

Accurate Decision Support System for Food Nutritional Analysis

Sandani Ishanki Kumarasinghe

169316C

Master of Science in Information Technology

Faculty of Information Technology

University of Moratuwa

Sri Lanka

February 2019

Accurate Decision Support System for Food Nutritional Analysis

Sandani Ishanki Kumarasinghe

169316C

Thesis submitted in partial fulfillment of the requirements for Master of Science in
Information Technology

Faculty of Information Technology

University of Moratuwa

Sri Lanka

February 2019

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Name of Student: S. I. kumarasinghe

Signature of Student:

Date:

Supervised by

Name of Supervisor: Mr S. C. Premaratne

Signature of Supervisor:

Date:

Acknowledgements

I would like to thank my supervisor, Mr S C Premaratne, for the patient guidance, encouragement and advice he has provided throughout my project, making it success. I have been extremely lucky to have a supervisor who cared so much about my project, and who responded to my questions and queries so promptly.

I'm grateful to my husband and my parents, for their enormous support given to me and who experienced all the ups and downs of my research.

Finally, I thank all the staff members of Faculty of Information Technology, University of Moratuwa.

Abstract

Calculate nutrient intake from food consumption details is an uninteresting, time consuming, error-prone task when the data is fed by manually. People check the label of foods for various purposes. Many customers would like to know how to use this food label details more easily as well as effectively. So, there should be an easy methodology to identify users who need nutritional care as well as with measurements of nutritional facts. Food nutrients appear together in foods as well as dietary patterns, and the general effects of dietary choices according to patterns are not well understood. Overweight and obese are major problems in the world. Obesity occurs from weight gains over long time. Therefore, correct and accurate nutritional analysis is very significant to the current society. Personal nutrient guide by calculating nutrient intake, predict nutritious foods/meals by considering the health and willingness, predict an exercise schedule for the user and forecast health condition for individuals are the main modules of this research project which are made accurate decision support system for food nutritional analysis.

According to implementation of this decision support system for food nutritional analysis, using Fatsecret API calls, fetched food items with their nutrient amounts as well as exercise details. Client can be a registered user and enter his weight, age and sex with his health condition to have a diet plan. Not only that if someone need to lose his weight or gain his weight, there also diet plan and exercises schedules. Admin can access these records and keep history data to get the most frequent food items. For the history data, use the advanced data generator for MySQL, populate thousands of records to develop the data set for associate data mining.

Outputs are when a user has two diseases, then recommended frequent food items for both diseases. Not only that, if a user wants to lose weight or gain weight, then give suggestions for relevant food items and give three or four exercises with time durations to control his body weight.

Keywords— nutrient intake, dietary patterns, nutritional analysis

Table of Contents

Declaration	i
Acknowledgements	ii
Abstract	iii
List of Figures	vii
Chapter 1	1
Introduction	1
1.1 Chapter Introduction	1
1.2 Background	1
1.3 Problem.....	2
1.4 Aim and Objectives	3
1.5 Chapter Summary	3
Chapter 2	4
Review of Literature	4
2.1 Chapter Introduction	4
2.2 Nutrient analysis methods and limitations	4
2.3 Chapter Summary	9
Chapter 3	10
Technology Adapted	10
3.1 Chapter Introduction	10
3.2 MySQL	10
3.3 CodeIgniter Framework	10
3.4 Unsupervised Learning	10
3.5 Weka Tool.....	11
3.6 Advanced data generator for MySql	12
3.7 Chapter Summary	12
Chapter 4	13
My Approach	13
4.1 Chapter Introduction	13
4.2 Fetch data and process to get most frequent food items	13
4.3 Outputs.....	19
4.4 Chapter Summary	19

Chapter 5	20
Analysis and Design	20
5.1 Chapter Introduction	20
5.2 Data Collection	20
5.2.1 Keep user Details	20
5.2.2 Retrieve relevant data by calling Fatsecret API	20
5.3 Decision making Process	23
5.3.1 Frequent Food items	23
5.3.2 Recommend a food plan for two identified non-infectious diseases	23
5.3.3 Using exercises and recommended food plan, achieve weight gain or weight loss	27
5.4 Chapter Summary	28
Chapter 6	30
Implementation	30
6.1 Chapter Introduction	30
6.1 Collection of Data and basic functionalities	30
6.1.1 Fetch food details based on food category	30
6.1.2 Create user accounts	31
6.1.3 Predict food plan using most frequent food items	31
6.1.4 Admin access of user records	33
6.1.5 Calculate daily calorie intake	34
6.1.6 Give an exercise schedule	34
6.1.7 Facilities for users	35
6.2 Special functions	36
6.2.1 Recommend a food plan for two identified non-infectious diseases	36
6.2.2 Using exercises and recommended food plan, achieve weight gain or weight loss	36
6.4 Chapter Summary	37
Chapter 7	38
Evaluation	38
7.1 Chapter Introduction	38
7.2 Test results of the process of weight loss	38
7.3 Chapter Summary	39
Chapter 8	40
Conclusion and Further Works	40

8.1 Chapter Introduction	40
8.3 Further Works	41
8.4 Chapter Summary	41

List of Figures

Figure 2.1: Nutrient content claims.....	9
Figure 4.1: Fatsecret food items with food_ids	13
Figure 4.2: Exercises with burned calories for specific time duration	14
Figure 4.3: History data records of several users.....	15
Figure 4.4: Food details of users.....	15
Figure 4.5: Health condition and food details taken by users.....	16
Figure 4.6: Set values for associate rule mining using Apriori Algorithm.....	19
Figure 5.1: Data integration	21
Figure 5.2: Food related Data from Fatsecret API.....	21
Figure 5.3: Exercise related Data from Fatsecret API.....	22
Figure 5.4: Nutrient content claims.....	23
Figure 5.5: Generate email addresses of users.....	25
Figure 5.6: Generate passwords of users	25
Figure 5.7: Generate nicknames of users	26
Figure 5.8: Generate age of users	26
Figure 5.9: Populated data records for specific user by considering meal and date ..	27
Figure 5.10: High Level design of the proposed system depicting the flow of steps	29
Figure 6.1: Food list with nutrient details	30
Figure 6.2: User accounts	31
Figure 6.3: Frequent foods after applying associate mining.....	32
Figure 6.4: Features of recommended food list page.....	33
Figure 6.5: Calculation of daily calorie intake.....	34
Figure 6.6: Exercise schedule for relevant customer	34
Figure 6.7: Food plan with calorie amount of certain criteria	35
Figure 6.8: Food plan details of a person who need a weight loss	35
Figure 6.9: Filtering recommended foods for 2 diseases	36
Figure 7.1: Weight over the date for a user who is willing to weight loss	38
Figure 7.2: Weight loss progress with the indicator of exercises are done or not	39

Introduction

1.1 Chapter Introduction

This chapter provides as an introduction to the study explaining the background which cause to the research of this study. Problem definition, aim and objectives are clearly mentioned in this chapter.

1.2 Background

Making the correct healthy food choices are essential part of our day-to-day life. As we know, Nutrition facts are needed to make healthy food choices. Food nutrition labels help in choosing a healthy foods. If someone needs to eat less fat, salt or sugar, the food labels can directly help to make decisions. Details of the amount per serving of nutrients such as cholesterol, sugar, and fiber, very important for people who has different health concerns. By comparing the nutrition labels, customer can determine which food is more nutritious.

Checking the ingredients list will help to avoid foods that contain ingredients which should be avoided. Maintaining a healthy weight is a big problem in current society. People who are trying to lose weight for health reasons have to consumed fewer calories, more fiber and less saturated fat, sugar and carbohydrates.

Although nutrition facts labels can be very helpful, not everyone understands the information included of it. People who had low levels of literacy and numeracy didn't understand these labels. Therefore people don't consume enough fiber, iron, calcium, and vitamins A and C[1]. Too much weight feels uncomfortable, and also damage to the health. Adults were considered obese by considering body mass index (BMI) of 25 or higher. Limit the number of calories taken in through their diet and burn extra calories with exercise lead to maintain a proper weight.

It is not easy for most general hospitals to provide effective nutritional care for all admitted patients. So, there should be a methodology to identify patients who need nutritional care as well as with measurements of nutritional facts. Generally nutritional

assessment is time consuming process because it requires some measurements, such as value of BMI, body weight and height with the patient's nutritional status[2].

By comparing with existing nutritional analysis systems this approach has these distinct advantages:

- Identify recommended food items for specific disease
- Suggest recommended foods for prevent/control diseases if a person is suffering more than one diseases
- Suggest foods which are helping to lose weight or gain weight if necessity
- Give exercises to burn calories with time duration to maintain/reduce body weight

1.3 Problem

Most of the people suffering of finding relevant food plan which provides the best solution by considering their health condition. When they become 30s, most frequently suffering from non-infectious diseases due to high cholesterol, high blood pressure and blood sugar problems. As we know imbalanced diet plans lead poor health condition and there is such a probability to face non – infectious diseases.

In packet foods, we can see the nutrient label but most of the people do not consider the amount of the nutrients. As an example, some of the instant drinks such as fruit juices have high amount of sugar which is not recommended for diabetes patients. Some people are suffering more than one non – infectious diseases. If a person has both diabetes and kidney disease, there are very few foods categories which can recommended.

Nutritional assessment is much time-consuming process because it needs some accurate measurements. So, there should be an easy way to suggest recommended foods by considering health condition and their purpose of gaining or losing body weight.

Body weight is also a major concern of people. Some of them want a weight gain and some want weight loss. Therefore, there should be a specific methodology for

recommended food plans and exercises for people who are suffering from specific disease and planning to weight gain or loss.

1.4 Aim and Objectives

Major objective is identifying most frequent food items for a specific non – infectious diseases or the purpose of weight gain or loss by considering daily calorie intake and health condition of a person by evaluating history records of client meals.

In meeting the above aim, these are the identified objectives.

1. Populate the history records for users who are indicating their food plans according to their health condition.
2. Identified recommended amount of nutrient for specific non-infectious disease by considering nutrient content claims.
3. Using data mining techniques, find most frequent food items which are recommended for specific disease.
4. Recommend a food plan for a specific person who is suffering from more than one non-infectious disease.
5. Let users to reduce or increase their body weight smoothly.
6. Those who want loss their body weight from both food controlling and exercises, also recommended an exercise schedule.

1.5 Chapter Summary

This chapter was discussed background and motivation, problem definition, aim and objectives of this research.

Review of Literature

2.1 Chapter Introduction

This chapter looks back at the past studies which has used different methods to analysis food nutrients and collecting food nutrient data to predict a suitable diet plan. Findings and limitations are discussed here.

2.2 Nutrient analysis methods and limitations

Calculate nutrient intake from food consumption details is an uninteresting, time consuming, error-prone task when the data is fed by manually. It is very useful that computerized food analysis system for help with diet records. Diaries and recalls are most helpful for clinical trials, because diet in trial is interested, and not the ‘usual’ diet. In their system, handled data which is collected as diet diaries or as food frequency data. In this system, DIARYAN has been called as handling diet-diary data and FREQUAN for food frequency.

Input data for DAIRYAN is four digits’ numerical food codes and equivalent quantities. Main advantage is that the program is flexible for feeding input data as well as output data. Item list, questions regarding food preparation, cooking habits and eating habits are the major categories of questionnaire. Measure of alcohol consumption is another advantage of this system. Food frequency data is very adjustable and modified according to study requirements. It is an advantage that this questionnaire targets a large population to design the feedback and not with specialized area[3].

The U.S. Department of Agriculture’s (USDA) Nutrient Data Laboratory (NDL) has designed a system Features are data acquisition, compilation, and distribution. It is developed a module to advance the assessment of analytical data value. USDA’s early data access procedures based on manually operated system to measure excellence of data. Manually data access process is a disadvantage and then these procedures furthermore expanded for multi-nutrient data assessment.

Food databases of USDA's are accessed commonly by worldwide community of science. The major purposes are to observe food intakes, to follow up nutrition researches and to develop policies for food and nutrient. Various resources are used to collect nutrient values. Because of that, data is rough and lacking details. So, maintain data quality is very important. At the end, suitable data from several resources used to provide the most descriptive evaluations for nutrition of foods.

There are five evaluation categories of assessment. They are plan of sampling, count of samples, handling samples, analytical methodology and analytically quality control. The sampling plan evaluated food samples by considering the type of food, food brand etc. This system provides the maximum sampling ratings of plan for designs of sampling consumed by the people who are interested.

Linking of sample units leads low analysis cost. The large number of samples has been analyzing is critical for the calculation of the mean. The less number of samples analysis limits the capability of calculating the mean value. Proper handling of sample units must be properly handled but it is critical. For gaining accurate data of nutrient, analytical methodology is critically evaluated. Analytical methods can be divided in two parts in the system. First method is assessed by optimal criteria and it is used to decide process validation. Second one is generated data must be performing well to use method to find precise data.

Manually or electronically data can be entered. The most efficient entry is the electronic entry. During making precise assessments, data from various resources can lead problems. Reasons are insufficient documentation of the samples of food, report detail failure and lack of details. These are the major problems. Frequently complete correct information is not documented always. Contact authors are badly time-consuming process. Authors documented of nutritional data must be always with relevant details[4].

USDA's Dietary Intake Data System is another approach. To increase the quality and accuracy of researches, automated methods are used for gathering and handling data of food intake. Dietary Intake Data System contains nutrient details database and web base system. This article clearly discussed how this system gives high-quality data of

food intake. Automated multiple-pass method (AMPM) is a new recall method that used in this system. Shortened the task of editing and updating the details of foods by asking questions to identify details of different foods, or variations by documenting specifications in a database.

Nutritional foods are ordered in such methods that respondents are willingly report them. Foods with brands are generally listed as part of the brand. Unfamiliar foods also listed to identify them. Various questions used to gather details about nutritional foods. These details matched with food details and descriptions with codes in the database.

Single response is required for most questions, but some people have more than one choices. Questions about ingredients and their quantities in such kind of foods like sandwiches or salads can have more than one responses. So, these are the disadvantages. The steadiness keeps up from these interviews is one of the main advantage. Questions and relevant responses are recorded. Details of food sometimes it may be unfinished because the respondent or the community does, so the interviewer fails to ask the right questioner.

Once food intake data have been collected Post-interview processing system (PIPS) is the following phase. This system followed up this process. Automatically assigning food codes saves time and increases excellence of by reducing coding discrepancies. It is really prized advantage. Data collected by AMPM can be developed coding pathways[5].

Frank B. Hu recently reported that article Dietary pattern analysis, a new direction in nutritional epidemiology. According to Frank B. Hu's view, that patterns of dietary are acting a big role of food and nutrient consumption. It is accurately predicting the risks of disease. Single nutrient analysis may be confused and wasting time.

Technique is the factor analysis. It is involving two or more variable quantities that means it is multivariate technique, which is used details collected in dietary records. It helps to classify issues or patterns of food intake. Cluster analysis also a multivariate method which can used to describe dietary patterns. Cluster analysis means divided individuals into similar subgroups with their similar food diets.

Using factor analysis, Frank B. Hu identified patterns of eating, which are like to the questionnaires and the detailed records of diet. The main factor is considered by a higher intake of fruits and vegetables, whole grains, and fish. The other factor was categorized by a higher consumption of processed meat, butter, high-fat dairy products, eggs, and refined grains. Dietary pattern is analyzed the overall diet. There are numerous alterations in nutrients between various dietary patterns, so this cannot be exact for the specific nutrients that reasonable for detected changes in disease risk. It is not much useful about organic relationships between disease risk and patterns of dietary. Dietary patterns are differed because of gender, current status and habited culture. This is a disadvantage[6].

In Unique Dietary Patterns and Chronic Disease Risk Profiles of Adult Men: The Framingham Nutrition Studies, the patterns of dietary were considered by analyzing of cluster. Cluster analysis is concerned to food nutrient details. Major task is clustering foods by considering their occurrence of consumption.

Link between male dietary patterns, nutritional risk, and chronic disease profiles propose the importance of targeted protective nutrition interventions to promote health in adult men. The nature of the identified male dietary patterns offers distinct, possible foci for preventive nutrition education and counseling[7].

Factor analysis is the base and introduced two basic patterns for eating habits. They are similar to the food frequency questionnaires and the records of diet.

As the conclusion, Factor analysis helped to develop patterns of food and created on the numerous groups of food for one by one-food frequency questionnaires and the records of diet. In contrast, it is better to consider not a single nutrient, but it is consider overall patterns of eating. Data must be assessed the reproducibility and validity[8].

Eric B. Rimm and Edward L. Giovannucci were assessed the validity and reproducibility of an expanded 131-item semi quantitative food frequency questionnaire among Male Health Professionals. The questionnaire was planned to categorize persons by considering levels of daily consumption of selected nutrients by getting average. Means and standard deviations were calculated on total nutrient intakes from food frequency questionnaires and from the two one-week diet records.

The questionnaire amounts of total energy intake were similar to intake measured by diet records. Intake of most nutrients was slightly lower when measured by questionnaire compared with diet records. However, total vitamin A intake was overvalued by questionnaire, as were intakes of many vitamins with supplements. This overestimation of intake by the questionnaire was primarily caused by extreme values for few persons[9].

PK Newby, Denis Muller, Judith Hallfrisch, Ning Qiao, Reubin Andres, and Katherine L Tucker have reported that Dietary patterns and changes in body mass index and waist circumference in adults. As a result, Obesity has increased more than 20% in the past 10 years in the United States. Because of that, majority of US adults are overweight or obese. Getting a nutritious diet filled with fruit and vegetables reduced saturated fat. Low consumption of processed meat and fast food help to maintain BMI and waist circumference. A restriction of the methods of dietary pattern is the incapability to separate nutrient-specific biologic effects, because dietary patterns contain numerous nutrients of foods which can behave as inversely to affect hunger and satiety etc[10].

Dietary Guidelines for the Prevention of Obesity provides nutrition content claims on food labels such as low calorie, low fat, sugar free which can help consumers to make healthier food choices [11].

Component	Claim	Not More Than
Energy	Low	40 kcal (170 kJ) per 100g (solids) or 20 kcal (80kJ) per 100 ml (liquids)
	Free	4 kcal per 100 g (100 ml)
Fat	Low	3 g per 100 g (solids) or 1.5 g per 100 ml (liquids)
	Free	0.15 g per 100 g (100 ml)
Saturated Fat	Low	1.5 g per 100g (solids) or 0.75 g per 100 ml (liquids) and 10% of total energy of the food
	Free	0.1 g per 100 g (solids) or 0.1 g per 100 ml (liquids)
Cholesterol	Low	0.02 g per 100g (solids) or 0.01 g per 100 ml (liquids)
	Free	0.005 g per 100 g (solids) or 0.005 g per 100 ml (liquids)
TFA	Low	1.5 g per 100g (solids) or 0.75 g per 100 ml (liquids) and 10% of total energy of the food
	Free	0.1 g per 100 g (solids) or 0.1 g per 100 ml (liquids)
Sugars	Free	0.5 g per 100 g (or 100 ml)
Sodium	Low	0.12 g per 100 g
	Very Low	0.04 g per 100 g
	Free	0.005 g per 100 g

Figure 2.1: Nutrient content claims

Recommended daily calorie intake for men and women can get from this chart. Recommended daily caloric requirements (RDCI) can be fairly well estimated from the age, height and gender of an individual. As we age, to maintain a constant BMI of 25 we must consume increasingly fewer calories. For example, 80-years old people must eat almost 20% less than they could at age 20. For women this is about 316 Calories and for men about 389 Calories. RDCI should be used as a standard for sedentary, overweight adults [12].

2.3 Chapter Summary

This chapter was discussed others work related to nutrient analysis of foods.

Technology Adapted

3.1 Chapter Introduction

This chapter provided the main technologies adapted to implement this accurate decision support system successfully.

3.2 MySQL

MySQL is an open-source relational database management system used on the web. It runs on a server and very fast, reliable, and easy to use. In my application, stored client records which were generated by advanced data generator for MySQL tool and food items with nutrient details from Fatsecret API in the database. Also, the exercises with calorie burned for specific time duration also get from the Fatsecret API and stored in the MySQL database.

3.3 CodeIgniter Framework

Generally, CodeIgniter is a powerful PHP framework which is built for programmers who need a simple and elegant toolkit to create full-featured web applications. In here I created a simple web application for the admin and the clients and they can register and access the web site.

3.4 Unsupervised Learning

Data mining techniques can divide in two main forms. Supervised and unsupervised learning. Both are facilitating of finding different kind of hidden patterns in large data sets. To find the most frequent food items, unsupervised learning is the technique that can apply. Unsupervised data mining does not focus on predetermined attributes, nor does it predict a target value. Additionally, unsupervised data mining finds hidden structure and relation among data[13].

Under the unsupervised learning, association rule mining is the most applicable function to find most frequent food items. Association models identify common co-

occurrences among a list of possible events. It is evaluating all the items available in the food list and finding the items that are frequently used.

This rule can be applying to select most frequent food items which are recommended for specific disease, combination of more than one disease and help to weight gain or loss. To select interesting rules from the set of all possible rules, constraints on various measures of significance and interest can be used. The best-known constraints are minimum thresholds on support and confidence[14].

3.4.1 Apriori data mining algorithm

Apriori algorithm is a classical algorithm in data mining. It is used for mining frequent items. The Apriori algorithm learns association rules and applied to a large data set.

The basic Apriori algorithm is a 3-step approach:

1. **Join.** Scan the whole database for how frequent 1-itemsets are.
2. **Prune.** Those item sets that satisfy the **support** and **confidence** move onto the next round for 2-itemsets.
3. **Repeat.** This is repeated for each itemset level until we reach our previously defined **size**.

Apriori is well understood, easy to implement and has many derivatives. The algorithm can be quite memory, space and time intensive when generating item sets[15][16].

3.5 Weka Tool

Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can be applied directly to a dataset. Weka contains tools for these areas such as data pre-processing, classification, regression, clustering, association rules, and visualization. In my application Weka is used for getting the associate mining results of frequent food items for specific disease[17][18].

3.6 Advanced data generator for MySql

Advanced Data Generator tool is allowing to generate test data in an easy-to-use and reusable manner. The Advanced Data Generator can connect to any data source. It can generate data directly into the database by using tables, stored procedures and custom SQL statements (INSERT INTO ...). In my application this tool is used to create history data/past records of clients who track their details of meals by considering their health condition and as well as the client records with meal details that who want weight gain or weight loss.

3.7 Chapter Summary

This chapter provided technologies that adapt to solve the research problem.

My Approach

4.1 Chapter Introduction

This chapter provides how to implement this decision support system using data from Fatsecret API calls with dummy records of users.

4.2 Fetch data and process to get most frequent food items

In my approach, here it is going to implement a decision support system for food nutritional analysis to identify most frequent food items. It is needed to identify most frequent food items which are recommended for specific non-infectious disease or combination of more than one non-infectious disease. Some users only want a weight gain or a weight loss, so they also need recommended most frequent foods for their purpose.

By calling Fatsecret API methods, collected the food items with their nutritional values and saved these data into the database table.

<input type="checkbox"/>	food_id	fatsecret_food_id	food_name	
<input type="checkbox"/>	2294	39784	Dry Somen Japanese Noodles	26B
<input type="checkbox"/>	2295	39785	Somen Japanese Noodles (Cooked)	31B
<input type="checkbox"/>	2296	39786	Enriched Dry Pasta	18B
<input type="checkbox"/>	2297	39787	Spaghetti (Without Added Salt, Enriched, Cooked)	48B
<input type="checkbox"/>	2298	39788	Dry Spaghetti (Protein Fortified, Enriched)	43B
<input type="checkbox"/>	2299	39789	Spaghetti (Protein Fortified, Enriched, Cooked)	47B
<input type="checkbox"/>	2300	39790	Dry Whole Wheat Pasta	21B
<input type="checkbox"/>	2301	39791	Whole Wheat Spaghetti (Cooked)	30B
<input type="checkbox"/>	2302	39792	Dry Spinach Pasta	17B
<input type="checkbox"/>	2303	39793	Spinach Spaghetti (Cooked)	26B
<input type="checkbox"/>	2304	39794	Bread Wheat Flours (Unenriched)	31B
<input type="checkbox"/>	2305	39795	Barley Flour or Meal	20B
<input type="checkbox"/>	2306	39796	Barley Malt Flour	17B
<input type="checkbox"/>	2307	39797	Oat Flour (Partially Debranned)	31B
<input type="checkbox"/>	2308	39798	Dry Rice Noodles	16B
<input type="checkbox"/>	2309	39799	Rice Noodles (Cooked)	21B
<input type="checkbox"/>	2310	39800	Dry Macaroni (Protein Fortified, Enriched)	42B
<input type="checkbox"/>	2311	39801	Macaroni (Protein Fortified, Enriched, Cooked, 2)	49B
<input type="checkbox"/>	2312	39802	Egg Noodles (with Added Salt, Enriched, Cooked)	47B
<input type="checkbox"/>	2313	39803	White Corn	10B
<input type="checkbox"/>	2314	39804	White Whole Grain Corn Flour	28B
<input type="checkbox"/>	2315	39805	Yellow Masa Corn Flour (Enriched)	33B
<input type="checkbox"/>	2316	39806	White Whole Grain Cornmeal	26B

Figure 4.1: Fatsecret food items with food_ids

Not only that, it is needed the details of exercises from Fatsecret API. So, from the API, collected the exercises with burned calories for specific time duration (burn calories per one hour). Using exercises and recommended foods, let the users to reduce their weight.

<input type="checkbox"/>	exercise_id	exercise	calories_hr
<input type="checkbox"/>	76	Conditioning exercise (health club)	404
<input type="checkbox"/>	77	Floorball	441
<input type="checkbox"/>	78	Bicycling (slow) - 11/mph	441
<input type="checkbox"/>	79	Gym	441
<input type="checkbox"/>	80	Hiking	441
<input type="checkbox"/>	81	Basketball	441
<input type="checkbox"/>	82	Weight Training (moderate)	441
<input type="checkbox"/>	83	Circuit Training	441
<input type="checkbox"/>	84	Shoveling Snow	441
<input type="checkbox"/>	85	Exercise machine (slow)	441
<input type="checkbox"/>	86	Pole Dancing	441
<input type="checkbox"/>	87	Power Walking	441
<input type="checkbox"/>	88	Scuba Diving	441
<input type="checkbox"/>	89	Trampolining	441
<input type="checkbox"/>	90	Tree Climbing (Ropes)	441
<input type="checkbox"/>	91	Netball	441
<input type="checkbox"/>	92	Skiing (water)	441
<input type="checkbox"/>	93	Dance (fast step, aerobic)	441
<input type="checkbox"/>	94	GRIT	514
<input type="checkbox"/>	95	Stairs (Climbing Stairs)	514
<input type="checkbox"/>	96	High Intensity Interval Training (HIIT)	514
<input type="checkbox"/>	97	Jiu Jitsu	514
<input type="checkbox"/>	98	Insanity	514

Figure 4.2: Exercises with burned calories for specific time duration

In this system, let the users to keep their daily meal records. These records contained their health condition too. According to the health condition, there are recommended foods for specific non-infectious disease as well as the purpose of weight gain or weight loss. With these recommended foods, created history records of several users. History records are generated using data generation tool.

user_id	user_role_id	email	password	nickname	age	sex	user_status
2	2	W.Weggemans@ukmail.net	\$2y\$10\$khHLinGkTDn7	Hai Kolk van der	36	2	1
3	2	V.Rivard@alloymail.ru	\$2y\$10\$khHLinGkTDn7	Jenifer Ault	59	1	1
4	2	J.Costantini@address.co.uk	\$2y\$10\$khHLinGkTDn7	Joye Rosario	22	2	1
5	2	Y.Becerril@go.co.uk	\$2y\$10\$khHLinGkTDn7	Edelmiro Orcajo	29	2	1
6	2	D.Quinn@hotmail.info	\$2y\$10\$khHLinGkTDn7	Giovanna Raamsdonk van	55	1	1
7	2	W.Frecha@netvigator.co.kr	\$2y\$10\$khHLinGkTDn7	Alvin Sandler	51	2	1
8	2	A.Melanson@latino.br	\$2y\$10\$khHLinGkTDn7	Philippina Reigada	68	2	1
9	2	N.Morrow@yahoo.co.jp	\$2y\$10\$khHLinGkTDn7	Chu Alaña	60	2	1
10	2	J.Zuri@latino.se	\$2y\$10\$khHLinGkTDn7	Alida Elwell	60	2	1
11	2	O.Aguayo@latino.de	\$2y\$10\$khHLinGkTDn7	Dwight Zijpp van der	52	2	1
12	2	K.Dick@icqmail.cz	\$2y\$10\$khHLinGkTDn7	Verlie Muse	40	1	1
13	2	Q.Garzon@go.se	\$2y\$10\$khHLinGkTDn7	Jitske Rivet	34	1	1
14	2	R.Hutton@loja.tk	\$2y\$10\$khHLinGkTDn7	Justus Kelly	25	1	1
15	2	A.Olander@uol.com.sg	\$2y\$10\$khHLinGkTDn7	Rosemary Fuertehijar	20	1	1
16	2	N.Vega@freemail.to	\$2y\$10\$khHLinGkTDn7	Adina Iban	33	1	1
17	2	Y.Emmelkamp@usmail.au	\$2y\$10\$khHLinGkTDn7	Aranta Arreola	23	1	1
18	2	K.Jarrell@icqmail.com.mx	\$2y\$10\$khHLinGkTDn7	Thomasina Farthing	36	2	1
19	2	A.Mcclure@popmail.hr	\$2y\$10\$khHLinGkTDn7	Wybke Keeris	56	1	1
20	2	M.Booth@address.tk	\$2y\$10\$khHLinGkTDn7	Kuno Wilborn	52	2	1
21	2	U.Severson@mailhost.org	\$2y\$10\$khHLinGkTDn7	Mercedes Villeneuve	29	2	1
22	2	Y.Osselaer van@emailplanet.i	\$2y\$10\$khHLinGkTDn7	Pleuni Munguira	42	1	1
23	2	Z.Linares@popmail.in	\$2y\$10\$khHLinGkTDn7	Ferne Wooden	46	1	1
24	2	N.Terroba@sina.de	\$2y\$10\$khHLinGkTDn7	Ernst Yan	63	1	1
25	2	D.Michaelson@mail.excite.com	\$2y\$10\$khHLinGkTDn7	Sixta Dineen	48	1	1
26	2	X.Hersey@mailhost.is	\$2y\$10\$khHLinGkTDn7	Rufus Tompkins	57	2	1
27	2	X.Albright@yupi.to	\$2y\$10\$khHLinGkTDn7	Kalle Velting	66	1	1
28	2	C.Erickson@euromail.hr	\$2y\$10\$khHLinGkTDn7	Jane Damico	50	2	1
29	2	P.Keele@excite.co.za	\$2y\$10\$khHLinGkTDn7	Alexei Yugueros	61	1	1

Figure 4.3: History data records of several users

meal_food_id	user_id	meal_id	food_id	meal_date
1	2	1	9031	2018-09-20
2	2	1	4730	2018-09-20
3	2	1	10877	2018-09-20
4	2	2	8854	2018-09-20
5	2	2	7451	2018-09-20
6	2	2	5675	2018-09-20
7	2	3	10977	2018-09-20
8	2	3	10510	2018-09-20
9	2	3	779	2018-09-20
10	2	1	7149	2018-09-19
11	2	1	8613	2018-09-19
12	2	1	10969	2018-09-19
13	2	2	8862	2018-09-19
14	2	2	4925	2018-09-19
15	2	2	10585	2018-09-19
16	2	3	8804	2018-09-19
17	2	3	6709	2018-09-19
18	2	3	10398	2018-09-19
19	2	1	10313	2018-09-18

Figure 4.4: Food details of users

<input type="checkbox"/>	user_disease_id	user_id	disease_id	food_added
<input checked="" type="checkbox"/>	1	2	1	1
<input type="checkbox"/>	2	3	1	1
<input type="checkbox"/>	3	4	1	1
<input type="checkbox"/>	4	5	1	1
<input type="checkbox"/>	5	6	1	1
<input type="checkbox"/>	6	7	1	1
<input type="checkbox"/>	7	8	1	1
<input type="checkbox"/>	8	9	1	1
<input type="checkbox"/>	9	10	1	1
<input type="checkbox"/>	10	11	1	1
<input type="checkbox"/>	11	12	1	1
<input type="checkbox"/>	12	13	1	1
<input type="checkbox"/>	13	14	1	1
<input type="checkbox"/>	14	15	1	1
<input type="checkbox"/>	15	16	1	1
<input type="checkbox"/>	16	17	1	1
<input type="checkbox"/>	17	18	1	1
<input type="checkbox"/>	18	19	1	1
<input type="checkbox"/>	19	20	1	1

Figure 4.5: Health condition and food details taken by users

The nutrient content claims in food labels such as low calorie, low fat and sugar free with the measurements are used for making claims for specific non-infectious disease. With these nutrient claims, logically identify the nutrient levels for selected non-infectious diseases. Diabetes, heart diseases, high blood pressure, obesity and kidney diseases are the selected non-infectious diseases[19].

For the diabetes, this logic selects foods which are in this range. Foods within this range of nutrient measurements are identified as recommended foods for diabetes. There should be, Saturated fat amount is less than 1.5 mg, sodium amount is less than 140 mg, calories are less than 120 and sugar is less than 5 mg.

Diabetes

```
SELECT *
FROM
  pro_nutrition_per_cup
WHERE saturated_fat < 1.5
  AND sodium < 140
  AND calories < 120
  AND sugar < 5 ;
```

Heart Diseases

```
SELECT *
FROM
  pro_nutrition_per_cup
WHERE calories < 400
  AND sodium < 140
  AND saturated_fat < 1.5
  AND monounsaturated_fat > 0
  AND polyunsaturated_fat > 0 ;
```

High Blood Pressure

```
SELECT *
FROM
  pro_nutrition_per_cup
WHERE sodium < 40
  AND sugar < 2
  AND `potassium` > 0
  AND `fiber` > 0 ;
```

Obesity

```
SELECT *
FROM
  pro_nutrition_per_cup
WHERE sugar < 2
  AND `fat` < 0.15
  AND `calories` < 40;
```

Kidney Diseases

```
-----  
SELECT *  
FROM  
  pro_nutrition_per_cup  
WHERE `calories` < 200  
  AND `protein` < 20  
  AND fat < 4  
  AND `sodium` > 40  
  AND `sodium` < 200  
  AND `potassium` = 0 ;  
-----
```

This is used to identify recommended foods for manipulating history records of users. By considering above logic, filtered records of food items which are most suitable to prevent specific non-infectious disease. Using these food items, then populate data records of users. These records are keeping as the data set for applying the associate rule mining.

By populating records, get all the recommended food items for specific disease as well as random no of less suitable and not suitable foods. In the records, the amount of most suitable foods is high, and others are less. To populate data records, here used advanced data generator for MySQL tool and saved in a table of the database. Then using weka tool, Associate rule mining, get the recommended, most frequent food items for specific non-infectious disease[20]. Apriori algorithm is used to get the frequent food items.

This figure shows how to set the parameters for doing associate rule mining using apriori algorithm. When we set lowerBoundMinSupport as 0.001, then we can get 1000 occurrences of most frequent food items.

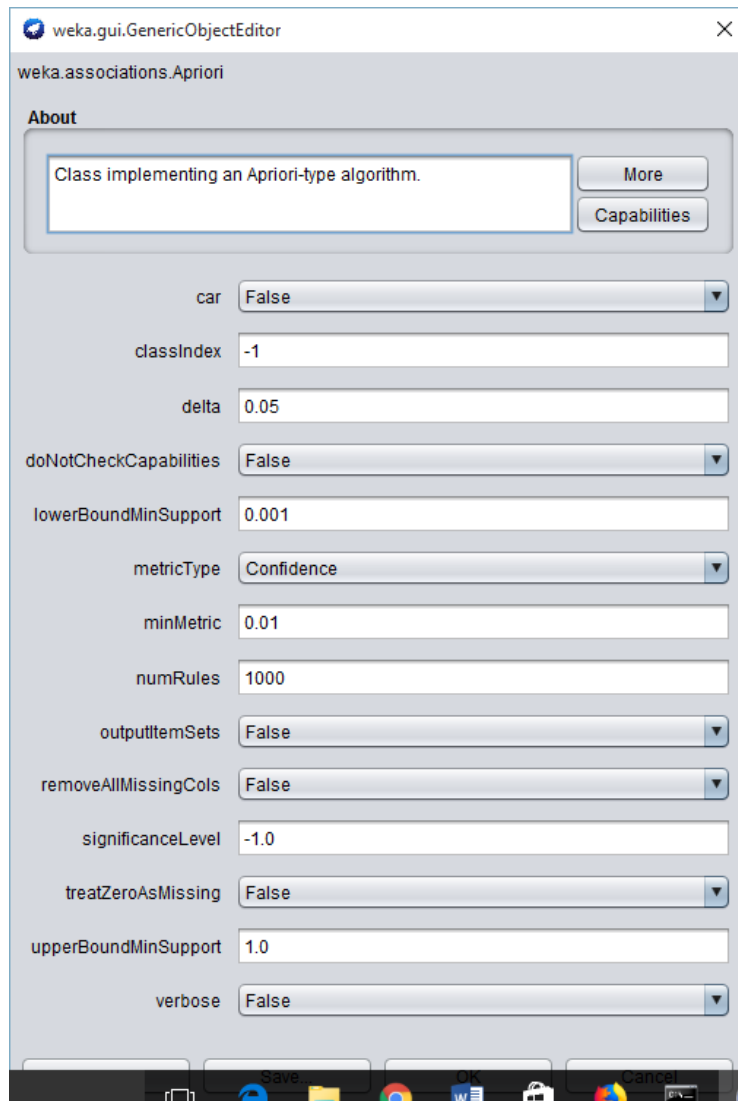


Figure 4.6: Set values for associate rule mining using Apriori Algorithm

4.3 Outputs

Outputs are when a user has more than one diseases, then recommended relevant food plan which is needed for preventing both non-infectious diseases. Not only that, if a user wants to lose weight or gain weight, then give suggestions for relevant food items and give three or four exercises with time durations to control his body weight.

4.4 Chapter Summary

This chapter was discussed how to implement this decision support system.

Analysis and Design

5.1 Chapter Introduction

This section described detail analysis and design of my accurate decision support system.

5.2 Data Collection

5.2.1 Keep user Details

Client can register by providing his age, weight, height and gender details and then he can access the features of the site as registered user. Registered user can enter his meal records daily as well as including his health condition. It means that if he is suffering such kind of non-infectious disease as well as he is planning to weight gain or weight loss, then he can mention it and update the database. These records are keep as history records to recommend foods for specific disease or the purpose of weight gain or weight loss. Before analysis the history records of users, it is needed to preprocess the data by using preprocessing techniques. History records are manipulated using Advanced Data Generator for MySql tool.

5.2.2 Retrieve relevant data by calling Fatsecret API

Using Fatsecret API, get food items with nutrient details and exercises with burn calories for specific time duration. So, there is no irrelevant and redundant information as well as noisy and unreliable data because using API calls, retrieve only relevant data. So, no need to do data cleaning steps such as filling missing values, smoothing noisy data and removing inconsistencies. It is needed the data integration which combines data from multiple sources into single data store. Store data from API calls and store data by generating the Advanced Data Generator for MySql tool. Data reduction is no need because able to query the relevant data set from the database.

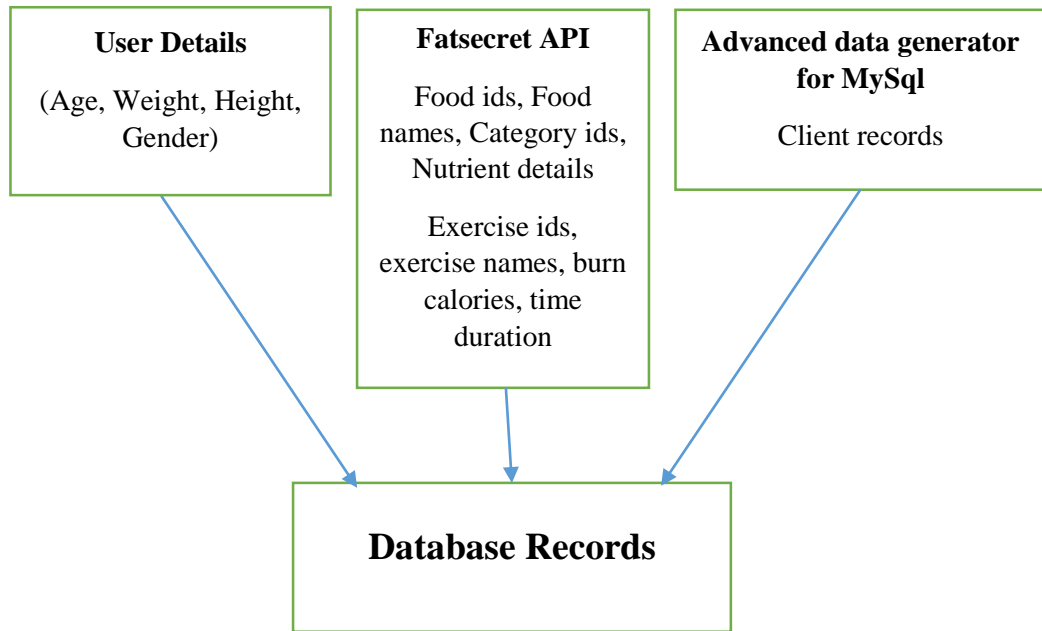


Figure 5.1: Data integration

By calling Fatsecret API, collected the food items with their nutritional values and saved these data into the database table.

<input type="checkbox"/>	food_id	fatsecret_food_id	food_name	
<input type="checkbox"/>	2294	39784	Dry Somen Japanese Noodles	26B
<input type="checkbox"/>	2295	39785	Somen Japanese Noodles (Cooked)	31B
<input type="checkbox"/>	2296	39786	Enriched Dry Pasta	18B
<input type="checkbox"/>	2297	39787	Spaghetti (Without Added Salt, Enriched, Cooked)	48B
<input type="checkbox"/>	2298	39788	Dry Spaghetti (Protein Fortified, Enriched)	43B
<input type="checkbox"/>	2299	39789	Spaghetti (Protein Fortified, Enriched, Cooked)	47B
<input type="checkbox"/>	2300	39790	Dry Whole Wheat Pasta	21B
<input type="checkbox"/>	2301	39791	Whole Wheat Spaghetti (Cooked)	30B
<input type="checkbox"/>	2302	39792	Dry Spinach Pasta	17B
<input type="checkbox"/>	2303	39793	Spinach Spaghetti (Cooked)	26B
<input type="checkbox"/>	2304	39794	Bread Wheat Flours (Unenriched)	31B
<input type="checkbox"/>	2305	39795	Barley Flour or Meal	20B
<input type="checkbox"/>	2306	39796	Barley Malt Flour	17B
<input type="checkbox"/>	2307	39797	Oat Flour (Partially Debranned)	31B
<input type="checkbox"/>	2308	39798	Dry Rice Noodles	16B
<input type="checkbox"/>	2309	39799	Rice Noodles (Cooked)	21B
<input type="checkbox"/>	2310	39800	Dry Macaroni (Protein Fortified, Enriched)	42B
<input type="checkbox"/>	2311	39801	Macaroni (Protein Fortified, Enriched, Cooked, 2)	49B
<input type="checkbox"/>	2312	39802	Egg Noodles (with Added Salt, Enriched, Cooked)	47B
<input type="checkbox"/>	2313	39803	White Corn	10B
<input type="checkbox"/>	2314	39804	White Whole Grain Corn Flour	28B
<input type="checkbox"/>	2315	39805	Yellow Masa Corn Flour (Enriched)	33B
<input type="checkbox"/>	2316	39806	White Whole Grain Cornmeal	26B

Figure 5.2: Food related Data from Fatsecret API

Not only that, it is needed the exercises from Fatsecret API. So, from the API collected the exercises with burned calories for specific time duration it means burn calories per one hour.

<input type="checkbox"/>	exercise_id	exercise	calories_hr
<input type="checkbox"/>	76	Conditioning exercise (health club)	404
<input type="checkbox"/>	77	Floorball	441
<input type="checkbox"/>	78	Bicycling (slow) - 11/mph	441
<input type="checkbox"/>	79	Gym	441
<input type="checkbox"/>	80	Hiking	441
<input type="checkbox"/>	81	Basketball	441
<input type="checkbox"/>	82	Weight Training (moderate)	441
<input type="checkbox"/>	83	Circuit Training	441
<input type="checkbox"/>	84	Shoveling Snow	441
<input type="checkbox"/>	85	Exercise machine (slow)	441
<input type="checkbox"/>	86	Pole Dancing	441
<input type="checkbox"/>	87	Power Walking	441
<input type="checkbox"/>	88	Scuba Diving	441
<input type="checkbox"/>	89	Trampolining	441
<input type="checkbox"/>	90	Tree Climbing (Ropes)	441
<input type="checkbox"/>	91	Netball	441
<input type="checkbox"/>	92	Skiing (water)	441
<input type="checkbox"/>	93	Dance (fast step, aerobic)	441
<input type="checkbox"/>	94	GRIT	514
<input type="checkbox"/>	95	Stairs (Climbing Stairs)	514
<input type="checkbox"/>	96	High Intensity Interval Training (HIIT)	514
<input type="checkbox"/>	97	Jiu Jitsu	514
<input type="checkbox"/>	98	Insanity	514

Figure 5.3: Exercise related Data from Fatsecret API

In this web application, Client can register as a new user and he can enter his meal details with his health condition as records to the database daily. When a customer has a specific disease and he takes meals to prevent that disease then he can update his details regularly. Not only that if someone need to lose his weight or gain his weight, then he can track his records by updating his meals according to his purpose. Admin can access these records and keep history data to get the most frequent food items. For the history data, use the advanced data generator for MySQL, populate thousands of records to develop the data set for associate data mining.

5.3 Decision making Process

5.3.1 Frequent Food items

Using associate mining rule, get the frequent food items for the specific disease. And also get food items for specific user to gain weight or lose weight by considering relevant calorie amounts that he wanted. Weka tool is used to do the association filtering.

5.3.2 Recommend a food plan for two identified non-infectious diseases

Identify the nutrient content claims on food labels such as low calorie, low fat, sugar free are used for the conditions for making claims for specific disease.

Component	Claim	Not More Than
Energy	Low	40 kcal (170 kJ) per 100g (solids) or 20 kcal (80kJ) per 100 ml (liquids)
	Free	4 kcal per 100 g (100 ml)
Fat	Low	3 g per 100 g (solids) or 1.5 g per 100 ml (liquids)
	Free	0.15 g per 100 g (100 ml)
Saturated Fat	Low	1.5 g per 100g (solids) or 0.75 g per 100 ml (liquids) and 10% of total energy of the food
	Free	0.1 g per 100 g (solids) or 0.1 g per 100 ml (liquids)
Cholesterol	Low	0.02 g per 100g (solids) or 0.01 g per 100 ml (liquids)
	Free	0.005 g per 100 g (solids) or 0.005 g per 100 ml (liquids)
TFA	Low	1.5 g per 100g (solids) or 0.75 g per 100 ml (liquids) and 10% of total energy of the food
	Free	0.1 g per 100 g (solids) or 0.1 g per 100 ml (liquids)
Sugars	Free	0.5 g per 100 g (or 100 ml)
Sodium	Low	0.12 g per 100 g
	Very Low	0.04 g per 100 g
	Free	0.005 g per 100 g

Figure 5.4: Nutrient content claims

By considering these nutrient claims, logically identify the nutrient levels for specific diseases[21][22]. Here two examples.

Diabetes

```
-----  
SELECT *  
FROM  
  pro_nutrition_per_cup  
WHERE saturated_fat < 1.5  
  AND sodium < 140  
  AND calories < 120  
  AND sugar < 5 ;  
-----
```

Heart Diseases

```
-----  
SELECT *  
FROM  
  pro_nutrition_per_cup  
WHERE calories < 400  
  AND sodium < 140  
  AND saturated_fat < 1.5  
  AND monounsaturated_fat > 0  
  AND polyunsaturated_fat > 0 ;  
-----
```

Then populated data using advanced data generator for MySQL, recommended foods for specific disease by considering above logic for nutrient levels.

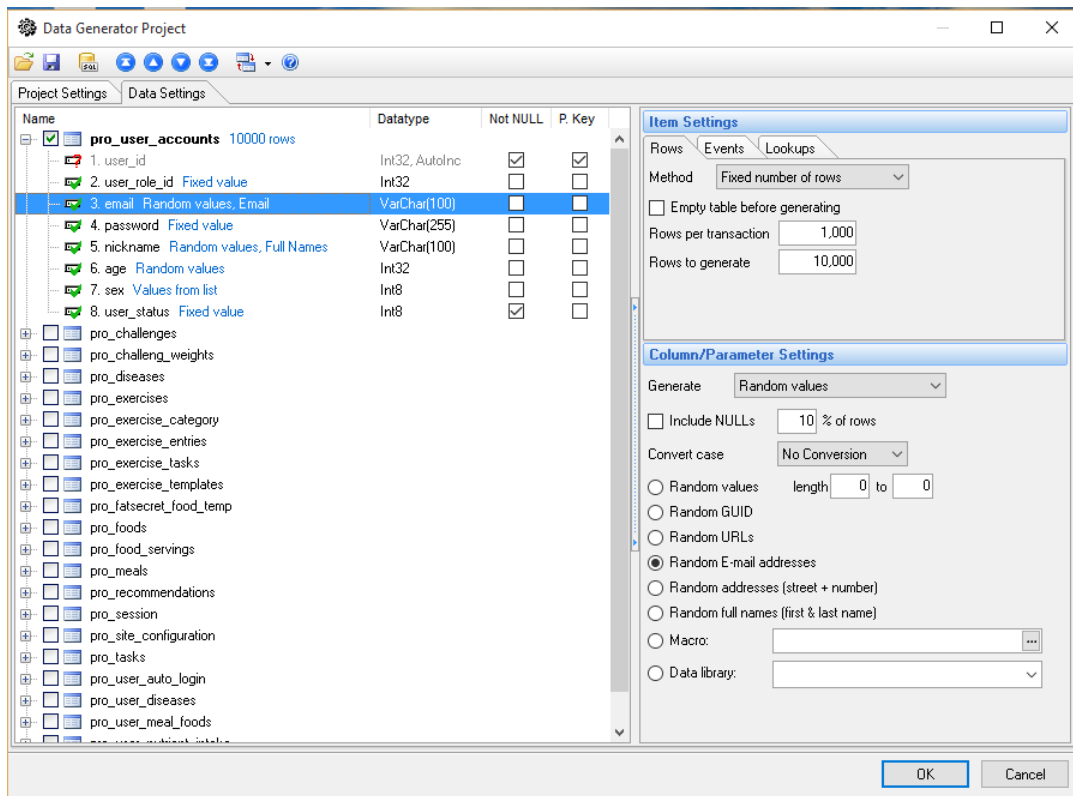


Figure 5.5: Generate email addresses of users

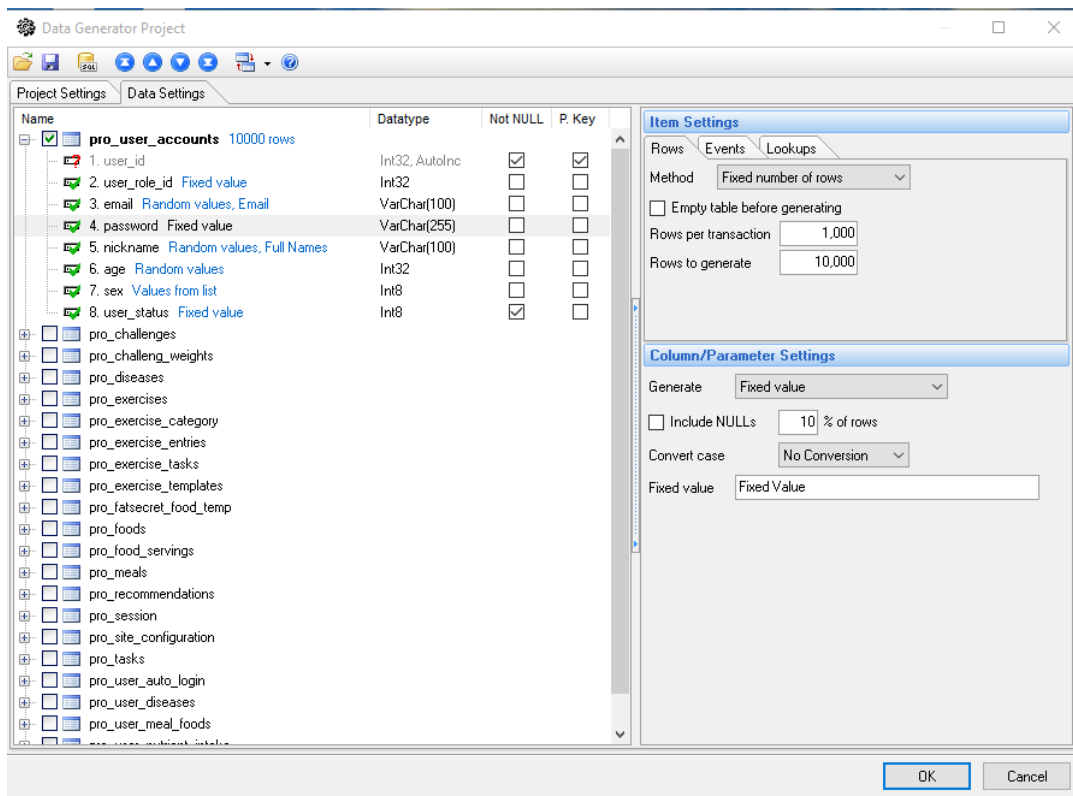


Figure 5.6: Generate passwords of users

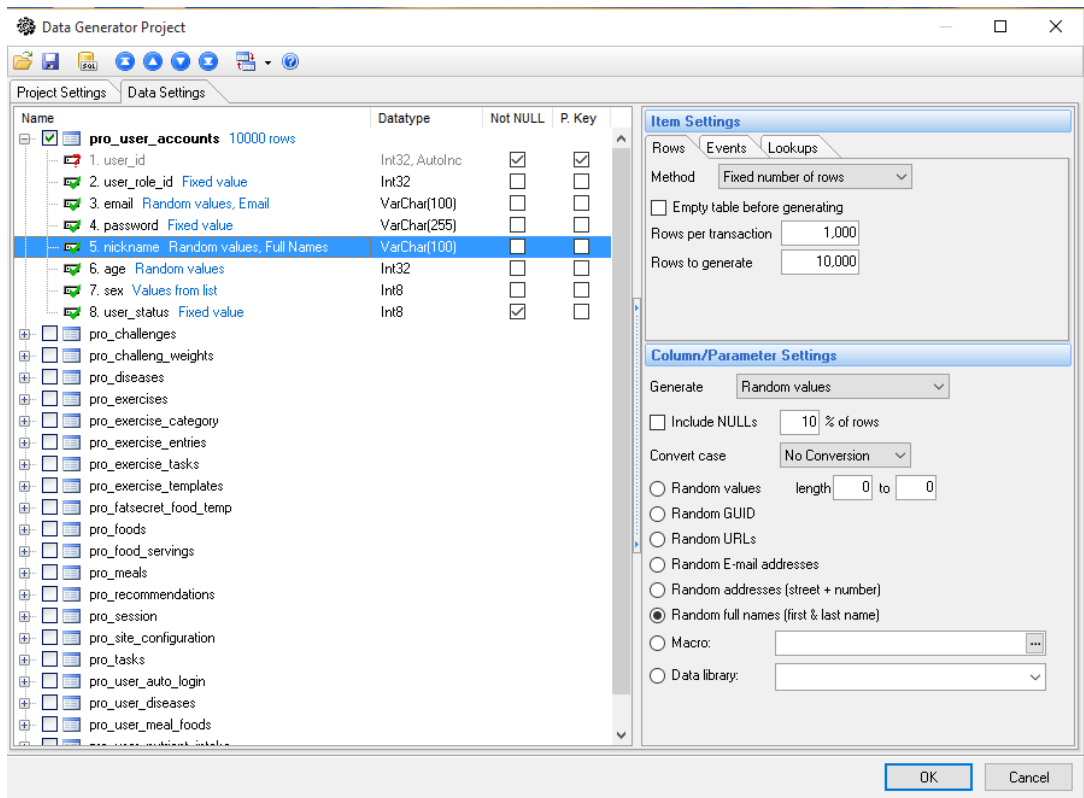


Figure 5.7: Generate nicknames of users

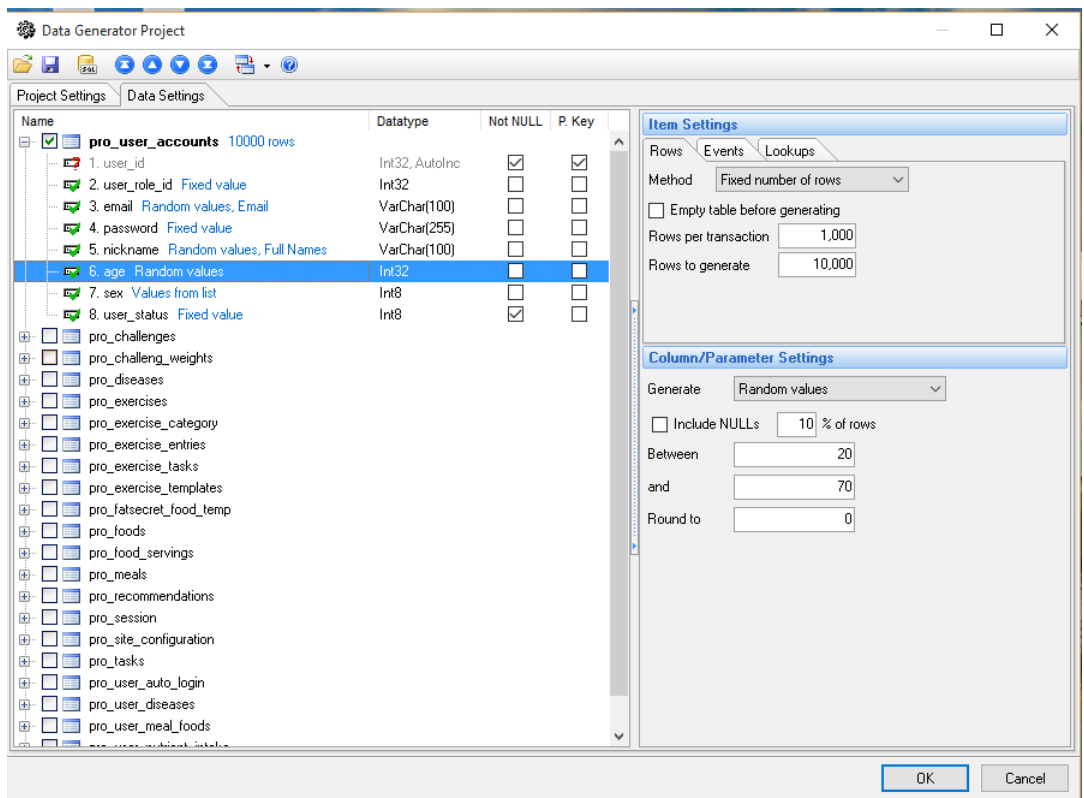


Figure 5.8: Generate age of users

History Records

ID	Name	Food Name	Age	Sex	Meal	Meal Date
1	Hai Kolk va	Cream Of Wheat	32	Male	Breakfast	2018-09-20
2	Hai Kolk va	Green Cabbage	32	Male	Breakfast	2018-09-20
3	Hai Kolk va	Cheese Muffin	32	Male	Breakfast	2018-09-20
4	Hai Kolk va	Iced Coffee with Cream	32	Male	Lunch	2018-09-20
5	Hai Kolk va	Radish Seeds (Sprouted)	32	Male	Lunch	2018-09-20
6	Hai Kolk va	Sugar Substitute (Herbal Extract Sweetener, Liquid)	32	Male	Lunch	2018-09-20
7	Hai Kolk va	Sweetened Almond Milk	32	Male	Dinner	2018-09-20
8	Hai Kolk va	Apple Cider Vinegar	32	Male	Dinner	2018-09-20
9	Hai Kolk va	White Beans (Mature Seeds, Canned)	32	Male	Dinner	2018-09-20
10	Hai Kolk va	Mushrooms	32	Male	Breakfast	2018-09-19
11	Hai Kolk va	Tea (Brewed)	32	Male	Breakfast	2018-09-19
12	Hai Kolk va	Baby Spinach	32	Male	Breakfast	2018-09-19
13	Hai Kolk va	Iced Coffee (Small)	32	Male	Lunch	2018-09-19
14	Hai Kolk va	Low Sodium Green String Beans (Canned)	32	Male	Lunch	2018-09-19
15	Hai Kolk va	California Rolls	32	Male	Lunch	2018-09-19

Figure 5.9: Populated data records for specific user by considering meal and date

Then using associate mining rule, get the frequent food items for the specific disease. Weka tool is used to do the association filtering[23].

5.3.3 Using exercises and recommended food plan, achieve weight gain or weight loss

Using advanced data generator for MySQL tool, generate history data with clients' meal details who kept records by considering weight gain and weight loss daily. Then identify the frequent food items for weight gain or weight loss and give them as suggestions. Get the recommended daily calorie intake of a person using the chart by considering person's sex, height and age. Those who want a weight gain he must add fixed calorie values. 250 cal(Male) or 125 cal(Female) to the recommended daily calorie intake. If a person wants weight loss he must reduce 20% of calories from recommended daily calorie intake.

Recommended daily calorie intake = 2500 cal

Reduction = 500 cal(250 cal from foods and 250 cal from exercises)

Then total 2000 cal from recommended frequent food items as a food plan and 250 cal burn from exercises. Also able to get Exercise details from the Fatsecret API with exercise name, time duration and burned calories.

5.4 Chapter Summary

This chapter described how to predict a food plan for a specific user by considering his health condition.

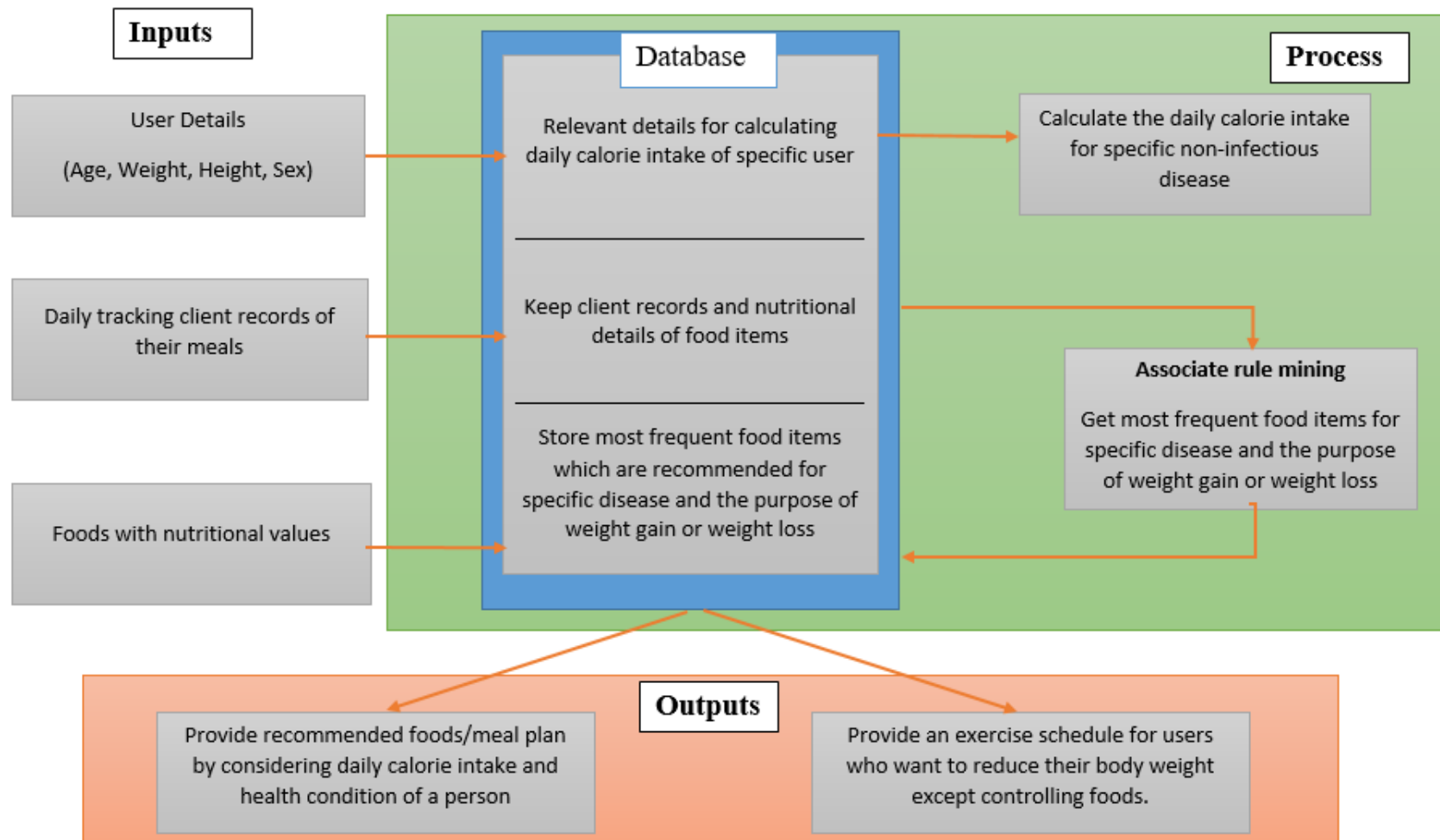


Figure 5.10: High Level design of the proposed system depicting the flow of steps

Implementation

6.1 Chapter Introduction

This chapter is described how to achieve the objectives after completing the system. User interfaces clearly described the progress.

6.1 Collection of Data and basic functionalities

6.1.1 Fetch food details based on food category

Using Fatsecret API calls, fetched food items with their nutrient amounts. Then saved them in database table. Category id of a food item is needed to populate history records of clients with their meals.

For the main categories, find the category id manually. Then grab the category id of sub categories. Foods of sub categories fetch one by one and saved in a temp table. Then parse food ids one by one (10 at a time) to the Fatsecret API and grab the food names and measuring details. After grabbing time to time, update the process status.

Food List

Food ID	Food Name	Metric Serving Amount	Calories	Carbohydrate	Protein	Fat	Saturated Fat	Polyunsaturated Fat	Monounsaturated Fat
11301	Lamb Shoulder	14.00	39	0.00	3.14	2.78	1.177	0.226	1.139
11300	Stewed Pork Chop	50.00	120	0.00	13.62	6.81	2.555	0.585	3.030
11299	Fried Breaded or Floured Pork Chop	52.00	165	7.15	11.08	9.91	2.941	2.227	3.901

Figure 6.1: Food list with nutrient details

6.1.2 Create user accounts

Each meal has one main category food item with other two food items. 3 food items per one meal and 3 meals per day. To create history records, used advanced data generator for MySQL tool and populate thousands of records for 100 users. This is how user accounts were created.

Then created dummy data for users with identified non-infectious diseases. Then selected recommended foods for specific non-infectious disease using nutrient claims. Take a specific user with identified disease and assign meals and foods randomly for one-year time duration which are recommended for his health condition.

Users

























User ID	Email	Nickname	Role Name	User Status	Action
1	msc_admin@gmail.com	MSC	Admin	Active	  
2	W.Weggemans@ukmail.net	Hai Kolk va	User	Active	  
3	V.Rivard@alloymail.ru	Jenifer Ault	User	Active	  
4	J.Costantini@address.co.uk	Joye Rosario	User	Active	  
5	Y.Becerril@go.co.uk	Edelmiro Orcajo	User	Active	  
6	D.Quinn@hotmail.info	Giovanna Raamsdonk van	User	Active	  
7	W.Frecha@netvigator.co.kr	Alvin Sandler	User	Active	  
8	A.Melanson@latino.br	Philippina Reigada	User	Active	  

Figure 6.2: User accounts

6.1.3 Predict food plan using most frequent food items

Before predicting good food meals for new user, identified most frequent food items for specific non-infectious disease. Used the dummy data of users which was created with user details, health conditions and their daily meals with number of foods. These user records taken to feed the weka tool to get the most frequent food items for specific non-infectious disease[24].

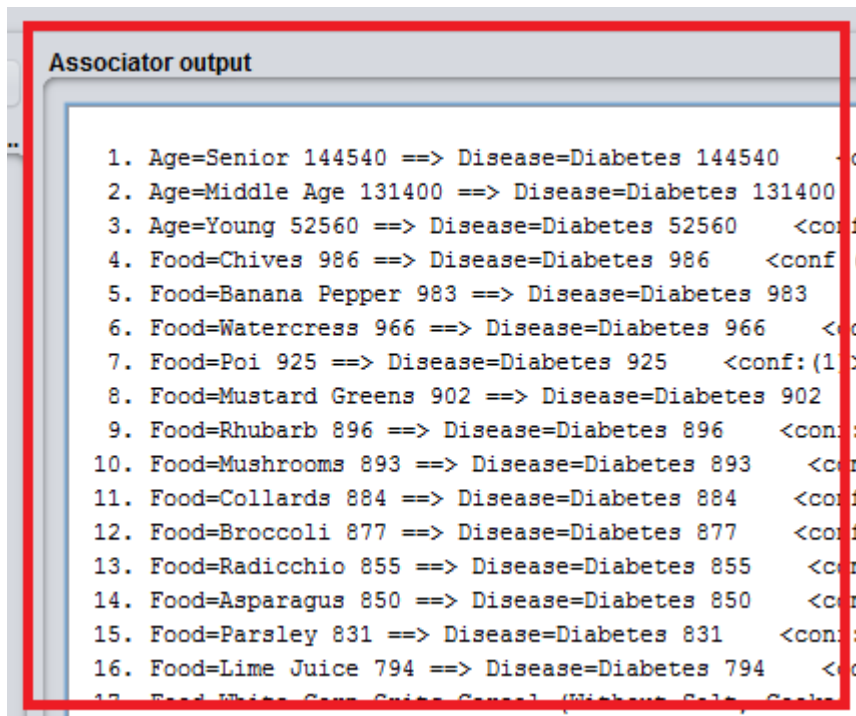
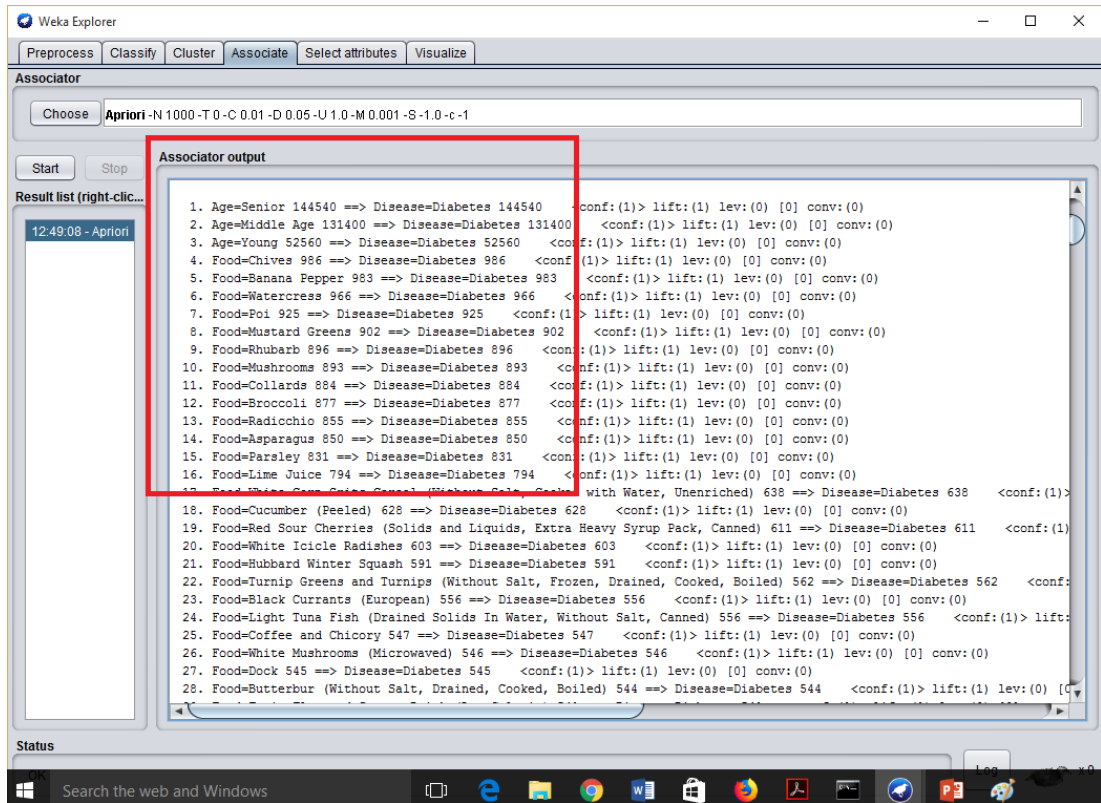


Figure 6.3: Frequent foods after applying associate mining

6.1.4 Admin access of user records

Admin can view the records of users belong to specific non-infectious disease and able to create csv file with relevant columns which was created for feeding the weka tool to generate most frequent food items for specific disease. This file can be directly uploaded to the weka tool and applied association rule mining[25][26].

History records of Diabetes Patients

Export CSV for Weka

History Records

ID	Name	Food Name	Age	Sex	Meal	Meal Date
1	Hai Kolk va	Cream Of Wheat	32	Male	Breakfast	2018-09-20
2	Hai Kolk va	Green Cabbage	32	Male	Breakfast	2018-09-20
3	Hai Kolk va	Cheese Muffin	32	Male	Breakfast	2018-09-20
4	Hai Kolk va	Iced Coffee with Cream	32	Male	Lunch	2018-09-20
5	Hai Kolk va	Radish Seeds (Sprouted)	32	Male	Lunch	2018-09-20
6	Hai Kolk va	Sugar Substitute (Herbal Extract Sweetener, Liquid)	32	Male	Lunch	2018-09-20
7	Hai Kolk va	Sweetened Almond Milk	32	Male	Dinner	2018-09-20
8	Hai Kolk va	Apple Cider Vinegar	32	Male	Dinner	2018-09-20
9	Hai Kolk va	White Beans (Mature Seeds, Canned)	32	Male	Dinner	2018-09-20

Figure 6.4: Features of recommended food list page

1. Button that facilitate to generate csv file with food items according to disease which can directly feed the weka tool for associate mining
2. History records of diabetes patients with meal and date

6.1.5 Calculate daily calorie intake

Using the details of the users such as age, weight, height and sex, able to calculate the daily calorie intake of specific user. It is displayed directly to the user.

Your daily calorie intake should be: **1462.248** 1

Weight (Kg)	Height (Cm)	Age (Years)
<input type="text" value="81.00"/>	<input type="text" value="159"/>	<input type="text" value="32"/>

Figure 6.5: Calculation of daily calorie intake

1. Daily calorie intake of specific user

6.1.6 Give an exercise schedule

Then integrate data from Fatsecret API calls and the data generated/populated using advanced data generator for mysql into the same database. From Fatsecret API calls, stored exercises with the amount of burned calories per one hour. These exercises and the amount of burn calories provided an exercise schedule for specific user who need a weight loss.

Exercise List **(146.2248 Calories)** 1

Exercise Name	Time	Burn Calories
Calisthenics (light, e.g. home exercise)	5Min	21.416666666667
Showering	5Min	12.25
Conditioning exercise (health club)	5Min	33.666666666667
Teaching	5Min	12.25
Boating or Canoeing	5Min	18.333333333333
Weight Training (Bodybuilding)	4.93Min	48.3081

Figure 6.6: Exercise schedule for relevant customer

1. Total calories that should burn from exercises
2. Each exercise with the burning amount of calories within 5 minutes

6.1.7 Facilities for users

Developed a web front end for customers by letting them to enter their age, weight, height and gender as basic inputs. According to them, calculate the daily calorie intake of that user. If a user suffering a non-infectious disease or more than one, also able to select it from the system. Other than that, if he need weight gain or weight loss, let him to select this from a drop-down list. After giving these parameter inputs, user can get recommended food plan with an exercise schedule. The total calorie count of the food plan is equal to his calculated daily calorie intake.

Foods List

Food Name	Metric Serving Amount	Calories	Carbohydrate	Protein	Fat
Farina Cereal (Without Salt, Cooked with Water, Enriched)	100g	48.0687	10.472103	1.420601	0.068670
Puffed Rice Cereal (Fortified)	100g	400.0000	89.785714	6.285714	0.500000
Great Northern Beans (Mature Seeds, Canned)	100g	114.1221	21.030534	7.370229	0.389313
Kidney Beans (Without Salt, Cooked, Boiled)	100g	127.1186	22.802260	8.672316	0.497175
White Beans (Mature Seeds, Canned)	100g	117.1756	21.938931	7.259542	0.290076
Baby Lima Beans (Mature Seeds, Without Salt, Cooked, Boiled)	100g	125.8242	23.307692	8.038462	0.379121
Chinese Noodles (Mung Beans) (Cellophane or Long Rice)	100g	350.7143	86.092857	0.157143	0.057143
Spanish Peanuts	31.45g	1021.3341	28.3568	46.8684	88.9003

Figure 6.7: Food plan with calorie amount of certain criteria

1. Calorie count is most probably equal to the daily calorie intake

Your daily calorie intake should be: 1624.72

Weight (Kg) Height (Cm) Age (Years) Want to Disease 1 Disease 2

Foods List

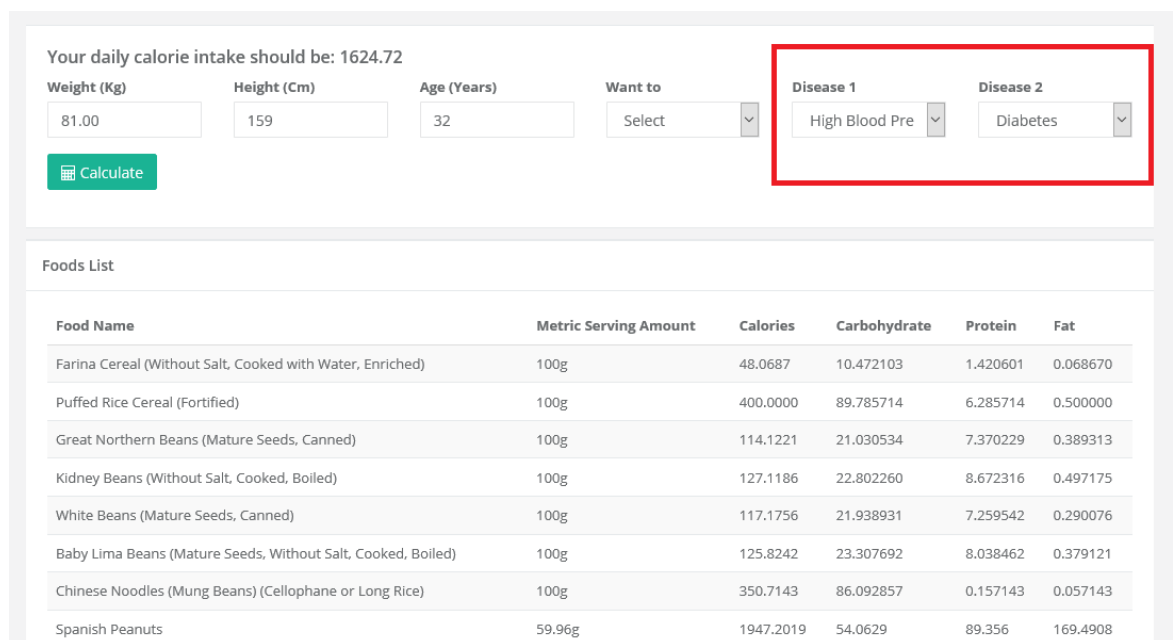
Food Name	Metric Serving Amount	Calories	Carbohydrate	Protein	Fat
Farina Cereal (Without Salt, Cooked with Water, Enriched)	100g	48.0687	10.472103	1.420601	0.068670
Puffed Rice Cereal (Fortified)	100g	400.0000	89.785714	6.285714	0.500000
Great Northern Beans (Mature Seeds, Canned)	100g	114.1221	21.030534	7.370229	0.389313
Kidney Beans (Without Salt, Cooked, Boiled)	100g	127.1186	22.802260	8.672316	0.497175
White Beans (Mature Seeds, Canned)	100g	117.1756	21.938931	7.259542	0.290076
Baby Lima Beans (Mature Seeds, Without Salt, Cooked, Boiled)	100g	125.8242	23.307692	8.038462	0.379121

Figure 6.8: Food plan details of a person who need a weight loss

6.2 Special functions

6.2.1 Recommend a food plan for two identified non-infectious diseases

Identify most frequent food items which are recommended to the specific non-infectious disease using associate rule mining. The associated output is saved in a text file and from this text file get the relevant data and saved in a database table. Then get the most frequent food items for another specific non-infectious disease by applying same associate mining rules. Finally get the most frequent food items which are recommended for both of them.



Your daily calorie intake should be: 1624.72

Weight (Kg): 81.00 Height (Cm): 159 Age (Years): 32 Want to: Select

Disease 1: High Blood Pre Disease 2: Diabetes

Calculate

Foods List

Food Name	Metric Serving Amount	Calories	Carbohydrate	Protein	Fat
Farina Cereal (Without Salt, Cooked with Water, Enriched)	100g	48.0687	10.472103	1.420601	0.068670
Puffed Rice Cereal (Fortified)	100g	400.0000	89.785714	6.285714	0.500000
Great Northern Beans (Mature Seeds, Canned)	100g	114.1221	21.030534	7.370229	0.389313
Kidney Beans (Without Salt, Cooked, Boiled)	100g	127.1186	22.802260	8.672316	0.497175
White Beans (Mature Seeds, Canned)	100g	117.1756	21.938931	7.259542	0.290076
Baby Lima Beans (Mature Seeds, Without Salt, Cooked, Boiled)	100g	125.8242	23.307692	8.038462	0.379121
Chinese Noodles (Mung Beans) (Cellophane or Long Rice)	100g	350.7143	86.092857	0.157143	0.057143
Spanish Peanuts	59.96g	1947.2019	54.0629	89.356	169.4908

Figure 6.9: Filtering recommended foods for 2 diseases

6.2.2 Using exercises and recommended food plan, achieve weight gain or weight loss

If a user wants a weight gain or weight loss he must do it according to daily calorie intake. After calculating the daily calorie intake for specific person, then identify the purpose of weight gain or weight loss. If it is a weight gain, then calculate the amount of calorie which are added to the daily calorie intake. If it is a weight loss, then lose specific number of calories from the daily calorie intake. After that, recommend a food plan which is almost nearest to the daily calorie intake of that user.

This whole process is summarized below.

Evaluation

7.1 Chapter Introduction

This chapter is explained the test results of users who followed up the process of weight gain or weight loss with recommended food plan and exercise schedule.

7.2 Test results of the process of weight loss

This web application helps to get most probable diet plans for people who are suffering non-infectious diseases and who are planning to reduce or gain their weights. Those who need to weight gain or weight loss, can enter their weight details weekly. When they enter their weights with the date, then visualize a graph of weight over the date.

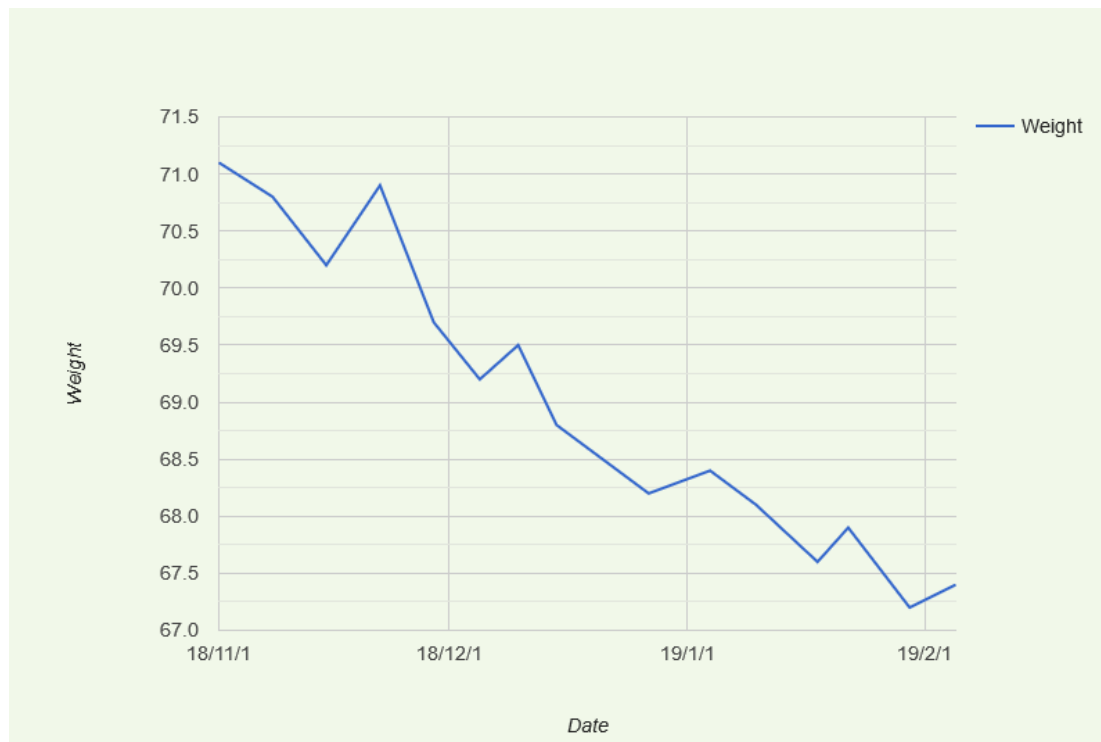
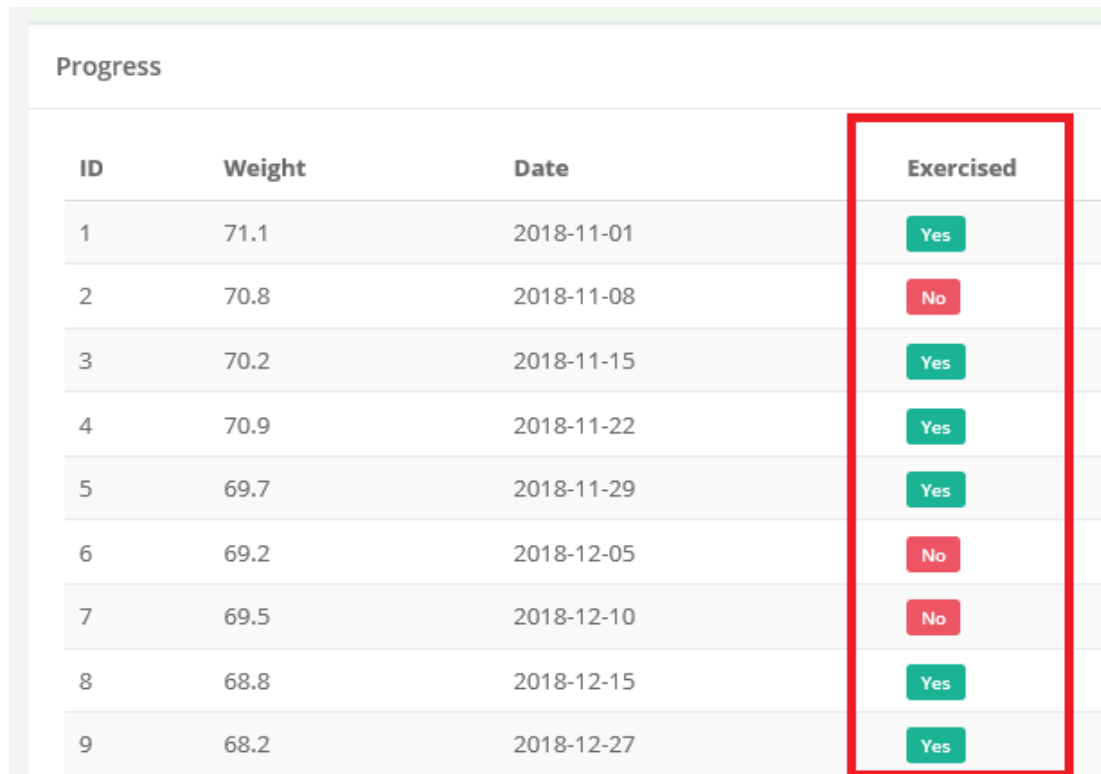


Figure 7.1: Weight over the date for a user who is willing to weight loss

Each user can clearly see their weight loss or weight gain progress within few months. Although there are not continuous weight loss or weight gain, overall weight gain or weight loss can be seen through the graph.

When given a diet plan for a specific user, also he got a recommended exercise schedule too. Some of them skip them and only take the recommended food plan. Therefore it is needed to have an indicator whether they have done their exercises. So, let the user to indicate that he is done the exercises correctly or not. Then through the system indicate it as red or green color beside the date he added his weight.



ID	Weight	Date	Exercised
1	71.1	2018-11-01	Yes
2	70.8	2018-11-08	No
3	70.2	2018-11-15	Yes
4	70.9	2018-11-22	Yes
5	69.7	2018-11-29	Yes
6	69.2	2018-12-05	No
7	69.5	2018-12-10	No
8	68.8	2018-12-15	Yes
9	68.2	2018-12-27	Yes

Figure 7.2: Weight loss progress with the indicator of exercises are done or not

7.3 Chapter Summary

This chapter was discussed the results of user weights after following recommended food plans.

Conclusion and Further Works

8.1 Chapter Introduction

This chapter is briefly described the achievements of the objectives and further works.

8.2 Conclusion

Most nutritional analysis of foods covers a small range of nutrients because lack of information is provided on the wide scope of nutrients. The food composition data should be representative, and high analytical quality. The analysis foods and nutrients must be comprehensive. Food details and descriptions must be totally clean and clear, and data should consistently with unambiguity. Tables of databases should be simple to use, and it should be filled with low missing values.

Data of Food composition are critically applied for the evaluation and the designing of human strength and nutritional intakes. This is much important when related to the groups of people more than evaluate for individuals. A person's intake of food nutrient can be simply calculated by using data of food and nutrient composition as well as details of food intake. This information can review adequate or inadequate of dietary and dietary imbalance.

According to implementation of this decision support system for food nutritional analysis, using Fatsecret API calls, fetched food items with their nutrient amounts as well as exercise details saved in database table. Client can register as a new user and he can get his diet plan by providing specific details. When a customer has one or more non-infectious disease, then he can have a diet plan according to that scenario. Not only that if someone need to lose his weight or gain his weight, then he can enter his weight, age and gender details and calculate his daily calorie intake to recommend a diet plan. Admin can access these records and keep history data to get the most frequent food items. For the history data, use the advanced data generator for MySQL, populate thousands of records to develop the data set for associate data mining.

Outputs are when a user has two diseases, then recommended frequent food items for both diseases. Not only that, if a user wants to lose weight or gain weight, then give suggestions for relevant food items and give three or four exercises with time durations to control his body weight.

8.3 Further Works

For the further works, let the users to give their feedbacks about telling the progress. The system should be improved to keep details of diagnosis of certain non-infectious diseases. By considering the diabetes as non-infectious disease, then let these users to update their monthly blood report details. When the sugar level is in the relevant range, let the user knowing that he is going in the correct path.

8.4 Chapter Summary

This chapter was discussed the conclusion of my research and future works.

References

- [1] “Importance of Nutrition Labels.” [Online]. Available: <http://healthyeating.sfgate.com/importance-nutrition-labels-5212.html>. [Accessed: 20-Oct-2017].
- [2] “Importance of Reading Nutrition Facts and Food Labels.” [Online]. Available: <http://www.sampateek.com/en/articles/importance-reading-nutrition-facts-and-food-labels>. [Accessed: 19-Oct-2017].
- [3] M. Arch and B. Graetz, “Community health studies volume xiii, number 2, 1989,” vol. XIII, no. 2, pp. 177–185, 1989.
- [4] J. M. Holden, S. A. Bhagwat, and K. Y. Patterson, “Development of a Multi-nutrient Data Quality Evaluation System,” *J. Food Compos. Anal.*, vol. 15, no. 4, pp. 339–348, 2002.
- [5] N. Raper, B. Perloff, L. Ingwersen, L. Steinfeldt, and J. Anand, “An overview of USDA’s Dietary Intake Data System,” *J. Food Compos. Anal.*, vol. 17, no. 3–4, pp. 545–555, 2004.
- [6] F. B. Hu, “Dietary pattern analysis: a new direction in nutritional epidemiology,” *Curr. Opin. Lipidol.*, vol. 13, no. 1, pp. 3–9, 2002.
- [7] B. E. Millen *et al.*, “Unique dietary patterns and chronic disease risk profiles of adult men: The Framingham nutrition studies,” *J. Am. Diet. Assoc.*, vol. 105, no. 11, pp. 1723–1734, 2005.
- [8] F. B. Hu *et al.*, “Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire,” *Am. J. Clin. Nutr.*, vol. 69, no. 2, pp. 243–9, 1999.
- [9] E. B. Rimm, E. L. Giovannucci, M. J. Stampfer, G. A. Colditz, L. B. Litin, and W. C. Willett, “Reproducibility and Validity of an Expanded Self-Administered Semiquantitative Food Frequency Questionnaire among Male Health Professionals,” *Am. J. Epidemiol.*, vol. 135, no. 10, pp. 1114–1126, 1992.
- [10] P. K. Newby, D. Muller, J. Hallfrisch, N. Qiao, R. Andres, and K. L. Tucker, “Dietary patterns and changes in body mass index and waist circumference in adults 1 – 3,” 2003.

- [11] M. F. Pyramid, "APPENDIX A1 : Dietary Guidelines for the Prevention of Obesity Maintain healthy body weight by balancing food intake with regular Minimize fat in food preparation and choose foods that are low in fat," pp. 68–87.
- [12] P. Bernardin and A. Moller, "Recommended Daily Caloric Intake," no. 854, p. 2005, 2005.
- [13] Q. V. Le, "Building high-level features using large scale unsupervised learning," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 8595–8598, 2013.
- [14] T. Hofmann, "Unsupervised learning by probabilistic Latent Semantic Analysis," *Mach. Learn.*, vol. 42, no. 1–2, pp. 177–196, 2001.
- [15] C. Mihaescu, "Mining Frequent Itemsets – Apriori Algorithm," *Lab. Modul. 8*, 2012.
- [16] R. Agrawal and R. Srikant, "Fast Algorithms for Mining Association Rules," *the 20th Int. Conf. Very Large Data Bases*, pp. 487–499, 1994.
- [17] Limo A.K *et al.*, "The WEKA Data Mining Software: An Update," *20th Congr. IADR East South. Africa Div. (Sept. 28-30, 2006)*, vol. 11, no. 1, pp. 10–18, 2006.
- [18] I. Russell and Z. Markov, "An Introduction to the Weka Data Mining System (Abstract Only)," *Proc. 2017 ACM SIGCSE Tech. Symp. Comput. Sci. Educ. - SIGCSE '17*, vol. 18, no. 3, pp. 742–742, 2017.
- [19] S. Barquera, J. A. Rivera, J. Espinosa-Montero, M. Safdie, F. Campirano, and E. A. Monterrubio, "Energy and nutrient consumption in Mexican women 12-49 years of age: Analysis of the National Nutrition Survey 1999," *Salud Publica Mex.*, vol. 45, no. SUPPL. 4, 2003.
- [20] Agrawal, "5-721: Database Management Systems Fast Algorithms for Mining Association Rules," pp. 1–2, 2005.
- [21] T. N. England, "Number 16 a Clinical Trial of the Effects of Dietary Patterns on Blood Pressure," pp. 1117–1124, 2015.
- [22] Currie; *et al.*, "Comparison of overweight and obesity prevalence in school-aged youth from 34 countries and their relationships with physical activity and dietary patterns," *Obes. Rev.*, vol. 6, no. 2, pp. 123–132, 2005.

- [23] R. Srikant, Q. Vu, and R. Agrawal, "Mining association rules with item constraints," *Kdd*, vol. 97, pp. 67–73, 1997.
- [24] B. Liu, W. Hsu, and Y. Ma, "KDD98-012.pdf," 1998.
- [25] S. Kotsiantis and D. Kanellopoulos, "Association Rules Mining: A Recent Overview," *GESTS Int. Trans. Comput. Sci. Eng.*, vol. 32, no. 1, pp. 71–82, 2006.
- [26] Q. Zhao, "Association Rule Mining : A Survey," *Surv. Pap.*, no. 2003116, pp. 1–20, 2003.
- [27] H. Greenfield and D. a. T. Southgate, "Food Composition Data," 1992.