ADDIS ABABA UNIVERSTY
SCHOOL OF GRADUATE STUDIES
COLLEGE OF NATURAL SCIENCES
DEPARTMENT OF INFROMATION SCIENCE

APPLICATION OF A HYBRID RECOMMENDER SYSTEM FOR ETHIOPIAN LOCAL FOOD SELECTION TO DIABETES TYPE 2 PATIENTS

A Thesis submitted to the school of Graduate Studies of Addis Ababa University in Partial fulfillment of the requirements for the Degree of Master of Science in Information Science

By

SimenehYesigat

October 2015

ADDIS ABABA UNIVERSTY
SCHOOL OF GRADUATE STUDIES
COLLEGE OF NATURAL SCIENCES
DEPARTMENT OF INFROMATION SCIENCE

This is to certify that the thesis prepared by SimenehYesigat, ***entitled: Application of hybrid recommender system for Ethiopian local food Selection to Diabetes type 2 patients***, and submitted to the college of Natural Sciences of Addis Ababa university in partial fulfillment of the requirements for the degree of Master of Science in Information Science.

By

SIMENEH YESIGAT

Signed by the Examining Committee:

Examiner: ___________________Signature________Date__________

Examiner: _______________Signature____________Date__________

Advisor: _________________Signature____________Date__________

Declaration

I declare that the thesis is my original work and has not been presented for a degree in other university.________________________

This thesis has been submitted for examination with my approval as university advisor

____

Advisor

Acknowledgements

First and for most I would like to give a special Gratitude to the Glory of God and who provided me the courage to finish my study.

I am deeply indebted to my Advisor Mr. Getachew Jemaneh and my former instructor Dr. Million Meshesha for their dedication, encouragement, inspiring guidance and valuable comments throughout my work.
I am extremely thankful to the organizations that showed interest in my research and collaborate with me to the success of the thesis.

My Special thanks go to my wife, Medhanit Habtamu, for her care, support and inspiration. Without her full support neither the research nor the study would have been possible.

My greatest gratitude is extended to all my family members for their encouragement and supports throughout my study. Finally, I would like to thank sincerely all my friends who helped me with their valuable support during the entire process of my study.

Table of Content

List of Figures

List of Tables

List of Acronyms

ADA: America Diabetes Association

CF: Collaborative

CN: Content Based

DM: Demographic

EDA: Ethiopian Diabetes Association

ENHRI: Ethiopian Nutrition and Health Research Institute

FRS: Food Recommender System

IDF: International Diabetes Federation

JCOLIBRI: java Cases and Ontology Libraries Integration for Building Reasoning Infrastructure

SDLC: System Development Life Cycle

Abstract

Diabetes is becoming one of the rapidly increasing non-communicable diseases and an important public health problem all over the world. Ethiopia is at a risk of increased diabetes incidence. The number of deaths attributed to diabetes reached over 21,000 in 2007. Among this number, Type-2 diabetes constitutes about 85 to 95% of all diabetes. Dietary management is considered to be one of the cornerstones of diabetes care. Good diabetes management is a balance between healthy eating, exercise and medication. The problem, however, is that most diabetic patients have difficulty of identifying the recommended quality and quantity of food that they have to eat in order to control their blood glucose.

The aim of the study is to assess the applicability of a hybrid filtering approach to create a personalized food recommender system that can promote healthier eating habits. So that, the patients can safely chooses a healthier food that can promote their self-dietary management.

To achieve this goal, the study attempts to design and develop a prototype hybrid recommender system which is a combination of both collaborative, content base and demographic filtering methods with a mixed, cascade and switching hybrid strategies. The knowledge is acquired using unstructured interviews from domain expert and relevant documents analysis method is also applied to capture explicit knowledge. The data about Ethiopian foods is collected from the food composition table from ENHRI booklet. The prototype is developed by using SQL Server 2008 database with a visual Basic programming language.

Moreover, in testing and evaluating the prototype system five volunteer patients were involved. The experiment conducted in five iterations for each user. So the result shows that, the recommender system has a good performance and achieved the desired goal. But further research is needed to be conducted of exploring different options of hybrid strategies.

Chapter One

1. Background of the Study

1.1 Introduction

According to IDF Diabetes Atlas Diabetes is recognized as a group of heterogeneous disorders with the common elements of hyperglycaemia and glucose intolerance, due to insulin deficiency, impaired effectiveness of insulin action, or both. If there is too little insulin in the body, or if the cells are insulin resistant, the body fails to transport glucose to the cells. This causes glucose to build up in the blood and makes the blood glucose level high. High blood glucose levels increase the risk on a wide variety of other diseases.

Diabetes is classified on the basis of aetiology and clinical presentation of the disorder in to two. Type 1 diabetes is generally discovered early in life and is characterized by a total insulin deficiency; the pancreas fails to produce insulin .These people need to inject insulin regularly. Type 2 diabetes usually develops at a later age. Unlike type 1 that is caused by genetic factors, in type 2 the disease is usually caused by a combination of genetics and lifestyle. Often diabetes type 2 is found in people that are obese. The pancreas does not produce enough insulin anymore or the body cells have become (partly) insulin resistant. In the first years after diagnosis with diabetes type 2, patients can sometimes postpone the need for medication by changing their lifestyle in terms of eating habits and physical activity. Even in a later stage, a good lifestyle is important for these people, because it makes the disease more controllable (Harriset al, 2003). For the remainder of this thesis we will focus on the food nutrition aspects associated with diabetes type 2. When diabetes is mentioned, it refers to type 2, unless the type is specified otherwise.

Nutrition therapy is a major solution to control diabetes by managing the nutrition based on the belief that food provides vital medicine and maintains a good health. Although medicine is our food, food is also our medicine. Some health problems require specific medication; many conditions of problems can be relieved effectively with nutrition therapy. Not only a single disease, but also a group of autoimmune diseases is included in the diabetes. The human body usually needs sugar for energy; however, too much sugar in blood can vitally damage the body especially the diabetes. Therefore, the diabetes prevention would be the proper nutrition and healthy diet which balance sugar to the optimal level and maintain a healthy weight, respectively (Alison et al, 2013).

Diabetic patients, who have been suffering from diabetes, are advised to be aware of their blood glucose level when the diagnosis shows that the level was higher than the limit and can cause more chance to develop a complication. It is also suggested to be more careful in their diet but, usually, the patients are reluctant to rely on the suggestion because their daily food consumption behavior is barely changed.

A recommender system is an intermediary program (or an agent) with a user interface that automatically and intelligently generates a list of information which suits the needs of users (Haruechaiyasak et al, 2004). It provides advice to users about items or services they might wish to purchase or examine. Recommendations made by such systems can help users navigate through large information spaces of product or service descriptions, news articles or other items (Burke, 2006).

When people have to make a choice without any personal knowledge of the alternatives from huge amount of options, a natural course of action is to rely on the experience and opinions of others. We seek recommendations from people who are familiar with the situation, who have been helpful in the past, whose perspectives we value, or who are recognized experts. We might turn to friends or colleagues, the owner of a neighborhood bookstore, movie reviews in a newspaper or magazine, or Consumers

Union product ratings. And we may find the social process of meeting and conversing with people who share our interests as important as the recommendations we receive (Terveen et al, 2001).Recommender systems intend to provide people with recommendations of products they will appreciate, based on their past preferences, history of purchase, and demographic information (Ziegler et al, 2005).

Recommender systems can now be found in many modern applications that expose the user to huge collections of items both commercially and in the research community, where many approaches have been suggested for providing recommendations. Such systems typically provide the user with a list of recommended items they might prefer, or predict how much they might prefer each item. It helps to match the users with items to make ease information overload and sales assistance (Shani et al, 2010). It is the web based application that uses the opinions of the community of users to help customers in the decision making and product selection process to identify content of interest from vast set of choices (Zanker, 2010). These choices may be books, music CDs, movie DVDs, flight tickets, hotel reservations, websites, field of studies, and they may be people (Ning , 2007).

1.2 Statement of the Problem

According to International Diabetes Federation (IDF) report in 2013, Diabetes is becoming one of the rapidly increasing non-communicable diseases and an important public health problem all over the world and an estimated number of 246 million people, or 5.9% of the world's populations, in the age group 20-79 have diabetes worldwide in 2007. More than 70% of the total estimated number is live in the developing countries. As the Diabetes Atlas elucidates the worldwide estimate is expected to increase to some 380 million, or 7.1% of the adult population by 2025 and the largest increases will take place in regions found in developing countries.

Sub-Saharan Africa, like the rest of the world, is experiencing an increasing prevalence of diabetes alongside other non-communicable diseases (Hall et al. 2011; Gill et al.

2009). Ethiopia, which is one of the developing nations, is at a risk of increased diabetes incidence. The number of deaths attributed to diabetes reached over 21,000 in 2007. This estimate has increased to about 25,000 in 2011 (Haregu et al 2012). Among this number, Type-2 diabetes constitutes about 85 to 95% of all diabetes in high-income countries and accounts for an even higher percentage in low and middle-income countries (Sicree et al. 2010).

Dietary practice refers to patients' choices in food consumption based on diabetes nutrition knowledge that gives emphasis to intake of food with lower fat, higher fiber (Shamsi et al. 2013). A study done in Bahrain indicated that lack of proper professional dietary assessment, follow-up and advice by health care providers were the main influencers on dietary practice of type 2 diabetic patients (Shamsi et al. 2013). Another study done at Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia, indicated that the majority of the patients had poor adherence to self-care practices, especially in diet management practices (Berhe et al. 2012).

However, in Ethiopia incomplete routine health information and lack of data on the proper dietary practice of diabetic patients affect the long term management of diabetes. These diabetic patients are facing difficulty in choosing food items when they feel like eating. They also fail to decide how much to eat whenever necessary. At the same time their care givers also fail to identify food items to be included in the diabetic meal and how to prepare them. According to the research conducted on the dietary practice and diabetes in Ethiopia, The prevalence of poor dietary practice was observed as a major factor that influence the effective management of diabetes in Ethiopia; it is therefore a major public health problem. Not getting diabetic nutrition education at hospitals, being despondency, difficulty of choosing relevant foods for their specific health problems, non-availability of fruits and vegetables, and thinking about the high cost of foods were important factors affecting dietary practices of type 2 diabetic patients. To deal with this problem, Health care services should also empower patients to heal themselves by addressing the causes of their disease and facilitating lifestyle changes like promoting self-administered dietary management through health promotion (Solomon et.al.2015).

This study developed a prototype food recommender system that can help to promote diabetes patients healthier eating habits. These improved eating habits can prevent or slow down the development of diabetes type 2 with its adverse health effects or help people that already suffer from diabetes type 2 to manage their disease through a proper dietary practice. This study will also describe the development of food recommender system algorithm that will help diabetes type 2 patients adopt a healthier food based on their personal ratings or likes / dislikes of ingredients as well as the their individual nutritional goal. The recommended food should be healthy and available to the patient.

1.3. Research Questions

The study tried to investigate and address the following research questions.

- How to develop an effective prototype of recommender system algorithm that can help diabetes type 2 patient in their daily food selection?
- What are the most important attributes that influence the food choice of diabetes type 2 patients?
- How can we acquire, model, represent and implement the required knowledge for a recommender system?

1.4. Objective of the Study

The following general and specific objectives are formulated towards solving the research problems.

1.4.1 General Objective

The general objective of the study is to design a prototype recommender system that can promote healthier eating habits for type 2 Diabetes patients using their personal ratings of the ingredients and nutritional preset goal.

1.4.2 Specific Objectives

With the above general goal in mind the research tried to fulfill the following specific objectives.

- To design conceptual model comprising relevant attribute that have a direct impact on diabetes patients recommended eating habits.
- To build a prototype recommender system model that can facilitate daily base food selection activities of diabetic type 2 patients.
- To develop a hybrid food recommender system algorithm typical to diabetic's type 2 patients.
- To evaluate the performance of recommender system using accuracy measures.

1.5. Scope of the Study

This thesis describes an alternative method to promote healthful eating habits amongst people with type 2 diabetes by offering personalized food recommendations in a computer system. For this purpose it is important to find out which characteristics of type 2 diabetic patients and which characteristics of the food stuffs are important for recommending healthier foods. Since, it is necessary to be able to identify these characteristics automatically by the system; this paper will deal with all the food items with their corresponding ingredients and nutritional values. Moreover, the user's interest must be collected by the system in a certain manner in this case rating for the ingredients. The food recommender system (FRS) approach that will be used in this paper will let the patient help themselves on choosing healthier foods. In addition to that the system will consider the geographical location of the patient to recommend foods that are available nearby the residence of the patient. On top of that the system has self-learning capabilities that will develop a personal profile for the patients that will grow as the patient ratings of the ingredients grow. The scope for the recommendation is bounded for Ethiopian local foods only.

1.6 Significance and Application of the Study

Awareness of unhealthy eating behaviour and knowledge about healthful behaviour does not seem to be enough to actually change the behaviour. This shows that it is important to guide people in their food decision process in another way in. Guidance in food decision would be beneficial to people with diabetes or other lifestyle-related diseases in particular. This thesis has a significant contribution to help the patients to safely select the healthier food and also provide more choices in consuming. In general, it will have a paramount importance on promoting a healthier eating habit and better diabetic's self-management practices.

In a contemporary world that we live in, information overload is a big issue, i.e. patients who are willing to change their eating habits couldn't find an easy ride to get trustworthy and personalized dietary information on the internet even though there are a bunch of sites on food recommendation. None of them are more specific for Ethiopian food cultures. So this thesis will play a big role and can be a spark for further researches on a computer based Food recommender systems in Ethiopia context.

1.7 Organization of the thesis

This thesis is organized into six chapters: the first Chapter briefly discusses about the introduction part of the study, background of the problem area, the general and specific objectives of the study, the scope and limitation, programming tools used to develop the prototype, evaluation procedure and significance and application of the results of the research. Chapter two gives an overview of related research work and background introduction about recommender system. Chapter three focuses methodology and Chapter four dedicated to creation of knowledge modeling to develop recommender system. Chapter five discusses on designing and implementation of the prototype which discuss the detail of developed system using selected programming tool. Testing and evaluation of the system by domain area is discussed in the sixth chapter and finally chapter final chapter presents the conclusions of the research done and forward recommendations for future studies.

Chapter 2

2. Literature Review

It is vital to review several literatures that have been written so far on the relevant topics to have through understanding on the problem at hand. For this reason, related literature such as books, journal articles, proceeding papers, magazines, manuals and some other sources have been consulted so as to understand the domain knowledge, concepts, principles and methods that are important for developing recommender system and for achieving the research objective.

2.1 Overview of Recommender system

Recommender systems were originally defined as ones in which "People provide recommendation as inputs, which the system then aggregates and directs to appropriate recipient" (resenick & varian 1997). The term now has a broader connotation, describing any system that produces individualized recommendations as output or has the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options. Such systems have an obvious appeal in an environment where the amount of on-line information vastly outstrips any individual's capability to survey It. (Schafer, Konstan and Riedl, 1999).

As Bruke stated it is the criteria of "individualized and interesting and useful" that separate the recommender system from information retrieval systems or search engines. The semantics of a search engine are "matching": the system is supposed to return all those items that match the query ranked by degree of match. Techniques such as relevance feedback enable a search engine to refine its representation of the user's query, and represent a simple form of recommendation.

When people have to make a choice without any personal knowledge of the alternatives from huge amount of options, a natural course of action is to rely on the experience and opinions of others. We seek recommendations from people who are familiar with the choices we face, who have been helpful in the past, whose perspectives we value, or who are recognized experts. We might turn to friends or colleagues, the owner of a neighborhood bookstore, movie reviews in a newspaper or magazine, or Consumers Union product ratings. And we may find the social process of meeting and conversing with people who share our interests as important as the recommendations we receive (Terveen and Hill, 2001).

Recommender systems intend to provide people with recommendations of products they will appreciate, based on their past preferences, history of purchase, and demographic information (Ziegler et al, 2005).

2.2 Architecture of a recommender system

The architecture of recommender system is represented in figure below.

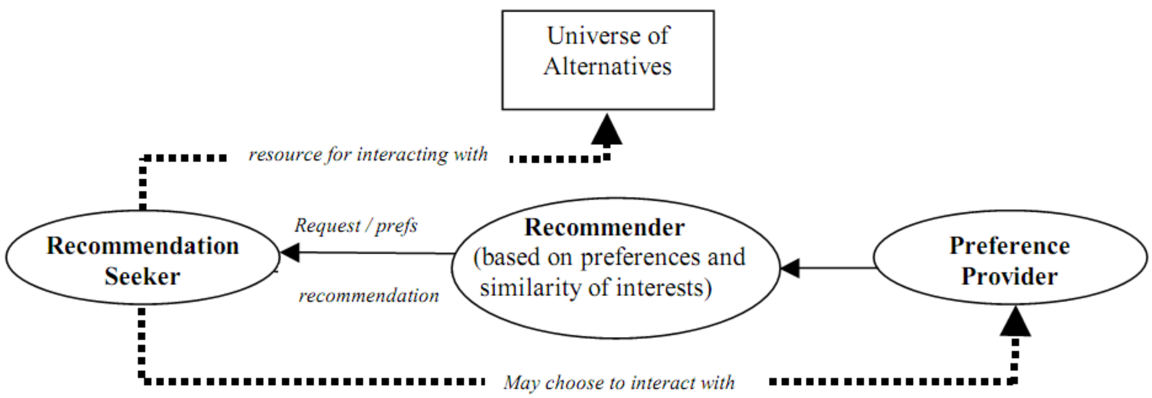

Figure 2.1 Architecture of recommender system (source: Terveen and Hill, 2001)

As shown in the above figure, the recommendation seeker asks recommendation with the help of recommender system from preference provider and universe of alternatives by entering the required queries by them. Based on the items listed in the preference provider, the system provides solution or recommendation by measuring similarity between the users profile and the item similarity with the query, which is provided by the seeker.

In recommender systems, the utility of an item is usually represented by a rating, which measures how much a specific user is interested in the item. Depending on the application, the ratings can either be specified by the users, or be computed by the system. Each element of the user space U can be described with a profile that may include several demographic characteristics, such as gender, age, nationality, marital status, etc., and/or some information about the user's tastes, interests and preferences (Adomavicius et al, 2005).

The way in which such user profiles and item descriptions are defined is a key point in any recommender system. However, it is not the only factor that influences the efficiency and effectiveness of the recommendation processes (Adomavicius & Tuzhilin, 2005). Analogously, each element of the item space I may be described with a set of characteristics or features. For example, in a food recommender system, food can be described not only by their descriptions, but also by their ingredient, nutritional value, etc.

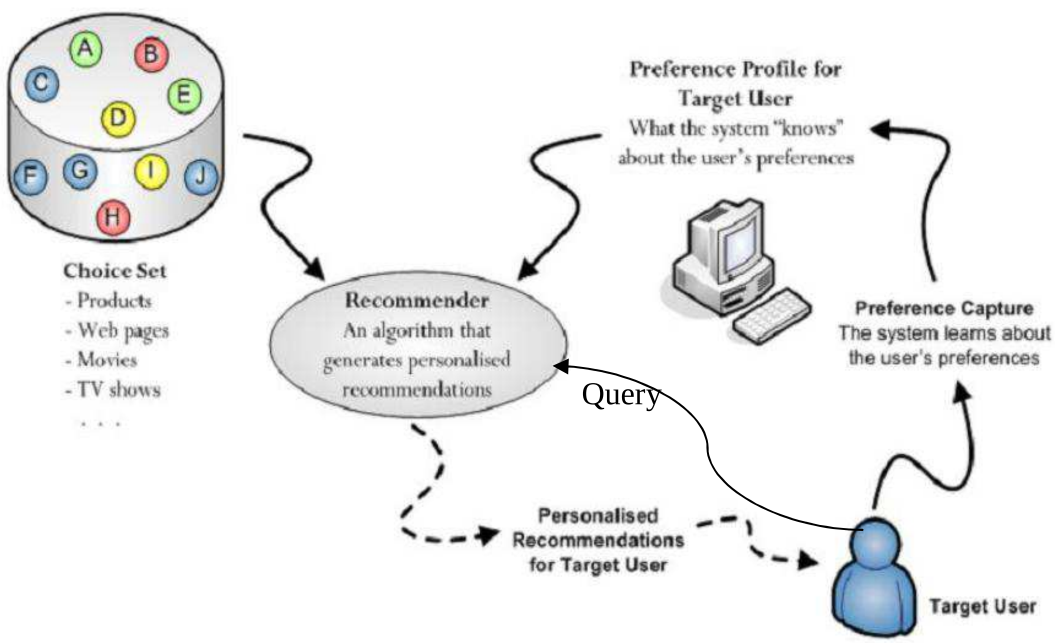

Figure 2.2 Components of recommender system

As shown above, the basic components of a recommender system make interaction to give recommendation for the user. First, a user profile learning module (explicitly or implicitly) captures the preferences from the user. Once the system “knows” about the user’s tastes and interests, it compares and/or combines user profiles and item descriptions. Then item with a maximum gain to the user is recommended (Schafer et al, 2001).

2.3 Recommendation Techniques

Recommendation techniques have a number of possible classifications (Resinick et al 1991; Schafer et al, 1991; Terven & Hill, 2001). The recommender systems have:

i. **Background data:** the information that the system had before the recommendation process begins,
ii. **Input data:** the information that user must communicate to the system in order to generate a recommendation.
iii. **An algorithm** that combines background and input data to arrive at its suggestion
On this basis we can distinguish four major different recommendation techniques as shown below

- ***Collaborative*** recommendation is probably the most familiar, most widely implemented and most mature of the technologies. Collaborative recommender systems aggregate ratings or recommendations of objects, recognize commonalities between users on the basis of their ratings, and generate new recommendations based on inter-user comparisons. A typical user profile in a collaborative system consists of a vector of items and their ratings, continuously augmented as the user interacts with the system over time. Some systems used time-based discounting of ratings to account for drift in user interests (Billsus & Pazzani, 2000; Schwab.et al. 2001). In some cases, ratings may be binary (like/dislike) or real-valued indicating degree of preference. Some of the most important systems using this technique are GroupLens/NetPerceptions (Resnick

.et. al. 1994), Ringo/Firefly (Shardanand & Maes, 1995), Tapestry (Goldberg et al. 1992) and Recommender (Hill et al. 1995). These systems can be memory based, comparing users against each other directly using correlation or other measures, or model-based, in which a model is derived from the historical rating data and used to make predictions (Breese et al. 1998). Model-based recommenders have used a variety of learning techniques including neural networks (Jennings & Higuchi, 1993), latent semantic indexing (Foltz, 1990), and Bayesian networks (Condliff, et al. 1999).

The greatest strength of collaborative techniques is that they are completely *Independent* of any machine-readable representation of the objects being recommended, and work well for complex objects such as music and movies where variations in taste are responsible for much of the variation in preferences call it as "people-to-people correlation." (Schafer. et.al 1999)

- ***Content-based*** recommendation is an outgrowth and continuation of information filtering research (Belkin et al, 1992). In a content-based system, the objects of interest are defined by their associated features. For example, text recommendation systems like the newsgroup filtering system NewsWeeder (Lang, 995) uses the words of their texts as features. A content-based recommender learns a profile of the user's interests based on the features present in objects the user has rated. Schafer, Konstan & Riedl call this "item-to-item correlation." The type of user profile derived by a content-based recommender depends on the learning method employed. Decision trees, neural nets, and vector-based representations have all been used. As in the collaborative case, content-based user profiles are long term models and updated as more evidence about user preferences is observed.

- ***Demographic*** recommender systems aim to categorize the user based on personal attributes and make recommendations based on demographic classes. An early example of this kind of system was Grundy (Rich, 1979) that recommended books based on personal information gathered through an interactive dialogue. The user's responses were matched against a library of manually assembled user stereotypes. Some more recent recommender systems have also taken this approach. (Krulwich 1997), for example, uses demographic groups from marketing research to suggest a range of products and services. A short survey is used to gather the data for user categorization. In other systems, machine learning is used to arrive at a classifier based on demographic data (Pazzani 1999). The representation of demographic information in a user model can vary greatly. Rich's system used hand-crafted attributes with numeric confidence values. Pazzani's model uses Winnow to extract features from users' home pages that are predictive of liking certain restaurants. Demographic techniques form "people-to-people" correlations like collaborative ones, but use different data. The benefit of a demographic approach is that it may not require a history of user ratings of the type needed by collaborative and content-based techniques.

- ***Knowledge-based*** *recommendation* attempts to suggest objects based on inferences about a user's needs and preferences. In some sense, all recommendation techniques could be described as doing some kind of inference. Knowledge-based approaches are distinguished in that they have functional knowledge: they have knowledge about how a particular item meets a particular user need, and can therefore reason about the relationship between a need and a possible recommendation. The user profile can be any knowledge structure that supports this inference. In the simplest case, as in Google, it may simply be the query that the user has formulated. In others, it may be a more detailed representation of the user's needs (Towle et al, 2000). The knowledge used by a knowledge-based recommender can also take many forms.

- ***Utility-based*** approaches calculate a utility value for objects to be recommended, and in principle, such calculations could be based on functional knowledge. However, existing systems do not use such inference, requiring users to do their own mapping between their needs and the features of products, either in the form of preference functions for each feature in the case of Tête-à-Tête or answers to a detailed questionnaire in the case of PersonaLogic. *Utility-based* and *knowledge-based* recommenders do not attempt to build long-term generalizations about their users, but rather base their advice on an evaluation of the match between a user's need and the set of options available. Utility-based recommenders make suggestions based on a computation of the utility of each object for the user. Of course, the central problem is how to create a utility function for each user. The user profile therefore is the utility function that the system has derived for the user, and the system employs constraint satisfaction techniques to locate the best match. The benefit of utility-based recommendation is that it can factor non-product attributes, such as vendor reliability and product availability, into the utility computation, making it possible for example to trade off price against delivery schedule for a user who has an immediate need.(Bruke,2005)

Assuming that **I** is a the set of items over which recommendation might be made, **U** is the set of users whose preference are known, ***u*** is the user for whom recommendation need to be generated, and ***i*** is some item for which we would like to predicted **u's** preference, all the techniques mention above can be summarized in the table below.

Technique	Background	Input	Process
Collaborative	Rating from U of items in I	Rating from u of items in I	Identify users in U similar to u, and extrapolate from their ratings of i.
Content-based	Features of items in I	u's ratings of items in I	Generate a classifier that fits u's rating behavior and use it on i.
Demographic	Demographic information about U and their ratings of items in I.	Demographic information about u.	Identify users that are demographically similar to u, and extrapolate from their ratings of i.
Knowledge-based	Features of items in I. knowledge of how these items meet a user's needs.	A description of u's needs or interests.	Infer a match between I and u's need.
Utility-based	Features of items in I	A utility function over items in I that describes u's preference	Apply the function to the items and determine I's rank.

Table 2.1 Summary of recommender system techniques

2.5 Comparing Recommendation Techniques

All recommendation techniques have strengths and weaknesses discussed below and summarized in Table 2.2. Perhaps the best known is the "ramp-up" problem (Konstan, et al. 1998). This term actually refers to two distinct but related problems.

New User: Because recommendations follow from a comparison between the target user and other users based solely on the accumulation of ratings, a user with few ratings becomes difficult to categorize.

New Item: Similarly, a new item that has not had many ratings also cannot be easily recommended: the "new item" problem. This problem shows up in domains such as news articles where there is a constant stream of new items and each user only rates a few. It is also known as the "early rater" problem, since the first person to rate an item gets little benefit from doing so: such early ratings do not improve a user's ability to match against others (Avery and Zeckhauser, 1997). This makes it necessary for recommender systems to provide other incentives to encourage users to provide ratings.

Collaborative recommender systems depend on overlap in ratings across users and have difficulty when the space of ratings is sparse: few users have rated the same items. The sparsity problem is somewhat reduced in model-based approaches, such as singular value decomposition (Strang, 1988), which can reduce the dimensionality of the space in which comparison takes place (Foltz, 1990; Rosenstein & Lochbaum, 2000). Still sparsity is a significant problem in domains such as news filtering, since there are many items available and, unless the user base is very large, the odds that another user will share a large number of rated items is small.

These three problems suggest that pure collaborative techniques are best suited to problems where the density of user interest is relatively high across a small and static universe of items. If the set of items changes too rapidly, old ratings will be of little value to new users who will not be able to have their ratings compared to those of the existing users. If the set of items is large and user interest thinly spread, then the probability of overlap with other users will be small.

Collaborative recommenders work best for a user who fits into a niche with many neighbors of similar taste. The technique does not work well for so-called “gray sheep” (Claypool, et al. 1999), who fall on a border between existing cliques of users. This is also a problem for demographic systems that attempt to categorize users on personal characteristics. On the other hand, demographic recommenders do not have the “new user” problem, because they do not require a list of ratings from the user. Instead they have the problem of gathering the requisite demographic information. With sensitivity to on-line privacy increasing, especially in electronic commerce contexts (USITIC,1997), demographic recommenders are likely to remain rare: the data most predictive of user preference is likely to be information that users are reluctant to disclose.

Content-based techniques also have a start-up problem in that they must accumulate enough ratings to build a reliable classifier. Relative to collaborative filtering, content-based techniques also have the problem that they are limited by the features that are explicitly associated with the objects that they recommend. For example, content based movie recommendation can only be based on written materials about a movie: actors’ names, plot summaries, etc. because the movie itself is opaque to the system. This puts these techniques at the mercy of the descriptive data available. Collaborative systems rely only on user ratings and can be used to recommend items without any descriptive data. Even in the presence of descriptive data, some experiments have found that collaborative recommender systems can be more accurate than content-based ones (Alspector, et al. 1997).

The great power of the collaborative approach relative to content-based ones is its cross-genre or “outside the box” recommendation ability. It may be that listeners who enjoy free jazz also enjoy avant-garde classical music, but a content-based recommender trained on the preferences of a free jazz aficionado would not be able to suggest items in the classical realm since none of the features (performers, instruments, repertoire) associated with items in the different categories would be shared. Only by looking outside the preferences of the individual can such suggestions be made.

Both content-based and collaborative techniques suffer from the “portfolio effect.” An ideal recommender would not suggest a stock that the user already owns or a movie that has already seen. The problem becomes quite tricky in domains such as news filtering, since stories that look quite similar to those already read may in fact present some new facts or new perspectives that would be valuable to the user. At the same time, many different presentations of the same wire-service story from different newspapers would not be useful. The DailyLearner system uses an upper bound of similarity in its content-based recommender to filter out news items too similar to those already seen by the user (Billsus et al, 2000).

Utility-based and knowledge-based recommenders do not have ramp-up or sparsity problems, since they do not base their recommendations on accumulated statistical evidence. Utility-based techniques require that the system build a complete utility function across all features of the objects under consideration. One benefit of this approach is that it can incorporate many different factors that contribute to the value of a product, such as delivery schedule, warranty terms or conceivably the user’s existing portfolio, rather than just product-specific features. In addition, these non-product features may have extremely idiosyncratic utility: how soon something can be delivered may matter very much to a user facing a deadline. A utility-based framework thereby lets the user express all of the considerations that need to go into a recommendation. For this reason, (Guttman ,1999) describes Tête-à-Tête as “product and merchant brokering" system rather than a recommender system. However, under the definition given above, Tête-à-Tête does fit since its main output is a recommendation (a top-ranked item) that is generated on a personalized basis.

All of the learning-based techniques (collaborative, content-based and demographic) suffer from the ramp-up problem in one form or another. The converse of this problem is the stability vs. plasticity problem for such learners. Once a user's profile has been established in the system, it is difficult to change one's preferences. A beef eater who becomes a vegetarian will continue to get beef recommendations from a content-based or collaborative recommender for some time, until newer ratings have the chance to tip the scales. Many adaptive systems include some sort of temporal discount to cause older ratings to have less influence, but they do so at the risk of losing information about interests that are long-term but sporadically exercised (Billsus & Pazzani, 2000; Schwab, et al. 2001). For example, a user might like to read about major earthquakes when they happen, but such occurrences are sufficiently rare that the ratings associated with last year's earthquake are gone by the time the next big one hits. Knowledge- and utility-based recommenders respond to the user's immediate need and do not need any kind of retraining when preferences change.

The ramp-up problem has the side-effect of excluding casual users from receiving the full benefits of collaborative and content-based recommendation. It is possible to do simple market-basket recommendation with minimal user input: "people who bought X also bought Y" but this mechanism has few of the advantages commonly associated with the collaborative filtering concept. The learning-based technologies work best for dedicated users who are willing to invest some time making their preferences known to the system. Utility- and knowledge-based systems have fewer problems in this regard because they do not rely on having historical data about a user's preferences. Utility-based systems may present difficulties for casual users who might be unwilling to tailor a utility function simply to browse a catalog.

Tradeoffs between Recommendation Techniques

Technique	**Pluses**	**Minuses**
Collaborative Filtering	A,BC,D	I,J,K,L,M
Content- based(CN)	B,C,D	I,L,M
Demographic(DM)	A,B,C	I,K,L,M,N
Utility-based(UT)	E,F,G	N,O
Knowledge-based(KB)	E,F,G,H	P,Q

Keys:

A. Can Identify cross-genre niches
B. Domain knowledge not needed
C. Adaptive: quality improves over time
D. Implicit feedback sufficient
E. No ramp-up required
F. Sensitive to changes of preference
G. Can include non-product features
H. Can map from user needs to products
I. New user ramp-up problem
J. New item ramp-up problem
K. "Gray sheep" problem
L. Quality dependent on large historical data set
M. Stability Vs. plasticity problem
N. Must gather demographic information
O. User must input utility function
P. Suggestion ability static(does not learn)
Q. Knowledge engineering required.

Table 2.2 Tradeoff between recommendation techniques

2.6 Hybrid Recommender System

Hybrid recommender systems combine two or more recommendation technique to gain better performance by minimizing the ramp-up problem, usually a cold start problem.

2.7 Strategies of Hybrid recommendation

Burke provides a thorough analysis of hybrid recommender systems, grouping them into seven classes:

• ***Weighted*** recommenders take the scores produced by several recommenders and combine them to generate a recommendation list (or prediction) for the user.

• ***Switching*** recommenders switch between different algorithms and use the algorithm expected to have the best result in a particular context.

• ***Mixed*** recommenders present the results of several recommenders together. This is similar to weighting, but the results are not necessarily combined into a single list.

• ***Feature-combining*** recommenders use multiple recommendation data sources as inputs to a single meta-recommender algorithm.

• ***Cascading*** recommenders chain the output of one algorithm into the input of another.

• ***Feature-augmenting*** recommenders use the output of one algorithm as one of the input features for another.

• ***Meta-level*** recommenders train a model using one algorithm and use that model as input to another algorithm.

The table below summarizes some of the most prominent combination of hybrid recommender system,

	Weighted	Mixed	Switching	Feature Combination	Cascade	Feature Aug.	Meta-level
CF/CN							
CF/DM							
CF/KB							
CN/CF							
CN/DM							
CN/KB							
DM/CF							
DM/CN							
DM/KB							
KB/CF							
KB/CN							
KB/DM							

CF=Collaborative CN=Content-based DM=Demographic KB=knowledge based

Redundant
Not posssible
Possible

Table 2.3: Combination of Hybrid Recommender System

2.8 Diabetes and Nutrition

According to Diabetes Research and Wellness Foundation, Diabetes is a defect in the body's ability to convert glucose (sugar) to energy. Glucose is the main source of fuel for our body. When food is digested it is changed into fats, protein, or carbohydrates. Foods that affect blood sugars are called carbohydrates.

Carbohydrates, when digested changed to glucose. Examples of some carbohydrates are: bread, rice, pasta, potatoes, corn, fruit, and milk products. Individuals with diabetes should eat carbohydrates but must do so in moderation. Glucose is then transferred to the blood and is used by the cells for energy. In order for glucose to be transferred from the blood into the cells, the hormone - insulin is needed. Insulin is produced by the beta cells in the pancreas (the organ that produces insulin).

In individuals with diabetes, this process is impaired. Diabetes develops when the pancreas fails to produce sufficient quantities of insulin – Type 1 diabetes or the insulin produced is defective and cannot move glucose into the cells – Type 2 diabetes.Either insulin is not produced in sufficient quantities or the insulin produced is defective and cannot move the glucose into the cells.

2.8.1. Types of Diabetes

According to MNT (Medical News Today) Knowledge Center there are two types of diabetes:

1) Type 1 Diabetes

The body does not produce insulin. Some people may refer to this type as **insulin-dependent diabetes**, **juvenile diabetes**, or **early-onset diabetes**. People usually develop type 1 diabetes before their 40th year, often in early adulthood or teenage years. Type 1 diabetes is nowhere near as common as type 2 diabetes. Approximately 10% of all diabetes cases are type 1. Patients with type 1 diabetes will need to take insulin injections for the rest of their life. They must also ensure proper blood-glucose levels by carrying out regular blood tests and following a special diet.

2) Type 2 Diabetes

The body does not produce enough insulin for proper function, or the cells in the body do not react to insulin (insulin resistance). **Approximately 90% of all cases of diabetes worldwide are of this type.**

Some people may be able to control their type 2 diabetes symptoms by losing weight, following a healthy diet, doing plenty of exercise, and monitoring their blood glucose levels. However, type 2 diabetes is typically a progressive disease - it gradually gets worse - and the patient will probably end up have to take insulin, usually in tablet form.

Overweight and obese people have a much higher risk of developing type 2 diabetes compared to those with a healthy body weight. People with a lot of fat, also known as central obesity, belly fat, or abdominal obesity, are especially at risk. Being overweight/obese causes the body to release chemicals that can destabilize the body's cardiovascular and metabolic systems.

2.8.2 Diabetes and Diet in Ethiopia

Dietary management is considered to be one of the cornerstones of diabetes care. It is based on the principle of healthy eating in the context of social, cultural and psychological influences on food choices (Ekore et al. 2008). Good diabetes management is a balance between healthy eating, exercise and medication (Control CfD et al. 2011). The problem, however, is that most diabetic patients have difficulty of identifying the recommended quality and quantity of food that they have to eat in order to control their blood glucose level (association SAd: South African diabetes association. In.: South Africa Diabetes Association 2001).

According to the **Diabetes Education Program for People with Diabetes** publication of Ethiopian Diabetes Association (EDA), Diabetic patients are not supposed to eat differently from the rest of the family but it is important to keep a healthy eating plan.To eat healthy meals means

- Diabetic patients are not supposed to eat more food than their body needs.
- They should eat **one** starchy food with each meal. It will be good to eat starchy foods which raise blood sugar slowly, like injera, potatoes, pasta, darker breads, rice and high fibre unsweetened cereals
- They need to know the amount of starchy food with each meal, like 2 pieces of injera, or 3 potatoes. Using a pesonal plate helps to achieve that.
- Advisable to eat more vegetables. Vegetables and fruit are protects body. Good to eat 2 times more vegetables than starchy foods and meat.
- Must reduce the amount of fat. Using more unsaturated fats, like sunflower and olive oils and avoiding saturated fat like butter and cream is preferable.
- Reduce sugary foods like candy, cookies, sweetened ready-made cereals, pastries and all kinds of soft drinks (Cola, Fanta, etc.)
- Drink very little alcohol (less than 1 drink per day)

2.8.2.1The plate method

Dietary management is considered to be one of the cornerstones of diabetes care. It is based on the principle of healthy eating in the context of social, cultural and psychological influences on food choices (Ekore et al. 2008). Good diabetes management is a balance between healthy eating, exercise and medication (Control CfD et al. 2011). The problem, however, is that most diabetic patients have difficulty of identifying the recommended quality and quantity of food that they have to eat in order to control their blood glucose level (association SAd: South African diabetes association. In.: South Africa Diabetes Association 2001).

It is a way to eat healthier meals without measuring. You fill your plate (20 centimeters in diameter) to match the amount of vegetables, starches and meat and then add a piece of fruit and/or a glass of low fat milk

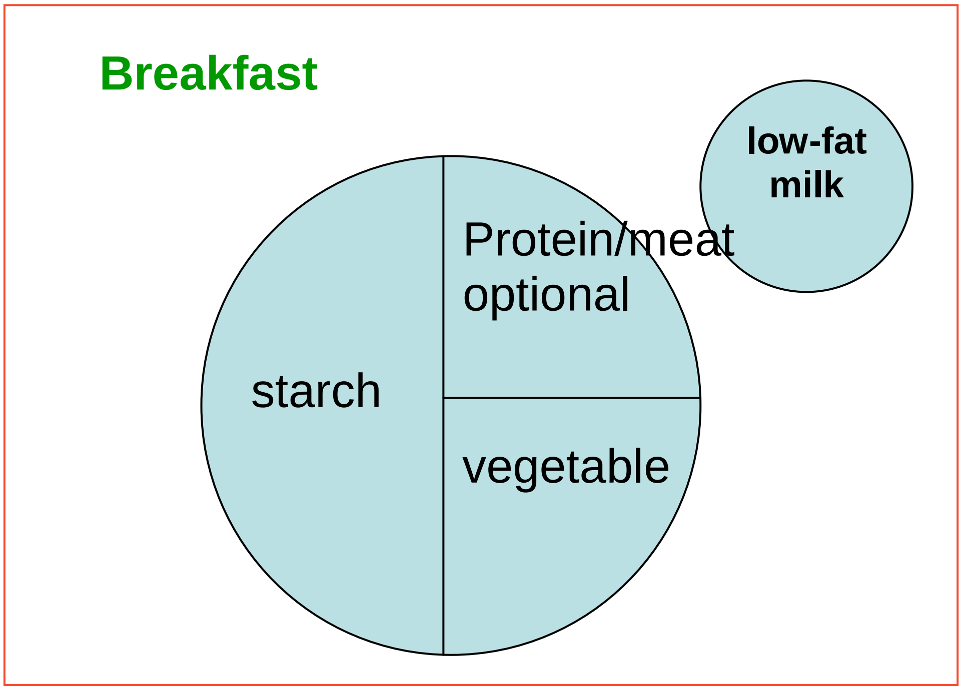

Figure 2.3 A Plate Sample for Breakfast

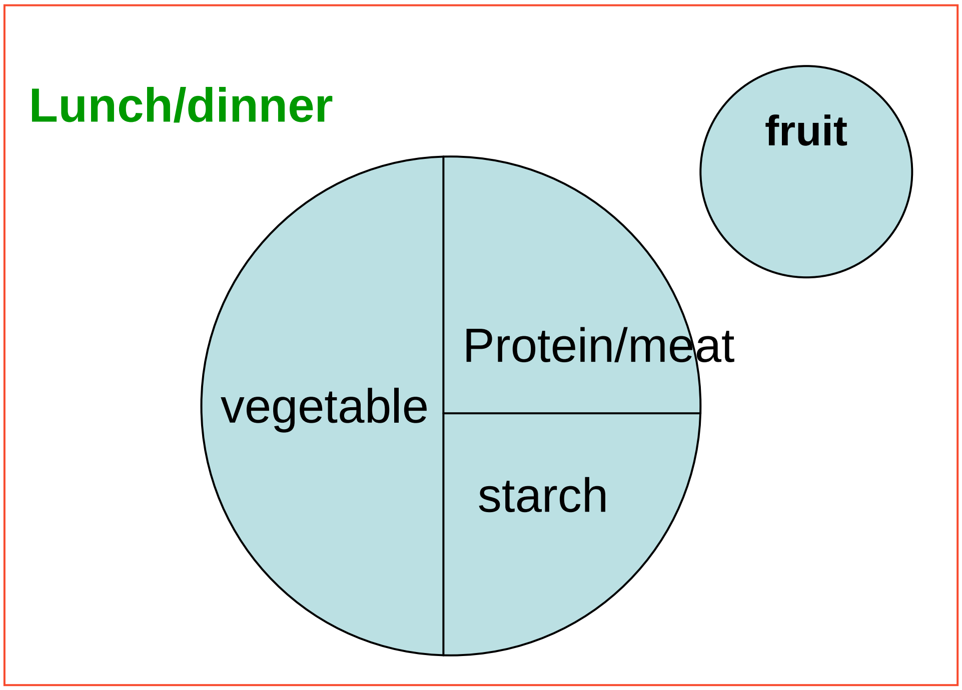

Figure 2.4 : A Plate Method for dinner/Lunch

2.9 Related Works Review

Due to the lack of local related works the researcher forced to look somewhere else to find research papers on recommender systems which is related to my specific area of interest like nutrition recommender system, recipe and meal planning systems .Some of related works are reviewed as follows.

Maiyaporn et.al (2008) proposed a food recommender system using food clustering analysis for diabetic patients. The study was conducted using Self-Organizing Map (SOM) and K-means clustering for food clustering analysis which is based on the similarly of eight significant nutrients for diabetic patients. The system also recommends the proper substitute food in the context of nutrition and food characteristic. At the end the system is evaluated by nutritionist and performed very well and became useful in nutrition areas. Finally the paper recommends improving the system by including proposed services, technology and algorithm for diabetes diet care.

Thienne et.al(2013).designed a mobile food recommendation system based on the traffic light diet. The research is aimed to Innovate and find out a real-time solutions to address the mismatch between the demand and supply of critical information in the context of informing and motivating diet and health-related behavior change. Under the umbrella of the above motivation the researchers stated the objectives on presenting insights related to the development and testing of a mobile food recommendation system targeting fast food restaurants. The system is designed to provide consumers with information about energy density of food options combined with tips for healthier choices when dining out, accessible through a mobile phone.

The researcher developed a mobile application, which is HIPAA (Health Insurance Portability and Accountability Act) complaint, called “Foodtracker” for the Android platform. The choice of opting for native development was due to the need to access to the GPS sensor on the device. The app uses the Traffic Light Diet as basis for the food intake recommendation system. The Traffic Light Diet: Aspects of Epsteins Traffic Light Diet (TLD) were adapted to categorize fast food selections based on energy density. TLD broadly categorizes foods on the basis of their macronutrient content into the colors of a traffic light with similar meaning:

- Green foods are low in calories and fat (< 2g of fat per serving);
- Yellow foods are medium calorie and fat foods (between 2 and 5g of fat per serving); and
- Red foods are high in calories and fat (> 5g of fat per serving).

The TLD categories were chosen because they are common and widespread symbols, which intuitively translate to nutrition guidelines that suggest eat more (Green) eat moderately (Yellow) and eat less (Red), without additional explanations or prior nutrition knowledge. The proposed system is still in its initial testing and evaluation phases, but preliminary data suggest it can be used for members of the community. The prototype was functional, and user's evaluation suggested that the product is useful. The contribution of the research work include new insights on mobile computing applications for health behavior change, such as the use of location-awareness combined with a recommendation database to assist users in making informed choices when eating out.

Sumedh et.al(2002) developed a yelp food recommendation system. The aim of the research is to build a recommender system that can make sophisticated food recommendation system for the users by applying a learning algorithm. The study was conducted by applying the principles and techniques of recommendations systems and developed a predictive model of customer rating for restaurant. The researcher implemented a singular valued decomposition, hybrid cascade of K-nearest neighbor clustering, weighted bi-partite graphs projection and several other algorithms. The system was tested and evaluated using root metrics mean squared error and mean absolute error and it performed best. The researchers also recommended include a review text and user rating evaluation as features in the perdition.

Napat et.al(2010) developed a knowledge-based framework for development of personalized food recommender system. The paper focused on the development of a personalized food recommender system that can provide dietary recommendations, which are based on both individual diet needs and preference. The design of the system uses a knowledge based framework specifically; the knowledge engineering approach is used in modeling the relevant user profile as well as food and nutrition knowledge in ontology form. The paper described development of a personalized food recommender system but it is not evaluated by nutrition experts.

Ipec tatli (2009) performs a Yemek Septi Recommendation system. The project on food recommender system which is focused on "What should I eat?”. The system used a content-based recommendation technique for producing food recommendation. It is based on the similarity of foods, basically, by constructing used profiles from the previously rated features and food profiles from the ingredients of the food and the system expected to recommend the most appropriate foods according to the preference of the users. The system is evaluated and worked accordingly.

Piyaporn et.al has done ontology based personalized dietary recommendation for weightlifting. The paper describes a food and nutrition ontology working with a rule based knowledge framework to provide specific menus for different times of the day and different training phases for the athlete’s diary nutritional needs and pesonal preference.

2.10 Research Gap

Though there are literatures on food recommender system globally on various specific aspects like recommender system for weight lifting, recommender system for infants etc. Food recommender system is still one of the tasks which are context dependent like culture, religion, season, economy and availability. Due to this and the specific eating culture and socio economy behavior of Ethiopia, the study requires being conducted specific Ethiopian context.

This study will develop a food recommender system that can help to promote diabetes patients healthier eating habits. This study will also describe the development of food recommender system algorithm that will help diabetes type 2 patients adopt a healthier food based on their personal ratings or likes / dislikes of ingredients as well as the their individual nutritional goal. The recommended food should be healthy and available to the patient.

Chapter Three

3. Research Methodology

3.1 Overview

Methodology refers to the analysis of the methods used appropriate to a field of study. It is a systematic way of accomplishing certain tasks and is defined as a collection of procedures, techniques, tools and documentation aids that helps a software developer to speed up and simplify the process. As a result, this step needs much attention on choosing the appropriate method which can provide the desired output.

3.2 Research Design

Research Design is the blue print or plan of procedure that cover the decision from the wide assumption to detailed methods of data collection (Creswell, 2009). Depending on the kind of problem situations, the existing knowledge and the resource availability on can use research design that suits. The research methodology used for this study, the initial requirement analysis is originated from the literature review. Prototyping methodology for SDLC is selected. Once the initial requirements are ready, the next step is to design the system, coding, testing and evaluating.

3.3 Data collection

In this study explicit and tacit knowledge is acquired from both codified (documented) sources and non-codified (non-documented) sources respectively. Non documented sources are addressed by consultation with medical doctors; nutritionist and domain experts at the EDA and critique knowledge elicitation methods to filter the acquired knowledge To acquire the required tacit knowledge from the selected domain, the researcher has employed semi-structured interview technique which focuses on the concept, procedures, and guidelines as well as experience which domain experts used while advising in diabetes meal. The researcher selects this technique because it makes the interview more flexible to add new questions based on the response and to change the order of the interview questions.

Relevant literature such as books, journal articles, conference papers and resources from internet that are related to recommender system especially on food recommender system and diabetes type 2 were reviewed. In addition to that, documents related to diabetes and diet will be highly considered in the study. Moreover, to assess what others have done in the area and to understand the problem better, related works assessed. In general, comprehensive investigations of available literature on different techniques and approaches of recommender system will be reviewed. Knowledge inside them is acquired by using document analysis technique.

Data about Ethiopian traditional foods are collected from the food composition table which is published by the Ethiopian Health and Nutrition research Institute (EHNRI). And the general principles and recommendation explored from the articles and consultation with the domain experts.

3.4 Data preparation

Since the quality of input data has a big impact on the quality of the data analysis and the result of the study in general, preparing data has a paramount importance. Data preparation is an important preprocessing step for data analysis that involves transforming raw data in to a format that can be understood. In real world data is often inconsistent, incomplete and error prone. So, data preparation is useful tool to tackle such issues. Data preparation includes various subtasks like selection, integration, transformation and cleaning of data. As depicted on the annex 1, the data collected from the food composition table was not in a format which is ready to be used for this research purpose. So some modifications or preprocessing was made to make it useful for the aimed task.

Moreover, from the literature reviews on type 2 diabetes the researcher understands that, only some of the nutritional attributes that are, energy, carbohydrate, protein, fiber and fat are considered as mandatory for diabetes type 2 patients to be weighted in their food choice activities. So, only those attributes are considered in this study.

3.5 Data representation

We need to define which aspects of foods are important. Since the goal was to advice patients with better choice of food. Therefore, we need to present them with healthy food that is close to their preferences and availability of the ingredients of the food.
The researcher has chosen to represent the foods and the user as a vector combination of numerical and binary features.

3.6 Approaches and Development Tools

Recommender systems can be classified into the following categories, based on how recommendations are made: (Adomavicius, 2005):

1. **Content-based recommender systems**, in which the user is recommended items and services similar to those the user preferred in the past,

2. **Knowledge based recommender system** in which either on explicit domain knowledge about the items or knowledge about the users is used to derive relevant recommendation,

3. **Collaborative filtering systems**, in which the user is recommended items that people with similar tastes and preferences liked in the past.

4. **Demographic Recommender system**: A demographic recommender provides recommendations based on a demographic profile of the user. Recommended products can be produced for different demographic niches, by combining the ratings of users in those niches

For the purpose of this study, a hybrid approach of demographic, content based and collaborative recommender system techniques with a switching, cascade and mixed hybrid strategy will be used since the recommender system to be developed in this study depends on the rates or the likes/dislikes of the ingredients in the food item as well as the nutritional goal of the patient and the demographic information of the individual.

3.6.1 System design and Development Implementation Tools

For the purpose of this study, the approach to be implemented is prototyping. The evolutionary prototyping methodology for SDLC is preferred for better management in testing of the proposed algorithm.

3.6.1.1 System Development Life Cycle

Like a traditional software development, the process of recommender system prototype development can be divided in to logical models that exist to develop software. There are many different life cycle steps. This can be done through methodology that will be adopted into the research. The System development life cycle (SDLC) is the entire process of formal, logical steps taken to develop software. There are many life cycle models that exist to develop the system; the researcher is at liberty to choose a model that fit this project, time and constraint. The phases of SDLC can vary somewhat but generally include the following: requirements specification, software design, coding, testing and evaluating.

Due to the circumstance surrounding this research, in particular time constraint, the SDLC that was chosen for the system development is Evolutionary Prototyping.

3.6.1.2 Knowledge Acquisition

Knowledge acquisition is the most important process and vital stage in knowledge based system development. How knowledge is obtained and where it is obtained determines the usefulness of the system. Knowledge acquisition process has many steps. Some of them are: selecting a problem to be solved by the program, interviewing an expert, observation, record reviews, codifying the knowledge in some representation language, and refining the knowledge base by testing it and extending its capability (Clancey, 1984).

3.6.1.3 Knowledge Modelling

After the knowledge is acquired from different sources, the next step was organizing and structuring of knowledge. Knowledge modelling is the representation of information in the form of logic for the purpose of processing knowledge to simulate intelligence (Makfi, 2011).

3.7 Testing and evaluation

Once the prototype is developed, the functionality of the system should be and extensively tested and evaluated to make sure whether the system works with an acceptable level of accuracy. The evaluation processes focus on the performance of the system. Performance measurement determine if the system perform the required task successfully. In addition to this, the standard measures of relevance (performance of the system) in the information retrieval (precision) have been used to evaluate the performance of the prototype.

The researcher evaluated the prototype system by experimenting on five diabetes patients after setting the standard nutritional goal based on their demographic information and international guidelines for the nutritional composition of diabetic patient's meal per day. Each patient got a recommendation five times and rate the recommended food based on their relevance. An experiment was conducted to know how new recommendation are matched with the patients desire and to know how the system precision progress at each iteration. 1 Testing and evaluating procedures

Precision and recall accuracy test and coverage test used to evaluate the prototype of the recommender system and relevant feedbacks collected to calculate the performance.

Chapter Four

4. Knowledge base System Concepts

The concept of knowledge base system is derived from the field of artificial intelligence. Artificial intelligence intends understanding of human intelligence and building of computer programs that are capable of simulation or acting one or more of intelligent behavior. Intelligent behavior includes cognitive skill like thinking, problem solving, learning understanding, emotion, consciousness, intuition and creativity, etc. Nowadays, some of the behaviors such as problem solving, learning and understanding are handled by computer programs (Sajja et al, 2010). Knowledge base system is a computer program that can solve problems like human expert by using knowledge about the application domain and problem solving techniques (Speel et al, 2001). Human experts use their knowledge about the domain and techniques that lead how to use the knowledge to solve problems. Computer knowledge base systems handle problems in the same way as humans do. The system represents knowledge about a specific application domain and uses one or more techniques that guide to use knowledge to solve problems (Sajja et al, 2010).

Knowledge base system is the general term used for the process of eliciting, structuring and representing knowledge from some knowledge source mostly from human experts and developing a computational problem solving model, specifically a program to be used in some consultative or advisory role (Sagheb-Tehrani, 2009, Fredlund et al, 1996). Such programs are generally called expert system.

Industries and societies are becoming knowledge oriented and dependent of decision making ability of expert. Knowledge base system can act as an expert on demand; can save money by leveraging experts; allowing users to function at higher level and promoting consistency (Sajja et al, 2010). Knowledge base systems increase productivity, document when there is shortage of knowledge for future use, enhances problem solving capability and this leads to increase quality in problem solving process.

4.1 Knowledge Acquisition in KBS Development

It refers computer files and transferring to the knowledge base using knowledge representation techniques used in knowledge-based system; namely, logic, production rules, semantic nets, frames and cases (fatonad, 2015). The acquired knowledge may be specific to the problem domain and the problem solving procedures, or it may be general knowledge (e.g., knowledge about business), meta-knowledge (knowledge about knowledge).

The acquisition of knowledge is the most important and decisive phase in building knowledge-based systems. However, it is an extremely hard and capable of making an error task that knowledge engineer does while developing a knowledge-based system. As a result of the challenges and difficulties confronted in the transfer of expertise knowledge, knowledge acquisition has been depicted as the obstruction of knowledge based systems development (Pearson, 2015).

4.2 Steps in Knowledge Acquisition

There are two main steps in knowledge acquisition process that are accomplished by the knowledge engineer so as to develop knowledge-based system. These are knowledge elicitation and knowledge structuring (Jones, 2015).

4.2.1 Knowledge Elicitation

It involves extracting knowledge from human experts, and/or written documents to build a knowledge-based system. For the purpose of this research, Knowledge about the food and diabetes type 2 acquired using the following knowledge acquisition methods:-

- Eliciting data and information from the domain expert (both nutritionist and the diabetic patients) using semi structured interviews to get the insight about food and diet in Ethiopia context.
- Acquiring knowledge from relevant documents
- Acquiring knowledge about food choice characteristics

- Acquiring knowledge from Ethiopian Diabetes Association and Ethiopian Nutrition and Health Research Institute(ENHRI)
- Analyzing and interpreting the collected data or information that gathered from various resources

4.2.2 Knowledge Structuring

It involves using concepts discovered during the knowledge elicitation session to build a model or representation of the domain experts. It is a process where knowledge engineer uses concepts discovered during the knowledge elicitation phase to build a conceptual model of the domain.

4.2.2.1 Knowledge Acquisition from Domain Experts

For this study, tacit knowledge from domain experts from EDA, ENHRI and Balance PLC was interviewed since interview is one of the knowledge elicitation techniques which involve asking the experts how the perform their daily tasks. 4 domain experts one from balance plc, 2 from EDA and one from ENHRI were selected using a purposive sampling technique to obtain the required knowledge on the domain area.

The question in the interview encompasses issues like. How the expert interact with the patients, what are the major issues to be considered to advice for diabetes meal, that are factors that influence food choice (see the full interview question in the appendix II). The researcher tries to make an through discussion with the experts and recorded the acquired knowledge from domain experts using pen and paper sheet.

Since identifying the factors that influence food choice has a big importance for underlying research the researcher forwarded few questions to the experts about "what are the factors that influence the food choice of a patient?" The domain expert's answer for the above question is summarized as the age, gender, availability; price, religion, season, health impact and taste are among the top factors that great impact on daily food choices. The other question that was raised by the researcher was "which nutrient need to considered at the patients daily meal" the domain response for this question was carbohydrate, fat, calorie, fiber and protein are the major nutrients to be under control in daily meal of a diabetes type 2 patient. The other question what are the challenges in advising for diabetes type 2 patients? And the answer was that the lack of standard in Ethiopian traditional foods so that they are forced to use their own traditional method of measurement to recommend a meal.

At the end, all the domain experts suggest that having a recommender system that can assist on food choice of the patient will have a vital importance for the patients.

4.2.2.2 Knowledge Acquired from the relevant documents

Reviewing related documents is vital in order to elicit knowledge from documents. These documents include literature, brochures, diabetes hand books, reports, publications and food composition tables. Literatures are reviewed for the purpose to clearly identify the food choice process, the food recommendation for diabetes and pin point the major factors that affect the food choice behavior of an individual with and without diabetes type 2

In literatures, several influences on food choice are distinguished, grouped into two categories: internal and external. Internal influences come from the person making the food decision, such as motivation to eat healthfully, while external influences are coming from the external world, such as availability of food in nearby. However, most influences are a combination of both external and internal factors. Several of these food choice factors are described as follows

- In terms of the importance of food properties, (Glanz et al 1998) investigated that **taste and cost**s are perceived as more important factors than the healthfulness of the diet. Therefore, healthful diets should be promoted as being tasty and inexpensive in order to induce people to eat healthier.

- Another obvious food choice factor is the religion of a person. Many religions pose constraints on the diet of believers, such as Muslims and Jews not being allowed to eat pork. Besides these food laws, (conformity to) religion also influences food choice in a social way. Just, (Heiman et al 2007) investigated the influence of religion on family members' food decisions. They conclude that children do play a role in family decisions but that this role is dependent on religious observance, age of the children, and the gender of the gatekeeper. This shows that there is an interaction between social situation and religion and that conformity to religion indirectly influences food choice.

- Monetary considerations are another factor that can influence food choice. (Maitland et al 2010) investigated monetary considerations are also influenced by availability in supermarkets (i.e. products that are scarce are more expensive) and by seasonal information (i.e. some products are cheaper in one season than in another).

- According to (maya ,2011), among all the general factors that influence food choice, the factors that seem to be most important in food choice for people with diabetes are the amount of vegetables, the amount of carbohydrates and of course the taste. Additionally, fat, calories, seasonal aspects, familiarity, time and price should also be taken into account by a recommender system. The paper by Maya also confirmed that liking of an ingredient, complexity of the dish, time necessary for preparation and healthfulness of the dish are important factors for people in determining whether they want to eat a meal or not. This implies that these factors should play a dominant role in a recipe recommendation system. The author clearly find out that the most important reason for rejecting a recipe is the fact that it contains an ingredient that is not liked. Other important reasons are the complexity, preparation time and healthfulness of the dish. For the recommender system this implies that it should account for these reasons.

The knowledge that acquired from the reviewed literatures about factors that influence food choice are summarized as follows

1. Availability: Individuals diets are restricted to the types of foods and amounts of food available nearby. For example, the climate has a huge influence on what is available, what foods are in season during that climate and the climate even affects the price of that food at the time.

2. Income, Food Prices and Convenience: Having a balanced diet is the one way to ward off chronic disease and risk of other conditions. One reason we may make food choices is cost.

4. Societal and Traditional Aspects: Depending on traditions and the society an individual grew up in, choice of food is a result of that. Social groups have an influence on what we eat as well. Family and friends are so influential when it comes to food choices. Family is primary source of a social structure so it only makes sense that they influence an individual food choice so much. The traditions that an individual pick up from his/her families have the most powerful and lasting impact.

5. Beliefs and Personal Values: Many people adopt a certain way of eating based on their personal values and this has a direct impact on health. There are many religious and spiritual practices that have guidelines that affect what foods go into an individual diet and what foods do not.

6. Other Factors That Affect Food Choice: One of the main reasons to choose between foods is taste. Taste determines preference and influences our food choices.

4.2.2.3 Nutrition Recommendation for Diabetes type 2

According to the American Diabetes Association position statement in 2004, the following nutrition recommendation is set for patients with diabetes type 2

4.2.2.3.1 Nutrition Recommendations for the Management of Diabetes

4.2.2.3.1.1 Carbohydrate in diabetes management

Recommendations

- A dietary pattern that includes carbohydrate from fruits, vegetables, whole grains, legumes, and low-fat milk is encouraged for good health.
- Monitoring carbohydrate, whether by carbohydrate counting, exchanges, or experienced-based estimation remains a key strategy in achieving glycemic control.
- Sucrose-containing foods can be substituted for other carbohydrates in the meal plan or, if added to the meal plan, covered with insulin or other glucose-lowering medications. Care should be taken to avoid excess energy intake.
- As for the general population, people with diabetes are encouraged to consume a variety of fiber-containing foods. However, evidence is lacking to recommend a higher fiber intake for people with diabetes than for the population as a whole.
- Sugar alcohols and nonnutritive sweeteners are safe when consumed within the daily intake levels.

3.6.1.2 Fiber Recommendation

As for the general population, people with diabetes are encouraged to choose a variety of fiber-containing foods such as legumes, fiber-rich cereals (≥5 g fiber/serving), fruits, vegetables, and whole grain products because they provide vitamins, minerals, and other substances important for good health.

4.2.2.3.1.2 Dietary fat and cholesterol in diabetes management

Recommendations

- Limit saturated fat to <7% of total calories.
- Two or more servings of fish per week are recommended.

4.2.2.3.1.3 Protein in diabetes management

Recommendations

- In individuals with type 2 diabetes, ingested protein can increase insulin response without increasing plasma glucose concentrations. Therefore, protein should not be used to treat acute or prevent nighttime hypoglycemia.
- High-protein diets are not recommended as a method for weight loss at this time. The long-term effects of protein intake >20% of calories on diabetes management and its complications are unknown. Although such diets may produce short-term weight loss and improved glycaemia, it has not been established that these benefits are maintained long term, and long-term effects on kidney function for persons with diabetes are unknown.

4.2.2.3.1.4 Alcohol in diabetes management

Recommendations

- If adults with diabetes choose to use alcohol, daily intake should be limited to a moderate amount (one drink per day or less for women and two drinks per day or less for men).
- To reduce risk of nocturnal hypoglycemia in individuals using insulin alcohol should be consumed with food.
- In individuals with diabetes, moderate alcohol consumption (when ingested alone) has no acute effect on glucose and insulin concentrations but carbohydrate congested with alcohol (as in a mixed drink) may raise blood glucose.

The following summary highlights nutrition guidelines for people with diabetes (CDAN,1999).

Carbohydrate	• Total carbohydrate: 50-60% of daily energy requirements, which can include added sugars up to 10% of daily energy requirements.
Total dietary fiber	• Adults: at least 25-35 g/day. • Children: 5 g plus 1 g/year of age as a guide. • Should include both soluble and insoluble fibre.
Protein	• Adults: at least 0.86 g/kg/day. • Children: RNI for age and gender.
Fats	• Total fat: < 30% of daily energy requirements. • Saturated and polyunsaturated fats: each < 10% of daily energy requirements. • Use of monounsaturated fats should be encouraged where possible. • Fish rich in omega-3 fatty acids should be consumed at least once per week.
Alcohol	• Alcohol consumption should be limited to 5% of total energy requirements or two drinks per day, whichever is less. • Regular alcohol intake can contribute to weight gain, poor glycemic control, and elevated lipids.

Sweeteners	• Nutritive and nonnutritive sweeteners may be used moderately as part of a well-balanced diet. • Use of saccharin and cyclamate is not recommended during pregnancy and lactation. • Aspartame is contraindicated in individuals with phenylketonuria.
Micronutrients (vitamins & minerals)	• Routine use of vitamin or mineral supplements is not recommended for people with diabetes except in cases of inadequate food consumption or other special needs. • Daily vitamin and mineral requirements should be obtained from a well-balanced diet.

Table 4. 1: Recommendations for the Nutritional Management of Diabetes Mellitus.

4.2.2.3.2 Food Choice for people with Diabetes

Since the aim of the study is to develop a system that provides tailored food suggestions. However, to find a food that an individual will to eat, knowledge about factors that influence individual food choice is vital. For this purpose it is mandatory to review literature on factors involved in food choice.

4.2.3 Knowledge models

Conceptual modeling is extensively acknowledged as the critical stage of knowledge acquisition. Before a knowledge-based system can be constructed, knowledge should be identified and collected, and a model of domain knowledge should be built. Models are applied to acquire the important characteristics of problem domains by decomposing them into more controllable parts that are simple to know and to use. Models are very related with the domain they denote (Savolainen et.al, 1995). “A model is a simplification of reality” (Booch.et.al 1999).

After the knowledge acquisition phase is completed from domain experts and relevant documents, the next step to be done is modeling the acquired knowledge. Knowledge modeling has a significant importance in the development process of a knowledge-based system to know the operational means. It provides a means to understand the source, input and output of knowledge and the designation of other parameters.

A hierarchical tree structure modeling technique is selected for this particular study because it easily models concepts and clearly explains the concepts in the problem area. The model starts with the main concept at the highest level of the hierarchy and other sub concepts put next to down ward in the hierarchy. It has modeled based on two main concepts: Person and Food concepts.

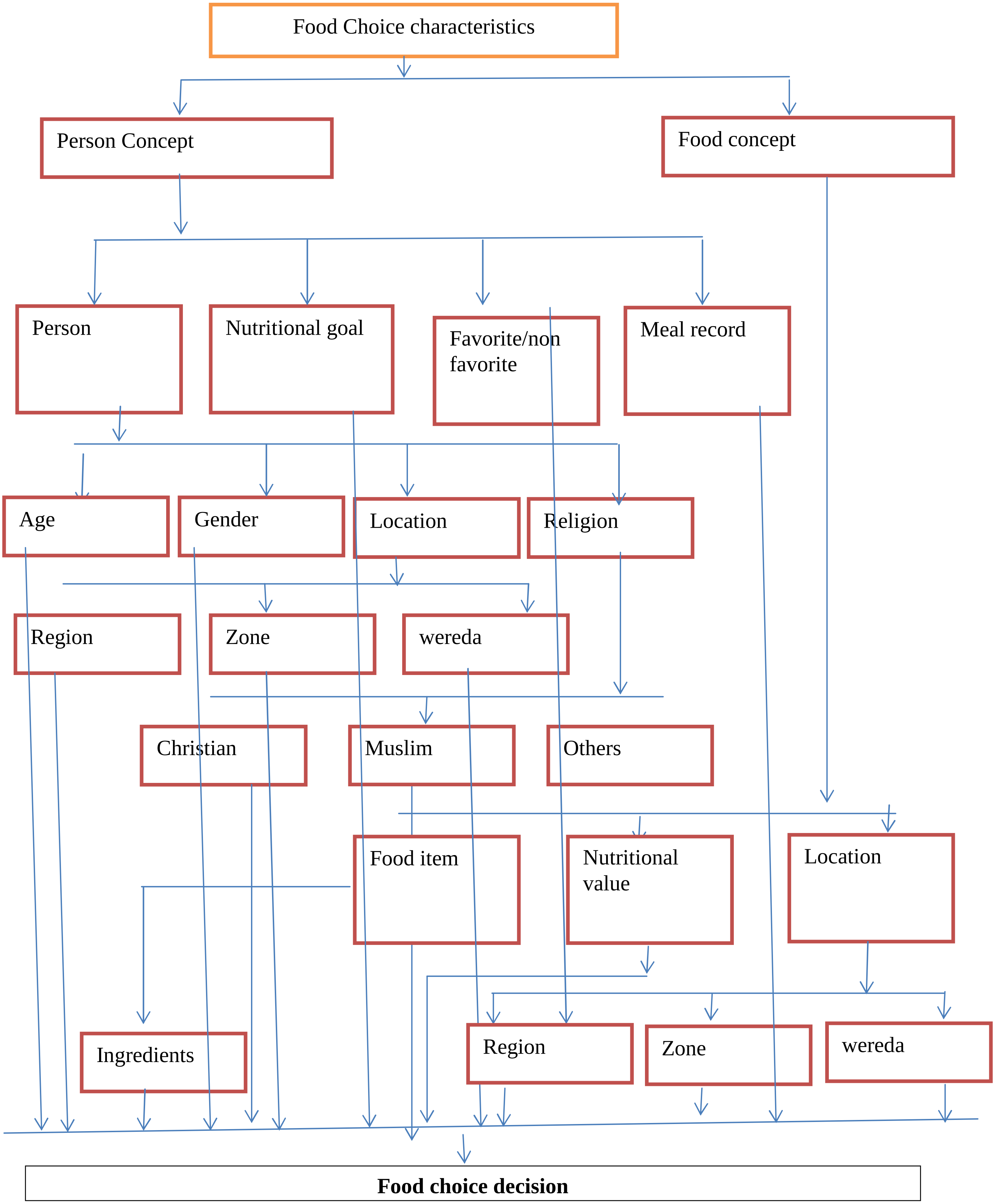

Figure 4.1 Hierarchical structure of food choice decision

Person Concept: The "Person" class represents an abstract concept of personal profile which involves both person's related status and food-related preference.

- **"Person"** is a major class which contains user profile and preference. Its attributes include personal data such as age, location, etc.
- **"Goal"** represents person's nutrition goals.
- **"Favorite Ingredient"** represents person's preference for some ingredients containing in a food dish.
- **"Un-favorite Ingredient**" represents person's dislike for some ingredients containing in a food dish.
- **"Meal Records"** it contains the food choice records of the a person with in a week

Food Concept: The "Food" class represents an abstract concept of food item which involves food types, food group and nutrition level.

- "**Food Item"** is a major class of food concept which is related to ingredients,
- **"Nutrition value"** represents value of nutrition which are groups based on nutrition value of (i.e., Fiber, calorie, carbohydrate, protien,fat)
- **"Location**" represents the origin of the food item

.

.

Chapter five

5.1 Design and Implementation

The design and implementation part of this section involves the actual development of a prototype food recommender system.

5.2 Recommender Approach

In the previous chapters it has been discussed in detail about different taxonomy of Recommender system. As depicted there, content based and collaborative recommender systems are the dominant research paradigms on information filtering. Content based filtering selects the right information for users by comparing representations food items to contents of user's profiles which express interests of users.

Collaborative filtering is the technique of using peer opinions to predict the interest of others. A target user is matched against the database to discover neighbors, who have historically similar interests to target user. Although collaborative filtering has been successful in research there is still remain some challenges for it as an efficient information filtering.

Combining collaborative filtering and content based approaches can be used to alleviate or compensate their shortcoming and produce a better recommendation. Many hybrids approaches use two recommendation algorithms and combine their results in some manner, such as combining the results by their relevance, mixing the output of the two algorithms, switching from content based into collaborative once the cold start phase is over, or using the output of one algorithm as input to the second algorithm. Using hybrid approaches we can avoid some limitations and problems of other recommender systems, like the cold-start problem (Burke, 2002). The combination of approaches can proceed in different ways (Adomavicius, 2005).

- Separate implementation of algorithms and joining the results.
- Utilize some rules of content-based filtering in collaborative approach.
- Utilize some rules of collaborative filtering in content based approach.
- Create a unified recommender system that brings together both approaches.

The researcher prefers to use a hybrid approach of content based and collaborative approach that use a separate implementation of both approaches and use the mixed hybrid strategy to combine the candidates of each method. On top of that, knowledge based recommender approach that can filter foods based on the similarity of the food location of the active user residence and the guideline food recommendation to diabetes type 2 patients. The intuition behind a hybrid model is that food choice is combinations of observed ingredients plus nutritional values and users preferences over these ingredients. A user's preference for a food is a function of his/her preferences for the ingredients and respective nutritional goal as well as the recommended guidelines for the patients.

5.2.1 Food representation

The first step is to create a representation of the food. The food was breakdown in to their ingredients and nutritional values. From the document review and interviews with the domain in the previous chapters the nutritional contents that needs to be under a scrutiny of the diabetes patient are calorie amount, fiber, carbohydrate and protein so in food representation there need to be a way to represent the nutritional attribute of a food ***fi*** F. Besides that there must be a way to represent the ingredients I_n of each food. There for the researcher represents foods in a vector representation

Definition1:- Let F = be the set of available food. A food ***fi*** can be written as a vector of feature values fi=(I_0,$I_{1,,}$,.,.,.,I_m,calorie,carbohydrate,protein,fat,fiber)

Where I_m represents the mth ingredient of the food and its value will be 1 if the food item encompasses the ingredient and 0 if the ingredient is not in the food.

Calorie, carbohydrate, protein, fat, fiber represents the nutritional value of the specific food and its value will be its value per 100gm of the specific food.

	I_1	I_2	I_3	………	I_n	Calorie	………	fiber
F_1	(0/1)	(0/1)	(0/1)		(0/1)	Value	….	Value
F_2	(0/1)	(0/1)	(0/1)		(0/1)	Value	…	Value
F_3	(0/1)	(0/1)	(0/1)		(0/1)	Value	…	Value
…	….	…	…	…	…	…	…	…
F_n	(0/1)	(0/1)	(0/1)	….	(0/1)	value	….	Value

Table 5.1 Food representation Table

5.2.2 User Representation

People in general and people with diabetes in particular, seem to find it difficult to deviate from their normal patterns in eating. In this paper the user is represented by a vector of ingredients with respect to the ratings for each ingredient. Therefore the user is represented by a vector of U X I.

Definition 2: Let U= a set of available. A user u_i can be written as a vector of ingredients.

U_i=(i_1,i_2,i_2,,.,.,.,i_n,caloriegoal,carbohaydrategoal,portiengoal,fatgoal,fibergoal)

Where U_i : represents user i

i_n :Ingredient n

r_{nm} :rating of user n for ingredient m

Value: represents the personal goal of a specific user

	I_1	I_2	I_3	…	I_m	Caloriegoal	….	fibergoal
U_1	r_{11}	R_{12}	R_{13}	…	r_{1m}	Value	…	Value
U_2	r_{21}	R_{22}	R_{13}	…	r_{2m}	Value	…	Value
U_3	r_{31}	R32	R_{33}	…	r_{3m}	value	…	Value
…	…	…	…	…	…	…	…	…
U_n	r_{n1}	r_{n2}	r_{n3}		r_{nm}	value	Value	Value

Table 5.2: User Representation

Each food item in the set of all foods can either be liked or disliked on a specific moment .The user profile will contain the rating of the ingredients that he/she rated on a particular occasion.

5.3 The hybrid filtering model

Recommender systems can also be classified according to their algorithmic technique (Baret.et.al, 2001).Two different classes can be identified:

Memory-based techniques calculate recommendations in real-time over the entire data set. In a nutshell, memory-based techniques rely heavily on simple similarity measures (Cosine similarity, Pearson correlation, Jaccard coefficient… etc) to match similar people or items together. If we have a huge matrix with users on one dimension and items on the other, with the cells containing votes or likes, then memory-based techniques use similarity measures on two vectors (rows or columns) of such a matrix to generate a number representing similarity.

Model-based techniques learn a model from previous user activities and use this model to classify new content according to its interestingness.

Model-based techniques on the other hand try to further fill out this matrix. They tackle the task of “guessing” how much a user will like an item that they did not encounter before. For that they utilize several machine learning algorithms to train on the vector of items for a specific user, and then they can build a model that can predict the user's rating for a *new* item that has just been added to the system.

In this study, memory based recommender technique is selected since the user profile will be generated on real time. As we discussed on previous chapter, there are different strategies to use a hybrid recommendation that are weighted, mixed, feature augmentation, cascade and feature combination etc. A switching, cascade and mixed hybridization strategies are used at different stages of the algorithm.

The content-based component can be relied on recommending new foods on the basis of the similarity between the user ratings and the ingredients of the food as well as nutritional values of the food item and the nutritional goal set by the user.

The collaborative component of the hybrid system will perform the recommendation based on the cosine similarities of the active user with the rest of the user. After selecting the most N neighbor users with the active user the system will look for food items that are rated by those users and not in the food records of the user.

The mixed hybrid system receives the candidates of the both the content based and collaborative recommender system and join their respective candidate lists with their respective rank. The candidates will be joined and ranked based on the sum of the ranks that they acquired from individual recommender systems. Then the new rank will be set based on the combined ranks of the food item.

The new user problem is addressed by demographic recommender system in a switching combination in such a way that if a user is new to the system, the system will list the candidates that is produced by the rule based demographic recommender the user, whereas if the user is not a new one, the system will use the candidates generated by demographic recommender system as an input for the next phase of the algorithm.

Both the content based and the collaborative filtering uses a similarity measures and neighborhood algorithms to predict the user tastes.

5.3.1 Similarity Measure

There are different similarity computations techniques that can be used to compute the correlation between users in ***memory-based*** algorithms (Breese et al, 1998). Some of the techniques are

Pearson Correlation: A weighted average of deviation from the neighbors' mean is calculated. Pearson correlation measures the degree to which a linear relationship exists between two variables and is defined as follows

$$w_{a,u} = \frac{\sum_{i=1}^{m} \left[\left(r_{a,i} - \overline{r_a} \right) \left(r_{u,i} - \overline{r_u} \right) \right]}{\sqrt{\sum_{i=1}^{m} \left(r_{a,i} - \overline{r_a} \right)^2 \sum_{i=1}^{m} \left(r_{u,i} - \overline{r_u} \right)^2}}$$

Where $w_{a,u}$ is the similarity weight between the active user and neighbour *u,* $r_{a,i}$ is the rating given to item *i* by active user *a*; r_a is the mean rating given by user *a*; and *m* is the total number of items

The vector similarity**:** uses the cosine measure between the user vectors to calculate correlation. The formulation according to (Breese et al, 1998)

Mean-squared difference: *is another alternative that was used in the Ringo music* recommender (Herlocker, 2000). Mean-squared difference gives more emphasis to large differences between user ratings than small differences.

(Herlocker, 2000) found that the mean-squared difference algorithm does not perform well compared to Pearson correlation. (Breese et al, 1998) has found that vector similarity does not perform as well as Pearson correlation in collaborative filtering. Due to these reasons; Pearson correlation is used as a similarity computation technique for this research.

5.3.2 Neighborhood

The system need to select a subset of neighbor users at prediction computation time in order to assure acceptable response time. Since, most of the users profile does not have similar tastes to the active user, using them as predictors will only result sparsity. So, determining which users are more close to the active user, and to be used in the computation of a prediction for the active user is very mandatory. Selecting the neighborhood to be used in computing a prediction instead of the entire user database is useful, both for accuracy and performance (Herlocker, 2000). There are two techniques

1. ***Correlation thresholding*****:** In this case, all neighbors with absolute correlations greater than a specified threshold are selected. A high threshold means that only good correlates that can generate more accuracy in the prediction will be selected. (Griffith et al, 2000).

2. *Best-n Correlations*: This is where the best N correlates are picked. In this case using a large value of N may result in too much noise for those with high correlates whereas setting a small N can cause poor predictions for users with low correlates (Griffith et al, 2000).

In this paper, Best N-neighborhood, which is recommended by (Herlocker, 2000) is selected.

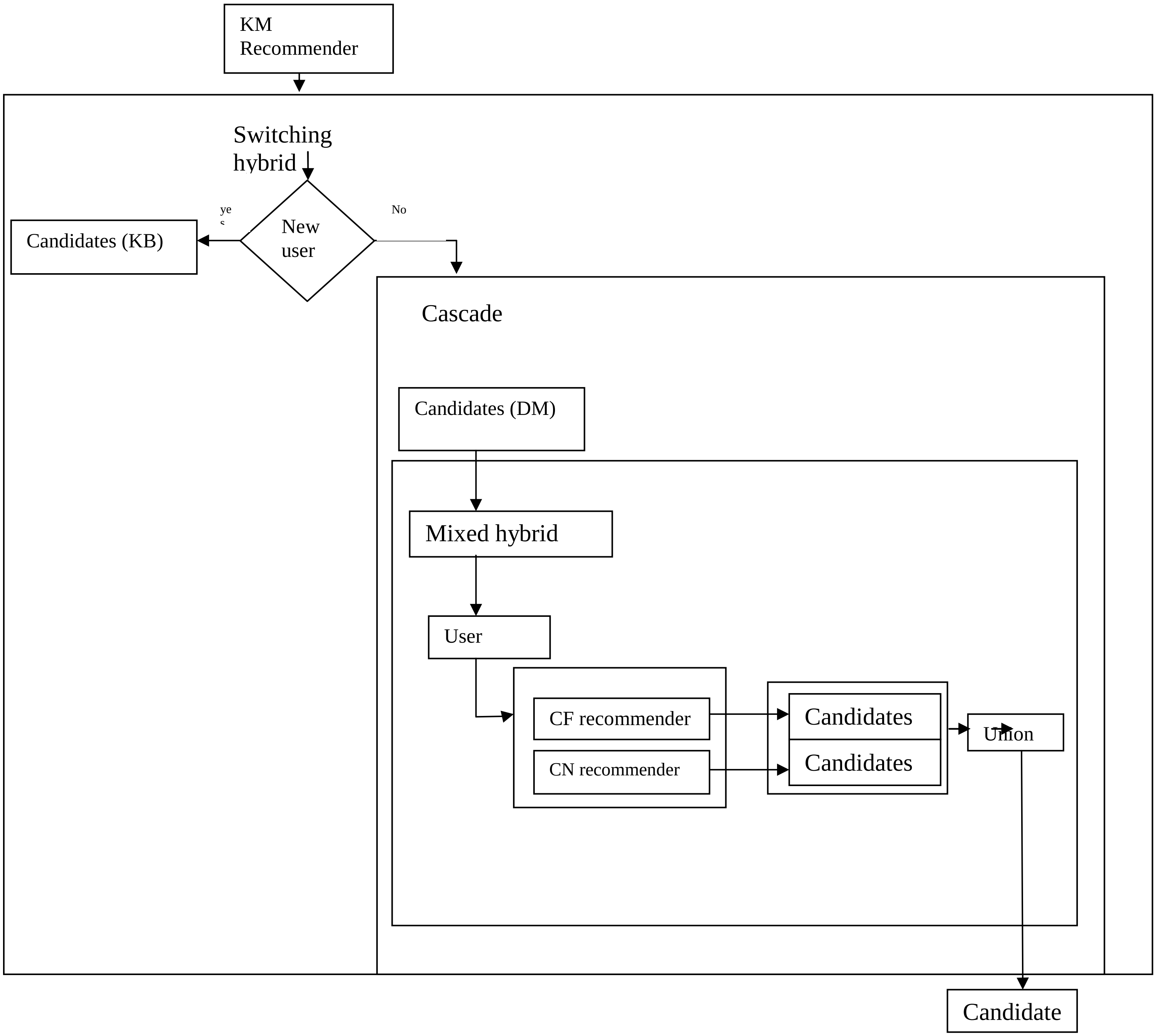

Figure 5.1 Proposed Frameworks

5.4 Proposed algorithm

1. The system will apply a demographic recommendation technique using a rule base reasoning and generates a candidates based on the location of the user and the food item.

Rule 1:

If the user location n is equal to food location

Then add the food item as a candidate for that specific user

Else

If the user location is not specified

Use the default guideline and generate the candidate

End

Rule 2:

Carbohydrate >=50% and <= 60% of Calorie

Rule 3:

Fiber

If user is Adult then fiber is between (25, 30) g/day

Otherwise 5g/day

End.

Rule 4:

Protein

If user is Adult then Protein>= .86g/kg/day

Rule 5:

Fat <=30% of calorie

2. Using a switching hybrid the system will check whether the user is new or not. If the user is new the system assigns the most frequent food from candidates which are generated by Demographic recommender system.
3. If the user is not the first time user then the system will proceed by using a cascade hybrid technique on the candidates generated by the Demographic recommender system.

Steps in collaborative approach

1. Weight all users with respect to similarity with the active user.
2. Select a subset of users to use as a set of predictors (possibly for a specific item)
3. Compute a prediction from a weighted combination of selected neighbors' ratings.

Steps in content based system

1. Calculate the similarity of all the foods with respect to the profile of the active user
2. Rank the subsets of the meals with respect to their similarity to the active user profile
3. Select meals that got a high similarity with the active user and not included in the meal records history of the user.

4. Mixed hybrid of the content based and collaborative filtering will be applied on the candidates of the demographic recommender.
5. The final candidate will be displayed.

Chapter Six

6.1 Experiment and Testing

Testing and performance evaluation of the prototype recommender system is the last step that assists the researcher to measure the achievement of the system against the objectives of the study. In this chapter the research experiment is carried out. This chapter presents performance evaluation of the prototype system. For the performance evaluation of this research has conducted testing of the prototype, case similarity testing, and retrieval performance evaluation using recall, precision and coverage.

6.2 Testing the Recommender System

For the purpose of testing the system, five volunteer diabetic patients are selected randomly and two domain experts from the Ethiopian diabetic association office are involved in the testing process. The test was conducted in such a way that the volunteer enter his/her name, gender, birth date, Weight and height using the interface but these are optional but the system requires the mandatory fields like location and nutritional goal(Calorie, Carbohydrate, protein, fiber and fat). The nutritional goal was made mandatory because the diabetic patients are supposed to consult a nutritionist to set their daily or per serving nutritional goal.

To test whether, the system improves as the number of users and their corresponding rating increases. Each user did the test five times. At each iteration the user was told to rank the recommended item and label relevant and not relevant recommendation.

Iteration	Number of Relevant items	Number of Not relevant items
1		
2		
3		
4		
5		

Table 6.1 User feedback table

6.2 The system performance evaluation

In this experiment accuracy and coverage are used as a performance measure.

6.2.1 Coverage

Coverage measures the percentage of items that the recommender system can provide a recommendation. In this experiment, when we consider all neighbors the coverage is 90%, that is, it has a maximum coverage. However, as the number of neighbors decrease coverage also decreases accordingly. Since the system recommends only the best five neighbors the coverage is found out to be 1.6% which is small value. But such value is expected since the decreasing the number of neighbors does mean limiting the number of recommended items. Limiting the size of items to be recommended has a significant advantage to increase the precision.

6.2.2 Accuracy

Precision and recall are the commonly used measures of performance of the retrieval process. Precision is the proportion of search results that are relevant to the query and recall is the ability of the retrieval system to retrieve all relevant cases to a given new query from the cases base.

Precision = number of relevant cases retrieved/ Total number of cases retrieved

Recall = number of relevant cases retrieved /Number of relevant cases in the case base

To evaluate the recall, the challenge was to find out all the relevant food items. In order to perform that the researcher provide all the food item lists to the user so that they can identify the total number of relevant food items in the database.

To start the evaluation process, for each recommendation the user rated the recommended lists as relevant and non-relevant in the separate paper which is provided to them by the researcher. And the researcher summed up all the relevant and not relevant counts of the five users for each iteration.

iteration 1		
	Relevant	not relevant
user 1	1	4
user 2	1	4
user 3	0	5
user 4	0	5
user 5	1	4

Table 6.2 User feedback for Iteration 1

Iteration 2		
	Relevant	not relevant
user 1	2	3
user 2	1	4
user 3	1	4
user 4	0	5
user 5	1	4

Table 6.3 User feedback for Iteration 2

Iteration 3		
	Relevant	not relevant
user 1	2	3
user 2	1	4
user 3	1	4
user 4	0	5
user 5	2	3

Table 5.4 User feedback for Iteration 3

Iteration 4		
	Relevant	not relevant
user 1	2	3
user 2	2	3
user 3	2	3
user 4	1	4
user 5	4	1

Table 6.5 Users feedback for Iteration 4

iteration 5		
	Relevant	**not relevant**
user 1	**3**	**2**
user 2	**2**	**3**
user 3	**2**	**3**
user 4	**2**	**3**
user 5	**4**	**1**

Table 6.6 user's feedback for Iteration 5

After summarizing the user's feedback as depicted on the above five tables. The next step is to further summarize the data in a format that is ready to calculate the precision and recall.

iteration 1			
	Retrieved	not retrieved	Total
Relevant	3	52	55
non relevant	22	228	250
Total	25	280	305

Table 6.7 Summary of user feedback for Iteration

iteration 2			
	Retrieved	not retrieved	Total
Relevant	5	45	50
non relevant	20	235	255
Total	25	280	305

Table 6.8 Summary of user feedback for Iteration 5

iteration 3			
	Retrieved	not retrieved	Total
Relevant	6	48	54
non relevant	19	232	251
Total	25	280	305

Table 6.9 Summary of user feedback for Iteration 3

iteration 4			
	Retrieved	not retrieved	Total
Relevant	11	48	59
non relevant	14	232	246
Total	25	280	305

Table 6.10 Summary of user feedback for Iteration 4

iteration 5			
	retrieved	not retrieved	Total
Relevant	13	53	66
non relevant	12	227	239
Total	25	280	305

Table 6.11 Summary of user feedback for Iteration 5

Using the above summary table the precision and the recall are calculated using their respective formulas. The result is depicted in the table below.

Iteration	Precision	Recall
1	12%	8%
2	20%	10%
3	24%	11.11%
4	44%	18.64%
5	52%	19.7%

Table 6.12 Precision and Recall Table

Recall in information retrieval measures ability of a retrieval system to obtain all or most of the relevant documents in the collection (Clarke & Willet, 1997). The higher recall value shows that the system obtains most of relevant items from the total relevant item which is retrieved and not retrieved. As we can see from the table, the recall value has a slight increase at each iteration. Therefore, the recommender system can retrieve relevant items that enable diabetic type 2 patients to make decisions. But the performance improves as the system acquires more ratings from the user and as more users use the system. On the other hand, the precision value of the system follows the same trend that it gets better as the number of iteration increases. That signifies that as the system gets more data about the user profiles the performance of the system also gets better.

Chapter seven

7. Conclusion and Recommendation

7.1 Conclusion

Nutrition Therapy is a major solution to control diabetes by managing the nutrition based on the belief that food provides vital medicine and good health. Good diabetes management is a balance between healthy eating, exercise and medication. The problem is that most diabetic patients have difficulty of identifying the recommended quality and quantity of food that they have to eat in order to control their blood glucose level. However, In Ethiopia incomplete routine health information and lack of data on the proper dietary practice of diabetic patients affect the long term management of diabetes.

A recommender system is an intermediary program or an agent with a user interface that automatically and intelligently generates a list of information which suits the needs of users. It provides advice to users about items or services they might wish to get.

In this study an effort has been made to design and develop a recommender system that can provide an advice for diabetes patients to facilitate dietary management. Knowledge is acquired from relevant documents and a conceptual model is developed accordingly. Then the frame work which is a hybrid of demographic, content based and collaborative filtering is applied for the study. Visual basic programming language with SQL database is used for the implementation of the prototype. The experiment was done based on the memory based algorithm. Pearson correlation coefficient was used as a similarity measure due to its wide spread acceptance. Top 5 items are then recommended to a user by calculating the predictability of those items to the active user and ranking the items based on their cumulative ranking of both the content based and collaborative result.

Even though the developed system did not get a chance to be fully experimented and tested, there are indications that hybrid recommendation can be a successful technique for Food recommender systems.

The experiment was done using 5 users and 305 food items; each user did the experiment five times and recalls, precision and coverage was used to measure the performance of prototype. The highest coverage of the system was 99% when all other users in the database are considered as best neighbors. And the precision and recall gets better as the number of iteration or dataset increases.

The challenge in the study was to find a large data set that can be used for the model based approach that incorporate and user and their respective food selection which is recorded for some time since there is no food recommendation system which is implemented and has got a bid log file such as allrecipe.com. So the approach is forced to a memory based. The other challenge was that Ethiopian foods does not have a standard for example “Shiro wet or Beyaninetu” is sauce but there is not any standard regarding the amount of water, flour or oil to be used as an ingredient it is all dependent on the personal interest. There was also a challenge to find a thorough documented food item lists with their respective ingredients and nutritional value.

Although there are still issues to be addressed like the standards for Ethiopian traditional foods and documentation for food items. I believe that personalizing food recommendation using the hybrid approach is promising for the promotion of health eating.

Furthermore, the following conclusions are drawn from the findings with regard to the research questions:

- The basic attribute that influence the food choice of a diabetes type 2 patients are identified as culture, religion, taste, availability and cost of the food.

- The applicability of a hybrid recommender system for diabetes type 2 patients has been proved.
- Tacit and explicit knowledge diabetes type 2 and food choice factors has been acquired and modeled.
- To what extent the recommender system helps the diabetes type 2 patients to promote a healthy daily food selection?

7.2 Limitations of the study

In Ethiopia, one of the major problems in developing a food recommender system is lack of standards in our local foods. For instance, If we consider Injera there is no a well-defined standard on the size and thickness. So, as the researcher understood through discussion with the domain experts (nutritionist) there is no enough research done on standardizing and labeling the local foods with their respective nutritional value and ingredients.

Since there is little or no recommender system which is widely implemented on food recommender system in general and on diabetic patients in particular in Ethiopia context, it was not possible to find a dataset to try out a model based recommender system and focus on the memory based recommender system.

7.3 Recommendation

Even though, this research has put a corner stone for the food recommendation system in Ethiopia context. This research paved the way for new researchers to further work on the area and upgrades the system so that the prototype can fully implement. The following recommendations are made based on the findings and limitations of this research.

1. A more thorough experiment which considers all the important determinant factors (such as neighborhood sizes) that influence the performance of the system should be made.

2. The performance of the system should be tested with reasonable amount of user samples. A way of getting better number of feedbacks and using experts or users to judge all the relevant items for the purpose of calculation recall should be brought to mind at the earlier stages of the research. This will ensure reliable generalization and conclusion to be made on the results.
3. The potentiality of the other similarity measuring techniques should also be discovered. That is, similarity measures other than Pearson correlation should be used to compare results and pick the best measure..
4. Model based algorithms such as clustering model and Bayesian network model must also be empirically tested in the pursuit for better performance.
5. Researches have to be undertaken to using different combination of the hybrids two systems for better result and efficiency.
6. Researches need to work on standardization on Ethiopian food items and to label of food stuffs so that any diabetic patients can understand the contents easily.
7. There is also some works to be done on life style changes that are required to be taken by diabetic patient in Ethiopian context.
8. The research can be strengthen by considering more food choice characteristics such as taste, price, availability, social factors on top of the ingredients of the food item.

Reference

1. Adomavicius, G. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Trans. Knowledge and Data Eng., vol. 17, no. 6, pp. 734-749.

2. Adomavicius, G. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Trans. Knowledge and Data Eng., vol. 17, no. 6, pp. 734-749
3. Adomavicius, G. and Tuzhilin, A. (2005). Query Language for Recommender Systems. Stern School of Business, New York University.
4. Adomavicius, G., A. Tuzhilin.(2005) "Toward the Next Generation of Recommender Systems: A Survey of the State-of the-Art and Possible Extensions," IEEE Trans. on Knowledge and Data Engineering (17:6), pp. 734-749.
5. Alison B. Evert, Jackie L. Boucher, Marjorie Cypress, Stephanie A. Dunbar, Marion J. Franz, Elizabeth J. Mayer-Davis, Joshua J. Neumiller, Pharmd, Cde, Cgp, Fascp7 Robin Nwankwo, Mph, Rd, Cde8 Cassandra L. Verdi, Patti Urbanski William S. Yancy Jr.,(2013) , Nutrition Therapy Recommendation For Managements Adults With Diabetes
6. Alspector, J., Koicz, A., and Karunanithi (1997), Feature-based and Clique-based User Models for Movie Selection: A Comparative Study. *User Modeling and User-Adapted Interaction* 7, 279-304.
7. Asanov, D. (2011). *Algorithms and Methods in Recommender Systems; Berlin Institute of Technology, progetto didattico. Berlino, Germania*.
8. Balabanovic, M. and Shoham, Y. (1997). *Content-based, collaborative recommendation. Communications of the ACM, 40(3).*
9. Barry, S. (2009). *Case-Based Recommendation*. *The Adaptive Web Methods and Strategies of Web Personalization, University of Warwick*.

10. Belkin, N. J. and Croft, W. B.: (1992), Information Filtering and Information Retrieval: Two Sides of the Same Coin? *Communications of the ACM* 35(12), 29-38.
11. Ben Schafer, Joseph Konstan, John Riedl, (1999) Recommender system in e-commerce,Grouplens research project
12. Berhe KK, Demissie A, Kahsay AB, Gebru (2012) Diabetes self-care practices and associated factors among type 2 diabetic patients in Tikur Anbessa specialized hospital, Addis Ababa, Ethiopia-a cross sectional study. International Journal of Pharmaceutical Sciences & Research , 3(11)
13. Billsus, D. and Pazzani, (2000). 'User Modeling for Adaptive News Access'. *User-Modeling and User-Adapted Interaction* 10(2-3), 147-180.
14. Breese, J. S., Heckerman, D. and Kadie,(1998), 'Empirical analysis of predictive algorithms for collaborative filtering'. In: *Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence*, pp. 43-52.
15. Brin, S. and Page, L.(1998), 'The anatomy of a large-scale hypertextual {Web} search engine'. *Computer Networks and ISDN Systems*, 30(1-7), 107-117.
16. Burke, R. (2002). Hybrid recommender systems: survey and experiments. User Modeling and User-Adapted Interaction 12, no. 4.
17. Burke, R. (2006). Knowledge based recommender systems, University of California, Irvine.
18. Burke, R. (2007). Hybrid Web Recommender Systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.). The Adaptive Web: Methods and Strategies of Web Personalization.
19. Burke, R., Hammond, K., and Young, B.(1997), 'The FindMe Approach to Assisted Browsing'. *IEEE Expert*, 12 (4), 32-40.
20. Burke, R.: (1999), 'The Wasabi Personal Shopper: A Case-Based Recommender System'. In: *Proceedings of the 11th National Conference on Innovative Applications of Artificial Intelligence*, pp. 844-849.

21. Burke, R.: (1999), ‘Integrating Knowledge-Based and Collaborative-Filtering Recommender Systems’. In: *Artificial Intelligence for Electronic Commerce: Papers from the AAAI Workshop (AAAI Technical Report WS-99-0 1),* pp. 69-72.
22. Burke, R.(2000), ‘Knowledge-based Recommender Systems’. In: A. Kent (ed.): *Encyclopedia of Library and Information Systems*. Vol. 69, Supplement 32.
23. C. S. Lee, M. H. Wang, H. C. Li, and W. H. Chen, (2008) IntelligentOntological Agent for Diabetic Food Recommendation, Proc. IEEEWorld Congress Computational Intelligence, pp. 1803-1810, June 2008
24. Centers for Disease Control and Prevention (CDC), and Centers for Disease Control and Prevention (CDC) (2011) “National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States, 2011”. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention 201.
25. Clancey, W. (1984). Knowledge Acquisition for Classification Expert Systems. Proceedings of ACM National Conference on computer science, Stanford University.
26. Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. and Sartin, M.: (1999) Combining Content-Based and Collaborative Filters in an Online Newspaper. *SIGIR ’99 Workshop on Recommender Systems: Algorithms and Evaluation*. Berkeley, CA. <URL: http://www.cs.umbc.edu/~ian/sigir99-rec/papers/claypool_m.ps.gz>
27. Condliff, M. K., Lewis, D. D., Madigan, D. and Posse,(1999), Bayesian Mixed-Effects Models for Recommender Systems . *SIGIR ’99 Workshop on Recommender Systems: Algorithms and Evaluation*. Berkeley, CA. <URL: http://www.cs.umbc.edu/~ian/sigir99-rec/papers/ condliff_m.ps.gz>
28. Creswell, J.W. (2007). Qualitative Inquiry and Research Design: Choosing among five approaches, Sage Publications, Thousand Oaks.
29. Ekore RI, Ajayi IO, Ekore JO (2008) Dietary management of diabetes: a practical approach for primary care physicians in Nigeria. Diabetes 16:13-14

30. Ethiopian Diabetes Association Website avaialable at: http://www.diabetesethiopia.org.et/ visited august 16, 2014
31. Foltz, P. W(1990), 'Using Latent Semantic Indexing for Information Filtering'. In: R. B. Allen (ed.): *Proceedings of the Conference on Office Information Systems*, Cambridge, MA, pp. 40-47.
32. G. Booch, J. Rumbaugh, and I. Jacobson, *"The Unified Modelling Language User Guide"*, Addison Wesley, Reading, Massachusetts, 1999. O.V. Fatonade, "The Application of Knowledge Based Systems to the Abstraction of Design and Costing Rules in Bespoke Pipe Jointing Systems", Available at: http://dspace1.isd.glam.ac.uk/dspace/bitstream/10265/583/1/fatonademphil.pdf, Accessed date: October 16, 2012.
33. G.Adomavicius and A. Tuzhilin, (2005)Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Transactions on Knowledge and Data Engineering, pp. 734-749, June,
34. Gill G, Mbanya J-C, Ramaiya K, Tesfaye S (2009) A sub-Saharan African perspective of diabetes. Diabetologia 52(1):8-16
35. Global estimates of diabetes prevalence for 2013 and projections for 2035 for the IDF Diabetes Atlas. Diabetes Research and Clinical Practice 2013
36. Guariguata L, Whiting D, Hambleton I, Beagley J, Linnenkamp U, Shaw J: Treatment of Diabetes
37. Guttman, R. H., Moukas, A. G. and Maes, (1998), 'Agent-Mediated Electronic Commerce: A Survey'. *Knowledge Engineering Review*, 13 (2), 147-159.
38. Hailu E, Mariam WH, Belachew T, Birhanu Z (2012) "Self-care practice and glycaemic control amongst adults with diabetes at the Jimma University Specialized Hospital in south-west Ethiopia: a cross-sectional study: original research". African Prim Health Care Family Med 4(1):1-6

39. Hall V, Thomsen RW, Henriksen O, Lohse N (2011) Diabetes in Sub Saharan Africa 1999-2011: epidemiology and public health implications. A systematic review. BMC Public Health 11
40. Haregu TN, Alemayehu YK (2012) Diabetes Management in Southwest Ethiopia: A Cross-Sectional Study. Public Health Res 2(5):162-166
41. International Diabetes Federation, Diabetes Atlas, fourth edition (2015)
42. Ipek TATLI (2009) Food Recommendation System,
43. J.E. Bartlett, J.W. Kotrlik, and C.C. Higgins, (2001), Organizational research: determining appropriate sample size in survey research, Information Technology, Learning and Performance Journal 19 no. 1, 43–50.
44. Jennings, A. and Higuchi, H.(1993), 'A User Model Neural Network for a Personal News Service.' *User Modeling and User-Adapted Interaction*, 3, 1-25.
45. Konstan, J. A., Riedl, J., Borchers, A. and Herlocker, J. L.: (1998), Recommender Systems: A GroupLens Perspective.In: *Recommender Systems: Papers from the 1998 Workshop (AAAI technical Report WS-98-08)*.
46. Krulwich, B.(1997), 'Lifestyle Finder: Intelligent User Profiling Using Large-Scale Demographic Data'. *Artificial Intelligence Magazine* 18 (2), 37-45.
47. Lang, K.:(1995), 'Newsweeder: Learning to filter news'. In: *Proceedings of the 12th International Conference on Machine Learning*, Lake Tahoe, CA, pp. 331-339.
48. M. A. El-Dosuky1, M. Z. Rashad1, T. T. Hamza1, A.H. EL-Bassiouny, *Food Recommendation using Ontology and Heuristics*
49. Maiyaporn Phanich, Phathrajarin Pholkul, and Suphakant Phimoltares, *Food Recommendation System Using Clustering Analysis for Diabetic Patients*
50. Maiyaporn Phanich, Phathrajarin Pholkul, and Suphakant Phimoltare (2010) ,Food Recommendation System Using Clustering Analysis for Diabetic Patients",
51. Makfi, P. (2011). Introduction to knowledge modeling and neural network. Available at: http://www.makafi.com/KCM_intro.html.

52. Napat Suksom, Marut Buranarach, Ye Myat Thein, Thepchai Supnithi,Ponrudee Netisopakul (2010) A Knowledge-based Framework for Development of Personalized Food Recommender System,
53. P.H. Jones, “Knowledge Acquistion: Knowledge Engineering in Agriculture”, Available at: https://engineering.purdue.edu/~engelb/abe565/knowacq.htm, Accessed date: june 23, 2015.
54. Pazzani, M. J (1999), A Framework for Collaborative, Content-Based and Demographic Filtering. *Artificial Intelligence Review*, 13 (5/6), 393-408.
55. Pearson Education, (2014) Chapter 18 - Knowledge Acquisition, Representation, and Reasoning, Available at: http://wps.prenhall.com/wps/media/objects/3778/3869053/Turban_Online_Chapter_W18.pdf, Accessed date: June 21 2014
56. *Recommender Systems Handbook* (2011) ISBN 978-0-387-85819-7 e-ISBN 978-0-387-85820-3 DOI 10.1007/978-0-387-85820-3, Springer New York Dordrecht Heidelberg London Springer © Science + Business Media, LLC
57. Resnick and H. R. Varian. (1997). Recommender systems. Commun. ACM 40, 3, 56-58. 1997
58. Resnick, P. and Varian, H. R. (1997), Recommender Systems. *Communications of the ACM*, 40 (3), 56-58.
59. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J.(1994), GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In: *Proceedings of the Conference on Computer Supported Cooperative Work*, Chapel Hill, NC, pp. 175-186.
60. Rich, E.(1979), ‘User Modeling via Stereotypes’. *Cognitive Science* 3, 329-354.
61. Schafer, J. B., Konstan, J. and Riedl, J (1999), Recommender Systems in E-Commerce’. In: *EC ’99: Proceedings of the First ACM Conference on Electronic Commerce*, Denver, CO, pp. 158-166.

62. Schwab, I., Kobsa, A. and Koychev, I. (2001), 'Learning User Interests through Positive Examples Using Content Analysis and Collaborative Filtering'. *User Modeling and User-Adapted Interaction*. To appear.
63. Shamsi N, Shehab Z, AlNahash Z, AlMuhanadi S, Al-Nasir F: (2013) Factors Influencing Dietary Practice among Type 2 Diabetics. Bahrain Medical Bulletin, 35(3)
64. Shardanand, U. and Maes, P. (1995), Social Information Filtering: Algorithms for Automating "Word of Mouth"'. In: *CHI '95: Conference Proceedings on Human Factors in Computing Systems*, Denver, CO, pp. 210-217.
65. Sicree R, Shaw J, Zimmet P, Heart BI: (2010) The global burden. Diabetes and impaired glucose tolerance Baker IDI Heart and Diabetes Institute
66. Solomon Gebremariam (2015) A-self learning knowledge based systemfor diagnosis and Association SAd: South African diabetes association. In South Africa diabetes association;
67. Stefanie Mika,(2003) *Challenges of Recommender System*
68. Sumedh Sawant Gina pai (2002) Yelp Food Recommendation System
69. T. Savolainen, D. Beeckmann, P. Groumpos, and H. Jagdev (1995), Positioning of Modeling Approaches, Methods and Tools, Computers in Industry, Vol.25, pp.255-262,.
70. Terveen, L. and Hill, W. (2001). Beyond Recommender Systems. Helping People Help Each Other.
71. Thienne Johnson, Jorge Vergara, Chelsea Doll, Madison Kramer, Gayathri Sundararaman, Harsha Rajendran , Alon Efrat and Melanie Hingle (2013) A Mobile Food Recommendation System Based on The Traffic Light Diet,
72. Towle, B. and Quinn, (2000), 'Knowledge Based Recommender Systems Using Explicit User Models'. In *Knowledge-Based Electronic Markets, Papers from the AAAI Workshop*, AAAI Technical Report WS-00-04. pp. 74-77. Menlo Park, CA: AAAI Press.

73. U.S. Department of Health and Human Services NATIONAL INSTITUTES ,*What I need to know about Eating and Diabetes*

74. U.S. Information Technology Industry Council, ‘The Protection of Personal Data in Electronic Commerce’, *PublicPolicy Document*, Nov. 20, 1997. URL:htttp://www.itic.org/iss_pol/ppdocs/pp-privprin.html

75. Wikipedia Avialable at: https://en.wikipedia.org/wiki/Diabetes_mellitus visited at September 25 2014.

76. Ziegler, C. N., S. M. McNee, J. A. Konstan, G. Lausen (2005) of the 14th International World Wide Web Conference, 2005, pp. 22-32.

Appendix 1 : Tables used in the SQL Server

ALEMNEH-PC\E...bo.FoodItem*

Column Name	Data Type	Allow Nulls
FoodItemID	nchar(10)	☐
Location	char(50)	☑
FoodItemName	nchar(50)	☑
Barley	bit	☑
Maize	bit	☑
beef	bit	☑
broadBean	bit	☑
Milk	bit	☑
chees	bit	☑
chickPea	bit	☑
emmerWheat	bit	☑
falseBanana	bit	☑
Fenugreek	bit	☑
fish	bit	☑
goatMeat	bit	☑
grassPea	bit	☑
GreenOnion	bit	☑
GreePepper	bit	☑
Lentil	bit	☑
Millet	bit	☑
Mushroom	bit	☑
Pea	bit	☑
PeanutButter	bit	☑
Potato	bit	☑
Pumpikin	bit	☑
sesame	bit	☑

ALEMNEH-PC\EC...o.MealRecords

Column Name	Data Type	Allow Nulls
FoodItemId	nchar(10)	☑
UserID	nchar(10)	☑
Date	date	☑
Rate	int	☑
		☐

ALEMNEH-PC\EC...NutritionValue | ALEMNEH-PC\EC...o.MealRe

Column Name	Data Type	Allow Nulls
FoodItemID	nchar(10)	☐
calorie	decimal(18, 0)	☑
carbohydrate	decimal(18, 0)	☑
protien	decimal(18, 0)	☑
fiber	decimal(18, 0)	☑
fat	decimal(18, 0)	☑
		☐

ALEMNEH-PC\EC...erInformation | ALEMNEH-PC\EC...NutritionV

Column Name	Data Type	Allow Nulls
userID	nchar(10)	☑
FullNAme	nvarchar(50)	☑
Location	nchar(15)	☑
Gender	nchar(10)	☑
Birthdate	date	☑
CalorieGoal	decimal(18, 0)	☑
CarbohydrateGoal	decimal(18, 0)	☑
ProtienGoal	decimal(18, 0)	☑
FiberGoal	decimal(18, 0)	☑
FatGoal	decimal(18, 0)	☑
		☐

ALEMNEH-PC\EC...idientRatings | ALEMNEH-PC\EC...erInformat

Column Name	Data Type	Allow Nulls
UserId	nchar(10)	☑
Barley	int	☑
beef	int	☑
maize	int	☑
broadBean	int	☑
Milk	int	☑
Chees	int	☑
chckpea	int	☑
EmmerWheat	int	☑
falsebanana	int	☑
fenugreek	int	☑
fish	int	☑
goatMeat	int	☑
grassPea	int	☑
GreenOnion	int	☑
GreenPepper	int	☑
lentil	int	☑
Millet	int	☑
mushroom	int	☑
pea	int	☑
peanutButter	int	☑
Potato	int	☑
pumpkin	int	☑
sesame	int	☑
sorghum	int	☑

Appendix II

Interview questions to domain experts

1. What kind of advising system is given when a patient com to your office?
2. What are the criteria that you follow in advising a patient?
3. What is the basic information you required from the patient?
4. Which nutrients are most relevant for your recommendation?
5. What are the factors that directly influence a food choice?
6. What are the challenges you faced in doing so?
7. What do you think if a recommender system that can assist a patient and the expert in doing so?