@heritageit.edu

Madhurima Chattopadhyay
Dept. of Applied Electronics and Instrumentation Engineering,
Heritage Institute of Technology
Kolkata: 7000107, India
madhurima.chattopadhyay@heritageit.edu.com

***Abstract*—This paper deals with real time development of a machine learning based portable system for determining eating habits of a human being using six point calibrated wearable MEMS tri-axial accelerometer. The rise of obesity as a global epidemic makes it immensely important to monitor food habits of a modern day person. In this proposed system, we have derived an easy to adopt algorithm based on a training based model for identifying the amount of calories consumed and burnt by a person. The proposed system consists of a wrist worn MEMS accelerometer that calculates calories burned per step which is directly sent over to the user's smart phone and a cloud based machine learning algorithm that does the prediction of health habit (i.e. healthy, unhealthy or undernutrition) based on the data obtained from the wrist worn device. In order to calculate the health habit of the user, the cloud uses logistic regression with calories burnt (from MEMS accelerometer) and calories consumed (daily manual input)to predict health habit of the user. The wrist worn device extracts calories burnt per step from the change in Y-axis acceleration data of the accelerometer in the wearable device, which after self-calibration is sent over to the user's smart phone through Wi-Fi. Thus, this cloud based food habit detection not only decreases the risk of obesity in a person but also introduces a low cost alternative device with reduced power consumption of (<13.5mW) and minimal covering size (12.56cm2) that can improve people's life.**

Keywords*—*MEMS Accelerometer; six point calibration; calories burnt; calories consumed; fitness tracker; learning algorithm; prediction; wearable band

I. Introduction

According to 2014 data 1.9 billion adults in the world are overweight out which 650 million are obese. Studies further suggest that by 2030, 515 of the world might be obese [1, 2]. United Nation (UN) Food and Agriculture Association reports suggest that 12.9% of the developing world is hungry [3]. The above mentioned conditions and many others are mainly caused by an imbalance between the calories consumed and calories burnt by an individual. In today's hectic and unhealthy lifestyle these problems are amplified as people do not have the time or means to track their fitness and calorie levels. The platform presented in this paper is a wrist worn band that can alert the user if their lifestyle is unhealthy and is susceptible to one of the various diseases caused by a severe calorie imbalance.

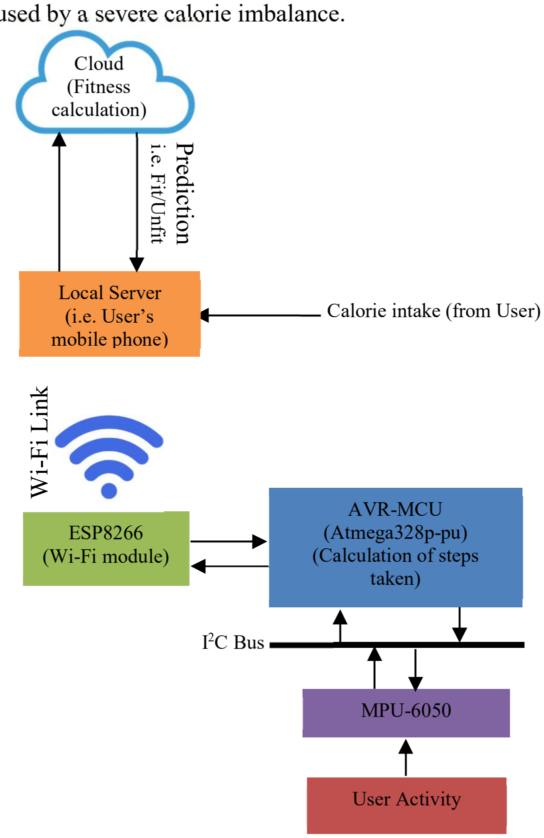

Fig.1. Block diagram of developed system

The developed platform consists of an off the shelf MEMS accelerometer (MPU6050), an AVR microcontroller (Atmega328p-pu) and a Wi-Fi module (ESP8266). The device tracks the user's daily calories/step count and hence the calories burnt with the help of the MEMS accelerometer. This data about calories burnt is sent over to a cloud via a Wi-Fi link. The user is also asked to provide, on a daily basis, their calorie intake. This data about calories burnt and calories consumed is used to predict how unhealthy the user's lifestyle is, using a machine learning algorithm, namely logistic regression. The basic block diagram of the overall system is shown in Fig.1.

The paper is divided into the following parts. Section 1 introduces the work, section 2 talks about the MEMS accelerometer calibration and how the calories burnt is calculated, section 3 describes the learning algorithm used for the classification of the user's lifestyle on the cloud, section 4 talks about the real time implementation of the consolidated device and finally section 5 puts forward the conclusion.

II. System Design And Modelling

The MEMS accelerometer, as mentioned earlier, is used to measure the acceleration, thereby giving the step count and calories burnt by the user. The presented device uses MPU-6050, which has a 3-axis gyroscope and a 3-axis accelerometer.

Part A of this section delineates the mathematical model of the accelerometer. Part B describes the calibration technique used by the accelerometer, which nullifies the error caused due to the displacement of the sensor in the wearable device. Part C contains the algorithm used to extract the step count and the calories burnt from the acceleration values given by the accelerometer.

A. *Mathematical Model of MEMS Accelerometer*

At still conditions the output from the X-axis of the MEMS accelerometer is represented by Eq.1 as shown below.

$$A_I = a_{I0} + K_{aI} \times a_I + K_{aII} \times a_I^2 \tag{1}$$

A_1 is the value of acceleration output from the accelerometer (mV), a_{10} is the zero drift (mV), K_{a1} is scaling factor (mV/(m/s^2)), a_1 is the actual value of output (mV) and finally K_{a11} is the quadratic non-linear coupling co-efficient (mV/(m/s^2)^2), all for the X-axis. The last term in (1) can be ignored as it causes an insignificant amount of error. By doing so the output model generated is given by Eq. 2.

$$A_I = a_{I0} + K_{aI} \times a_I \tag{2}$$

The error model involving the output model of all the three axes (X, Y, Z) can be written in matrix form as shown in Eq. 3.

$$\begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \\ a_{30} \end{bmatrix} + \begin{bmatrix} K_{a1} & S_{a11} & S_{a12} \\ S_{a21} & K_{a2} & S_{a22} \\ S_{a31} & S_{a32} & K_{a3} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \tag{3}$$

S_{a11}, S_{a12} are installation error coefficients for the X-axis, S_{a21}, S_{a22} for Y-axis and S_{a31}, S_{a32} for the Z-axis. A_1, A_2, A_3 are the acceleration values output by the accelerometer for the X, Y and Z axes respectively.

The error model based on the actual output values of the accelerometer is shown in Eq. 4 [4].

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} K_{a1} & S_{a11} & S_{a12} \\ S_{a21} & K_{a2} & S_{a22} \\ S_{a31} & S_{a32} & K_{a3} \end{bmatrix}^{-1} \begin{bmatrix} A_1 - a_{10} \\ A_2 - a_{20} \\ A_3 - a_{30} \end{bmatrix} \tag{4}$$

B. *Six Point Calibration*

The six point calibration technique is described here [5].

The six positions as shown in Table I, are used to obtain the calibration coefficients of the error model of the accelerometer.

TABLE I. ACCELEROMETER ORIENTATION FOR SIX POINT CALIBRATION

S.no	Orientation			Gravity		
	X-axis	Y-axis	Z-axis	X-axis (a_1)	Y-axis (a_2)	Z-axis (a_3)
1.	Down	East	South	1	0	0
2.	Up	North	West	-1	0	0
3.	West	Down	South	0	1	0
4.	North	Up	East	0	-1	0
5.	South	West	Down	0	0	1
6.	East	North	Up	0	0	-1

Using values of a_1, a_2 and a_3 from the Table I, the error coefficients are obtained by six point calibration of the three axes (X, Y, Z) by using Eq. 5, Eq. 6 and Eq. 7 respectively as shown below. Note that A_{11} implies the reading for the first orientation of the X-axis, A_{21} implies the reading for the first orientation of the Y-axis.

$$\begin{cases} A_{11} = a_{10} + K_{a1} \; A_{14} = a_{10} - S_{a11} \\ A_{12} = a_{10} - K_{a1} \; A_{15} = a_{10} + S_{a12} \\ A_{13} = a_{10} + S_{a11} \, A_{16} = a_{10} - S_{a12} \end{cases} \tag{5}$$

$$\begin{cases} A_{21} = a_{20} + S_{a21} \; A_{24} = a_{20} - K_{a2} \\ A_{22} = a_{20} - S_{a21} \, A_{25} = a_{20} + S_{a22} \\ A_{23} = a_{20} + K_{a2} \; A_{26} = a_{20} - S_{a22} \end{cases} \tag{6}$$

$$\begin{cases} A_{31} = a_{30} + S_{a31} \; A_{34} = a_{30} - S_{a32} \\ A_{32} = a_{30} - S_{a31} \; A_{35} = a_{30} + K_{a3} \\ A_{33} = a_{30} + S_{a32} \; A_{36} = a_{30} - K_{a3} \end{cases} \tag{7}$$

The values a_{10}, a_{20}, a_{30} are the mean of the six raw output values for X, Y and Z axes respectively.

The coefficients of the X-axis are given by Eq. 8

$$\begin{cases} K_{a1} = \frac{A_{11}-A_{12}}{2} \\ S_{a11} = \frac{A_{13}-A_{14}}{2} \\ S_{a12} = \frac{A_{15}-A_{16}}{2} \end{cases} \tag{8}$$

The coefficients of the Y-axis are given by Eq. 9

$$\begin{cases} S_{a2} = \frac{A_{21}-A_{22}}{2} \\ K_{a2} = \frac{A_{23}-A_{24}}{2} \\ S_{a2} = \frac{A_{25}-A_{26}}{2} \end{cases} \tag{9}$$

The coefficients of the Z-axis are given by Eq. 10

$$\begin{cases} S_{a31} = \frac{A_{31}-A_{32}}{2} \\ S_{a32} = \frac{A_{33}-A_{34}}{2} \\ K_{a3} = \frac{A_{35}-A_{36}}{2} \end{cases} \tag{10}$$

These coefficients when put back in Eq. 4 give us the actual values of acceleration.

C. Conversion of Acceleration to Calories Burnt

The acceleration obtained from the MEMS accelerometer after calibration needs to be converted into calories burnt so that it can be fed into the learning algorithm described in section III, which determines the fitness of the person. First, the acceleration values are converted into steps taken, which gives us calories burnt by the user.

The step count is incremented every time A_v goes past a certain threshold [6].

$$A_v = \sqrt{a_1^2 + a_2^2 + a_3^2} \tag{11}$$

The amount of calories burnt per step taken by an individual is determined with the help of Table II as shown below [7, 8].

TABLE II.

Activity	Average Number of Steps Taken Per Km	Calories/Step
Walking	2200	0.57*Weight/2200
Running	1400	0.5*Weight/1400

III. Learning Algorithm And Results

A person's susceptibility to obesity is based mainly on calories consumed and burnt by the person. Therefore, the learning algorithm deployed here has two features, calories burnt and calories consumed. Based on the two values of a certain user, the algorithm labels the user as fit or unfit (1 or 0), using logistic regression [9].

There are 150 training examples in the training set on which the logistic regression classifier is trained. Each training example has the two features as mentioned above and they are labeled either fit or unfit. Features used were not exactly calories burnt and consumed for a single day of the user, instead the features are the mean values of the calories consumed and burnt, over a span of 30 days.This way the learning algorithm incorporates the long term dietary habits and activities of the user rather than the user's activity for a single day, which might turn out to be a mere exception to the general pattern in the user's long term diet and activity. It is to be noted that factors like genetics, digestion and family history have been ignored in the following learning algorithm. The following figure, Fig. 2, is a plot of the training set used.

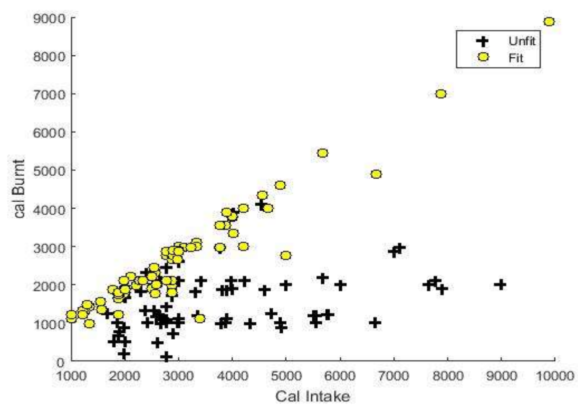

Fig.2 Training Set Plot

The plot shown by Fig. 2 has 75 examples of fit and unfit data each. It is used to train a logistic regression classifier with 1-degree hypothesis as shown below in Eq. 12.

$$H(x) = sigmoid(\theta_o + \theta_1 \times x_1 + \theta_2 \times x_2) \tag{12}$$

In (12), H(x) is the hypothesis. Θ_o, Θ_1, Θ_2 are the parameters and x_1, x_2 are the mean values of the calories burnt and consumed respectively over a span of 30 days. Sigmoid is a function defined as shown below in Eq. 13.

$$sigmoid(z) = \frac{1}{1+e^{-z}} \tag{13}$$

After training the classifier, using the training set shown, the classifier divides the training example into two parts with the help of a decision boundary which basically separates the unfit and fit training examples. The divided plot with the decision boundary is shown in Fig. 3.

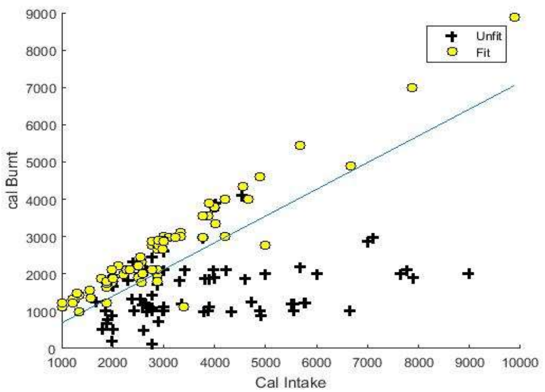

Fig.3 Decision Boundary

The decision boundary as depicted in Fig. 3 separates the unfit and fit examples in a way that most of the unfit examples are below the boundary and most of the fit examples are above it. The trained classifier when encountered with unlabeled data will simply give a probability of the unlabeled data lying in either of the two parts of the decision boundary [10-12]. This means that the user gets a probability of whether his or her dietary habits will lead them to an unfit or a fit lifestyle in the future.

The trained algorithm works with an accuracy of 85.42%. The block diagram, as shown in Fig. 4, explains the learning and prediction process.

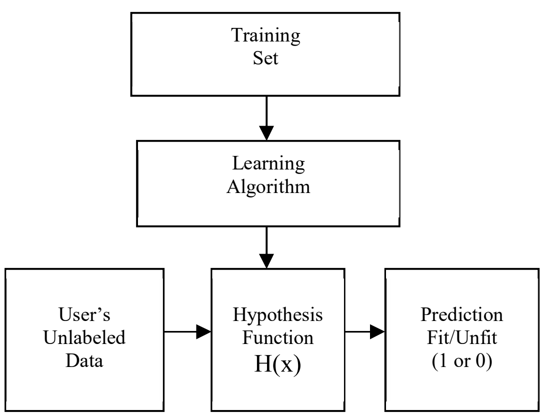

Fig.4 Flow Chart of Learning and Prediction Process

IV. Real Time Implementation

The consolidated device is a wearable band, which consists of a power supply, an Arduino Pro-Mini from which the Wi-Fi module and the MPU-6050 is connected. The Atmega328p-pu in the Arduino takes the data from the accelerometer, calibrates it, converts it into steps taken and calories burnt and sends it over to the user's phone through the Wi-Fi module, which in turn sends it to a cloud. The user inputs his or her daily calorie intake in the phone which is also sent over to the cloud where the logistic regression classifier classifies the data as fit or unfit based on the average number of calories burnt and consumed over a period of 3o days. Figure 5 gives the general structure of the device as seen by the user.

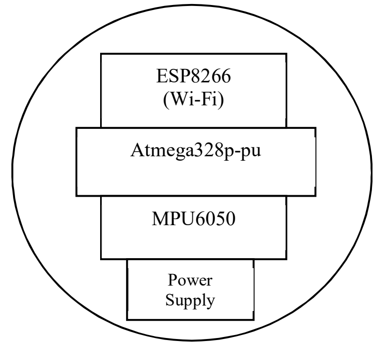

Fig.5 Device As Seen By the User

V. Conclusion

The platform presented in this paper is to be used as a wearable device to be worn on the wrist. The device has several advantages over other conventional fitness trackers, the biggest being that it can predict the advent of a condition in the future, rather than trying to detect an already existing condition. Apart from that the device also gives a percentage value or probability of a person developing a condition, because of the use of logistic regression as a classifier. Both the above advantages may also result in better engagement of the user with the device [13, 14]. Furthermore, the accuracy of this classifier can be improved drastically if more training data is available. With more specific data available about specific conditions the user can get the probability for encountering each individual condition. Apart from that a repository of calorific values of different kinds of food products can be made in the application itself so that it is easier for the user to find out calories burnt by them. The learning algorithm can be made to include recommender systems which can recommend a dietary plan for the users based on their activity patterns over the course of time in a way that will keep them fit. The calculation of acceleration is also very accurate because of the use of the six point calibration. Thus, we see how the device has tremendous scope for further research and development owing to its miniature size and low power consumption and has the potential to be a major product in the consumer electronics sector.

References

[1] WHO, Obesity and overweight fact sheet, June 2016, http://www.who.int/mediacentre/factsheets/fs311/en/

[2] Finkelstein, Eric A., "Obesity and severe obesity forecasts through 2030." *American journal of preventive medicine* 42.6 (2012): 563-570.

[3] United Nation Sustainable Development, Hunger and Food Security, http://www.un.org/sustainabledevelopment/hunger/

[4] Yu Liu, Gaolin Xiang, Yongle Lu, Yang Cao, Yunmei Li, Lin Lv, "Calibration of MEMS Accelerometer Based on Kalman Filter And the Improved Six Point Position Method", Journal of Communication Vol. 11, No. 5, May 2016.

[5] Chowdhury D., Banerjee S.J., Sanyal K., Chattopadhyay M. (2015) A Real Time Gesture Recognition with Wrist Mounted Accelerometer. Information Systems Design and Intelligent Applications. Advances in Intelligent Systems and Computing, vol 340. Springer, New Delhi

[6] L. Vinh, S. Lee, H. Le, H. Ngo, H. Kim, M. Han, and Y.-K. Lee, "Semimarkov conditional random fields for accelerometer-based activity recognition," Applied Intelligence, vol. 35, pp. 226–241, 2011.

[7] Kawahara, Yoshihiro, Nanami Ryu, and Tohru Asami. "Monitoring daily energy expenditure using a 3-axis accelerometer with a low-power microprocessor." *International Journal on Human-Computer Interaction* 1.5 (2009): 145-154.

[8] Hausmann, Jody, and KatarzynaWac. "Activity level estimator on a commercial mobile phone: a feasibility study." *Proc. Int. Workshop Front. Act. Recognition using Pervasive Sens*. 2011.

[9] Janssen, Ian, "Comparison of overweight and obesity prevalence in school-aged youth from 34 countries and their relationships with physical activity and dietary patterns." *Obesity reviews* 6.2 (2005): 123-132.

[10] S Dreiseitl, L Ohno-Machado, "Logistic Regression and Artificial Neural Network Classification Models: A Methodology Review." Journal of Biomedical Informatics Vol. 35, pp 352-259, 2002, Elsevier.

[11] B.D. Riley, "Pattern Recognition and Neural Networks, Cambridge University Press, 1996, ISBN 0-521-46086-7

[12] C. Bishop, "Neural Networks for Pattern Recognition", Oxford University Press, 1995, ISBN: 0198538642.

[13] Harrison, D., Marshall, P., Bianchi-Berthouze, N., & Bird, J. (2015, September). Activity tracking: barriers, workarounds and customisation. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (pp. 617-621). ACM.

[14] Morrison, Geoffrey Stewart. "Tutorial on logistic-regression calibration and fusion: converting a score to a likelihood ratio." Australian Journal of Forensic Sciences 45.2 (2013): 173-197.