Machine learning enabled food journal

1 Sadab Qureshi, 2 Bradford Turcott, 3 Mohammed Athar, 4 Arjun Dubey

Dept. of Software Engineering, San Jose state University
Email Id:

sadab.qureshi@sjsu.edu
bradford.turcott@sjsu.edu
mohammedathar.na@sjsu.edu
arjun.dubey@sjsu.edu

***Abstract* -- In today's world, people are very health conscious and they want to remain fit. One of the main aspects of improving health is maintaining weight which is closely related to total calorie intake and burned by an individual. This problem offers a need for a tool that not only enables users to log foods but also helps individuals to evaluate their daily calorie surplus or deficit and they should be able to know how much they would weight in future.**

Our solution leverages image classification to identify foods and log nutritional info into user's journal. This simplifies calorie counting and provides a dataset to run analytics. Future improvements include pulling in health and activity data from wearables and smart scales. Predictive analytics could then be used to estimate weight loss and health based on real-time nutrition and fitness data. The application would leverage the cloud to score images using pre-trained DNNs, and store user's data. A smartphone based front end application would be required.

With help of the solution, individual will not only know about their daily calorie consumption and calculate the calorie burned in the day but also get predicted weight loss or gain for a future date.

***Keywords*-- Calories intake, Food logging.**

1. Introduction

Machine learning enabled food journal is an application which envisages towards enabling individuals to log what they eat and keep track of physical activities like activity level, sleeping hours, calorie burned etc. It enables users to take snap of the food and log the food to the database. It connects to Fitbit to pull the activity data which is used by the machine learning module to evaluate which feature impact his or her weight.

The application presents all the meaningful data to the user to understand the factors which positively or negatively affect their weight and which don't. Users get to evaluate all the factors and quantify their impact.

2. Analysis

Fitness trackers, fitness monitoring applications, low or zero fat food products, varieties of fitness centers are everywhere to see or talked about in the fitness world. Of all it is the fitness trackers which is getting the most buyers due to its proximity/availability to the user. Individuals can carry it with them all the time and it gives them some sense of control over their health via monitoring some of the physical expect of the body. Seeing this we decided to understand why these fitness trackers are being used and are they serving its purpose.

A. Need

Do we really need gadgets in our life like these fitness trackers or applications? This was the first we asked, and we found out many reasons and the primary most definitive reason was to find a way to track the calorie burned count so that a user can set a target such that he or she lose weight. Today due to change in lifestyle and availability of technology or facilities people eat more calories and burned very less of it, due to this imbalance most of the population is facing problem of obesity. Hence, people are trying to find ways to not be obese or in other terms not to become overweight. So yes these gadgets do provide a way to keep track of calories.

B. Options in market

There are lot of players in market already like Fitbit, Garmin, Xiomi, Apple, Samsung etc. catering the need of counting calories burned, manual food logging, sleeping time and other functionalities. The wearable fitness tracker helps you check current and historical data which are mostly accurate and does gives insights of the activities.

C. What we propose?

As captured in part A and B that we need trackers to help people monitor their health and there is great market for such devices. Also there are already big players catering the market, but when we see the primary need i.e. weight loss and what functionalities being offered we can see the direct disconnect. Users doesn't get insights which could guide them in taking actions that could accentuate weight loss or target the areas which strongly impacts losing of weight. Our applications would help users to easily log the food that they eat, this way user's calorie intake would get logged. Application also pulls the user's data from Fitbit like application, such as calorie burned and sleeping hours. Using all the above data and weight logged by user, machine learning module will draw the relationship between parameters and actual weight loss/gain and present it to user. This relationship will present clear view on the parameters on which user could work on to achieve his/her goal.

D. Are we achieving anything?

Using the camera (under future enhancement) user takes the picture of the food and with google vision API, application identifies the fruit and pulls the calorie count from 3rd party API which can be logged. This makes the food logging convenient.

Fig. 1 A sample food item

With the data from Fitbit and food logging and manual weight logging, we build training set for the machine learning module.

Below is the sample 5 entries of the training set

	Calorie_Intake	Calorie_burned	Hours_of_Sleep	Weight_loss
1	1811	1951	7	0.039989
2	1818	1996	8	0.051402
3	2007	1969	8	-0.011018
4	1791	1942	7	0.042026
5	1698	1935	6	0.067540

Fig. 2 Training set with 3 features

Features

- Calorie_Intake: Total amount of calories intake per day (Calculated by User food log)
- Calorie_burned: Total amount of calories burned per day
- Sleep: Total sleep duration per day

Output

- Weight loss: Measured change in weight in a day

From the below visualization user can see how the features impacts weight loss. This clearly shows that features Calorie_Intake and Calorie_burned significantly impact weight loss or gain whereas sleeping hours has no significant relationship with it.

Visualize the relationship between the features and the response using scatterplots

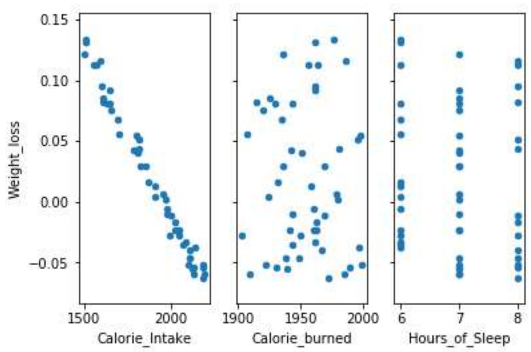


Fig. 3 Scatterplot with 3 features

Further information is also available to summarize how and to what extent features impact the output here weight gain or loss.

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-0.0347	0.024	-1.464	0.150	-0.082	0.013
Calorie_burned	0.0003	1.22e-05	24.427	0.000	0.000	0.000
Calorie_Intake	-0.0003	1.41e-06	-198.597	0.000	-0.000	-0.000
Hours_of_Sleep	-0.0005	0.000	-1.164	0.250	-0.001	0.000

Summary of fitted model

We can go further and add as many features as available and see the relationship. As shown below we added two more features based on diet i.e. fat diet or carbs diet.

Visualize the relationship between the features and the response using scatterplots

	Calorie_Intake	Calorie_burned	Hours_of_Sleep	Weight_loss	Diet_Carbs	Diet_fat
1	1796	1940	6	0.041365	1	0
2	1856	1923	7	0.017938	0	1
3	1822	1907	7	0.023377	0	1
4	1950	1902	7	-0.012662	0	1
5	1909	1927	7	0.004747	0	1

Fig. 4 Scatterplot with 3 features

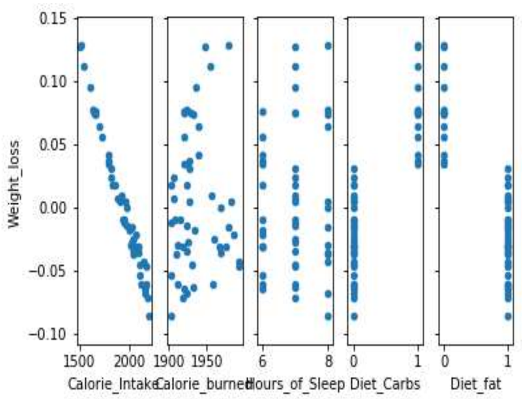


Fig. 5 Scatterplot with 3 features

3. Architecture

A. Database and dataset

We are logging the food in MySQL database and using Fitbit dataset to pull the data from to be used by machine learning module. Nutrinics maintains extensive database for foods which is being used to get calories count for the identified food.

B. Authentication

This project involves user key authentication for Fitbit API of OAuth2 type. Application in the back end uses Google Cloud Vision API project, so the application is authenticated by Google Cloud Platform API service is by using Application Default Credentials (ADC)

C. Server implementation

We are using Flask python web framework for image recognition and calorie count. Tomcat application server with spring boot application to connect to Fitbit and pull data from.

D. Data analysis and Building Model

We are using Jupiter notebook to demo and share the python code, where we use pandas package for providing expressive data structures. Matplotlib is being used to do scatter plot. Statsmodels api to generate fitted models that helps in calculating the weights for the features as well as intercept. Statsmodels also provide summary which shows the positive or negative association as well as how it relates to output and could be used for feature selection.

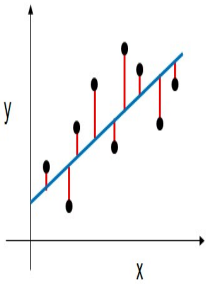
Scikit-learn library is used to apply linear regression on the training set to build the model and which is being used to predict future weight loss/gain.

Simple linear regression model take the form

$$y = \alpha + \beta x,$$

And then we use the algorithm to fit the model using least squares approach which minimizes the sum of squared residuals. In other words, α and β solve the following minimization problem:

$$\text{Find } \min_{\alpha,\beta} Q(\alpha,\beta), \quad \text{for } Q(\alpha,\beta) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2$$



$$SS_{residuals} = \sum_{i=1}^{N} (\hat{y}_i - y_i)^2$$

Fig. 6 Model diagram and cost function

E. Application deployment

Applications have been configured and deployed on Amazon EC2 ubuntu instance. The application could be scaled up with demand.

4. Future Enhancement

We need to integrate the application with mobile application. Machine learning needs to upgrade to capture features like identifying diets based on the user eating habits.

We can build a meal recommender system to suggest the next meal based on the current calorie count and daily calorie goal limit. Since we are capturing food user eats then we can suggest food which fits on calorie consumption as well as what user likes to eat.

5. Conclusions

The application successfully captures the image, process it and uses data from food database to calculate the calorie intake. It uses the data from wearables to fetch physical activity data. Machine learning module of the application uses the available data produce insights and presents it to the user.

6. References

[1] Google vision documentations https://cloud.google.com/docs/

[2] Fitbit API documentation https://dev.fitbit.com/docs/

[3] Resource to use jupyter and machine learning http://nbviewer.jupyter.org/github/justmarkham/DAT4/tree/master/notebooks/