



**San José State**  
UNIVERSITY

# **Hadoop based Obesity Monitoring System**

**Course Name :CMPE272, Team Name: Watson**

**Submission Date: May 9<sup>th</sup>, 2011**

**Team Members:**

**Roopa Govindarajan, Smarth Madan, Anupama Shetty**

## Table of Contents

1. Abstract:.....	3
2. Goal: .....	3
3. Introduction:.....	4
3.1 What is Obesity?.....	4
Figure 3.1.1: Trends in obesity among Children and Adolescents (US: 1963-2008) [1].....	5
3.2 How to Measure Obesity?.....	5
3.3 Factors affecting Obesity:.....	5
Figure 3.1.2 : The Energy balance is defined as the equilibrium between food intake and energy expenditure. ....	6
3.4. How is obesity monitoring achieved by our system?.....	7
4. Proposed Solution for Self Monitoring.....	7
4.1 System Details:.....	7
Figure 4.1.1: System Architecture for Self Monitoring Application using Hadoop to lower Obesity.....	8
4.2 Technologies Used :.....	9
4.3 Basic Work flow of the Application:.....	10
Figure 4.3.1: Basic Work flow of the Application.....	10
5.0 Use Cases .....	11
5.1 Use Case 1: .....	11
5.2 Use Case 2:.....	11
5.3. Use Case 3:.....	13
6. Conclusion:.....	13
7.0 References:.....	14

## 1. Abstract:

In the United States, Obesity is a serious health concern for children and adolescents. Results from the 2007-2008 National Health and Nutrition Examination Survey (NHANES), using measured heights and weights, indicate that an estimated 17 percent of children and adolescents ages 2-19 years are obese. Obese children and adolescents are at high risk for cardiovascular problems such as high blood pressure, high cholesterol, and Type 2 diabetes. Other consequences include Psychosocial Risks, Cardiovascular Disease Risks, Sleep Apnea issues etc which can deter a particular individuals overall growth and development.

As obesity is now becoming like an epidemic especially in the US, there is a heavy need to monitor one's individual growth especially parameters like **daily calorie intake, activity levels, body weight,height,BMI** etc in order to achieve constant monitoring and make serious changes to the lifestyle to lower the risk of Obesity. Currently, there is no one stop solution that provides an integrated approach to actively track the parameters mentioned above and make recommendations using Predictive Modeling based on the current input of an individual. Such high risks and the increasing rate of obesity calls for a way to monitor and control obesity.

## 2. Goal:

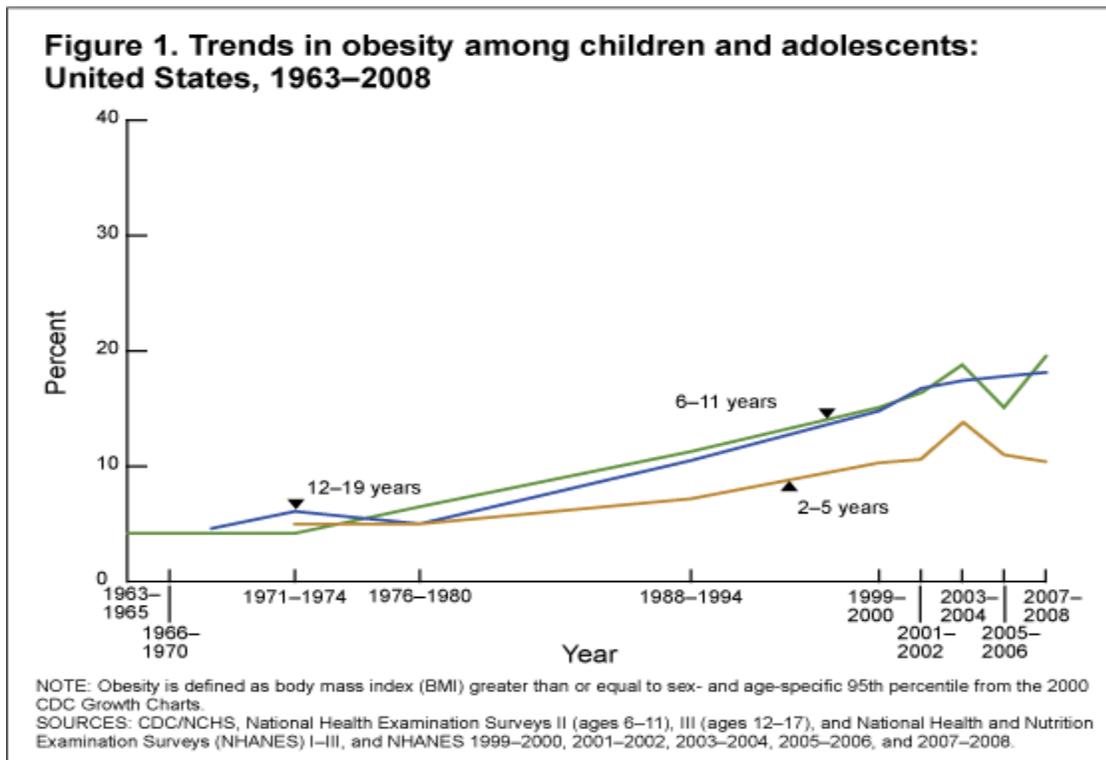
The Goal of this paper is to come with a **self monitoring solution** or an application that actively tracks an individual food intake and exercise levels and makes recommendation based on the inputs to lower the risk of Obesity. The initial emphasis will be to lower obesity in children. We plan to then extend this to Adults as a future road map for the application. Also, this application will be designed to support multiple platforms like online web, Smart phones (iPhone, Android etc). The intent is to make this information available both for the individual child as well as their respective parent so they can actively monitor their child's risk for Obesity at any given time with 24X7 accesses. Also, respective physicians can be given access when need be. Hence, this can create collaboration between the child, its parent and their physician to actively monitor the risk of obesity.

We have designed a system which enables easy monitoring of obesity level in any individual and also provides suggestions on the control measures that can be taken to control the obesity level. Energy balance is one of the primary reasons that lead to the accumulation of adipose tissues in body which in turn leads to weight gain in the form of obesity. Our research paper provides details about the system design and the various components used to achieve a monitoring system that helps in controlling obesity. Various use cases have been explained to provide an introduction to the various areas across which the system can be utilized in a very efficient way.

### 3.Introduction:

#### 3.1 What is Obesity?

**Obesity** is a medical condition in there is excess body fat accumulated in an individual which can pose adverse effect on health, leading to reduced life expectancy and/or increased health problems.



### **Figure 3.1.1: Trends in obesity among Children and Adolescents (US: 1963-2008) [1]**

As shown in **Figure 3.1.1**, Obesity is a serious health concern for children and adolescents in United States. Results from the 2007-2008 National Health and Nutrition Examination Survey (NHANES), measured based on heights and weights, indicate that an estimated 17 percent of children and adolescents ages 2-19 years are obese. Between 1976-1980 and 1999-2000, the prevalence of obesity increased. Among pre-school age children 2-5 years of age, obesity increased from 5 to 10.4% between 1976-1980 and 2007-2008 and from 6.5 to 19.6% among 6-11 year old. Among adolescents aged 12-19, obesity increased from 5 to 18.1% during the same period. Obese children and adolescents are at risk for health problems during their youth and as adults. For example, during their youth, obese children and adolescents are more likely to have risk factors associated with cardiovascular disease (such as high blood pressure, high cholesterol, and Type 2 diabetes) than are other children and adolescents. Other consequences include Psychosocial Risks, Cardiovascular Disease Risks, Sleep Apnea issues etc which can deter a particular individuals overall growth and development.

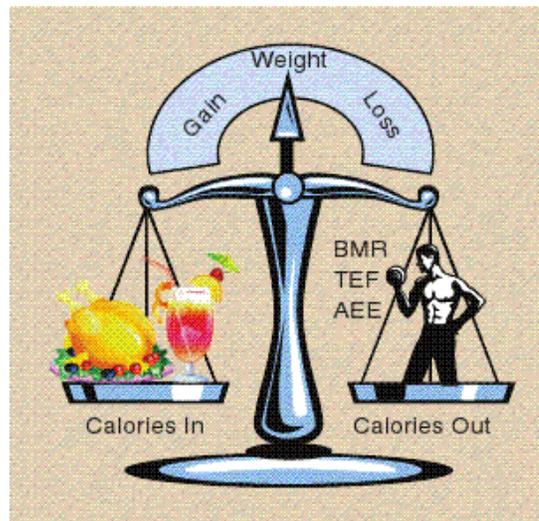
### **3.2 How to Measure Obesity?**

Body Mass Index (BMI) and body fat percentage are two indices used to determine obesity and overweight among individuals. Obesity is defined as excessive body fat while being overweight refers to excessive weight in comparison to height. It is however difficult to measure body fat percentage precisely and therefore BMI is the most commonly used index for measurement. BMI is defined as the weight of an individual over square of his/her height. Individuals having BMI value greater than 25 and 30 are termed as obese and overweight respectively.

### **3.3 Factors affecting Obesity:**

Obesity may be caused by various reasons related to physical, genetic, psychological or even behavioral issues. One of the fundamental reasons for obesity is excessive accumulation of adipose tissue in human body. Adipose tissues are commonly known as the fat tissue and have the primary function of providing heat insulation and to act as a source of energy. These tissues act as buffer for energy storage but their excessive storage leads to fat accumulation and thus results in weight gain. Energy imbalance is seen to cause excessive accumulation of white adipose tissues in body. Energy imbalance or the gap in energy is caused when the amount of energy intake and energy expenditure component are not equal. Thus one measure to control the

obesity levels seen in individual is to reduce the energy gap and maintain a considerable balance in body by monitoring.



**Figure 3.1.2** : The Energy balance is defined as the equilibrium between food intake and energy expenditure.

Source : A Review: Monitoring Energy Intake and Energy Expenditure in Humans. *Engineering in Medicine and Biology Magazine, IEEE*. Vol 29 (1). DOI: 10.1109/MEMB.2009.935470

As seen in the research conducted by the team of Sazonov.E.S and Schuckers.S, energy intake is provided by the release of chemical energy processed by the digestive system. Thus the energy intake for a body depends on the type of food consumed and can be controlled by making the right choice of diet. Energy expenditure on the other hand is composed of three primary components called the basal metabolic rate (BMR), thermal effect of food (TEF) and activity energy expenditure (AEE). BMR being the major contributor for energy expenditure refers to the energy spent to maintain body's temperature, cardiovascular, nervous and respiratory functions, and other vital functions of cells, tissues, and organs. TEF plays a minor role and refers to the energy spent on processing the food intake and storage for utilization by body. Proteins for example have higher TEF than fat and this indicates that a diet having high protein content helps in maintaining energy balance. AEE being the third factor refers to Exercise Energy Expenditure (EEE) and the Non Energy Expenditure Thermo Genesis (NEAT). NEAT could be considered as a result of the advancement in technology seen today. For example, an hour of sitting in front of a computer or a laptop could result in a net energy gain of nearly 100k calories.

Many researches have been conducted to determine the ideal energy gap to be maintained but no single answer has been found yet. Based some on the researches conducted in “Estimating the changes in energy flux that characterize the rise in obesity prevalence,” and “The magnitude of the energy imbalance in obesity is generally underestimate,”, an energy gap in the range of 300-400kcal/day in individuals with a BMI of 30 is seen to cause obesity. Estimating ideal energy gap for children is more complicated as the children need additional energy intake for their growth. Thus various parameters needs to be analyzed to detect the ideal energy gap.

### **3.4. How is obesity monitoring achieved by our system?**

The obesity monitoring health care system designed by us focuses on calculating the specific parameters involved in calculating the energy balance. This is done by providing the user with options to enter his/her personal details such as age, sex, height and weight and by collecting details about the various routine activities. BMI for every individual is calculated based on the initial profile inputs and stored in the system. To assess the energy intake component, diet details will have to be entered with the food item name and the quantity consumed. And to assess the energy expenditure, details of various physical activities performed by the individual such as jogging, walking will have to be entered into the system along with duration details. These data lead to the calculation of the EEE factor described above. Monitoring the other components such as TEF and BMR requires controlled laboratory environment for accurate value calculations. Hence our system focuses on the other components involved for the energy gap calculation which are the EEE and the energy intake. Energy gap calculation performed by our system is based on the formula stated below which involves energy intake and the energy expenditure components.

#### **Energy Gap = Energy Intake - Energy Expenditure**

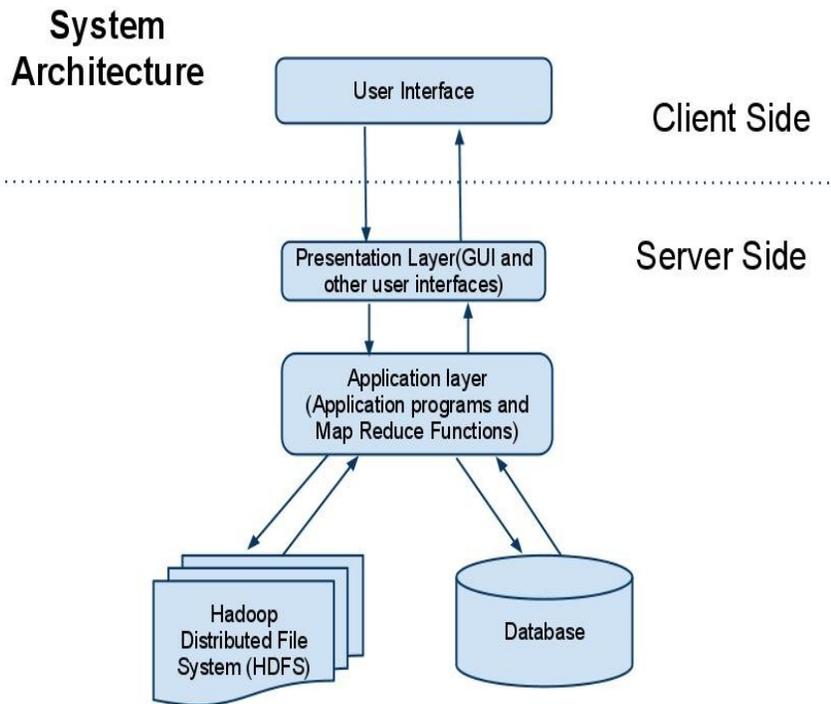
The system provides the options to either enter the specific values for the BMR and TEF components based on controlled laboratory made experiments or the system calculates the energy expenditure primarily based on EEE with some default values for the other components. The resulting energy gap will be further analyzed by the system based on the personal physical details entered by the user into the system. Suggestions or recommendations are then provided to the user based on the resulting value and may include specific physical exercises or the diet schedule required to improve the energy balance in body.

## 4. Proposed Solution for Self Monitoring

### 4.1 System Details:

A personalized obesity monitoring system has been designed by our team as a solution for obesity control. The system offers the personalized profile creation feature for every user which allows the user to store his physical record details as mentioned in the above section. Based on the entered data, system calculates the BMI and performs various predictive analysis resulting in calculation of parameters which may directly or indirectly effect the individual's obesity level. The system can be used either at an individual level or at a larger organizational scale like in a hospital or a school. One application which can greatly help in reducing the obesity level seen in large proportions in children today is by making the system available at schools. Teachers or dedicated professionals could use the system to monitor the obesity level of every individual children by monitoring their food items during recess and lunch time with cooperation from their parents.

The system consists of two primary storage resources namely HDFS(Hadoop Distributed File System) and MySQL database. These data sources store a list of food items and their corresponding calorie values per serving. Data analysis is performed by the Map Reduce module of the system. The various system components are as displayed in Figure 4.1.1. The various technologies used as part of the system is explained in the following section in detail.



**Figure 4.1.1:** System Architecture for Self Monitoring Application using Hadoop to lower Obesity

## 4.2 Technologies Used :

Below is the description of the various technologies used within the system developed by us.

### Web Services

Web Service is an application component to expose the services to the world of Internet users. It works on text based SOAP protocol. We can use web services to manage users to interact with our system for account management, data input and ultimately the analyzed data result. Once our application is deployed and published on the web, It could be accessed from Internet from anywhere.

The school authorities could use it for viewing the data at school level rather than individual child level, whereas the parents can access the data of their child food intake in the school. This would provide a level of abstraction to various users depending on their roles.

## **HDFS**

Hadoop Distributed file system is a distributed file system on which we can run jobs by distributing the load onto several clusters. This file system is fault tolerant i.e. if a job fails on a particular cluster, the HDFS replicates it on another cluster and the execution of the job is not halted. This makes the system quite reliable. In our application we are using HDFS to store the user data from various sources which could be later reduced and analyzed to give the right results back to the user. Input data can be stored in simple text files which are easier to handle and retrieve from the user. This eliminates the use of costly and bulky database solutions available in the market by several vendors at high prices.

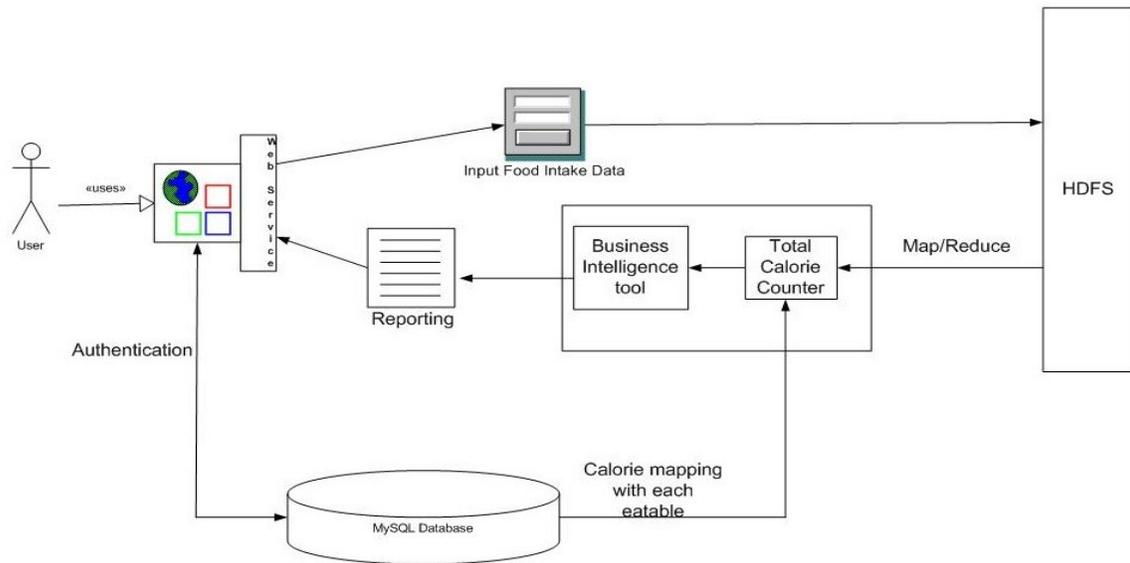
## **MapReduce**

MapReduce jobs can be scheduled on the data present on HDFS to get very specific output data required for analysis. The main advantage of using MapReduce jobs in our application would be that we could reduce a huge chunk of data which could be of size in Terabytes and a guaranteed output with the help of HDFS fault tolerant architecture. Our application would record the food intake of a child in a school located in particular city. We could use MapReduce jobs to reduce the data of total food items computed and their respective quantities. Depending on the input parameter, that could be the name of child, School name, City, State, or Country we can reduce the data to any level for further analysis.

## **Database**

Database would be a key feature where we can store all the static data such as User information and credentials. Also the calorie count of each known eatable available in the school would be stored in the database which would be used to sum up the total calorie intake. Since the data to be inserted in our database would be small as compared to constant input stream coming into HDFS and it does not require frequent updates, we can use any of the commercially available databases. For our application we have chosen MySQL database which is quite popular and open source as well.

## **4.3 Basic Work flow of the Application:**

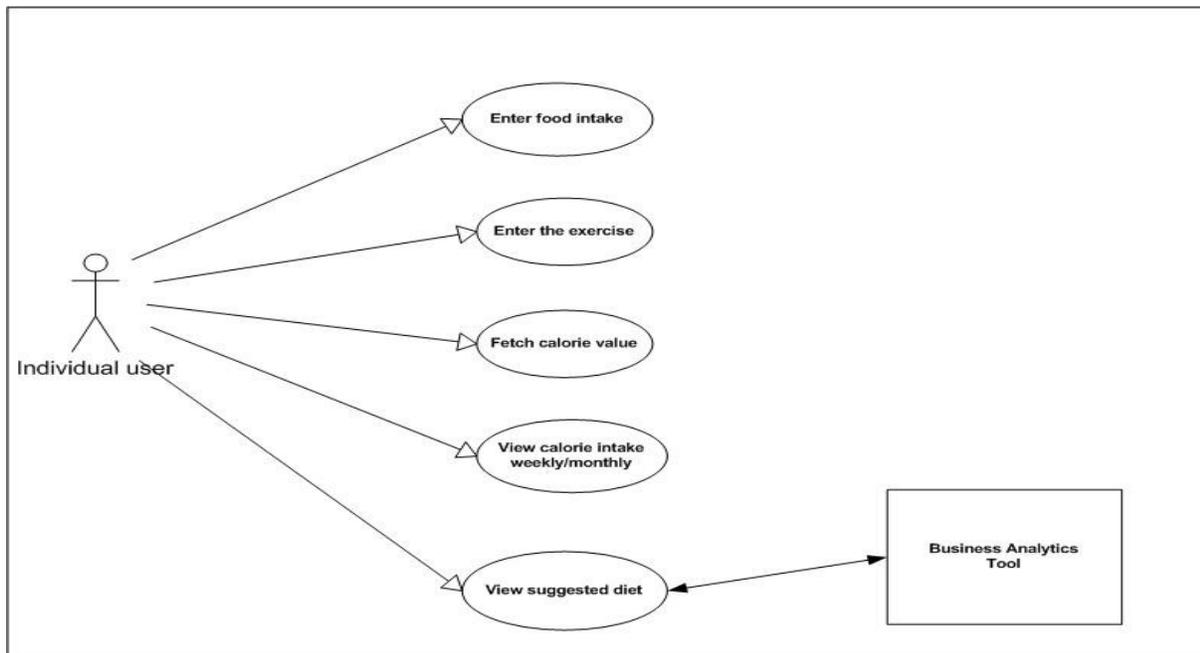


**Figure 4.3.1: Basic Work flow of the Application**

- User access the system either from a Smart phone or online using the Web.
- User then inputs daily calorie intake which includes values from Breakfast,Snack,Lunch,Snack and Dinner
- User also enters what kind of exercise he/she has done for the day .
- Data entered reaches the HDFS component and acts as an input to Map reduce component.
- MySQL database to used to fetch the static calorie mapping to the food item entered by the user.
- Total calorie intake is calculated which acts as an input to the energy gap calculation. This energy gap determined is further compared with expected and normal energy gap.
- Based on the intelligence tool function, the output indicates the current status for the user and provides suggestions regarding the diet and exercises to control the obesity level.

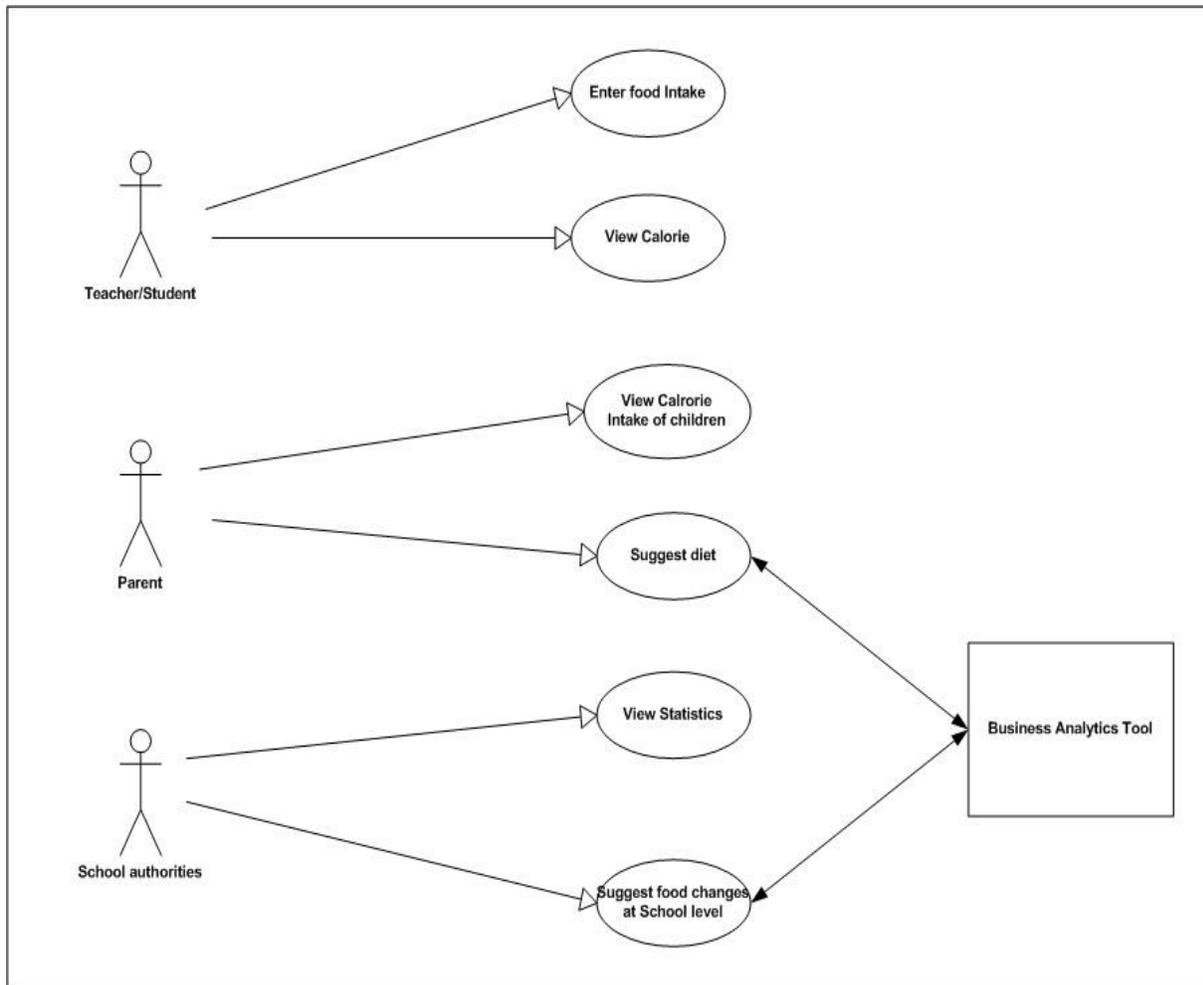
## 5.0 Use Cases

### 5.1 Use Case 1:



Any individual can use the system to monitor his personal obesity level. It can be used on mobile devices, or browser to increase the availability of this app. The person can then input his details such as the food intake he had and also see the calories value of the food item before consuming it. Secondly the person can input the amount of physical activity he is going through except for normal day to day routines. The system would use it measure the actual calories which are left for body to use after subtracting the portion of exercise from food intake. At the end of the day or week or month , the system can show him his report which would clearly state his Energy gap and also what that energy gap means. If this energy gap is more, it could result in adding fat to the person’s body. Hence the system would also suggest the type of food that should be cut down from his diet or what amount of increase in physical activity would be good for the individual.

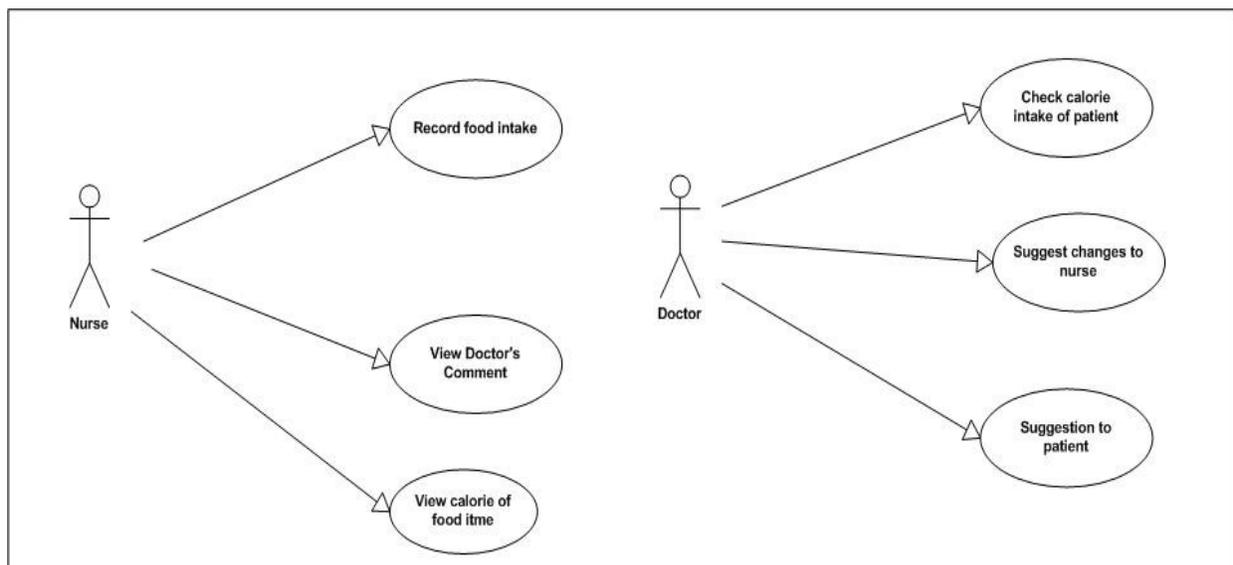
## 5.2 Use Case 2:



In schools the application can be used in canteens and mess to keep a track of food items consumed by children. At the end of the month, School authorities can access the system to monitor the obesity level of each student by creating and maintaining separate profiles for every student. These reports can also be provided to parents. Also the report can be consolidated to provide the complete picture of what food items are being consumed and in what quantities and what is average calorie consumption of the children. If it is high, the school authorities can take preventive measures to regulate the food served and also the physical activities of the students. This can help track and control the obesity causing factors within the school premises. Regular monitoring of the meal items consumed by every student helps in determining the suitable diet fit for majority students and referring students with obesity concerns to dedicated physicians with cooperation from their parents. With the use of HDFS in the system, huge amount of data can be stored and analyzed with ease.

### 5.3. Use Case 3:

The system can also be used in Hospitals, to monitor the food intake of the patients. Nurses can record the food intake of patient and doctors can view this data while sitting in their office on their PDA or web browsers. The doctors can also provide suggestions or recommendations that are to be followed by the nurses. The system can provide doctors and nurses at hospitals an easy way to monitor patients at regular intervals. With the use of web services, the system can access the patient records available in the hospital database to provide personalized suggestions and guidance in controlling the obesity levels or general health details.



The system can also be used in Hospitals, to monitor the food intake of the patients. Nurses can record the food intake of patient and doctors can view this data while sitting in their office on their PDA or web browsers. The doctors can also provide suggestions or recommendations that are to be followed by the nurses. The system can provide doctors and nurses at hospitals an easy way to monitor patients at regular intervals. With the use of web services, the system can access the patient records available in the hospital database to provide personalized suggestions and guidance in controlling the obesity levels or general health details.

## 6.Conclusion:

This paper embarked on the issue of Obesity which has become a global issue for all of us living in the 21st century today. Realizing the urgency of this issue, our team designed a self monitoring tool(using Hadoop and MapReduce framework)that allows individual to track their calorie and exercise levels and make recommendations using Predictive Modeling to lower their risk of Obesity. This tool is so versatile that it can be downloaded as a Mobile application or it can be directly accessed via the Online web. Also, potential users can vary from individual users to children at various schools. We are confident that users of our Tool can lower their risk for Obesity by incorporating it's usage as part of their daily routines. Further extensions can be added to the system to make it usable at a larger scale and for wider number of people. The system could also be enhanced with capability to monitor other components involved in the energy expenditure process.

## 7.0 References:

- 1.Working Group Report on Future Research Directions in Childhood Obesity Prevention and Treatment.(August 21-22, 2007). National Health Lung and Blood Institute. Retrieved from <http://www.nhlbi.nih.gov/meetings/workshops/child-obesity/index.htm>
- 2.Retrieved from ubergizmo.com(May 5<sup>th</sup>, 2010). <http://www.ubergizmo.com/2010/11/zamzee-activity-tracker-hopes-to-combat-obesity-in-children/>
- 3.Health Advantage(1996 - 2009 ).Retrieved from <http://www.ntwrks.com/~mikev/chart1.html>
- 4.Tsai.C.C & Lee.G & Raab.F & Norman.G.J & Sohn.T & Griswold.W.G & Patrick.W. (2007).Usability and Feasibility of PmEB: *A Mobile Phone Application for Monitoring Real Time Caloric Balance. Mobile Netw Appl (2007) 12:173–184.*DOI 10.1007/s11036-007-0014-4.
5. Sazonov, E & Schuckers, S.(Jan-Feb, 2010). The Energetics of Obesity: A Review: Monitoring Energy Intake and Energy Expenditure in Humans. *Engineering in Medicine and Biology Magazine, IEEE*. Vol 29 (1). DOI: 10.1109/MEMB.2009.935470.
6. Ershow.A.G & Hill.J.O&Baldwin . J. T(September, 2004).Novel Engineering Approaches to Obesity, Overweight, and Energy Balance:Public Health Needs and Research Opportunities. *Proceedings of the 26th Annual International Conference of the IEEE EMBS.*

7.Landi.A & Piaggi.P & Lippi.C & Santini.F & Pinchera. A(2010).Statistical Toolbox in Medicine for Predicting Effects of Therapies in Obesity.IEEE. DOI: 978-1-4244-4998-9/10