ONTOLOGY BASED PERSONALIZED DIET RECOMMENDATION SYSTEM FOR SRI LANKAN

Ms.D.M.L.M.Dissanayake

218522D

Master in Artificial Intelligence

Department of Computational Mathematics
Faculty of Information Technology

University of Moratuwa
Sri Lanka

April 2024

ONTOLOGY BASED PERSONALIZED DIET RECOMMENDATION SYSTEM FOR SRI LANKAN

Ms.D.M.L.M.Dissanayake

218522D

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree
Master in Artificial Intelligence

Department of Computational Mathematics
Faculty of Information Technology

University of Moratuwa
Sri Lanka

April 2024

DECLARATION

I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works.

Signature: Date: 21st July 2024

The above candidate has carried out research for the PhD/MPhil/Masters thesis/dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Dr. Thushari Silva
Department of Computational Mathematics
Faculty of Information Technology
University of Moratuwa

Signature of the Supervisor: Date:21/07/2024

DEDICATION

To my devoted family and steadfast supporters, this thesis is dedicated. Your constant love, encouragement, and belief in my abilities have been the driving force behind my success. Your unwavering support has provided me with the strength and motivation to overcome challenges and achieve my goals. I will always be appreciative of your presence in my life and the significant influence you have had on my academic career. Thank you for being my rock and standing by my every step of the way.

ACKNOWLEDGEMENT

The following people and organizations have helped me along the way to finishing my study, and I would like to sincerely thank them for their support.

First and foremost, I would like to express my sincere gratitude and appreciation to my supervisor, Dr. Thushari Silva, Senior Lecturer, Department of Computational Mathematics, Faculty of Information Technology, University of Moratuwa, for her wise counsel, motivating words, and helpful critiques provided during this research.

My special thanks to Dr.Bavindu Jayantha Ranasinghe for providing information about this study and for his excellent contribution.

I would also like to acknowledge the Department of Computational Mathematics at University of Moratuwa. My understanding of and proficiency in the field of artificial intelligence have greatly increased as a result of the lectures in this department.Their dedication to teaching and willingness to assist me throughout my academic journey are sincerely appreciated.

In addition, I want to sincerely thank my family for their unfailing love, support, and encouragement.. Their belief in me and their constant motivation have been instrumental in my achievements.

I want to express my gratitude to all of my friends for helping me along the way with this challenge. Their support, empathy, and words of encouragement have been a never-ending source of motivation.

Finally, I would like to acknowledge the contributions of all these individuals and institutions, without whom this study would not have been possible. Thank you for your invaluable support and encouragement throughout this journey.

ABSTRACT

The hectic life cycle of human beings and negligence of health are the main causes for non-communicable diseases. Healthy diets help to prevent various non- communicable diseases such as diabetics, pressure, cholesterol etc. However, due to busy life cycle, people do not have time to consult dietitians and get their meal plans. The diet recommendation system can solve this issue by providing a fingerprint distance response. In this study, we developed ontology based Recurrent Neural Network-LSTM model for diet recommendation in Sri Lankan context. This system recommends diet plans for breakfast, lunch and dinner based on user's calorie requirement, macronutrient needs, user preference, cultural background, dietary restriction and disease information. We developed ontology, which is consist with food related information including their nutrients values. Then we developed RNN –LSTM model to predict target macronutrient values based on user information such as age, BMI, BMR and diseases information. The RNN-LSTM model demonstrated high performance with an R-squared value of 0.9704, a Pearson correlation coefficient of 0.9869, a mean squared error of 0.0209, a mean absolute error of 0.1239, and a root mean squared error of 0.1703. These metrics indicate that the RNN-LSTM model provided accurate predictions with a strong positive linear relationship between the actual and predicted value. Additionally, the RNN-LSTM model was compared with the KNN model using the same dataset. The RNN-LSTM model outperformed the KNN model across all evaluation metrics. Using content-based filtering and cosine similarity between target nutrient values and food nutrient values, system recommends the food for three meals. This recommendation properly aligns with the nutrients needs and calorie needs. Finally, we can conclude that there is greater potential for future work in the field of food recommendation using Artificial Intelligence. Future work of this study involves further development of an ontology based RNN model combined with micro nutrient needs.

Keywords: Diet Recommendation, Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM),Ontology

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
BMI	Body Mass Index
BMR	Basal Metabolic Rate
WHO	World Health Organization
NCD	Non- Communicable Diseases
CF	Collaborative filtering
CBF	Content based filtering
ML	Machine Learning
OWL	Web Ontology Language
AI	Artificial Intelligence

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Prolegomena

The hectic life cycle of human beings and negligence of health are the main causes for non-communicable diseases. According to the World Health Organization(WHO), a healthy diet can aid in the prevention of diabetes, heart disease, stroke, and cancer like Non Communicable Diseases (NCDs) [1]. It is a challenge to figure out healthy food menus/dishes based on our preference, cultural background and disease information. However, with the busy life we do not have time to get instructions from nutritionist about healthy diets. This is the place where food recommendation systems come into play. The food recommendation system can overcome all these challenges by helping us find the right foods to eat. They provide a workable answer for people who struggle to strike a balance between a hectic schedule and the need for a balanced diet by bridging the gap between nutritional science and daily meal planning. Recommendation systems are information-filtering tools that help consumers locate products, services, or other material by gathering user preferences through implicit or explicit means[2]. A shared conceptualization is specified in formal, explicit terms by an ontology.

Entities, characteristics, and qualities associated with a domain form the foundation of ontology. Ontologies simplify the processes of parsing, reasoning, exchanging, and reusing information. Regardless of the application domain, it has demonstrated its usefulness in numerous recommender systems by enhancing the quality and customization of results.[3]. An ontology-based machine-learning model combines the principles of ontology and machine learning to leverage the structured knowledge representation provided by ontologies in the training and prediction processes [4]. To improve the machine-learning algorithm's capacity for learning and reasoning, it adds domain-specific ontologies.

When it comes to Sri Lankan context, there are few food recommendation systems are available. However, they do not align with user's specific needs and preferences. The main goal of this study is to develop a solution that can recommend Sri Lankan diet based on user preference, health condition and cultural background. In order to recommend diets in the context of Sri Lanka, we particularly developed an ontology-based machine-learning model. Here we developed ontology, based on food and nutrient information. Furthermore, user-specific data is included into the ontology to improve the relevance and accuracy of dietary suggestions catered to user preference, health condition and cultural background. We have developed a Recurrent Neural Network (RNN) model that calculates target nutrient values by considering factors such as Body Mass Index (BMI), Basal Metabolic Rate (BMR), age, and specific health conditions. We are able to create a customized list of food recommendations for breakfast, lunch, and dinner by applying content-based filtering to these target nutrient values along with comprehensive food information. This ensures that the user's dietary requirements and preferences are successfully satisfied.

1.2 Objectives

This research aims to develop an ontology based personalized diet recommendation system for Sri Lankan by considering individual preferences, disease information and cultural significance. The following objectives have been identified to develop ontology based personalized diet recommendation to address our research problem.

1. Critical review of existing diet recommendation systems.
2. Review of technology used in ontology based machine model in recommendation systems.
3. Design and implement the ontology based machine-learning model for Sri Lankan diet recommendation.
4. Evaluate the personalized Sri Lankan diet recommendation model.

1.3 Background and Motivation

Food is an essential component of human survival, acting as a source of energy and sustenance as well as being vital to people's general health and well-being. It gives the body the vital nutrients which needs for development, repair, and optimal operation. Food has cultural, social, and emotional value in addition to its biological relevance, impacting daily life and customs throughout. Every nation has their own food culture. Additionally, Sri Lanka has a rich culinary tradition. Food is a crucial factor in the context of health. The primary cause of an unhealthy life span is a diet pattern. Dietary advice is therefore crucial at this point. Consequently, because food recommendation systems are important for leading healthy lives, there has been a growing interest in food recommendation systems. Most of the systems can be implemented based on food preference or health problem. According to Thi Ngoc Trang Tran and his research team, there are different techniques of recommendation such as a knowledge based recommendation content-based filtering, collaborative filtering, hybrid recommendation, matrix factorization and graph-based recommendation. Previous research has proposed different food recommendation system by considering CBF, CF, hybrid recommendation and matrix factorization [5].

The integration of machine learning techniques in food recommendation systems is becoming increasingly sophisticated with the advancement of technology. Ontology based machine learning models help to recommend foods and diet based on different factors such as user preference, health information, and dietary restriction. There are different existing food recommendation systems using different machine learning techniques such as Self Organizing Map, K-Mean Clustering, k-nearest neighbors (KNN), Random Forest [6].

When it comes to Sri Lankan domain systems for recommending the Sri Lankan food menus-based diets that take into consideration individual preferences, disease information and cultural significance have been overlooked. This serves as the study's primary driving force. Our study is primarily motivated by this oversight.The main goal is to create a diet recommendation system that respects and incorporates the rich cultural legacy of Sri Lankan food in addition to individual health demands and user preferences. By improving the personalization of nutritional recommendations, this strategy hopes to encourage better eating practices and a heightened awareness of the variety of cuisines available in the country.

1.4 Problem Statement

Many people in Sri Lanka suffer from health problems such as diabetes, cholesterol and obesity due to their dietary patterns. Existing methods of food recommendation mechanisms are not properly success, and people are too busy to dedicate significant time to plan their meals effectively. In addition, systems for recommending the Sri Lankan food menus-based diets that take into consideration individual preferences, disease information and cultural significance have been overlooked. This research study proposes ontology based personalized diet recommendation system for SriLankan.

1.5 Proposed Solution

This research presents ontology based machine-learning solution to recommend diets based on individual preferences, disease information and cultural significance for breakfast, lunch and dinner. The system takes user profiles job and food information dataset as inputs. The model undergoes data collection, preprocessing, and target macronutrients value identification using neural network concepts. It constructs an ontology by integrating the neural network output with food information. Then within an ontology model, it performed the nutrient relevance analysis. Food recommendations can be made using content-based filtering and similarity scores via SPARQL queries, which take into account cultural significance, disease information, and personal tastes. The system is designed to be user-friendly, cost-effective, and incorporates evaluation and feedback. Validation with experts in food domain is conducted to refine the model, ensuring its accuracy and relevance.

1.6 Resources Requirements

Following equipment, resources and software were utilize in this study.

- Hardware: Computer with at least 6 GB NVIDIA GPU and 16 GB RAM
- Software and Tools: Jupyter notebook, Portege, Machine-Learning libraries
- Internet Access
- Data Sets: Personal information dataset
  Food related information dataset

1.7 Structure of the Thesis

The thesis is structures as follows. Chapter 2 is consist with critical literature review on recommendation systems in food domain using machine learning and ontology concepts. In depth, Study of technology adapted is presented in Chapter 3. The proposed approach is explained in Chapter 4. Design of the proposed solution presents in Chapter 5. The implementation of the proposed system discussed in Chapter 6. Chapter 7 presents the results and the evaluation of the developed system. And Finally in Chapter 8 discuss the conclusion of the research study with future works.

1.8 Summary

This chapter presents introduction to ontology based personalized Diet recommendation system for Sri Lankan. We have defined prolegomena, objectives background and motivation, problem in brief and proposed solution, resource requirements and the structure of the thesis. In this chapter, we have identified background of the study and problem statement. Chapter 2 presents the comprehensive literature survey on the review in the area of recommendation systems with machine learning techniques.

CHAPTER 2

LITERATURE REVIEW ON DEVELOPMENT AND CHALLENGES IN THE RECOMMENDATION SYSTEM

2.1 Introduction

In Chapter 1, we have presented the overall picture of this thesis covering the research problem and the essentials of the solution. This chapter gives a comprehensive literature review in the area of recommendation systems. The literature review has been structured to cover the recommendation systems, ontology development, machine learning techniques, and major developments in recommendation system with machine learning techniques and future directions of the ontology based recommendation system with machine learning techniques. Here we have also summarized the research challenges in recommendation system with machine learning techniques and defined our research problem.

2.2 Recommendation System

One useful technological solution to the problem of people getting too much information is the recommendation system. It allows the suggestion of products that are relevant to each user, gives each user a rating list of recommendations, and forecasts the quality of the items that will be recommended to them.[7].

In a service environment that has the ability to store or obtain many types of data, these systems function as information-filtering systems and offer a user a personalized suggestion for an item.. Recommendation systems, which make use of information filtering, only offer suggestions that are thought to be helpful to the user or that are tailored to their preferences.[8]. According to Thi Ngoc Trang Tran and his research group (2024) , there are different techniques of recommendation such as a knowledge based recommendation, collaborative filtering recommendation, contentt based filtering, recommendation, hybrid recommendation, matrix factorization and graph-based recommendation[5]. CBF is an approach that recommends items based on their content similarity to items the user has already liked or consume. Three parts make up a recommendation system: a filtering component, a profile learner and a content analyzer. User behavior or user evaluations of recommended things serve as the basis for CF recommendations. It recommends items liked by similar users and explores diverse possible content[9]. When individual predictions are combined into a single prediction, content information is added to the collaborative model, the weighted average of content and collaborative recommendations is calculated, or final recommendations are derived from the combined rankings, hybrid recommendation combines the best features of both approaches [10].

2.3 Ontology Development

Ontologies and the semantic web are two exciting new technologies that can be used to represent knowledge and draw conclusions for the creation of intelligent systems. The Similarity concepts serve as the foundation for the semantic web. This idea of similarity is applied in a number of study areas. A set of ideas or classes that represent the objects or entities found in the domain make up an ontology most frequently. A number of interconnects these ideas different relationships that show how they relate to one another [11].

According to Ying Ding (2002), the topic of ontology is one that is quickly expanding and holds enormous potential for improving information management, organization, and comprehension. It is necessary for the Semantic Web, the next phase of the Web's evolution, because it enables content-based access, communications, interoperability, and the provision of significantly higher service levels. Ontology engineers are still required to build the knowledge base and ontology for a specific job or area, as well as to manage and update the ontology to keep it current and relevant. Ontologies that are manually created take a lot of time, labor, and mistake. There are different ontology generation techniques. They can be summed up as follows: middle-out, which starts with the most crucial concepts and moves toward generalization and specialization (Eg: Enterprise and Methodology ontologies),bottom-up: from specification to generalization,top-down, which starts with generalization and ends with specification (Eg: KACTUS ontology) and bottom-up: from specification to generalization . An ontology can be created from scratch, utilizing just global or local ontologies that already exist, a corpus of data sources, or a combination of the latter two techniques. Three are three ways as fully manual, semi-automated and fully automated [12].

Logical theories like ontology are typically represented with knowledge representation languages like First Order Logic (FOL) and the widely used Web Ontology Language (OWL). This is according to computer science nomenclature. Based on a set of ideas with their semantic connections and limitations on a particular area of knowledge (domain), an ontology is a structured data repository data modeling technique. Ontologies' true power is found in their capacity to create relationships between classes and instances and provide the information needed for inference within those relationships. The most famous ontology development tool is Protégé[13].

2.4 Major developments in recommendation system with machine learning techniques.

The evolution of recommendation systems powered by machine learning techniques has witness significant advanced in recent years. This section aim to explore the major development of recommendation systems with machine learning techniques across various domains culminating in depth discussion of recommendation systems within the food domain.

Machine Learning (ML) is a one of the sub field available under artificial intelligence. It is a branch of research that allows computers to learn without explicit programming. The Learning from the data is the aim of machine learning. Supervised learning, Unsupervised learning, Semi supervised learning and reinforcement learning are the basic categories of Machine learning algorithms. An array of algorithms called a neural network is intended to mimic the functions of the human brain in order to find underlying patterns in a given amount of data. We can classify neural supervise neural network, Unsupervised neural network and Reinforced neural network. The output of the neural network is predetermined in a supervised neural network. The input and output are not known to the unsupervised neural network in advance. The network's primary function is to classify the data based on particular similarities. After establishing the correlation between the inputs, the neural network groups the data. Reinforcement learning algorithms are goal-oriented algorithms that figure out how to maximize along a certain dimension over a series of steps or accomplish a complex aim (goal)[14].

Koen Bothmer and Tim Schlippe (2022) presented a recommended method for skill recommendations by considering the interaction between job searchers, educational institutions, and employers. As was previously noted, they created distinct recommendation systems for the categories. Skill extraction, vectorization, clustering, and comparison make up their research workflow. They created methods to extract bullet points from these sources based on rules and keywords. Word embedding have been effectively clustered using K-Means clustering, which is flexible and scalable. As a result, they clustered their 768-dimensional vectors using K-means and the cosine distance as the distance metric. They used the NLP (Natural language processing) algorithm for the recommendation [15].

A team of researchers led by Jingyue Gao, Xiting Wang, Yasha Wang, and Xing Xie (2021)focused on creating a recommendation system using Attentive Multi-View Learning techniques. This study suggests creating an explainable recommendation system that combines the benefits of deep learning-based models and currently used explainable techniques to close the gap between accuracy and explainability. They provide a careful multi-view learning approach to ensure precise rating prediction. They managed sparse and noisy data by the framework's co-regularization of several feature levels and thorough combining of forecasts. Personalized explanation generation is described as a constrained tree node selection problem and a dynamic programming technique is proposed to solve it, allowing for the mining of legible explanations from the hierarchy. According to experimental results, their model performs better than previous recommendation systems in terms of explainability and accuracy when compared to state-of-the-art techniques [16].

Sundaresan Bhaskaran, Raja Marappan and Balachandran Santhi (2021) developed cluster based intelligent hybrid recommendation system for e-learning applications. In this study, they develop an intelligent recommender that uses clustering and the split-and-conquer technique to automatically adapt to the requirements, interests, and knowledge levels of the learners. Various learning styles are accommodated through clustering based on the split-and-conquer technique. The recommender automatically examines and picks up on the traits and styles of students. Using the recommended cluster-based linear pattern mining technique, the learners' functional patterns are extracted. Then, by analyzing the ratings of frequent sequences, the algorithm generates intelligent recommendations. The proposed model recommended key learning activities based on the learning style, interest classification, and talent aspects of the various learner groups and datasets that were the subject of experiments. Tests demonstrated that as compared to students in the no-recommender cluster group, learners using the suggested cluster-based recommender complete more lessons, improving recommendation performance . More than 65 percent of the students considered all of the factors while evaluating the recommended recommender[17].

A team of researchers led by Maiyaporn Phanich, Phathrajarin Pholkul, Suphakant Phimoltares (2010) focused on creating a recommendation system based on K-Mean Clustering technique and SOM (Self Organizing Map). They created this recommendation system for diabetic patients by substituting foods according to nutrition and food parameters. Eighteen (18) nutrient values are considered initially. Then based on nutritionist marks and rank, eight (8) nutrients were selected as most important nutrients for diabetics. The SOM is built and trained first, and then it is clustered using the K-mean method. The food clusters that show which foods in the same category contain roughly the same quantity of each of the eight nutrients are the end result. Nutritionists have assessed the suggested approach, and they find it to be very effective and beneficial for the field of nutrition [18].

M.I.M Jayarathna and Chinthanie weerakoon (2021) used genetic algorithms to implement healthy Sri Lankan meal plan. It's mobile application that calculates a person's daily energy demands and nutritional requirements based on their age, body mass index (BMI), and routine activities. It recommends eating wholesome meals that feature Sri Lankan cuisine and beverages. This application uses evolutionary computing on two levels. A dietary supplement containing 21 genes was used to build the chromosomes in the first level. The nutrients and amounts were used as the genes. Using food items of different lengths, the chromosomes in the second step were generated. The genes in this study, which rely on user inputs for several genes, represent the information on food products and quantities. The fitness function in the first level was created using sample personal data based on nutritional information and chromosomes, whereas the fitness function in the second level was created using foods that the user inserted, nutrition recommendations provided by the first genetic algorithm, and chromosomes. The system was assessed using the preferences provided by users and those obtained by the same set of users who used the Fitness Meal Planner app[19].

A DASH diet guidance system was presented for people with hypertension. It used content-based filtering in conjunction with machine learning algorithms, taking into account user preferences for food, allergies, smoking, alcohol consumption, blood pressure, and dietary intake. Proposed system was based on three modules such as estimate module, food classifier model using the multilayer perception model (Feed forward neural network) and content-based filtering module. The three components of the estimate model that are available are the alcohol calculator, the smoke calculator, and the sodium consumption calculator. A food classifier model's accuracy is 99%. A group of patients with hypertension examined the system, and they found it to be satisfactory. This is carried out by using a post test survey questionnaire. According to 70% of users, the app satisfies their dietary needs as a DASH diet app [20].

Dexon Mckensy-Sambola and his research team (2022) intended to create nutritional recommender system using ontology technology. They created the Ontology of Dietary Recommendations (ODR) with the goal of aiding those who suffer from obesity. A set of diets, recipes, and ingredients are represented by axioms. An algorithm for making recommendations that can determine a user's level of overweight and obesity and offer food plans to help reduce it. There are 2111 records with 17 attributes used for this research study. An actual group of people who have been anonymized is used by the suggested system to validate. A group of consultants assessed every single record and recommended the best diets for the individuals in the ontology. The mean accuracy of the proposed system is 87% [21].

Prof. Anand Ingle, Pinky Yadav, Raj Thakkar and Shubham Singh Kardam(2021) suggested a website that uses machine learning to recommend diets. In this study, recommendation is based on user’s physical aspect (age, gender, height and weight) and user’s end goal. In this study dataset is divided into three categories as breakfast dataset, lunch data set, dinner dataset. They used K – Mean algorithm and Random forest algorithm with this study. K-Means is used for clustering to cluster the food according to calorie. The food items are categorized using a random forest classifier, and the model predicts the food items depending on user input. This recommendation is for breakfast, lunch and dinner. The Users can choose from multiple recommended items and make their diet plan. After selecting food items the system will calculate selected food calories and show user's comparison between how much calories they chosen against how much they need to consume daily. Finally based on user’s profile healthy diet plan is generated [22].

NutriCure, a disease-based food recommender system utilizing the k-nearest neighbors (KNN) algorithm, was presented by Raksha Pawar and his research team (2021). The system is based on ailments like cardiovascular disease, eye diseases, renal disorders, obesity, scurvy, rickets, and issues during pregnancy. Here, the user can input past and present health-related data, and the result is a personalized diet plan that is adjusted to meet their specific needs. The illnesses and nutrient consumption of the users were the main subjects of this investigation. They evaluate dietary recommendations based on nutrient values for vegan diets, high-protein diets, Type A and Type O diets, Paleo diets, and Omni diets.In this study, recommendation is based k-nearest neighbors (KNN) algorithm by considering user preference and diseases history. In order to anticipate a diet plan, KNN measures the "distance" between each recipe in its database and the goal recipe. By applying this formula, they were able to locate the closest neighbor in order to choose the nutrients that would best meet the user's needs. This system was evaluated by using 100 demo users and received positive responses [23].

2.5 Future and Next Steps of ontology based recommendation system using Machine Learning

A number of possible developments, including improved data integration, hybrid recommendation techniques, contextual recommendation, explainable recommendations, active learning and user feedback, personalized learning paths, and cross-domain recommendations, are anticipated for the ontology-based recommendation system using machine-learning alogorithms in the future. One of the future direction is Matrix factorization methods. Using Matrix Factorization methods, Behnoush Abdollahi and Behnoush Abdollahi investigated a novel recommendation system strategy . Modern Matrix Factorization (MF) recommender systems call for additional data sources, like item content and rating data, in their most recent explanation methods. In order to deal with situations where further data is not accessible, this work introduces a novel Explainable Matrix Factorization (EMF) method that computes a precise top-recommended list of explainable objects.

Additionally, they present Mean Explainability Precision(MEP) and Mean Explainability Recall, two new explanation quality measurements (MER)[24].

Another future direction of recommendation system is integration with advanced ML models such as deep learning, reinforcement learning, hybrid approaches. By using deep learning models, like deep neural networks, it is possible to anticipate and comprehend user preferences more accurately, even when they are based on intricate patterns that the ontology may not have explicitly stated. And also Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) may be used in future systems to improve the interpretation of complicated user data, including pictures and sequential interaction data, and to make predictions that are more accurate and in line with the semantic information found in ontologies. And Subsequent systems may employ more advanced machine learning models to examine user-generated material, including food-related reviews, images, and videos. In order to enhance the user profiles that the ontology uses to provide more individualized suggestions, sentiment and particular dietary preferences or constraints could be extracted using natural language processing (NLP) techniques.

Cross-domain recommendation could be added to ontology-based systems in the future. For example, combining data on fitness and dietary habits to provide comprehensive lifestyle suggestions. This would necessitate using multi-task learning frameworks and creating intricate ontologies that map relationships between many domains. The relevancy of recommendations can be greatly increased by developments in context-aware systems that take into account the user's present circumstances (such as location, weather, and health status). Based on these contextual factors, machine learning algorithms could choose the best kinds of food or dishes to serve.

2.6 Summary

This Chapter provides a comprehensive analysis of the literature on ontology based machine-learning model for recommendation task by covering 100 research papers with very high citation records. We discuss the research domain, historical approaches, key achievements, current systems, challenges, and future directions in the field. The chapter 3 presents our in-depth study of Machine Learning techniques and ontology model by highlighting their relevance for diet recommendation based on user preference, diseases information and cultural significance.

CHAPTER 3

TECHNOLOGY ADAPTED FOR DIET RECOMMENDATION SYSTEM

3.1 Introduction

Chapter 2 offers an in-depth literature review focused on the domain of recommendation systems, particularly highlighting the integration of machine learning with ontology-based approaches tailored to dietary recommendations. This review meticulously explores the technological frameworks and methodologies adapted for developing an ontology-based machine learning system that caters specifically to the dietary needs of the Sri Lankan population. This chapter presents the technology adapted for ontology-based machine learning approach to Sri Lankan diet recommendation by considering individual preferences, disease information and cultural significance. Here, we present technology adapted for ontology development, tools used to build an ontology and technology used in machine learning model development.

3.2 Introduction to Technology in Diet Recommendation

In the field of food and nutrition, technological advancements have significantly transformed how dietary recommendations are formulated and delivered. With the development of numerous data-driven technologies like artificial intelligence (AI) and machine learning (ML), traditional dietary advising methodologies have changed. These technologies used vast amount of data such as nutritional content, user preference, individual health record to tailor dietary advice that aligns with personal health goals and different medical conditions. Machine Learning models are increasingly employed to predict nutritional needs and formulate diet plans. These models use algorithms to process complex datasets, learning from patterns in user preferences, nutritional content, dietary restrictions and health conditions. Another significant field that has shown significant advancement over time is ontology. Ontologies play a crucial role in the realm of diet recommendation systems by enhancing the semantic understanding and integration of diverse health and nutritional data. Ontology represents knowledge as a set of concepts and relationships within the domain. People can now more easily get individualized nutritional guidance at their fingertips thanks to the growth of health applications and online platforms. These applications frequently use artificial intelligence (AI) to assess user progress and deliver personalized recommendations in real-time, improving engagement and adherence to dietary recommendations.

3.3 Technology adapted for Ontology Development

The technology adapted for ontology development involves a combination of languages, tools, and frameworks designed to create, manipulate, and utilize ontologies effectively. One method of data modeling for structured data is ontology. It is made up of ideas such as domain-related entities, characteristics, and properties. Parsing, reasoning, sharing, and reusing knowledge are made easier by employing ontologies. The World Wide Web Consortium (W3C) created the Web Ontology Language (OWL), a Semantic Web language intended to communicate intricate and detailed information about objects and their relationships. Along with RDF, RDFS, SPARQL, and other technologies, OWL is a component of the W3C's Semantic Web stack[25].Web Ontology Language, or OWL, is a crucial semantic web language for ontology creation and exchange. OWL is intended to describe complex information about objects, groups of objects, and relationships between objects. It is based on the RDF (Resource Description Framework) and RDFS (RDF Schema) formats. The Resource Description Framework, or RDF, is a framework for online information representation. Although its main purpose is to represent metadata related to web resources, conceptual modeling also makes use of it. An addition to RDF that offers tools for characterizing collections of linked resources and the connections among them is called RDFS (RDF Schema). There are different ontology development tool such as protégé, TopBraid Composer, ontoStudio, Virtuoso, Blazegraph. Protégé was the ontology development tool used in this study.

Before developing ontology, it is mandatory to identify domain information. This study domain is the food. In here, we need to consider domain scope, key concepts, relationships and attributes and stakeholders' requirements. There are some terminologies associated with ontology development.

Class: Domain knowledge can be categorized into essential classes or concepts. Classes in ontologies are typically hierarchical groups of objects or instances. In a food ontology, for instance, "FoodIrtems" might be a class [26].

Subclass: A subclass represents a more specific class within a broader class. It inherits properties and characteristics of the broader class. For instance, "Fruits" might be a subclass of "Food Items"[26].

Individuals: The Real word objects that fall under the classes' classification are called instances, or individuals. For instance, "Water melon" and "Banana" could be instances of the subclass "Fruit"[26].

Object properties: To specify the relationships between two instances of a class, utilize the object properties. They serve as an efficient means of connecting an individual (instance) from one class to another[26].

Data properties: Relationships between class instances and RDF literals or data values (such as texts, integers, and dates) are defined using data properties. They are employed to give particular values, which can characterize qualitative or quantitative characteristics[26].

Taxonomy: A taxonomy is a hierarchical classification system that uses attributes to group entities within a specific domain. The hierarchy of classes is explicitly referred to as taxonomy in ontology[26].

Axioms:

Axioms are statements or assertions in ontologies, assumed to be true, that help with the inference of knowledge about the classes or the structure of the ontology[26].

Reasoner:

A software program known as a reasoner deduces logical conclusions from a collection of presumptions or stated facts. Ontologies employ it to draw inferences from the defined relationships, classes, and characteristics.[26].

SPARQL query language is a query language, which can be used to retrieve information from ontology. Triple patterns, filters, and optional patterns are put together within a query in SPARQL. To extract the results that satisfy the criteria, these patterns are compared to an RDF graph. Defining the data patterns of interest and having the query engine look for matches for these patterns in the RDF data constitutes the fundamental framework of a SPARQL query.

3.4 Technology used for Machine Learning Model Development

Both machine learning and deep learning make a substantial contribution to the recommendation system. . A deep learning model called a recurrent neural network is trained to process sequential data input and translate it into a particular sequential data output. Depending on the amount of inputs and outputs in the network, there are many varieties of RNNs. One to One, One to Many, Many to One, and Many to Many are the four types. For every time step in the recurrent neural network, there is a single fixed activation function unit. The hidden state of a unit is its internal state, which is unique to each one. At a given time step, the network's previously collected data is known as the concealed state. Every time step, the network updates this concealed state to take into account any changes to its historical knowledge.

The formula for calculating the current state: $h_t = f(h_{t-1}, x_t)$ where,

h_t :current state

h_{t-1} : previous state

x_t : input state [27]

Formula for applying Activation function (tanh): $h_t = \tanh(W_{hh} h_{t-1} + W_{xh} x_t)$ where,

w_{hh} : weight at recurrent neuron

w_{xh} : weight at input neuron[27]

The formula for calculating output: $Y_t = W_{hy} h_t$ where,

Y_t : output

W_{hy} : weight at output layer[27]

We used the Long Short-Term Memory (LSTM) version of Recurrent Neural Networks (RNNs) in this work because, in contrast to conventional RNNs, it is better at capturing long-term dependencies in sequential data. When examining time series or sequence data, this feature is very helpful because it allows for the analysis of prior knowledge that spans varying periods of time. We selected Jupyter Notebook as the development environment because of its interactive computing and visualization features, which make it a favorite tool among data scientists. Our productivity is improved by Jupyter Notebook because it enables us to combine live code, graphics, and descriptive prose into a single document. This makes the presentation of computational processes and outcomes more dynamic and comprehensible.

This initial phase included cleaning the data to ensure accuracy and relevance by removing outliers and handling missing values. Then BMR and BMI values are calculated. With this BMI, class is Assign for every individuals. There are many categorical variables available in the data set. Therefore, categorical variables convert into numerical variables by applying encoding techniques. There are different feature encoding techniques available in Machine Learning such as one-hot encoding, target-mean encoding, frequency encoding . One-hot encoding assigns vectors to each category. The vector indicates the presence or absence of the related feature (1), (0).Target-mean encoding substitutes the target variable's mean value for categorical values. The number of times a specific category value appears in connection to a feature is taken into consideration during frequency encoding. The typical RNN architecture comprises several layers, such as Input Layers (batch size, sequence length, number of features), Recurrent Layers (Simple RNN cells, LSTM cells, GRU cells) dropout layers, output layers (Dense Layer, Sequence output, activation functions) and bidirectional wrappers. The different types of layers incorporated into the model and their technological justifications are summarized in Table 3.1.

Table 3.1: RNN LAYER TYPES AND THEIR TECHNOLOGICAL JUSTIFICATION

Layer Type	**Terminologies Justification**
Input layer	Defines the input shape for the model as a 3D tensor (batch size, 1, 6), where each input sequence has 1 time step with 6 features. This layer is essential for preparing the model to receive the specific format of data.
LSTM Layer	Employs 64 LSTM units to process sequential data with memory capabilities, which are effective in capturing time dependencies within the sequence. LSTMs help the model understand patterns over time, making it ideal for series data, despite the sequence here being very short (1 time step).
Dropout Layer	Randomly drops units (20% dropout rate) during the training phase to prevent overfitting. This helps in making the model generalizable by reducing the reliance on any particular set of neurons.
Dense Layer	A fully connected layer with 3 units outputs a vector of 3 continuous values. As it lacks an activation function, it's used for regression, directly predicting real-valued outputs without transformations, suitable for many quantitative prediction tasks.

In this study, adaptive moment estimation is used as an optimization algorithm during the training phase. From this, it can updated the weights of the network to minimize the loss function. This is traditional stochastic gradient reduction method. Adam uses momentum and RMSProp in a parameter update technique similar to gradient descent. In addition to the exponential moving average of the gradient (m_t), it also employs the exponential moving average of the squared gradient (v_t), just like in RMSProp. Thus, it combines the momentum and the advantages of the RMS-Prop [28]. To update parameters, the following formulas are utilized [28]:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \longrightarrow 1$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) {g^2}_t \longrightarrow 2$$

$$\hat{m_t} = \left(\frac{m_t}{1-\beta_1^t}\right) \hat{v_t} = \left(\frac{v_t}{1-\beta_2^t}\right) \longrightarrow 3$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\in + \hat{v_t}}} \hat{m_t} \longrightarrow 4$$

β_1- exponential decay rate.

m^$_t$ - correction bias of m_t

v^$_t$ - correction bias of v_t

Default β_1 is 0.9

Default β_2 is 0.999

3.5 Technology Used in Recommendation System

Recommendation systems are integral to many modern technological platforms, helping to personalize user experiences in many domain such as health, food, agriculture, e- commerce etc. There are different techniques and technologies used in recommendation systems. Content based filtering, collaborative filtering, hybrid approach, matrix factorization techniques are some of them. In this study, we have used content based recommendation. By contrasting the representation of content that describes the targeted items with the representation of content that the user is interested in, content-based recommender systems generate personalized recommendations. [29].In this study RNN model, predict the target macronutrients values. Those values are used to recommend diet by considering nutrients values available in food ontology. Here we used content based filtering with cosine similarity calculation. Cosine similarity is a measurement that is used to determine how similar two vectors are to each other, regardless of size. What is measured theoretically is the cosine of the angle generated by two vectors projected in a multidimensional space[30]. The following formula can be used to determine cosine similarity once the feature matrix has been created:

$$\text{Cosine similarity of } (v1, v2) = (v1 \cdot v2) \,/\, (\|v1\| \; \|v2\|) \longrightarrow 5$$

Given two distinct vectors, v1 and v2, the magnitudes of those vectors are denoted by ‖v1‖ and ‖v2‖, respectively.

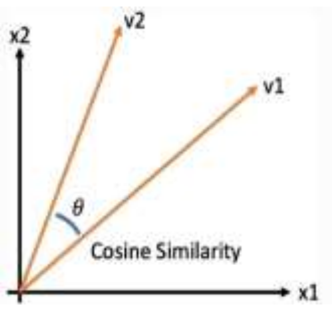


Figure 3.1: Cosine similarity between vector v1 and v2

This study cosine similarity is calculated consider target macro nutrient value vector and nutrient values vector of food groups. For compute cosine similarity score, It can use Numpy and scikit- learn libraries in python.

After that rank, the food based on their cosine similarity score with user profile. Then using similarity scores of food, diet recommend based on diseases information, cultural background and user preferences. In this recommendation diet recommend for three meals breakfast, lunch and dinner.

3.6 Summary

This chapter outlines technologies for developing an ontology based personalized diet recommendation System tailored specifically for the Sri Lankan population. Here, we detailed technologies adapted for ontology development, machine learning model development and user interface development. The next chapter describes an approach for developing an ontology based personalized diet recommendation System with reference to the technologies identified here.

CHAPTER 4

AN APPROACH TO DIET RECOMMENDATION SYSTEM FOR SRI LANKAN

4.1 Introduction

Chapter 3 justified the suitability of machine learning technique as the technology adapted for developing ontology based machine-learning model for Sri Lankan diet recommendation. This chapter presents our ontology-based machine learning approach to Sri Lankan diet recommendation by considering individual preferences, disease information and cultural significance. This chapter focus on ontology based machine learning approach by covering hypothesis, input, output, process ,features and users. Hypothesis section provides a comprehensive explanation of the underlying principles upon which recommendation system is base on. The input section outlines the data inputs required for the model, whereas output section describes the expected model outcomes. Among others, process specially identifies the task within the ontology based machine learning model, which provides the insight into design of the solution. The features section emphasizes the model's essential characteristics.Finally, the users section describes the solution's intended audience and how they willbenefit from it.

4.2 Hypothesis

The central hypothesis of the study is ontology based machine learning model can be implemented to diet recommendation for Sri Lankan. This model is designed to cater to individual user preferences, nutritional requirements, cultural background, and specific health conditions such as cholesterol, pressure and diabetics. By inputting user preferences, cultural background, and pertinent health information, the model generates personalized dietary suggestions tailored to the Sri Lankan context. Utilizing ontology-based machine learning, this system analyzes a vast array of data points, including traditional Sri Lankan cuisine, nutrient composition, dietary guidelines, and health outcomes. By leveraging ontologies, which define relationships between various dietary components, cultural practices, and health conditions, the model ensures a comprehensive and accurate understanding of dietary needs within the Sri Lankan population. Moreover, this approach incorporates user feedback and adapts recommendations over time, ensuring continuous improvement and relevance. By dynamically adjusting to evolving user preferences and health profiles, the model maximizes its effectiveness in promoting optimal nutrition and well-being.

4.3 Input

The ontology-based machine-learning model for dietary recommendations in Sri Lanka integrates a diverse array of inputs to provide personalized and culturally relevant dietary guidance. There are two input data set available with this study. First dataset is food related information. Vegetables, Fruits, Grains, Diary product and protein (fish and meat) food types were considered. Seventy five (75) food items available in this data set. Each food item consist with food name, portion size(g), calorie (kcal) available in the food, macro nutrient contents (Carbohydrate content(g), Protein content(g) and Fat content(g)), micronutrients (Sodium content(mg), Fiber content (mg), Calcium content(mg), Potassium content (mg), Vitamins (mg)), suitable meal type and suitability for diabetics, cholesterol and pressure. This dataset is used to create ontology. The essential inputs for ontology are dietary guidelines and recommendation for individuals with diabetes, high blood pressure, cholesterol, and obesity. Second data set consists with user information and there are 1000 records available in this dataset. User information data set consist with not only food preferences but also crucial demographic and health-related data such as age, weight(kg), height(cm), cultural background, disease history and parent diseases information. By incorporating this comprehensive set of user attributes, the model ensures that recommendations are fine-tuned to individual needs and circumstances. In addition to user profiles, the model consist with ontology encompassing foods and their nutritional profiles, dietary guidelines, and specific recommendations tailored to various health conditions prevalent in the Sri Lankan population. In this study, RNN-LSTM model is used for recommendation. Age, BMI, BMR and availability of diabetes, cholesterol and pressure are the input for the model.

4.4 Output

The objective of this study is to create an ontology based machine-learning model to recommend diets. The key output of this system is diet/foods for three meals. These recommendations are intricately aligned with the user's multifaceted health conditions, dietary preferences, nutritional requirements and cultural background. This comprehensive approach encompasses not only the individual's current health status but also considers their familial health history, recognizing the potential genetic influences on their well-being. In the subsequent subchapter, we describe the detailed methodology used to generate the desired outcome.

4.5 Process

The process of developing an ontology based personalized Diet Recommendation model is based on several steps. The first step is data collection. Data collection involves gathering various data that is essential for creating personalized diet recommendations. There are two types of data set available in this study. User information dataset and food related dataset. User information dataset includes food preference, cultural background, age, weight (kg), height (cm) and disease-related information of different individuals. There are two values consider with food preference: Vegetarian and Non Vegetarian. In addition, we considered individual health condition by considering the availability of blood pressure, diabetes and cholesterol. The second dataset comprises information about various foods along with their nutritional values and dietary guidelines tailored for individuals with specific diseases information such as high blood pressure, cholesterol issues, obesity, and diabetes. Next ontology is created using food information. Ontology is consist with seven main classes such as Dietary category, Diet plans, Diseases, Food Items, Meal Type, Nutrients and User profiles. Diseases class divided into four sub classes such as cholesterol, diabetes, obesity and pressure. Food Items class has five subclasses as vegetable, fruits, grains, meat fish and dairy products. Nutrient class consists with two subclasses such as macronutrients and micronutrients. User profile class has three subclasses such as BMI profile, BMR profile and Cultural background class. In this phase ontology reasoning is used.

Next step in developing an ontology based diet recommendation model involves generating user profiles. User profiles are essential for tailoring diet recommendations to meet individual needs and preferences. The user profile is created by combining demographic information with dietary preference and health related history. The next process is calculating Body Mass Index (BMI) and Basal Metabolic rate (BMR) using demographic information such as age, gender, height, and weight. Then the user profiles are integrated the ontology with mapping with user profile class in ontology. SPARQL Query language is used for achieved this task. It is a query langue used for querying and manipulating data which is available in Resource Description Framework format (RDF).Next process is machine learning model development. In this step machine learning model is developed by using Recurrent Neural Network (RNN) based on feature vectors age, BMI, BMR, availability of pressure, availability of cholesterol and availability of diabetes. There are three targets variables in this model, target protein content, target fat content and target carbohydrate content. In this study LSTM - RNN model is developed for regression task. The neural network architecture adopted in this model has various configurable parameters, each playing a significant role in influencing the model's behavior and performance. Below explains required parameters used in this study [31].

Input Layer - Input (shape= (1, 6)): This outlines the expected shape of the input data according to the model. In this model, there are six features available.

Layer STM: LSTM (64): This includes a 64-unit LSTM layer. After processing the input sequences, the LSTM layer analyzes the data to find patterns. The dimensionality of the LSTM layer's output space is represented by the number of units.

Dropout Rate (Dropout(0.2)): A regularization method called dropout keeps neural networks from overfitting. During training, it randomly deactivates a portion of the neurons, which forces the network to develop more resilient features and reduces reliance on particular neurons. Here, 0.2 indicates that 20% of the units will be randomly dropped during training.

Batch Size (batch_size): The dataset is split up into smaller batches during training, and the model modifies its parameters by calculating the average loss calculated for each batch. The amount of samples handled in each training iteration is determined by the batch size. Lower batch sizes may provide better generalization and convergence features, but larger batch sizes can speed training by utilizing parallel processing and optimizing memory usage. Batch size is equal to 32 in this study [31].

Number of Epochs: During the training phase, each epoch is a single run over the whole training dataset. The number of epochs denotes the number of times the model iterates across the dataset. In this study 20 epochs were considered [31].

‘Adam’ is the optimization algorithm used with the RNN model. RNN model was compiled, trained and finally evaluated the model’s performance. As the next step, recommendation system was developed by inputting three target nutrient content predicted by RNN model. Content based filtering approach is used for recommendation. Cosine similarity algorithm is used in content-based filtering. Cosine similarity used to compute the similarity between target value predicted by

RNN Regression model and nutrient contents available in food items. After that rank, the food based on their cosine similarity score with user profile. Then recommended to top three rank food items for each meal. Food related information is available in the ontology. This information is retrieved to python environment using the SPARQL query language.

Finally, through the integration of a food-related nutritional ontology into the RNN regression model, employing content-based filtering techniques, personalized Sri Lankan diet recommendations are generated. This process ensures a tailored approach to dietary suggestions, enhancing the efficacy and relevance of the recommendations provided. Evaluation is key phase available in this study. Evaluation of the RNN model based on mean squared error, root mean squared error, mean absolute error, R squared value and correlation coefficient. Then the recommendation system is evaluated by using test cases.

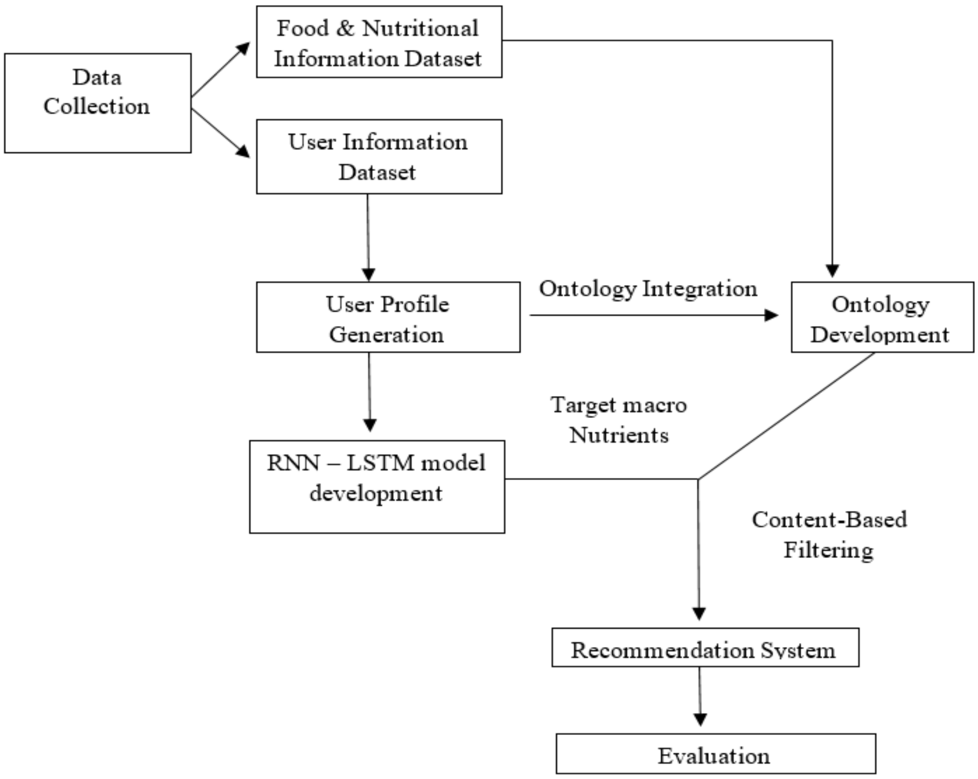


Figure 4.1: Process Diagram

4.6 Features

Main features of ontology based personalized Diet Recommendation System as follows:

- Sri Lankan food recommendation: Diet plans that are specifically tailored to the user preference and customs of Sri Lankan cuisine, guaranteeing cultural tolerance.
- Nutrient need identification: Analyzing a person's nutritional needs based on their age, gender, height, weight, and health issues can help determine which nutrients they specifically need.
- Food menu diversity: Provide a wide range of meal options to satisfy different nutritional needs and user preferences while encouraging adherence to suggested diets.
- Ontology driven reasoning: To enable individualized diet planning and decision-making, use ontology-driven reasoning to model links between nutritional information linked to food, dietary guidelines, and health issues. Semantic reasoning is used in this study
- Health condition management: Customized dietary guidelines for the management of particular diseases, such as diabetes, high blood pressure, cholesterol problems, and obesity. In addition, ensure that dietary recommendations meet the objectives and goals of each individual.
- Real-time adaptation : Continuous monitoring of dietary recommendation

4.7 Users

The ontology based based personalized Diet Recommendation System for Sri Lankan caters to a diverse range of users, each with distinct dietary needs and interests. These users include individuals with specific dietary requirements seeking personalized recommendations tailored to their health conditions or cultural preferences. . Additionally, tourists visiting Sri Lanka can benefit from the system by exploring authentic Sri Lankan cuisine recommendations aligned with their tastes and dietary preferences. Furthermore, the model serves anyone interested in exploring Sri Lankan cuisine.

4.8 Summary

This chapter outlines our approach for developing an ontology based personalized diet recommendation System tailored specifically for the Sri Lankan population, utilizing a combination of ontology and machine learning techniques. Here, we detailed our hypothesis, inputs, desired outputs, processes, features, and users. The next chapter describes the design of ontology based machine learning model with reference to the process identified here.

CHAPTER 5

RESEARCH DESIGN

5.1 Introduction

In Chapter 4, we have presented the approach for ontology based personalized diet recommendation System for Sri Lankan by using ontology and machine learning techniques. We have defined hypothesis, input, output, process, features and users. This Chapter presents the research design with reference to the process identified under approach. The design has been structured to cover the top-level architecture of the proposed model, ontology construction, machine-learning engine, and recommendation system and user interface development.

5.2 Top-level architecture

The top-level architecture of an ontology based personalized diet recommendation system for Sri Lankan based on individual preferences, disease information and cultural significance would typically involve several components. User data play major role in this research design. In constructing the user profile, various factors are taken into consideration to ensure comprehensive personalization. These factors encompass food preferences, distinguishing between vegetarian and non-vegetarian dietary inclinations, as well as cultural background, age, weight, and height, disease-related information, such as high blood pressure, diabetes, and cholesterol levels, along with parent diseases data, are considered. From this rich dataset, two derived attributes are computed: Body Mass Index and Basal Metabolic Rate, offering insights into the individual's nutritional status and metabolic needs.

Next component is a comprehensive ontology, which consists with food and nutritional information as well as vast dietary guidelines and diet plans .Ontology is the knowledge base of this study, which is extracted from nutritionists through interviews. The machine-learning engine is a crucial component of the ontology based personalized diet recommendation system for Sri Lankan. It plays a vital role in analyzing user profiles, ontology, and food related information to generate a personalized list of food menus. Figure 5.1 shows the proposed system architecture.

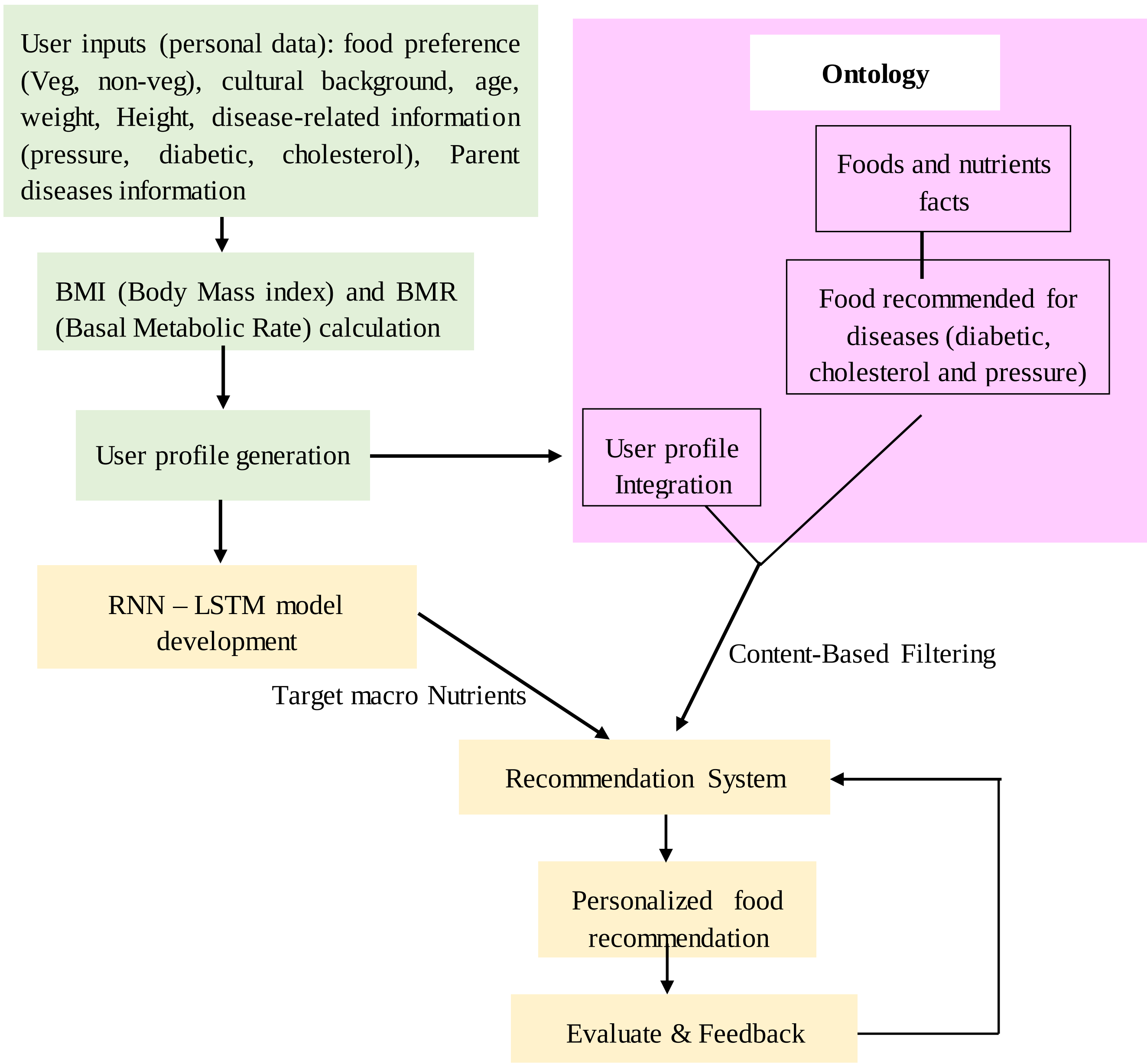


Figure 5.1: Proposed System Architecture

5.3 Ontology

Ontology development plays a crucial role in the modern technology by allowing anyone to share information in a specific domain. In this study, ontology acts as a knowledge base. Ontology contains all the information about foods. It is possible to map the relationship between food information inside the ontology. As the first step in ontology development, we collect information from nutritionists through interview, gather data from public database and Sri Lanka food websites and also use food based dietary guidelines published by ministry of health and World Health Organization (WHO) nutrients recommendations. Here we consider the food name, portion size, calorie available in the food, macro nutrient contents (Carbohydrate content, Protein content and Fat content), micronutrients (Sodium content, Fiber content , Calcium content, Potassium content, Vitamins), suitable meal type and suitability for diabetics, cholesterol and pressure under food information. There are five types of food groups such as grains, dairy products, fruits, vegetables and protein foods (fish and meat). Then we integrate user profiles information into ontology. User Id, age, sex, cultural background, user preference, BMI value, BMI category, BMR value and diseases information.

5.4 Machine Learning Engine

A key element of the ontology based personalized diet recommendation system for Sri Lankans is the machine-learning engine. It plays a vital role in analyzing user profiles and generate personalized diet recommendations. Here is a detailed explanation of the Machine Learning Engine. User data is collected through questionnaire. Data set size is 1000. Here data set is consist with user information such as food preference, cultural background, age, weight, Height, disease-related information (pressure, diabetic, cholesterol), Parent diseases information.

Next step is data preprocessing which involves cleaning and transforming the input data to ensure its suitability for further analysis. We can use one hot encoding technique to convert categorical variables as numerical values. Then by using age, height and the weight of the users, Basal Metabolic Rate and Body Mass Index is calculated. Then user profiles generation by integrating user information into ontology. Then by using nutritionist help target nutrient values calculated for each user. Next step is machine learning model development. First, split the dataset into training, validation and test sets. Model is developed by using Recurrent Neural Network (RNN) – Long Short-Term Memory (LSTM). Train the model using training dataset. Feature vectors of RNN model is age, BMI, BMR, availability of blood pressure, availability of diabetes and availability of cholesterol. RNN LSTM model predict daily target carbohydrate content, daily target protein content, and daily target fat content target variables. Finally evaluate, the model Evaluate the model performance by Mean squared Error, Mean Absolute Error, Root Mean Squared Error and R- squared value.

5.5 Recommendation System

Recommendation systems are pivotal and valuable in the design of personalized dietary guidance. In this particular design, content-based filtering is applied using a machine learning model, specifically a Recurrent Neural Network (RNN), which is integrated with an ontology to deliver a personalized list of food items to users. The RNN model plays a crucial role, as it calculates daily target values for carbohydrates, proteins, and fats based on individual user profiles. These target nutrient values are then used to query a well-structured ontology using SPARQL to retrieve comprehensive food item information. This integration allows the system to align the retrieved food data with the nutritional needs and preferences of the user. The recommendation process not only considers the basic nutritional content of foods but also incorporates the user's cultural background, personal food preferences, and specific health conditions such as diabetes, cholesterol levels, and blood pressure, alongside information about hereditary diseases that might impact dietary needs. By utilizing content-based filtering, the system employs similarity scoring to determine how closely the nutrient content of various food items matches the user's nutritional targets and preferences. This similarity scoring is essential as it underpins the mechanism through which the system assesses and ranks food items, ensuring that recommendations are closely aligned with the unique dietary requirements and preferences of each user, thereby providing a highly customized dietary plan. This comprehensive approach enhances user satisfaction and adherence to suggested dietary guidelines, making the recommendation system a critical component in managing health through diet.

5.6 User interface

The user interface is an important component of the ontology based personalized diet recommendation system for Sri Lankan. It serves as the primary means through which users interact with the recommended system and access personalized diet suggestions. The primary objective at this stage is to create a user-friendly, web-based interface utilizing Flask. User can easily access this web based interface and they have an ability to input their food preference, cultural background, diseases related information, parent disease information and food restriction . Then through this interface user can receive a list of Sri Lankan foods for breakfast, dinner and lunch. To achieve this task, web interface is integrated with ontology based machine learning model.

5.8 Summary

This chapter presented design of ontology based machine learning model for the diet recommendation. We have defined top-level architecture, which encompasses several components and their interactions. In addition, we have explained about ontology, machine learning engine, recommendation system and user interface in this design section. The next chapter describes the implementation of ontology based machine learning model with reference to the design identified her.

CHAPTER 6

IMPLEMENTATION

6.1 Introduction

Research design for ontology based personalized diet recommendation System for Sri Lankan by using machine-learning techniques and ontology is described in Chapter 5. We have defined top-level architecture of the proposed model, ontology construction, machine-learning engine, and recommendation system and user interface development. This Chapter presents the implementation with reference to the research design identified in the research design chapter. Implementation chapter is structure to cover ontology model implementation, machine learning model implementation, Recommendation system implementation, User interface implementation and overall implementation.

6.2 Overall Implementation

With an emphasis on the unique dietary preferences and nutritional requirements of the Sri Lankan population, the ontology based personalized diet recommendation system's overall implementation consists of a number of essential elements that function together to provide users with customized diet recommendations. Ontology Model Implementation, Machine Learning Model Implementation, Recommendation System Implementation and User Interface Implementation are the key components in this process. Figure 6.1 describes flow chart of system implementation

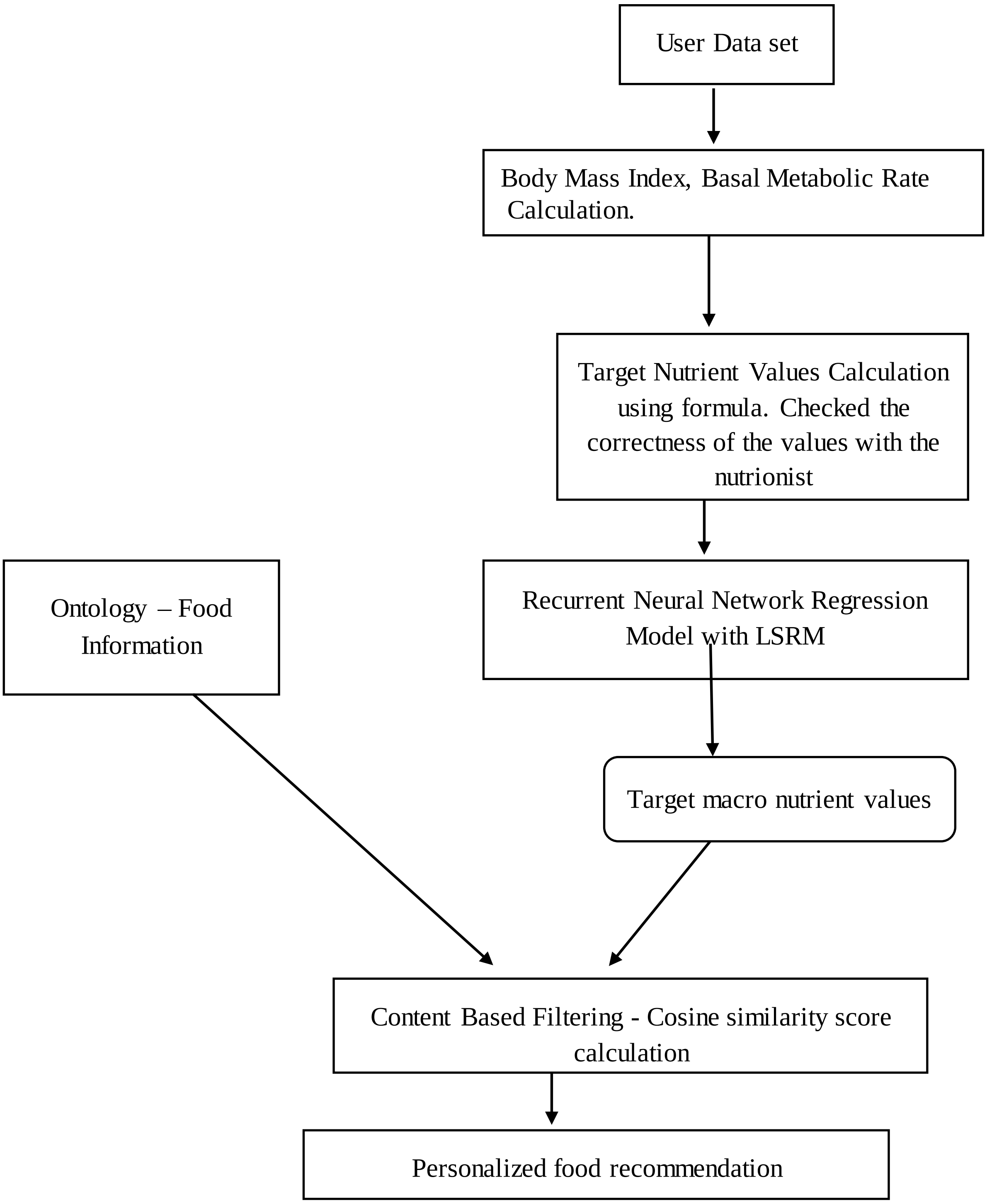


Figure 6.1: Flow chart of system implementation

6.3 Ontology Model Implementation

The ontology model is consisted with the information related to foods. Food related information collected from nutritionists through interviews and gathered some data from public database and Sri Lanka food websites. In addition, to develop this ontology, we use food based dietary guidelines published by the Ministry of health and World Health Organization (WHO) nutrient recommendations. Figure 6.2 represents collected food and their nutrient contents.

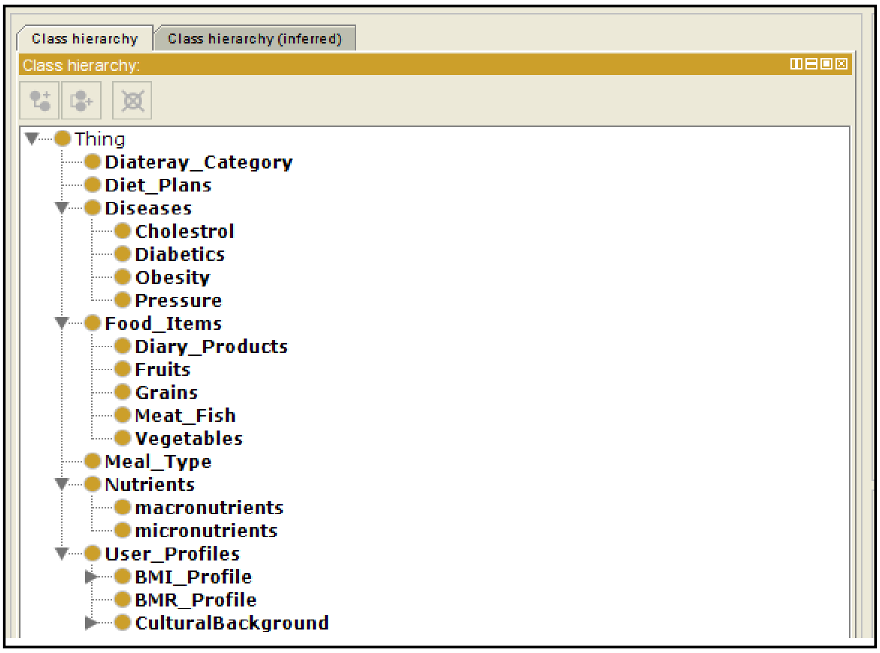
No	Food	Quantity	Nutrients											
			Energy (kcal)	Carbohydrate(g)	Fat(g)	Protein(g)	Vitamin C (mg)	Calcium(mg)	Fiber(g)	Water	Cholestrol(mg)	Potassium(mg)	Iron(mg)	Zinc
1	Red rice	100g	356	76	3.33	8.89	0	0	11.1		0	156	1.11	30
2	Rice	100g	130	28.7	0.3	2.7			0.4		0	35		
3	Hoppers with Lunumiris	45g	123	24.96	2.33	2.08	0.19	8.04					0.68	0.44
4	Halapa	85g	197		6.2	2	0.5	13					0.47	
5	Mung Milk rice with Lunumir	135g	213	32.1	7.2	6.9	2.7	38					1.63	
6	Imbul Kiribath	200g	425	75.6	13.2	5.6	1.3	41					1.11	
7	Kiriroti with kiri pani	35g	173	27.7	6.8	19	7	9					82	
8	Herbal Porridge with Juggery	200ml	232	44.1	5.3	44	9.4	279					1.59	
9	Pittu	100g	260	53	3.5	4	0	10	2		0	80	0.4	0
10	Roti	100g	300	62	9.2	7.85		36	7.1			239	2.2	
11	String hoppers	100g	130	30	0	2	0	10	1		0	0	0.4	0
12	Bada irigu (Maize)	100 g	338	23.54	4.8	8.47		9.2	3.46	8.79		288		
13	Bread	32g	82	13.8	1.1	4	0	42	1.9		0	37	1	
14	Naan (Plain nann)	90g	262	45	5	9			2					
15	Noodles	100g	137	25	2.1	4.5		12	1.2		29	38	1.5	

Vegetables

No	Food	Quantity	Nutrients												
			Energy (kcal)	Carbohydrate(g)	Fat(g)	Protein(g)	Vitamin C (mg)	Calcium(mg)	Fiber(g)	Water	Cholestrol(mg)	Potassium(mg)	Iron(mg)	Zinc(mg)	Soduim(mg)
1	Thuba Karawila	100g	17	3.7	0.2	1	84	19	2.8			296	0.43	0.77	6
2	Ladyfinger (Bandakka)	100g	35	7	0.2	2.2	23		3.2			299	0.62	0.6	7
3	Del	100g	113	26	0.4	1.5			3.8			200			
4	Kos (Jack Fruit)	100g	51	23.25	0.3	2.6		24	1.5			303	0.23	0.13	2
5	Puhul	100g	10	1.9	0.1	0.4	90		3.6			298			
6	Lotus root	100g	79	14.67	0.93	1.94		37.7	4.7	76.26		611			20.63

6.2 : Foods and nutrients content

Collected information is categorized into five types of food groups such as fruits, protein foods (fish and meat), vegetables, grains and dairy products . Here we consider the food name, portion size, calorie available in the food, macro nutrient contents (carbohydrate content, protein content and fat content), micronutrients (sodium content, fiber content , calcium content, potassium content, vitamins), suitable meal type and suitability for diabetics, cholesterol and pressure under food information. This information map into classes, data properties and object properties within the ontology. Ontology is implemented using tools Protégé and stored in a format Web Ontology Language (OWL). There are seventy-five food information available in the ontology. Ontology is consist with seven main classes such as Dietary category, Diet plans, Diseases, Food Items, Meal Type, Nutrients and User profiles. Diseases class divided into four sub classes such as cholesterol, diabetes, obesity and pressure. Food Items class has five subclasses as vegetable, fruits, grains, meat fish and dairy products. Nutrient class consists with two subclasses such as macronutrients and micronutrients. User profile class has three subclasses such as BMI profile, BMR profile and Cultural background class. Figure 6.3 shows the classes hierarchy of the ontology.



6.3 : Classes hierarchy of the ontology

Object properties in an ontology define relationships between two instances of classes. They are employed to convey the relationships between various entities (or individuals) in the ontology. Data properties assign data values to individual instances. There are few data properties and object properties available in our ontology. Domain and the range are the two main concepts associated with object properties and data properties. The following table represents the object properties available in the ontology with their domain and range.

Table 6.1 : Object properties in the ontology

Object Properties	**Domain**	**Range**
belongsToMealType	Food_Items	Meal_Type
foodPreference	User_Profile	Dietary_Category
hasBMIValue	User_Profile	BMI_profile
hasCulturalBackground	User_Profile	Cultural_Background
hasDiateoryCategory	Food_Items	Dietary_Category
hasDiseases	User_Profile	Diseases
hasNutrients	Food_Items	Nutients
hasSuitableforCholestrol	Food_Items	Cholesterol
hasSuitableforDiabetics	Food_Items	Diabetes
hasSuitableforObesity	Food_Items	Obesity
hasSuitableforPressure	Food_Items	Pressure

The object properties hasBMIValue and foodPreference associated with the user information. After collecting user information from the questionnaire, those information is integrated into ontology. This will explain with machine learning model development. In ontology, data properties help to hold the food and nutrients values, calorie content available in the food, Glycemic index, suitable meal types, suitability for diabetes, pressure and cholesterol. Moreover, the ontology includes a variety of other data properties that are crucial for constructing comprehensive user profiles. These properties encompass age, Basal Metabolic Rate (BMR) values, BMI class, cultural background, gender, and specific disease-related information. Each of these data properties plays a vital role in the recommendation system by ensuring that all user-specific factors are considered when generating diet plans. The Figure 6.4 shows the data properties available in the ontology.

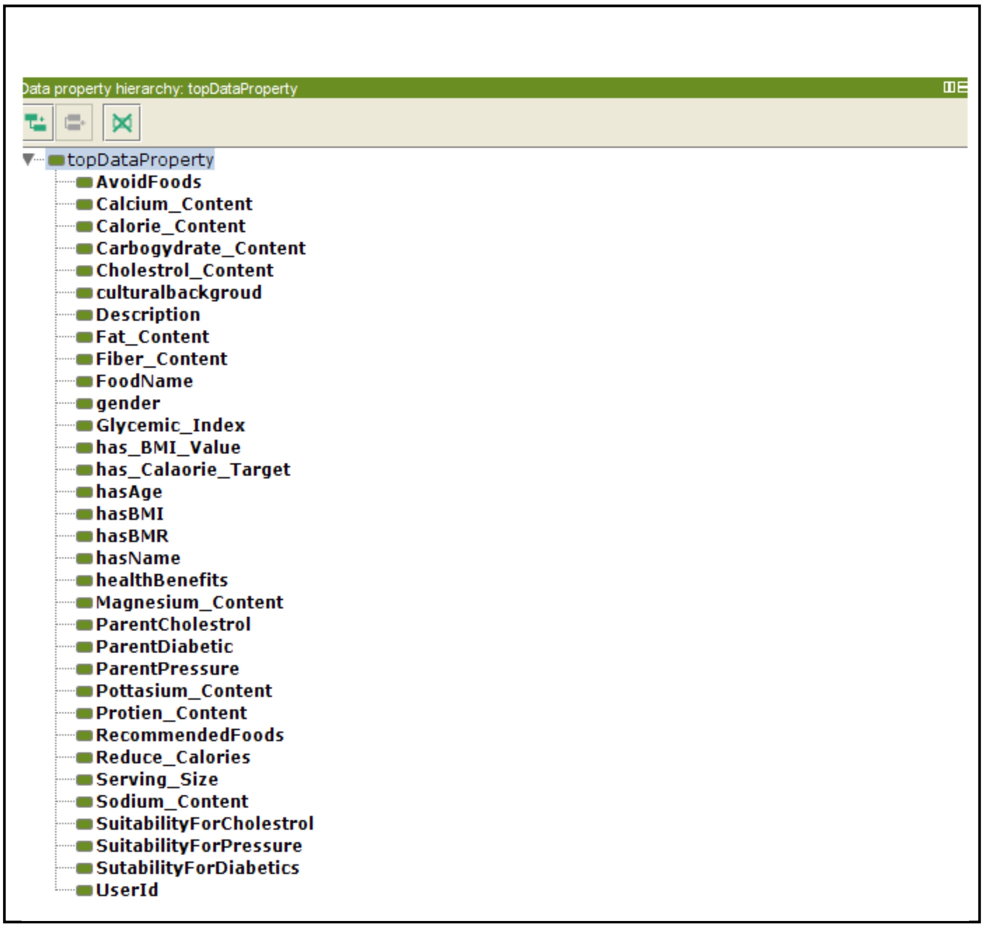


Figure 6.4: Data properties available in the ontology

Food items are represented using individual concepts available in ontology. For instance, consider the representation of "Halapa," a traditional Sri Lankan food, within the ontology. Figure 6.5 in the documentation illustrates how Halapa is conceptualized as an individual entity within the ontology framework.

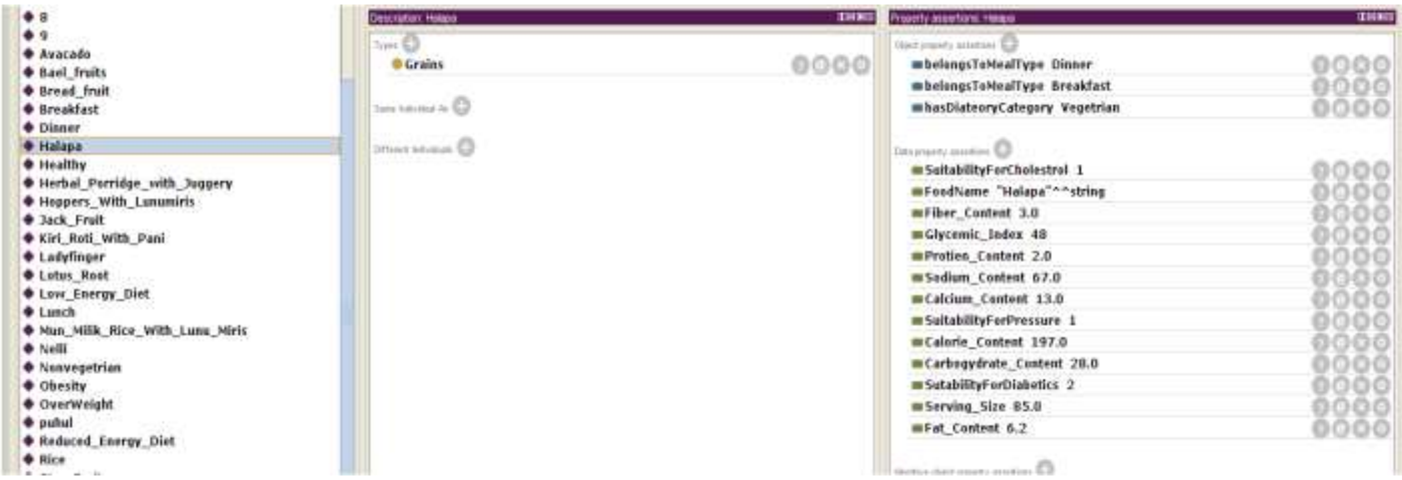


Figure 6.5 : “Halapa” individual

6.3 Machine Learning Engine

The most important aspect of this research is the creation of machine learning engines. As the first step, data is collected through questionnaire. 1000 records available. Then collected data preprocessed by handing missing values. Then categorical data convert into numeric values using One- hot encoding. By using age, sex ,height and weight BMI value is calculated.

BMI = weight (kg) / (height (m))^2

BMR for Men= 88.362 + (13.397 × weight (kg)) + (4.799 × height (cm)) - (5.677 × age)

BMR for Women = 447.593 + (9.247 × weight (kg)) + (3.098 × height in (cm)) - (4.330 × age)

Then by using nutritionist help target nutrient values calculated for each user by considering their BMI and health status. Then Next step is model selection we selected an RNN enhanced with LSTM layers for this study. After that, we performed feature extraction. Several important health indicators are included in our dataset throughout the feature extraction stage. We used programming to extract features from our data, including age, blood pressure, diabetes status, cholesterol levels, body mass index, and basal metabolic rate. These characteristics were combined into a multidimensional array called X, where each row and one of the previously described health indicators represent a user by each column. We simultaneously created the goal variables, which represent the estimated nutritional demands and contain the protein, fat, and carbohydrate targets. To make sure that each nutritional objective corresponds with the associated user data in X, these targets were arranged into another array, Y. In addition to guaranteeing accurate alignment between the input (x) and output (y) data, this organized approach to data preparation makes it easier to train the LSTM model effectively. This approach, which is vital for efficiently managing big datasets, is optimized for speed and scalability by utilizing pandas' vectorized operations. This setup is essential for the LSTM model to learn from intricate, multidimensional nutritional data and make precise predictions about individualized nutrient objectives. Before training the models, we have normalized the input and target data and we split the input and target data into training and testing sets. We set the test size to 20% of the whole dataset and then shuffled the data to ensure stable model training. The compiled the model with Adam optimizer. Model uses mean squared error loss function. The metrics used to evaluate the models during training are mean squared error and accuracy. The model has a batch size of 32 and is configured to train over 20 epochs.

```
model = Sequential([
    Input(shape=(1, 6)),
    LSTM(64, kernel_regularizer=regularizers.l2(0.01), return_sequences=True),
    Dropout(0.2),
    LSTM(32, kernel_regularizer=regularizers.l2(0.01)),
    Dropout(0.2),
    Dense(3, kernel_regularizer=regularizers.l2(0.01))
])
```

Figure 6.6: RNN model

6.4 Recommendation System Implementation

A significant part of this study is the recommendation system. We considered protein, fat, and carbohydrate targets values predicted by the RNN model, user preference, disease information and cultural background for this recommendation. Content based filtering used for this recommendation. System Recommend diets for breakfast, lunch and dinner. In recommendation process, the first step is user profile generation. Then user information integrated with ontology. Then after ontology information (food and nutrient information) loaded into the jupyter environment by using the following code segment.

```
from rdflib import Graph, Namespace

# Load the ontology
g = Graph()
g.parse("E:/Msc/Research/Development/model/ontology1.owl", format="xml")

# Define the namespace
food_ontology = Namespace("http://www.food.com/ontologies/food.owl#")
```

Figure 6.7 : Ontology loading function

In this study, retrieval of the data from the ontology using the SPARQL query language. The target macronutrient values predicted by the RNN model, food and nutrient data from the ontology, disease information, cultural background, and user preferences from user profiles are all considered in the recommendation process. Next, we compute the cosine similarity between the predicted target nutrition values and the nutrient content of available foods in the ontology. This computation allows us to assess how closely the foods match the user's nutritional needs. Then, using content-based filtering, we combine the cosine similarity scores with the user preference, cultural background, and diseases history to provide a personalized list of foods that are suggested for them. Here diets are recommended for all three meals breakfast, lunch and dinner. Through the incorporation of these components, we aim to improve the recommendation process and guarantee that the recommended diets closely correspond to user preference, diseases history and dietary needs. Figure 6.7 shows the recommendation function and the cosine similarity calculation function.

```
meal_targets = {nutrient: daily_targets[nutrient] / len(meals) for nutrient in daily_targets}
target_vector = np.array([meal_targets['Fats'], meal_targets['Proteins'], meal_targets['Carbohydrates']])
meal_recommendations = {meal: [] for meal in meals}

for meal in meals:
    recommended_foods = {}
     # Collecting all recommendations before filtering by meal type
all_foods = {}
for x in food:
    suitable_for_condition = (
        (not user_profile.get('Diabetes')) or x['diabetes']
    ) and (
        (not user_profile.get('Pressure')) or x['pressure']
    ) and (
        (not user_profile.get('Cholesterol')) or x['cholestrol']
        and (
    not (user_profile.get("Cultural Background") == "Muslim" and x['food_name'].lower() == 'pork')
    and not (user_profile.get("Cultural Background") == "Tamil" and x['food_name'].lower() == 'beef')
    )
    )


    if suitable_for_condition and user_profile.get("FoodType").lower() == x['foodtype'].lower():
        food_vector = np.array([x['fat_content'], x['protein_content'], x['carbohydrate_content']])
        similarity_score = cosine_similarity([food_vector], [target_vector])[0][0]
        all_foods[x['food_name']] = {
            'score': similarity_score,
            'mealtype': x['mealtype'].lower()
    }

# Generating recommendations
used_foods = set()
meal_recommendations = {meal: [] for meal in meals}
for meal in meals:
    recommended_foods = {}
    for food_name, details in all_foods.items():
        if details['mealtype'] != meal or food_name in used_foods:
            continue
        recommended_foods[food_name] = details['score']

    sorted_foods = sorted(recommended_foods.items(), key=lambda x: x[1], reverse=True)
    top_recommendation = [food[0] for food in sorted_foods[:top_n]]
    meal_recommendations[meal] = top_recommendation
    used_foods.update(top_recommendation)

return meal_recommendations
```

Figure 6.8: Recommendation Function

6.5 Web interface Implementation

Final step of this study is to integrating the RNN Regression model and recommendation engine with the web interface. Using Flask, we have created a user-friendly platform, enhancing the overall user experience. The development process was streamlined with Visual Studio Code. This integration makes a significant milestone, enabling users to interact with the recommendation system efficiently. The web interface is consists with two pages. First page is designed for user inputs. User can input their person information such as food preference, cultural background, age, weight, height, disease-related information (pressure, diabetic, and cholesterol), and parent diseases information through this web page. Moreover, the second page is designed to display list of foods recommended for breakfast, lunch and dinner.

Figure 6.9: Web interface

6.6 Summary

In this chapter, we have presented the implementation of our study. Specifically, we discussed the ontology development using protégé. Then we discussed about machine learning model development. Here we discussed the pre-processing steps we took to clean and process the user data set along with feature extraction techniques for both the input and target datasets. We then delved into the implementation of RNN models that we built for our study, providing a thorough account of their specifications.Then we outlined the process we followed to compile the model. Lastly, we discussed the recommendation system using content based filtering. Next chapter will present an extensive discussion of the results we obtained and themethods we used to evaluate the system.

CHAPTER 07

RESULTS AND EVALUTION

7.1 Introduction

As covered in the previous chapter, the implementation of the study under three categories ontology development, machine learning model development and recommendation engine. In this chapter, we have outlined the results obtained from the model and the result of the recommendation system. In addition, we will discuss the evaluation of the model and the recommendation system. Finally, we discuss the novelty of our research and findings.

7.2 Machine Learning Model Evaluation

Several evaluation metrics are used to evaluate the performance of the RNN LSTM model. Model uses mean squared error loss function and the metric used to evaluate the model during training is mean squared error. Model is set to train for 20 epochs with a batch size of 32.After training the model, evaluated the results by considering the model loss function over the epochs trained. Here model gave test Mean Squared Error of 0.0290. This indicates that there is a squared discrepancy of 0.0290 between the expected and actual value. Then test Root Mean Squared Error is 0.1703, which measures the average of the squared differences between the predicted and actual values. A lower MSE and RMSE indicates better model performance. In evaluation, results suggest that the model has achieved reasonably good performance. Model provides 0.1239 of Mean Absolute Error, which calculate the average errors between predicted and actual values. We represented the loss function vs number of epochs graphicallyand presented the results in Figure 7.1.

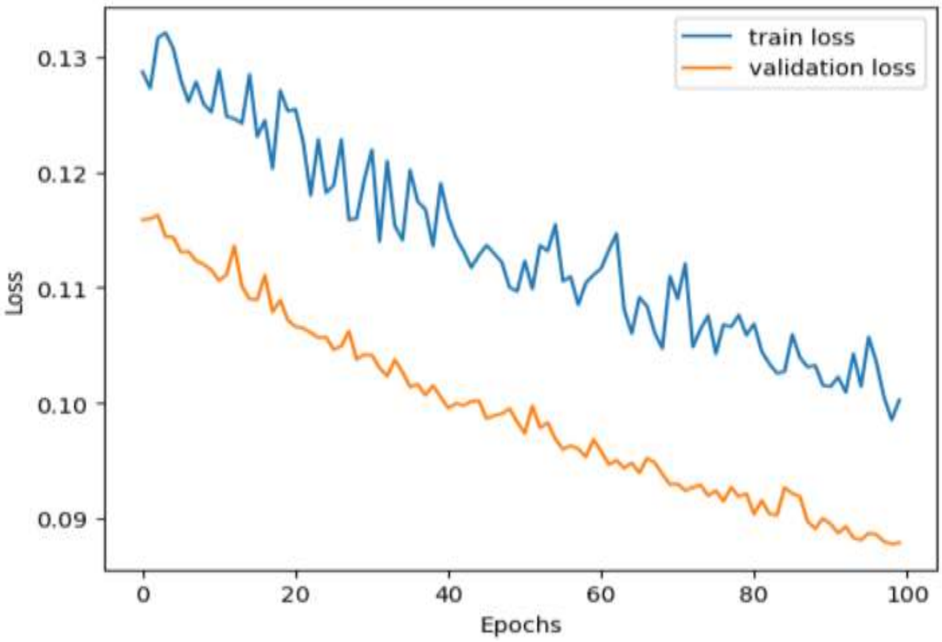


Figure 7.1: Training and Validation Loss over Epochs for RNN Regression

The model showed the relationship between actual and predicted values in order to calculate its R-Squared value, which came out to be 0.4490.It determines whether proportion of the variation of the dependent variable can be forecast using the independent variables. The R-squared value of 0.9704 indicates that the model explains 97.0% of the variance in the test data. Model obtain Pearson Correlation Coefficient of 0.98 for the test set. The Pearson correlation coefficient of 0.98720 indicates a significantly positive linear relationship between the predicted values and actual values. Overall, these evaluation metrics conclude that the model has performed reasonably well in predicting the target variable.

```
Mean Squared Error (MSE): 0.0290
Mean Absolute Error (MAE): 0.1239
Root Mean Squared Error (RMSE): 0.1703
Coefficient of Determination (R-squared): 0.9704
Pearson Correlation Coefficient: 0.9869512680790572
```

Figure 7.2: Model Evaluation output

The plotted the actual vs predicted values in a scatter plot to evaluate the performance graphically in Figure 7.3. In here if the model is performing well, then data points should be closed to the diagonal line. In our RNN model , a large number of data points were dispersed over the diagonal, but scattered some scatter points were far from the diagonal line. This implies that there is moderate relationship between the model's actual values and predicted values.

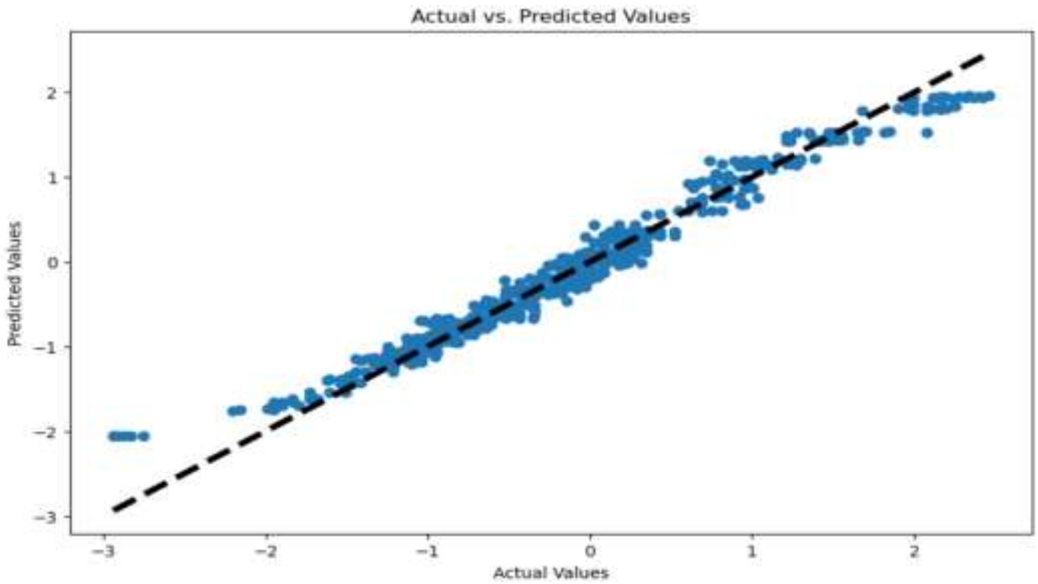


Figure 7.3: Scatter Plot of Predicted vs Actual Value

Finally, It is plotted the graph to compare the first 20 predicted values with the actual values. Figure 7.4 represents the nutrient target predictions with their actual values.

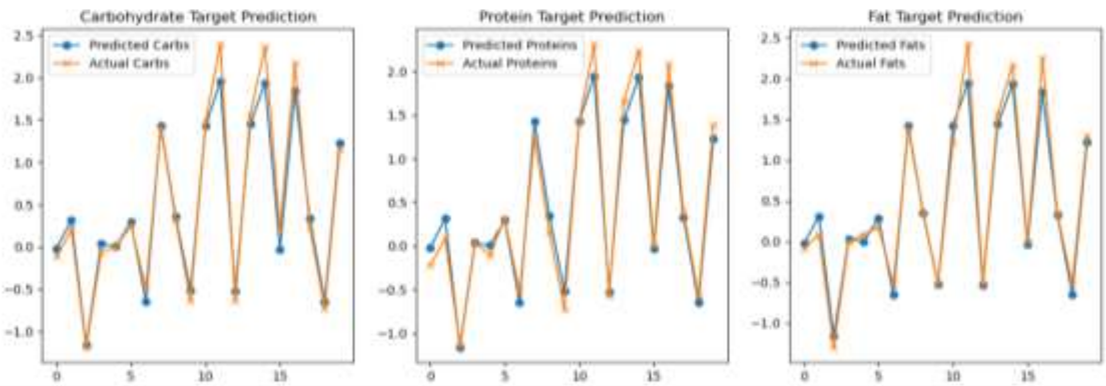


Figure 7.4: Nutrient target prediction

Raksha Pawar and his research team (2001) suggested the K-Nearest neighbors (KNN) algorithm to find the necessary nutrients based on user needs [23]. To evaluate the performance of the RNN-LSTM algorithm suggested by this study, we developed a KNN model using the same dataset for comparison. By implementing both models on identical data, we aim to rigorously compare their predictive accuracies and overall performance. The following table presents a comprehensive comparison between the performance metrics of the RNN-LSTM and KNN models.

Table 7.1 : RNN LSTM and KNN model Comparison

Metric	RNN- LSTM model	KNN model
Mean Squared Error (MSE)	0. 0209	0.0915
Mean Absolute Error (MAE)	0.1239	0.2181
Root Mean Squared Error (RMSE)	0.1703	0.3025
R-squared value (R^2)	0.9704	0.8067
Pearson correlation coefficient	0.9869	0.9054

The RNN-LSTM model has a significantly lower MSE compared to the KNN model, indicating that the RNN-LSTM model's predictions are closer to the actual values on average. In addition, lower RMSE of the RNN-LSTM model further confirms its higher prediction accuracy, as RMSE penalizes larger errors more significantly. The RNN-LSTM model explains 97.04% of the variance in the target variable, compared to 80.67% explained by the KNN model, indicating a much better fit to the data. While both models show strong positive correlations between predicted and actual values, the RNN-LSTM model has a higher correlation coefficient, indicating a better linear relationship with the actual data. Overall, the RNN-LSTM model demonstrates good performance compared to the existing KNN model across all evaluated metrics. By capturing complex patterns and relationships within the data, the RNN-LSTM model delivers more accurate and reliable predictions.

7.3 Results of Recommendation System

Recommendation is the most valuable part in this study. Content based filtering and cosine similarity function is used for recommendation. Implemented recommendation system was tested by using different test cases.

Test case 1:

Case description:

User preference: Non Vegetarian

Cultural Background: Sinhala

Dietary Restriction: No

Pressure: False

Diabetes: False

Cholesterol: False

Parent Pressure condition: False

Parent Diabetic condition: False

Parent Cholesterol condition: False

Recommendation output:

Figure 7.6: Recommendation output1

Test case 2:

Case description:

User preference: Vegetarian

Cultural Background: Sinhala

Dietary Restriction: No

Pressure: False

Diabetes: False

Cholesterol: False

Parent Pressure condition: False

Parent Diabetic condition: False

Parent Cholesterol condition: False

Recommendation output:

Figure 7.7: Recommendation output2

Test case 3:

Case description:

User preference: Non Vegetarian

Cultural Background: Muslim

Dietary Restriction: No

Pressure: True

Diabetes: False

Cholesterol: False

Parent Pressure condition: False

Parent Diabetic condition: True

Parent Cholesterol condition: True

Recommendation output:

Figure 7.8: Recommendation output3

Test case 4:

Case description:

User preference: Non Vegetarian

Cultural Background: Muslim

Dietary Restriction: Chicken

Pressure: True

Diabetes: False

Cholesterol: False

Parent Pressure condition: False

Parent Diabetic condition: True

Parent Cholesterol condition: True

Recommendation output:

Figure 7.9: Recommendation output4

7.4 Summary

This chapter focuses on the evaluation of an ontology-based personalized diet recommendation system for Sri Lanka. For this study, we utilized an RNN-LSTM model and content-based filtering. The performance of the RNN-LSTM model was evaluated using several evaluation indicators.To assess the model, mean squared error (MSE) and root mean square error (RMSE) were employed. The model achieved a test MSE of 0.0209 and a test RMSE of 0.1703. This indicates that the model performed exceptionally well. Additionally, other assessment metrics such as the Pearson correlation coefficient, mean absolute error (MAE), and R-squared value were computed to provide a comprehensive evaluation of the model's performance. The model obtained a Pearson correlation coefficient of 0.9869, a MAE of 0.1239, and an R-squared value of 0.9704. These results indicate that the model explained 97.04% of the variance in the test data, with a relatively small error in the absolute difference between the actual and predicted values. The high Pearson correlation coefficient signifies a significant positive linear relationship between the actual and predicted values.

However, since the scatter plot revealed certain data points dispersed away from the diagonal line, further investigation is required to enhance the model's performance. When discussing our recommendation system, the system recommends diet plans for three meals by considering user preferences, disease information, and cultural background through content-based filtering. In this regard, the system provides recommendations that align with nutrient needs, user preferences, disease information, and cultural background. However, due to a limited number of food items in the ontology, there is less diversity in the recommendations. The study successfully accomplished its objectives by conducting a thorough literature review and discussing the adoption of technologies to address the research problem. This chapter presented an evaluation of the study, specifically discussing the machine learning model evaluation and the results of the recommendation system. Lastly, we outlined the summary of the evaluation. The next chapter will present the conclusion, scope, and future works of our study.

CHAPTER 08

CONCLUSION AND FUTURE WORKS

8.1 Introduction

We emphasized in Chapter 1 how crucial it is to follow dietary guidelines designed with Sri Lankans in mind. We brought attention to the distinct dietary requirements and cultural quirks of Sri Lankan people, emphasizing the vital role that individualized dietary counseling plays in fostering the best possible health and wellbeing for this population. We defined our problem as generating diet plans using explainable AI and proposed our solution to address it. Chapter 2 provided a critical review of literature on recommendation systems in food domain. We have discussed ontology concepts and machine learning concepts and how those technologies use in recommendation context. Here we discussed historical approaches to diet/foods recommendation, current systems, challenges and future direction of recommendation systems. In Chapter 3, we provided a detailed description of the technology adopted, along with the existing and proposed solutions. In Chapter 4, we have discussed the entire approach of our study, by considering input, output, process, features, and users of the solution. We discussed the model's fundamental data input requirements, anticipated model results, methodical processes for producing the intended result, and model features and vital qualities. Chapter 5 covered the research study's design and included recommendations for an appropriate system architecture.

Then Chapter 6 discussed the implementation of the study in detail. We organized this chapter into several sections, including ontology development, data preprocessing, feature extraction in both input and target data representations, model development and recommendation system implementation. In Chapter 7, we evaluated our results for the developed models in detail, using mean squared error, root mean squared error, R-squared value, and mean absolute error and Pearson correlation coefficient. In addition, discussed the result of the recommendation system. In this section, we will go over the study's result and upcoming initiatives to raise the precision and effectiveness of the proposed solution.

8.2 Conclusion

The use of ontology concepts and machine learning techniques in recommendation system has made significant strides in recent years. The main objective of this study is to propose ontology based recommendation system for Sri Lankan. Modern technologies in artificial intelligence, particularly the RNN model, content based filtering and ontology development have been efficiently harnessed to attain this goal. The central hypothesis of this research study was to find the possibility to recommend diets/foods by considering user preference, diseases information, food restriction and cultural significance by using AI. In this study, we created comprehensive ontology consist with 75 food items from grains, vegetable, fruits, protein items (fish and meat) and diary product. We define food, nutritional values and suitability for diabetes, pressure, cholesterol. Then, in order to create unique user profiles, 1000 records of user data were gathered and examined. Then RNN Regression model was developed to find target nutrient values. Then based on target nutrient values, user preference, disease information and cultural background personalized list of foods recommend using content based filtering.

This study provides a comprehensive evaluation of an ontology-based diet recommendation system tailored for the Sri Lankan population. According to our assessment of the RNN-LSTM model, evaluation metrics such as mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE), Pearson correlation coefficient, and R-squared value showed impressive performance. The model achieved an R-squared value of 0.9704, a Pearson correlation coefficient of 0.9869, an MAE of 0.1239, a test MSE of 0.0209, and a test RMSE of 0.1703. These results indicate a positive linear connection between the actual and predicted values, with the model explaining 97.04% of the variance in the test data. However, further investigation is needed to address certain data points that deviate from the predicted trend in the scatter plot.

To further validate our approach, we compared the RNN-LSTM model with the existing KNN model using the same dataset. The comparison revealed that the RNN-LSTM model outperforms the KNN model across all evaluated metrics. Specifically, the RNN-LSTM model demonstrated lower mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE), indicating higher prediction accuracy. Additionally, the RNN-LSTM model achieved a higher R-squared value, demonstrating a better fit to the data, and a stronger Pearson correlation coefficient, indicating a more positive linear relationship between the actual and predicted values compared to the KNN model.

In addition, Evaluation of model, recommendation system represents its ability to generate personalized diet plans for breakfast, lunch and dinner based on user preference, disease information and cultural background. In conclusion, our study has advanced the development of an AI-driven strategy for tailored dietary recommendations for the Sri Lankan environment. Subsequent research endeavors have to concentrate on optimizing the model's functionality and augmenting the variety of food items within the ontology. We may further optimize the system to better suit the nutritional demands and preferences of the Sri Lankan population by addressing these areas.

8.3 Implications

Reflecting on the significance of the research, Ontology based personalized diet recommendation systems consequences for public health in the context of Sri Lanka. From today, many people suffer from non-communicable diseases due to their dietary patterns. In addition, people do not have time to meet nutritionist with this lifecycle. Sometimes meeting nutritionist and getting diet plan is very costly. Ontology based diet recommendation system has the ability to recommend a diet that aligns with user nutrient needs and user preference. This system helps user to find suitable food based on his or her health condition. People can try various Sri Lankan cuisines which is suitable for them. This system helps to improve the individual health condition and it will lay grater foundation towards the development of the country. Overall, the importance of this research resides in its ability to enable people to make knowledgeable food decisions, enhance health outcomes, and add to the larger conversation about personalized medicine and healthcare. The study has successfully achieved its primary objectives by developing an Ontology based personalized diet recommendation system and demonstrating its effectiveness through rigorous evaluation.

8. 4 Limitation

One of the most demanding components of this research was the collecting of data. While the system effectively aligns recommendations with nutrient needs and user preferences, cultural significance and disease information, the limited diversity of food items in the ontology remains a challenge. Therefore, ontology needs to have more food related information. Furthermore, the huge amount of user data required for training the model presented another issue. Handling such a massive dataset necessitates the employment of high-performance computers to handle the data effectively. In the recommendation process of this study, we only consider the macro nutrients need of the individuals. It is another limitation of this study. While macronutrients play a crucial role in dietary health, the omission of micronutrients is a gap in our dietary recommendations.

8. 5 Future Works

In this research study, we developed Ontology based RNN Regression model for diet recommendation. Huge amount of data required for training the model. So next step of the study is to train the model with amount of data and add more food related information into ontology. Reducing the scatter of data points away from the diagonal line in the scatter plot of actual vs. predicted values might enhance the performance of the model. This implies that the data might contain a few outliers or that some patterns may not be adequately captured by the model. Techniques such as outlier detection or feature engineering may help address these issues. Additionally, exploring alternative architectures or hyperparameters may also lead to better model performance. Omission of the micronutrient is another limitation available in the study. Future research could focus on incorporating micronutrient analysis into the recommendation process. Overall, there is great potential for further work in this area, and continued research and development could lead to significant advancements in the field of food recommendation domain.

8.6 Summary

In this chapter, we have presented the overall conclusion of our study on explainable AI based diet recommendation system for Sri Lankan. Our study has shown ontology based RNN regression model can recommend the diets /food based on user preference, cultural background and diseases information. We have provided a comprehensive summary of our study findings and discussed the recommendations and future works that can be done to improve our findings. In our opinion, more study and development in this topic may result in notable progress in the use of explainable AI in the food domain.

.

REFERENCES

[1] "Healthy diet." Accessed: Apr. 02, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/healthy-diet

[2] Q. Zhang, J. Lu, and G. Zhang, "Recommender Systems in E-learning," *J Smart Environ Green Comput*, 2022, doi: 10.20517/jsegc.2020.06.

[3] Y. Ding and S. Foo, "Ontology research and development. Part I - A review of ontology generation," *Journal of Information Science*, vol. 28, pp. 123–136, Apr. 2002, doi: 10.1177/016555150202800204.

[4] A. El-Rady, "An Ontological Model to Predict Dropout Students Using Machine Learning Techniques," Mar. 2020, pp. 1–5. doi: 10.1109/ICCAIS48893.2020.9096743.

[5] "T. N. Trang Tran, M. Atas, A. Felfernig, and M. Stettinger, 'An overview of recommender systems in the healthy food domain,' J Intell Inf Syst, vol. 50, no. 3, pp. 501–526, Jun. 2018, doi: 10.1007/s10844-017-0469-0. - Google Search." Accessed: Apr. 18, 2024. [Online].

[6] C. Trattner and D. Elsweiler, "Food Recommender Systems: Important Contributions, Challenges and Future Research Directions," Nov. 2017.

[7] "Electronics | Free Full-Text | A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields." Accessed: Apr. 18, 2024. [Online]. Available: https://www.mdpi.com/2079-9292/11/1/141

[8] R. van Meteren, "Using Content-Based Filtering for Recommendation," 2000. Accessed: Apr. 19, 2024. [Online]. Available: https://www.semanticscholar.org/paper/Using-Content-Based-Filtering-for-Recommendation-Meteren/4a57e0f0641b7a70fece89c14fbf5030869ededb

[9] C. Obeid, I. Lahoud, H. El Khoury, and P.-A. Champin, "Ontology-based Recommender System in Higher Education," in *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*, Lyon, France: ACM Press, 2018, pp. 1031–1034. doi: 10.1145/3184558.3191533.

[10] "A review of ontologies for describing scholarly and scientific documents." Accessed: Apr. 22, 2023. [Online]. Available: https://ceur-ws.org/Vol-1155/paper-07.pdf

[11] H. el Massari, S. Mhammedi, Z. Sabouri, and N. Gherabi, "Ontology-Based Machine Learning to Predict Diabetes Patients," Jan. 2022, pp. 437–445. doi: 10.1007/978-3-030-91738-8_40.

[12] Y. Ding and S. Foo, "Ontology research and development. Part I - A review of ontology generation," *Journal of Information Science*, vol. 28, pp. 123–136, Apr. 2002, doi: 10.1177/016555150202800204.

[13] "A Review on Ontology Development Methodologies for Developing Ontological Knowledge Representation Systems for various Domains - Google Search." Accessed: Apr. 19, 2024. [Online].

[14] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. da P. Francisco, J. P. Basto, and S. G. S. Alcalá, "A systematic literature review of machine learning methods applied to predictive maintenance," *Computers & Industrial Engineering*, vol. 137, p. 106024, Nov. 2019, doi: 10.1016/j.cie.2019.106024.

[15] K. Bothmer and T. Schlippe, "Skill Scanner: Connecting and Supporting Employers, Job Seekers and Educational Institutions with an AI-based Recommendation System," Jun. 2022.

[16] V. W. Anelli, A. Bellogín, T. Di Noia, F. Donini, V. Paparella, and C. Pomo, "Adherence and Constancy in LIME-RS Explanations for Recommendation," Sep. 2021.

[17] S. Bhaskaran, R. Marappan, and B. Santhi, "Design and Analysis of a Cluster-Based Intelligent Hybrid Recommendation System for E-Learning Applications," Mathematics, vol. 9, no. 2, Art. no. 2, Jan. 2021, doi: 10.3390/math9020197.

[18] M. Phanich, P. Pholkul, and S. Phimoltares, "Food Recommendation System Using Clustering Analysis for Diabetic Patients," *2010 International Conference on Information Science and Applications*, pp. 1–8, 2010, doi: 10.1109/ICISA.2010.5480416.

[19] "Healthy Sri Lankan Meal Planner with Evolutionary Computing Approach | SpringerLink." Accessed: Jul. 16, 2024. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-33-4355-9_17

[20] R. Sookrah, J. D. Dhowtal, and S. Devi Nagowah, "A DASH Diet Recommendation System for Hypertensive Patients Using Machine Learning," in *2019 7th International Conference on Information and Communication Technology (ICoICT)*, Kuala Lumpur, Malaysia: IEEE, Jul. 2019, pp. 1–6. doi: 10.1109/ICoICT.2019.8835323.

[21] D. Mckensy-Sambola, M. Á. Rodríguez-García, F. García-Sánchez, and R. Valencia-García, "Ontology-Based Nutritional Recommender System," *Applied Sciences*, vol. 12, no. 1, Art. no. 1, Jan. 2022, doi: 10.3390/app12010143.

[22] S. S. Kardam, P. Yadav, R. Thakkar, and A. Ingle, "Website on Diet Recommendation Using Machine Learning," vol. 08, no. 04, 2021.

[23] “Pawar et al. - 2021 - NutriCure A Disease-Based Food Recommender System.pdf.” Accessed: Jul. 16, 2024. [Online]. Available: https://ijisrt.com/assets/upload/files/IJISRT21JUN107.pdf

[24] Y. Zhang and X. Chen, “Explainable Recommendation: A Survey and New Perspectives,” *INR*, vol. 14, no. 1, pp. 1–101, Mar. 2020, doi: 10.1561/1500000066.

[25] “OWL - Semantic Web Standards.” Accessed: Apr. 19, 2024. [Online]. Available: https://www.w3.org/OWL/

[26] N. F. Noy and D. L. McGuinness, “Ontology Development 101: A Guide to Creating Your First Ontology”.

[27] “Introduction to Recurrent Neural Network,” GeeksforGeeks. Accessed: Apr. 19, 2024. [Online]. Available: https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/

[28] GeeksforGeeks (2024) *What is Adam Optimizer?*, *GeeksforGeeks*. Available at: https://www.geeksforgeeks.org/adam-optimizer/ (Accessed: 16 July 2024).

[29] C. S. C. Kwan, M. Q. Koh, and M. B. Jasser, “A Comparison Study Between Content- Based And Popularity-Based Filtering Via Implementing A Book Recommendation System”.

[30] M. S. Reddy, P. T. R. Kumar, L. M. Siddarth, and R. Mothukuri, “Designing Recommendation System for Hotels Using Cosine Similarity Function,” in *Soft Computing for Security Applications*, G. Ranganathan, Y. EL Allioui, and S. Piramuthu, Eds., Singapore: Springer Nature, 2023, pp. 1–15. doi: 10.1007/978-981-99-3608-3_1.

[31] S. Abdul-Jabbar, A. Farhan, and A. Luchinin, “Comparative Study of Anemia Classification Algorithms for International and Newly CBC Datasets,” *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 19, pp. 141–157, May 2023, doi: 10.3991/ijoe.v19i06.38157.

Appendix – A: Questionnarie

Questionnaire - Sri Lankan Food Recommendation Using Explainable AI

I am Maheshika Dissanayake a postgraduate student of Faculty of Information Technology, University of Moratuwa pursuing a Master's degree in Artificial Intelligence. As a part of the program, I am conducting a research on "Sri Lankan Food Recommendation Using Explainable AI". You have been identified as a key stakeholder on this research and kindly fill this questionnaire and share your experience with me. Your details entering here are solely confidential and using only for academic purposes.

Section 1: Demographic Information

1. Enter your age

2. Gender

- Male
- Female
- Other:

3. Cultural Background

- Sinhalese
- Tamil
- Muslim
- Other:

Section 2: Food preference

1. Food Preference

- Vegetarian
- Non-vegetarian

Section 3: Health Information

1. Height (cm):

2. Weight (kg):

3. Do you have any dietary restrictions or allergies?

- Yes
- No

If Yes, (please specify):

4. Do you have high blood pressure?

- Yes
- No

5. Do you have diabetes?

- Yes
- No

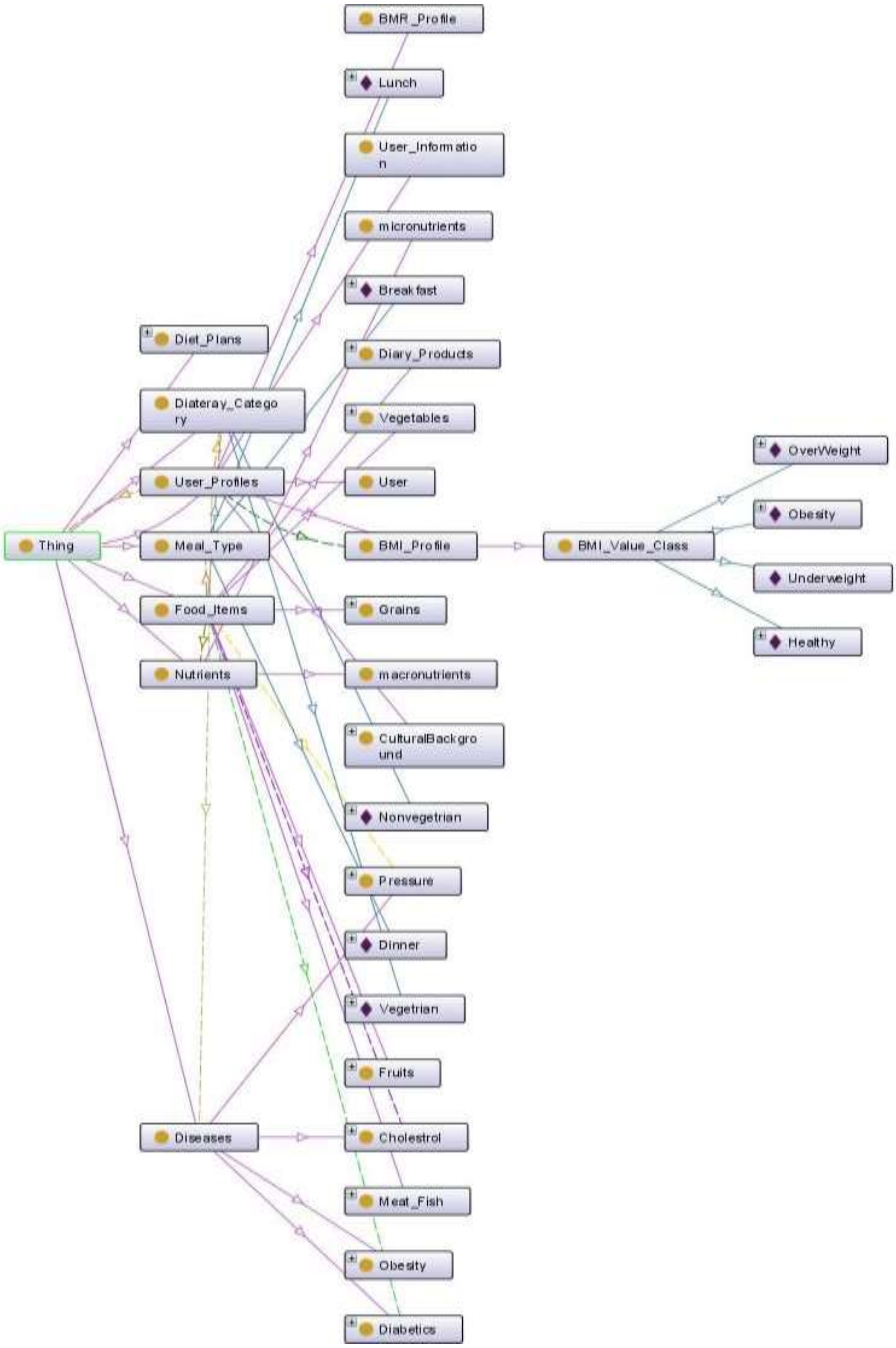
6. Do you have high cholesterol?

- Yes
- No

7. Does your father or mother have the following health conditions?

- [ ] Diabetes
- [ ] Cholesterol
- [ ] High blood pressure
- [ ] None of the above

Appendix – B: Ontology Graph



Appendix – C: Source code of model Architecture

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense, Dropout, Input


x =[]
for index, row in data.iterrows():
    age = row['Age']
    BMI = row['BMI']
    BMR = row['BMR']
    blood_pressure = row['High Blood Pressure']
    diabetes = row['Diabetes']
    cholestrol = row['Cholesterol']
    x.append([age,BMI,BMR,blood_pressure,diabetes,cholestrol])

y = []
for index, row in data.iterrows():
    carb_target = row['carb_target']
    protein_target = row['protein_target']
    fat_target = row['fat_target']
    y.append([carb_target,protein_target,fat_target])

scaler_x = StandardScaler()
scaler_y = StandardScaler()

x_scaled = scaler_x.fit_transform(x)
y_scaled = scaler_y.fit_transform(y)


X_train, X_test, y_train, y_test = train_test_split(x_scaled, y_scaled, test_size=0.2, random_state=42)


X_train = X_train.reshape((X_train.shape[0], 1, X_train.shape[1]))
X_test = X_test.reshape((X_test.shape[0], 1, X_test.shape[1]))


model = Sequential([
    Input(shape=(1, 6)),
    LSTM(64),
    Dropout(0.2),
    Dense(3)
])


model.compile(optimizer='adam', loss='mse', metrics=['mae'])
model.fit(X_train, y_train, epochs=20, batch_size=32)

model.summary()
```

Appendix – D: Source code of recommendation

```
meal_targets = {nutrient: daily_targets[nutrient] / len(meals) for nutrient in daily_targets}
target_vector = np.array([meal_targets['Fats'], meal_targets['Proteins'], meal_targets['Carbohydrates']])
meal_recommendations = {meal: [] for meal in meals}

for meal in meals:
    recommended_foods = {}

all_foods = {}
for x in food:
    suitable_for_condition = (
        (not user_profile.get('Diabetes')) or x['diabetes']
    ) and (
        (not user_profile.get('Pressure')) or x['pressure']
    ) and (
        (not user_profile.get('Cholesterol')) or x['cholestrol']
        and (
    not (user_profile.get("Cultural Background") == "Muslim" and x['food_name'].lower() == 'pork')
    and not (user_profile.get("Cultural Background") == "Tamil" and x['food_name'].lower() == 'beef')
    )
    )


    if suitable_for_condition and user_profile.get("FoodType").lower() == x['foodtype'].lower():
        food_vector = np.array([x['fat_content'], x['protein_content'], x['carbohydrate_content']])
        similarity_score = cosine_similarity([food_vector], [target_vector])[0][0]
        all_foods[x['food_name']] = {
            'score': similarity_score,
            'mealtype': x['mealtype'].lower()
    }

# Generating recommendations
used_foods = set()
meal_recommendations = {meal: [] for meal in meals}
for meal in meals:
    recommended_foods = {}
    for food_name, details in all_foods.items():
        if details['mealtype'] != meal or food_name in used_foods:
            continue
        recommended_foods[food_name] = details['score']

    sorted_foods = sorted(recommended_foods.items(), key=lambda x: x[1], reverse=True)
    top_recommendation = [food[0] for food in sorted_foods[:top_n]]
    meal_recommendations[meal] = top_recommendation
    used_foods.update(top_recommendation)
```