DEVELOPMENT OF A PRESCRIPTION RECOMMENDATION SYSTEM USING CASE-BASED REASONING

BERKAY KAAN KARATAŞ

May 2019

DEVELOPMENT OF A PRESCRIPTION RECOMMENDATION SYSTEM USING CASE-BASED REASONING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES OF
ÇANKAYA UNIVERSITY

BY
BERKAY KAAN KARATAŞ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
COMPUTER ENGINEERING
DEPARTMENT

MAY 2019

Title of the Thesis: **Development of a Prescription Recommendation System Using Case-based Reasoning**

Submitted by **Berkay Kaan KARATAŞ**

Approval of the Graduate School of Natural and Applied Sciences, Çankaya University.

Prof. Dr. Can ÇOĞUN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Erdoğan DOĞDU
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Murat SARAN
Supervisor

Examination Date: 23.05.2019

Examining Committee Members

Prof. Dr. Erdoğan DOĞDU	Çankaya Univ.	
Assoc. Prof. Dr. Gökhan ŞENGÜL	Atılım Univ.	
Assist. Prof. Dr. Murat SARAN	Çankaya Univ.	

STATEMENT OF NON-PLAGIARISM PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Berkay Kaan KARATAŞ

Signature:

Date : 28.05.2019

ABSTRACT

DEVELOPMENT OF A PRESCRIPTION RECOMMENDATION SYSTEM USING CASE-BASED REASONING

KARATAŞ, Berkay Kaan
M.Sc., Department of Computer Engineering
Supervisor: Assist. Prof. Dr. Murat SARAN

Mayıs 2019, 59 pages

In this thesis, a prescription recommendation system was developed based on past prescriptions in order to reduce the workload of physicians and increase the accuracy of written prescriptions. The case-based reasoning method used in this research is among the technological developments used in real life in different fields. In this study, the prescription recommendation system was developed using case-based reasoning method and research was performed to find out the performance of this system. In order to create a set of data to be used in the study, 7120 anonymous prescription information was collected from 300 volunteers through a website. The success rate of the system was then calculated by comparing the prescriptions (1) with the prescriptions prescribed by ten physicians in different branches, and (2) with the latest prescriptions in the data set by using the nearest neighbors' algorithm. The success rate of the system obtained by comparing the prescriptions of the real-life prescriptions for the patients and the recommended prescriptions is 0.78. Besides, the success rate of the system regarding the comparison of the last prescriptions of the 50 most common diseases in the system's data set and the prescriptions recommended by the system is 0.91. The results of this study indicate that health-care professionals can benefit from the recommendation system developed in this study. In general, with the

recommendation system designed in this study, health-care professionals are supported to make faster and more accurate decisions during the prescription writing process.

Keywords: Cased Based Reasoning, Recommendation, Problem-Solving.

ÖZ

VAKA TEMELLİ MUHAKEME YÖNTEMİ KULLANARAK BİR REÇETE TAVSİYE SİSTEMİNİN GELİŞTİRİLMESİ

KARATAŞ, Berkay Kaan
Yüksek Lisans, Bilgisayar Mühendisliği Anabilim Dalı
Tez Yöneticisi: Dr. Öğr. Üyesi Murat SARAN

MAYIS 2019, 59 sayfa

Bu tezde, hekimlerin iş yükünü azaltmayı ve geçmişte kullanılan vakalara dayanarak doğruluğu arttırmayı amaçlayan bireysel kullanım için reçete önerileri sisteminin geliştirilmesi ve performansının araştırılması amaçlanmaktadır. Bu araştırmada kullanılan vakaya dayalı akıl yürütme yöntemi, günümüzde farklı alanlarda gerçek hayatta kullanılan teknolojik gelişmeler arasındadır. Bu çalışmada geliştirilen reçete öneri sistemi vaka temelli akıl yürütme yöntemi kullanılarak geliştirilmiş ve bu sistemin başarım oranını bulmak için araştırma yapılmıştır. Öncelikle çalışmada kullanılacak veri setini oluşturmak amacıyla 300 gönüllüden bir web sitesi aracılığıyla yaklaşık 7120 anonim reçete bilgisi toplanmıştır. Ardından, sistemin başarım oranı, önerilen reçetelerin (1) farklı branşlardaki on hekimin yazdığı reçetelerle (2), en yakın komşu algoritması ile ilgili hastalık için belirlenen veri setindeki en yeni reçetelerle karşılaştırılmasıyla hesaplanmıştır. Hekimlerin hastaları için yazdıkları gerçek yaşam reçeteleri ile sistemin önerdiği reçetelerin karşılaştırılmasıyla elde edilen öneri başarı oranı 0.78'dir. Ayrıca, sistemin veri setinde en çok görülen 50 hastalık için yazılan son reçetelerle bu hastalıklar için sistem tarafından önerilen reçetelerin karşılaştırılmasıyla elde edilen öneri başarı oranı 0.91'dir. Bu çalışmanın sonuçları, sağlık uzmanlarının bu çalışmada geliştirilen öneri sisteminden

yararlanabileceğini göstermektedir. Genel olarak, bu çalışmada tasarlanan öneri sistemi ile sağlık uzmanlarının reçete yazma sürecinde daha hızlı ve daha doğru kararlar almaları sağlanabilecektir.

Anahtar Kelimeler: Vaka tabanlı akıl yürütme öneri sistemi.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Assist. Prof. Dr. Murat SARAN for his supervision, suggestions and encouragement through the development of this thesis. Also, I would like to thank to my family to their continued support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FORMULAS

LIST OF TABLES

LIST OF ABBREVIATIONS

CBR	Case-Based Reasoning
DBS	Database Management System
HTML	Hyper Text Markup Language
MVC	Model – View – Controller

CHAPTER 1

INTRODUCTION

1.1. Introduction to Case-Based Reasoning

Case-based Reasoning (CBR), which was firstly described by Schank, is the method of solving new problems based on the solutions of previous similar problems [1]. According to Kolodner [2], case-based reasoning is to adapt old solutions to meet new demands; use old solutions to explain new situations, or to interpret a new situation or to produce a solution equal to a new problem.

Case-based reasoning is considered as an appropriate methodology for conducting research in the medical domain. Commonly, it is known that the symptoms indicate the problem, and the diagnosis and the resultant prescription represent the solution in the medical domain. It is appeared that, compared to other applied methods, CBR was found to be more flexible when updating cases, had better explanation to the situations, and was better at handling incomplete data and more features than other methodologies.

In summary, the case-based reasoning methodology can be defined as retrieving solutions of the previous similar problems and use them in the current problems. The following section describes the CBR algorithm used in this study in detail.

1.1.1. CBR Algorithm

CBR is an essential method in many domains and plays a significant role in problem-solving. The CBR algorithm is a technique which solves new problems by retrieving similar problems among the previous ones and matching new one to fit their needs. It has commonly been assumed that the solution to the problems generated from similar solutions. In general, the findings from studies suggest that throughout the year's people experienced many problems, cases or situations. Most of them are related to each other in some ways. It is necessary here to match problems with CBR that calculates similarity rate among the past experiences.

According to Mantaras et al. [3], solving a problem with CBR, (1) obtaining the definition of the problem, (2) measuring the similarity of the problem with the previous problem, (3) attempting to reuse one or more similar situations, and (4) adapting the differences of these solutions. CBR has an advantage over improving efficiency and quality of problem-solving and continuous adding new solutions to knowledge have advantages for suggesting a better solution [3]. CBR techniques may be classified according to problem and methodology; thereby it can vary depending on the design.

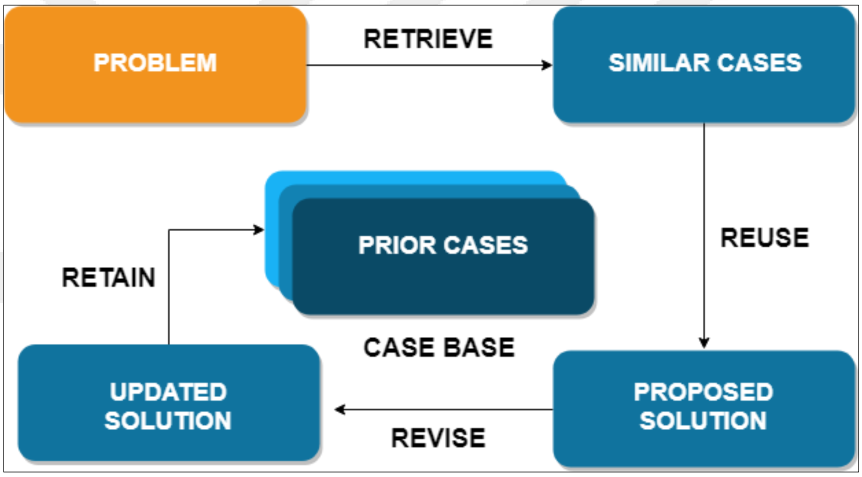

Figure 1 *- Classical "4REs" CBR cycle*

Figure 1 presents an overview of the classical model of problem-solving in CBR. The phases of CBR, called 4RE's, can be listed as follows: Retrieve, Reuse, Revise, Retain [5]. The graph shows that classical CBR cycle, research may be defined as which consists of four process cycle: (a) retrieve, (b) reuse, (c) revise, (d) retain.

Berghofer and Iglezakis compared the case-based reasoning systems and demonstrated that the four processes of the Aamondts framework is task-oriented and this is sufficient for explaining the mechanism [6]. The framework which is classical CBR 4REs was divided into two groups according to application and maintenance. Figure 2 shows the application phases steps and distribution of their task.

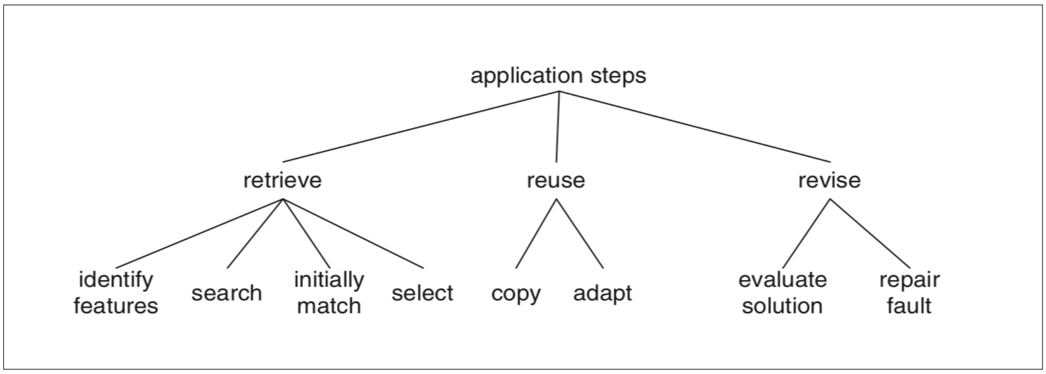

Figure 2 *- Application steps and their task decomposition*

Firstly, Retrieve can broadly be defined as a retrieving a most similar solution from previous cases. Secondly, the term reuse is generally understood to mean of pairing the knowledge and solution to the target problem. Thirdly, the term revise has been used to refer to situations in which used in case of necessity to the proposed solution. Lastly, in broad terms, retain can be defined as an adopted solution that is stored in a case suggesting a future problem.

Generally, a case defined as $c = (c^d, c^s)$ is a pair, $c^d \in D$ is a problem definition and $c^s \in S$ is a solution. Moreover, D denotes 'problem description space' and S denotes 'solution space.' 'The problem description' defined with the query $q \in D$ and the aim of the system find the solution for q [7].

a. Retrieve

Retrieve has a pivotal role and initial step in a CBR system. In this case, the most similar solution retrieved from the previous cases to solve the current problem. This phase aims to explore the relationship between the current problem and previously existing solutions which best-matched case among each other is our solution [7].

$$T_q = \{c_1, c_2, \dots, c_k : f(c_i^d, q) < \theta\}$$

Formula 1 *- The Case Retrieving Process*

The retrieval case represented by formula 1, where 'distance metric between two problems' defined as $f: D\ X\ D \rightarrow \mathbb{R}$ and threshold defined as θ. Also, T_q denotes the 'cases retrieved from the memory' and R_q denotes 'set of the solution' [8].

b. Reuse

The process of reuse could be identified as recommending a solution from previous solutions to our current problem [3]. So that, from the definition of reuse phase, it can be called a solution part of the problem where each case could likely to be represented as a proper solution. It attempts to solve the following issues, adopting a proper solution and how do differences at problem affects the solution.

As Kyrilov states, at reuse case, 'The case-based T_q in order to learn a function $g: S^k \rightarrow S$, which transforms a set of k solutions into a single solution, $s^\iota = g(R_q)$, which is then suggested as a solution to q.'

c. Revise

Revise is the latest process of an application phase, and it is necessary here to evaluate precisely the solution and repair the faults of the data. Since CBR only proposes solutions due to the inaccurate match, proof of correctness or external validation may be required. A specialist can only handle the verification of the solution via its application process. So, if the result is correct, the system retained the solutions and case learned [9]. Also, it is a stage for learning failure if the solution is applicable or not in the reuse phase; it is possible to learn from mistakes.

In the revised case, as Kyrilov declares, 's^ι may not be the acceptable one, the accepted solution is defined as $s* = h(s^\iota)$, where $h: S \rightarrow S$ that modifies a solution, evaluated again until an acceptable solution is found.'

d. Retain

The retain process is particularly valuable that the problem - solving property is incorporated into the existing case base in order to make the knowledge available for later reuse. This phase used to assess the feedback which is a necessary condition for a system to learn from knowledge. Whereas, if the system faced with a particularly similar problem, just as it controls the database and most similar proposal is submitted to the user otherwise problem recorded the database from future use [9].

1.2. Aim of the Study

There is a growing demand for health care services. One of the most significant challenges of this demand is the lack of health professionals, and observations have indicated a severe decline in the population of patient and health professional ratio [12]. As the role of technology appears to be more and more critical in the health care domain, incorporating technology in this domain has also become a necessity to be successful. The software-based approach employing decision support systems and recommendation systems is one of the main approaches to technology usage in the health care domain.

The primary purpose of this study is to develop a recommendation system using case-based reasoning methodology that can help health professionals in diagnosing and prescribing diseases. The system uses prescriptions that we created by using clinical prescription patterns, which was anonymized entirely, for health professionals use for finding a most similar prescription for a case. On the other hand, one of the most beneficial aspects of this system will be for new practitioners who can benefit from experienced physicians. The health professionals find out a proper drug in the large-scale distribution of drugs. The use of CBR provides a mechanism for health professionals to write a prescription accurately.

1.3. Research Questions

The purpose of this study is to develop a system using CBR for recommending a proper prescription to health professionals for their patients. This thesis answers how to build a new prescription recommender system and how we recommend the proper prescription. In this study, we aim to find out the answers of the following questions:

1. What is the success rate of the system when comparing the prescriptions recommended by the system and the real-life prescriptions written by thirteen physicians in different branches?
2. What is the success rate of the system when comparing the prescriptions

recommended by the system and the latest prescriptions in the data set for the 50 most frequent diagnosis?

1.4. Significance of the Study

Although CBR methodology has been used in the medical domain [10], the use of CBR in the prescription recommendation has not been sufficiently investigated yet and has not been widely used.

In the rapidly evolving pharmaceutical technology, selecting the right drug within some reimbursement regulations has become one of the main obstacles of the health-care ecosystem. In Turkey, since it was reported in 2018, the number of prescriptions has been started to increase each following year periodically [11].

Table 1: *The Prescription Statistics Number, Cost, and Amount Per Prescription*

Data	2016	2017	2018
Prescription Number (Thousand pieces)	257.545	265.481	283.625
Invoice Sum	₺15.415.296	₺18.435.461	₺22.528.171
Amount Per Prescription	₺59,86	₺69,44	₺55.21

Table 1 presents the summary statistics of the total number of prescriptions, costs, and amount per prescription according to the Social Security Institution annual report. From Table 1 shown above, we can see that the number of prescriptions and the prescription written by physicians are continually increasing year by year. Therefore, in this study, we have developed a case-based recommendation system for a prescription prediction aimed at reducing the workload of physicians and increasing accuracy based on the cases used in the past. One of the objectives of this study is to present the best recommendations to health-care professionals from similar prescriptions written in the past. Notably, this recommendation system might play a vital role for professionals who are beginning of their career. This prescription recommendation system can also give drug information such as reimbursement, active ingredients, drug price while offering a recommendation. This system helps the physician to identify and observe other drugs in the same diagnosis category.

This study aims to contribute to this growing area of the healthcare system by offering prescriptions for a specific diagnosis to ease physician's jobs. The main aim of this study is to develop software that helps healthcare professionals by recommending a prescription. In this study, we designed a Prescription Recommendation System Software running on Apache Server with Java via "vue.js."

1.5. Limitations of the Study

The results of this study are limited to 7.120 anonymous prescription information that was collected from 300 volunteers.

1.6. Organization of the Thesis

The overall structure of this thesis covers seven chapters. The structure as follows:

Chapter 2 begins by giving a brief history and development of Case-Based Reasoning. It moves on to provide an overview of previous literature of research.

Chapter 3 presents the methodology used for this study. It begins by describing research techniques that will be used in our study. Then, the overall structure of the research design, software requirement specification, and tools used on the application are illustrated. Moreover, the development of the interface and database model are also explained in this chapter. Finally, the test phases of an application are described.

Chapter 4 reveals the results by providing the findings of our CBR recommendation system concerning the doctors' real-life prescriptions for a disease and the prescription recommended by the system regarding the previous prescriptions, and the last prescription for a disease and the prescription recommended by the system regarding the previous prescriptions.

Chapter 5 includes a conclusion drawn from our prescription recommendation system using Case-Based Reasoning and discuss the findings on literature.

Chapter 6 states the suggestion for future work.

CHAPTER 2

LITERATURE REVIEW

The purpose of this chapter is to give brief information about CBR history, review the literature on CBR and investigate prior studies related to CBR. It starts with a brief overview of its history and continues with a review of the literature on CBR. This chapter concludes with a summary.

2.1. History of Case-Based Reasoning

CBR is a widely used methodology for solving activities, problems at a variety of fields such as manufacturing, medicine, pharmacy, law, design [13]. CBR was used to solve problems by adapting the previous experience to similar cases. It does not mean to solve the problems entirely but suggests a proper solution.

In 1992, an example CBR study was carried out by Kolodner in which he created a relationship between physicians and patients [2]. Traditionally, the physicians have shown a variety of patients whose symptoms were somehow like each other. If the physicians have seen similar symptoms on patients, they should have remembered the previous diagnosis and need to make some correlation between the symptoms. These numerous similarities could not be the exact solution to the new problem. However, it should be verified by physicians before considering a similar diagnosis.

Nevertheless, it focuses them to find a possible solution easily. From this point of view, the CBR methodology might help to physician's significant savings of time. Kolodner suggests an additional explanatory example for car mechanics. It just looks like the physician's evaluation procedure, and it can be used in any domain as well. Suggesting a similar diagnosis to the physician can lead to an increase in the effectiveness of medical exam of every patient as well as car mechanics. A positive correlation was found between using CBR suggestion system and efficiency.

As Richter [14] studied that associated with CBR are schema-oriented memory models, which have a long tradition. Around the early 1970s, small-scale research and case studies began to emerge linking with artificial intelligence and CBR. Roger Schank and his students conducted a study that explored ways in which the computer could understand the everyday language we speak and based their work on observations of the way people seem to understand everyday language. [15]. Researchers have explored that previous experiences could make a significant impact on learning methodology so that CBR has been carried out in museums, pre-school and school — basically, CBR process aiming to solve problems by using previous similar problems. Finding a relationship between problem and ex-problems is the essence of CBR [16].

However, CBR gathered with some concepts that are outside the computer science [17]. On the contrary, Bartletts schema theory, Schank proposes a dynamic memory model. In 1977, Bartletts and his students at Yale University offered probably the most cognizable work of CBR. Schank et al. set up a script that records the desired steps and performed them into action. Schank found dynamic memory theory and this theory have memory organization packet (MOP) which reminds both cases and patterns. So, these stereotypes qualified by using states, events, scenes, actors, and purposes. His dynamic memory theory had been much more useful on CBR. In 1990, the study by Kolodner reviewed dynamic memory with using E-MOPS in CYRYS system. However, that one was the first actual system created with CBR. The usage of dynamic memory has been an example and influenced among other systems, some of them are CYRUS, CHEF, and MEDIATOR, created by Riesbeck, Schank, Kolodner. Also, these systematic developments and contributions in CBR have led to the advancement of artificial intelligence.

In 1983, Rissland and colleagues from the University of Massachusetts examined their works on law domain using CBR methodology. The objectives of their research are to focus on the role of precedence reasoning in proper adjustment. They seek to obtain data which will help to interpret states and extends the argument to both sides.

Between 1988 and 1991, Bruce Porter from the University of Texas did several works at specific law application over CBR which are respectively PROTOS and GREEBE. Similarly, in 1991, Ashley proposed a work named a HYPO that inspired by Rissland. The purpose of this work is to explore the relationship between analyzing the law rules and producing a justification [18].

In 1992, it was the beginning of the new period for CBR, especially in Germany. Methods used in engineering started to be more profitable and productive.

In 1993, Kolodner applied CBR in the health domain. In his study which called CASEY, provides in-depth analysis of the work of heart attack. It was the first to use the system joining both CBR and deep model-based. A longitudinal CBR study by Kolodner was the most comprehensive publication in this field [19].

Europe began to work in the CBR area later than in America. In Europe, the research and development of CBR quality systems have become stronger. As a result, the MOLTKE system was developed by Richter at the Althoff University of Kaiserslautern. The development of this system triggered further work. Michael M. Richter developed a system named PATDEX. As Richter states "Patdex 1 is an expert system that uses case-based reasoning in the diagnosis of faults of complex machines" [20].

In the Blanes IIIA, Enric Plaza and Ramon Lopez de Mantaras developed the CBR system that diagnosed the disease for practitioners [21]. As a result of these studies, the REFINER was developed by Sunil Sharma and Oehlmann, respectively. Then, problem-solving mechanisms provided them to focus on the use of merged cases and the general knowledge of a domain. This approach resulted in the creation of a system called CREEK [21]. An analogical argument was carried out by Mark Keane Trinity College, and he provided a robust solution for this type of CBR. Gerhard Strube, Freiburg University, made a significant contribution to the research of cognitive models with the EVENTS project.

A conference about CBR, first organized in 2001, became constant at the Florida Artificial Intelligence Research Society Conference: FLAIRS. Since 2002, the annual

German workshop has been called the Workshop on Experience Management. This workshop was called mining of experience because CBR methods used more commonly in each domain and various of many problems. Following years between the second part of the 1990s and middle of the 2000s, it became one of the most important topics among the recommender systems.

2.2. Related Studied using Cased-Based Reasoning

There are studies investigating the use of CBR method in health domain. For example, Zhang et al. [22] proposed a fault diagnosis method for sophisticated equipment using CBR. He points out that CBR has used to diagnosis field, but efficient retrieval approach, not enough for defining the diagnosis. He developed an approach by using case retrieval for diagnosing 4135 Diesel engine faults by classifying. The results indicated that the fault diagnosis method is robust and practical, and it can be used in other areas for practical problem diagnosing.

In another study, Lawanna et al. [23] presented a study for improvement of the CBR system. The research states that the reusing phase increases the complexity of the mapping required for the selected problem to a proper solution. After a while, bigger datasets make the process slow down and hard to execute. In his study, he identified and focused on improving the suitable solution. Authors proposed a model by classifying problems according to their types and removed irrelevant cases from their datasets. Conclusively, the results showed that Using a proposed model is better than traditional models with 1.6-2 times on the other hand fixing problems are much preferable than other similar studies.

In their study, Feng et al. [24] conducted a research for exploring a new emergency cases retrieval method based on CBR. Lately, a substantial economic crisis, environmental disasters, wars have a severe effect on human beings' lives. Therefore, how to search for a highly relevant historical case in the more extensive case database is topic research in recent years. So, this study set out with the aim of using PSO and TS algorithm in attributes of CBR. The current study found that the problem of feature weight value was avoided. Finally, the authors said that Further studies, which these studies have not been verified, will need to be undertaken.

Gatzioura et al. [25] published a paper in which they examined a study for music playlist recommender system using CBR. In order to generate more accurate, more enthusiastic, more complete experience recommendations to a user, hybrid CBR approach combined with a graph model. The authors indicated that their hybrid CBR system is performed better and more accurately than other recommendation techniques.

Tsatsoulis et al. [26] published a paper in which they integrated CBR and Decision Theory. The purpose of the current study was to improve the ability of CBR. The research performed in designing pharmaceuticals and generating drugs. Development phases of a drug are very complicated and hard so that CBR and decision theory benefit from each other. The results of this investigation show that CBR and Decision Theory integration assists in handling their problems.

Zhang et al. [27] conducted a research using a framework of a hybrid recommender system for a personalized clinical prescription. Zhang and colleagues consider the difficulties when the practitioners decide to write prescriptions to patients. The hybrid framework was designed that combined artificial neural network and CBR to support the decision-making phase. Conclusively, the results of this study indicate that the system examines associations between other fields which need other expertise on these domains, the system should be analyzed very well before using in domains.

Jin et al. [28] demonstrated research for proposing clinical network and CBR method to ease clinical practitioners' duty on their daily activities with such a decision support system. He indicates that most of the studies on the medical domain use statistical methods for modeling. However, the author proposed a clinical network model. Finally, the findings contribute in several ways to our understanding of algorithms on used datasets.

Hsiao et al. [29] conducted a study for forecasting with using hybrid CBR system. The authors state that forecasting with CBR is a very effective way of prediction. They designed a hybrid system with granular computing for useful weather prediction. Also, the authors stated that forecasting has complexity because of the many variables which affect the predictions. This study noted that their similarity measure was based on interval-valued would be applicable for future work.

Gao et al. [30] investigated an intelligent fault diagnosis approach integrating cloud model and CBR. As noted by authors, the complexity of the technology and structure of the vehicle is more complicated than ever so that, the fault complexity of the problem is getting harder to produce a solution. This paper test and analyzed their new model towards this situation. In the final part of the article, the cloud model recovered some information from uncertainty and CBR which is improved with Euclidean Distance formula, used for them to find a proper solution to decision makers. Overall, the authors have approved the effectiveness of this method.

Chergui et al. [31] examined research using CBR approach to reusing Communities of Practice (COP) for a university student. According to the researchers, a relationship exists between social interaction with colleagues and academic success to motivate university students who were revealed in COP. In order to understand how COP with the CBR approach regulates academic success, they developed their COP environment using CBR cycles. In summary, these results show that the CBR approach for COP has positive effects on university students and they defined success elements for students in higher education.

Ajjouri et al. [32] conducted a study with developing a new model of Intrusion Detection Based (IDS) on Multi-Agent Systems with adopting CBR technique. In their introduction to the paper, the authors state that IDS is a crucial role in network systems in order to detect network attacks. Sometimes IDS might be triggered mistakenly. Also, they indicate that many IDS systems are monolithic and centralized in collecting data. Overall, the results indicated that their model has better scalability and accuracy in detecting new attacks.

Yu et al. [33] published a paper in which they described research with analyzing aircraft fault diagnosis system that combines CBR and Fault Tree Analysis (FTA). Aircraft systems are complicated systems, and when the fault happens, it affects variables units. Also, the authors state that mathematical models are not suitable for diagnosing a fault. Many other techniques and models influence aircraft fault diagnosis, but according to the experts, CBR is the most suitable technique for fault diagnosis expert system. The investigation of this study has shown that building a fault diagnosis model through CBR is supports maintenance personnel for finding an efficient and accurate solution by enriching and upgrading solution knowledge.

Deng et al. [34] studied for the similarity of equipment fault diagnosis algorithm with using CBR approach. In this investigation, the aim was to improve the efficiency of fault diagnosis by taking advantages of CBR. In general, therefore, the results of this study indicated that the algorithm more feasible and well performant. Also, the authors stated that further research should focus on reducing cases with the help of similarity.

Ashraf et al. [34] examined a literature review on CBR for matching composite sketches to facial photographs. The findings indicate that the regular process of identifying criminals or suspects could take a long time, so that, the authors proposed a CBR approach for improving the process. The results of this investigation show that CBR is an efficient way to recognize sketches.

Sharma et al. [36] published a paper in which they examined a study to diagnose HIV/AIDS detection with an ontology supported CBR system. The study proposes a system for HIV/AIDS detection primarily with comparing symptoms which used experiences stored in the database. With the help of this approach, medical assessment can be asked further test from patients and can perform quickly. The findings of this study have many important implications for future practice, and the authors reported that this model could be used in other diseases.

Further research such other similar functions may be required to determine efficiency for better analysis.

Sappagh et al. [37] published research that making a relationship between SNOMED and CBR in order to support the clinical decision support system to diabetes. The researchers use EHR data, standardized with SNOMED, collected from patients and used in CBR. The results showed that the new ontology with using CBR had been created for diabetes. Authors also declared that further research in this field should be concentrate on other conceptual fields such as hypertension, heart diseases.

Karim et al. [38] investigated research that explores a generic ontology for CBR systems. In their study, hybrid and generic recommender system wants to be created which could be reusable any application domain. In general, therefore, it seems that this paper presented some parts of ontology in domain-independent CBR recommender system and further work needs to be done to establish in other remaining steps.

Xiaopeng et al. [39] demonstrated research to prevent diseases from rice pest. The study highlights the climate change which closely related to diseases of pest and the authors designed a data mining CBR model for prevention and controlling rice pest. Furthermore, the present research aimed to examine the steps of the CBR model and further works needs to be improved for finding a correct solution.

Begum et al. [40] published a meta-analysis study in which they present a systematic literature review and the questionnaire on the medical CBR systems. They stated that, currently, the health domain had become one of the most common uses of CBR. The results of this study show that CBR has commonly applied to various medical domains.

2.3. Chapter Summary

In this section, studies on CBR has been presented in a variety of domains. This section began with a brief overview of the history of CBR. Then it continued by review of usage of CBR in the industry, health and the medical domain and other domains. In summary, CBR has been used to evaluate medical decision support systems and improve public health and health-care professionals' effectiveness.

CHAPTER 3

METHODOLOGY

Chapter 3 begins by laying out the data collection method and looks at how the system calculates the similarity case. The chapter continues by explaining the test environment and how it is implemented. Next, the test methodology will be explained. This chapter concludes with a summary. The research methodology was designed and details will be discussed in the forthcoming sections.

3.1. Data Collection Method

Based on the clinical prescriptions 300 volunteers were entered anonymous prescription information by the approval of the Cankaya University's ethics committee. Moreover, these anonymous data were enhanced using the diseases and equivalent drug databases with the permission of the Turkish Pharmacist Association. Icd-10 codes are international standards of the statistical classification of diseases and related health problems, and our icd-10 table was generated in the tree structure [41]. A total of 11.630 icd-10 codes were transferred to the table with sub-branches. Also, the drugs used in the market were added with the corresponding information. The drug table includes approximately eighteen thousand drug records. Then, we have linked the drugs with icd-10 codes in a table in our database. Approximately 82.200 units were matched with the icd-10 code. Prescriptions were reproduced by using indications by using the software, and then pharmacists controlled the accuracy of the prescriptions. Physicians' opinion was obtained for combined drug use in prescriptions. Permission was obtained from the Turkish Pharmacists' Association for the use of indications and drug information.

__Table 2:__ The data used in the prescription recommendation system

Data	Number
Drugs	18.457
Icd-10 code	11.630
Matched icd-10 and drugs	88.220
Prescription (used in system)	7.321
Prescription Drugs (used in system)	11.647
Prescription icd-10 (used in system)	102

Table 2 presents an overview of the data used in prescription recommendation system.

3.2. Calculating Similarity of Cases

The analysis was based on the conceptual framework proposed by Richter et al. [42]. In order to execute prescription recommendation system, the cases must be necessary to execute cluster function. Each object must have a U cluster [42]. To establish whether linked with a new case and previous case, every single case that belongs to the cluster must need an appropriate value.

$$O \in A$$

***Formula 2** – The Case Belonging*

The main aim of this calculation is to analyze the current situation and providing a proper solution to the situation. Resolving and assessing cases contribute to a better understanding of the problem. The similarity may have been an essential factor in specifying cases because the similarity of the cases and attributes plays a vital role to find similar results. An attribute-value relationship is a well-established approach in this system [2]. These cases can be classified into the similarity between attributes and cases.

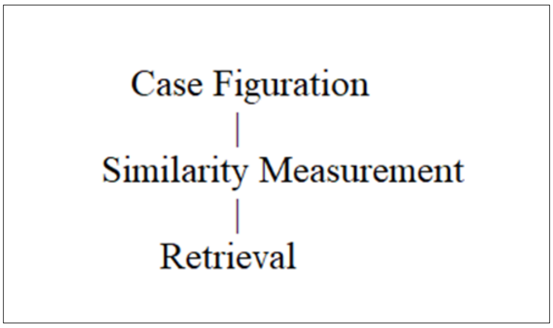

Figure 3 *– Similarity Process*

These cases, which are described in Figure 3, of similarity, are essential matching against the cases between each other. The following values (see Figure 4) were given to measure the similarity between the features:

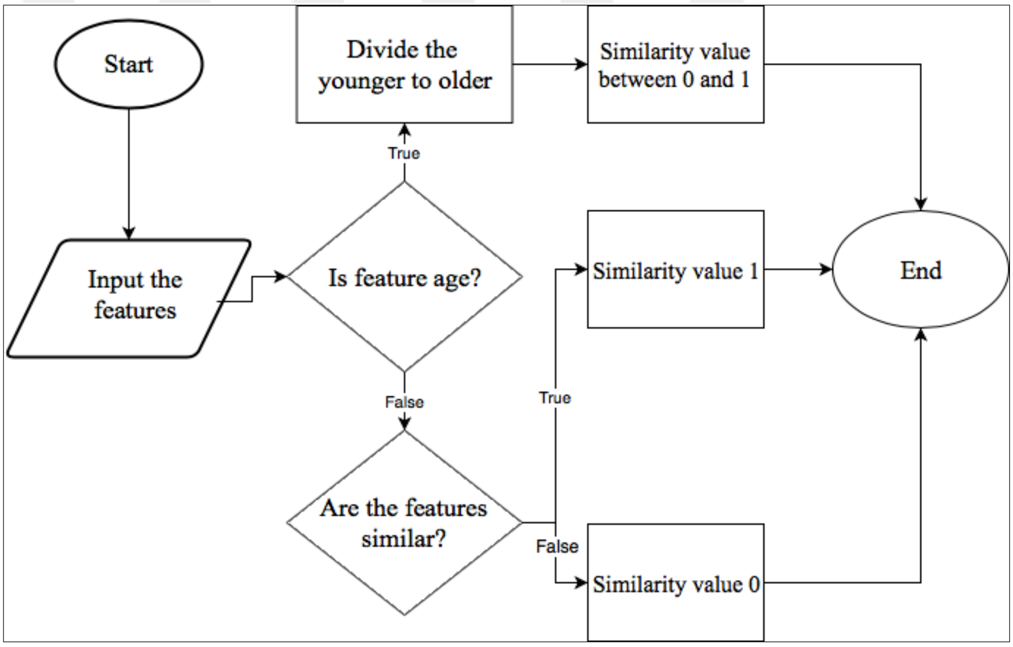

Figure 4 *– Measurement of Similarity of the Features*

Table 3 is an illustration of how the similarity process works by providing an example. Five different types of features are used to calculate the similarity between the cases as shown in Table 3. These are the diagnosis, gender, prescription type, age, and physician branch. As can be seen in Table 3, the diagnosis, gender and physician branch are the same, but the prescription type and age are different.

Table 3: *Example of Similarity Comparison between two cases*

Feature	1st Case	2nd Case	Value
Gender	Male	Male	1
Prescription Type	Green	White	0
Year	1988	1998	0.66
Branch	Orthopedics and Traumatology	Orthopedics and Traumatology	1
Diagnosis	Meniscus Tear	Meniscus Tear	1

There are several approaches to determining the similarity. A numerical value is given to the case by comparing the degree of similarities. Each feature represents one part of the problem set and describes the cases. The similarity between the two problems is presented as two points in an n-dimensional space. There are many standards measures when implementing the CBR system, which is based on domain and knowledge [43, 44].

$$sim: P\ x\ CBP\ \rightarrow\ [0,1]\ \in\ R$$

Formula 3 *– Similarity Measure*

“CBP denotes an input descriptions P for which a solution exists such that (P, S) is in the case base.” Formula 3 is a mapping for similarity measurement. There are some assumptions for reducing arbitrariness which is commonly used but generally predicted [45]. Therefore, we assumed the following in our similarity analysis:

i. $0\ \leq\ \mathrm{sim}(x,y)\ \leq\ 1$ (normalization)
ii. $sim(x,x)\ =\ 1$ (each object is itself the nearest neighbor)
iii. $sim(x,y)\ =\ sim(y,x)$ (symmetric property)

As Wangenheim states [45], “For an attribute-value representation, a simple similarity measure is the generalized hamming measure that combines the importance of each attribute of the problem description with its local similarity value and sums the values to create a global similarity value for each case.” In our case, the weight of each feature was considered equal. Many researchers have utilized this method to measure the similarity between the cases [46, 47]. Our approaches to the prescription recommendation system were the same. This method was selected for its reliability and validity.

Table 1 presents an overview of the results obtained from the user's analysis of the similarity between cases. It can be seen from the data in Table 1 that the similarity between the two cases was calculated by comparing each other. What stands out in the table is described as follows:

There are two different types of information in our datasets. One of them includes patient information (gender, age), and the others includes non-patient (icd-10, branch, prescription type). When the system tries to compare gender feature between two cases, the one is given if the genders are the same in both cases. The one is given if the age same with the compared case. However, if the ages are not the same, then, the younger age is divided into a older age. Following, the result of this calculation yields a value between 0 and 1.

Prescription type, physician branch, and icd-10 code are non-patient information from our dataset. The one value is given if the prescription types are the same. Otherwise, the zero value is given to that comparison. The same process is carried out in physician branch. If the physician branch is paired with the comparison, the one value is retrieved from the calculations. The comparison is based on the icd-10 code. For example, if the physician reported that the patient had a meniscus tear on his/her knee, the icd-10 code of that diagnosis is S83.2. In our approach, we apply all possible calculations of the icd-10 code up to its parent.

Thus far, this section has argued that how the similarity of features was calculated. In order to identify similarity rate, all features, collected above section, put together all coefficients and divide the number of features. Overall, these results represent that similarity rate between the two cases. The correlation between prescription is related to their diagnosis. Up to their parent code, the comparison continues. Thereby, the retrieval process can first try to find a match with the child code and upwards.

CBR is frequently used to find the closest match case and to obtain similarity. In a study conducted by Cinar [48], it was shown that indexes and weight are directly related through to matching, but the ranking is related to the total score. In order to find matched and similar data, CBR investigates whole database not only the compared cases. Many studies (e.g., Holt, Macdonell, Benwell) have shown that several popular statistical techniques had already been used to set designate similarity such as linear, exponential or logarithmic functions, fuzzy-logic, artificial neural network. Most researchers investigating CBR have utilized the nearest neighbor algorithm [13]. This nearest neighbor algorithm used to find the closest datasets. In this study, the nearest neighbor algorithm is used to determine the similarity between problems.

The nearest neighbor algorithm is likely to be shown as follows:

$$S(I,S) = \frac{\sum_{t=1}^{n} W_t \, x \;\; \mathrm{sim}(f_t^{input}\,, f_t^{stored}\,)}{\sum_{t=1}^{n} W_t}$$

Formula 4 *– Similarity Function*

In this function (Formula 4), S denotes the similarity score, Wi denotes the feature weight, n denotes the feature number, input denotes the feature value of input case, and history denotes the feature value of stored case-base. The S represents the sum of the similarity, w is the importance of the feature, and the sim is similarity assessing function for comparing feature, f_i, f_i^I, and f_i^Rvalues for feature $f_i$in the input new case and old case.

Similarity rate is calculated using this formula with different properties. Formula 4 presents our formula for calculating the similarity coefficient of our prescription recommendation system. Cinar also used the same similarity function in his study to make a correlation between the solution and problem [48]. This algorithm yields a result which is between 0 and 1. Therefore, the results of this calculation can be represented as a percentage as shown in Formula 5.

$$S = [0,1]\ or\ S(\%) = [0,1]\ x\ 100$$

Formula 5 *– Similarity Function in Percentage*

This above formula set out to investigate the similarity between prescription and patient, and the most similar result is retrieved according to the patient, diagnosis, and prescription features. In our case, the weight of each feature was considered equal. One primary importance of this approach is that the similarity of each feature is attended to the sum. The final five results are displayed to the users concerning their rate from high to low. The users can be picked among them.

3.3. Test Environment of Prescription Recommendation System

The literature has highlighted several various applications of CBR. For example, CBR is used for problem-solving (e.g., for design, for planning, for diagnosis, for explanation), and interpretation (e.g., justification and adversarial reasoning, classification and interpretation, projecting effects) [2]. In this study, we applied CBR for the problem-solving purpose to find the closest prescription by ordering them according to similarity rate for a specific diagnosis. In other words, our CBR recommendation system gives the most accurate prescription to physicians for their patients by using previous prescriptions in the database. The primary purpose of this application is to find out the similarity and relevance between the prescriptions and the drugs or compounds they contain. This application uses patient demographic information (gender and age) and medical diagnosis to recommend a most similar prescription to the physician during the examination. The whole system hosted at Apache server. The application implemented with Java. Bootstrap used for user interface implementation. Moreover, in order to store data, MySQL database is used.

Java – Model-View-Controller (MVC)

Java is an object-oriented programming language released by Sun Microsystems. Java has lots of advantages over the other languages such as easy to learn, compile, object-oriented, secure and platform independent. Hoff studied the Java programming language and indicated Java as a scalable, robust and high-performance tool [49]. Robustness, ease of use, the independent platform, security features make the Java eligible from the other programming languages. With Java, we use MVC as a pattern to develop our application. MVC divided into three classes as Model, View, and Controller. The view is split off from the structure and responsible for the representation, Controller implemented separately and responsible for rendering a middle-ware to model, the model is responsible for communication with data structure [50]. Component separation makes the code more comfortable to use and re-usable from different layers. Concerning these advantages, we decided to use Java and MVC in our application.

User Interface Development

Bootstrap is an open-source tool which contains interface components for creating responsive websites. Bootstrap includes HTML, CSS and JavaScript files; additionally, use fewer style-sheets for CSS and min JavaScript files that aimed to ease of web-development. Today, the usage of the wide screen comes with problems as well. Harb et al. indicated that problems and presented "Responsive web design" for different size of screens. He stated that, mobile and desktop resolution, resources, speed and needs precisely different from each other so that the usage of responsive web design unavoidable for the developer [51]. Natka demonstrated that growth in Internet technologies and increasing consume on technological smart devices involve responding to every size of a screen of a device [52]. Optimal user experience needs to adopt responsive web design technology so that bootstrap is used. Various size of screen resolution is an essential factor in designing interface. Avoiding such a problematic case, we implemented with bootstrap. The benefits of using Bootstrap in our application can be listed as:

– Higher user experience,

– Ease of use,

– Professional look and design.

Database

MySQL is the predominantly used open-source relational database system. Many websites and applications run on MySQL database as a consequence of free and prosperous open-source. MySQL may be divided into some crucial features such as;

– Scalability,

– High availability,

– Open-sourced,

– High performance,

– Easy management.

Open-source is an essential factor in why we have chosen the MySQL. Database diagram is shown in Figure 5.

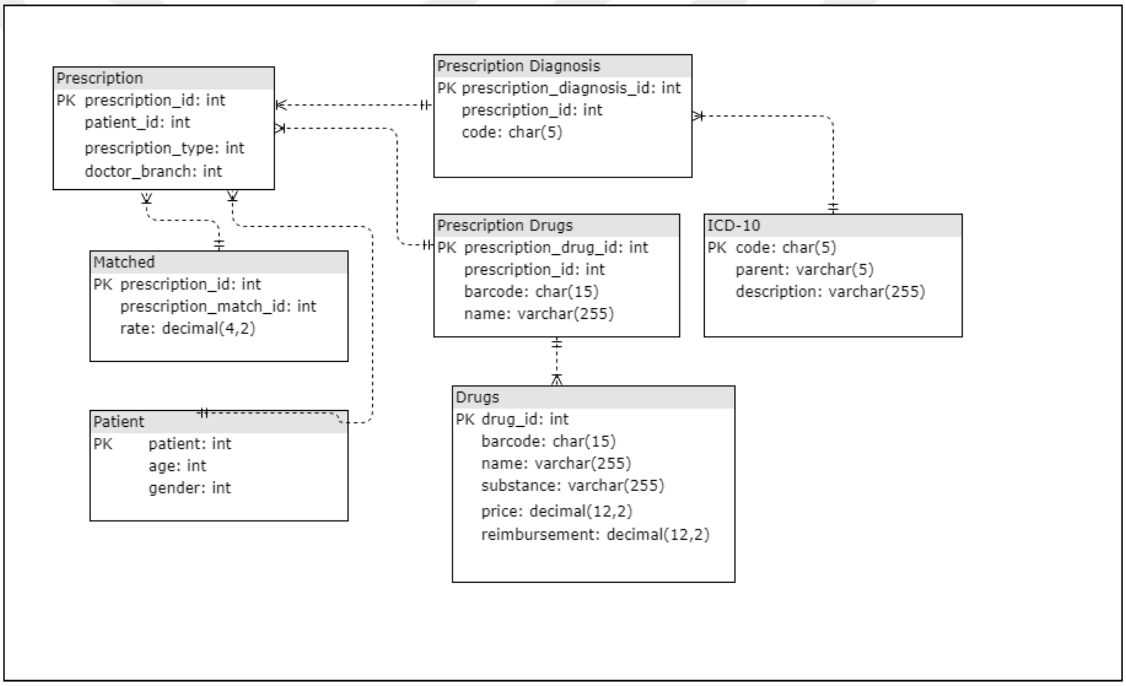

***Figure 5** – Database Diagram*

3.4. System Architecture

The system architecture explains the database and application structure, modules created for the prescription recommendation system. Our approaches during the implementation in this study will be discussed.

3.4.1. Database Structure

USERS: This table store user information includes name, surname, mail, user type. The ID is the auto-incremented primary key which differentiates users from each other, also ID set as a primary key. This table is related to the corresponding tables.

DRUGS: In this table, we stored drug information respectively drug id, barcode, name, price, public reimbursement condition, equivalent code, active pharmaceutical ingredient information. Also, drug id is an auto-incremented primary key.

PATIENT: In this table, we stored the patient's information that contains age and gender. Also, we stored data anonymously, and a relationship has been set with auto incremented patient id.

PRESCRIPTION: This table designed to store prescriptions anonymously. The prescription id is an auto-incremented primary key. This table consists of a patient id as a foreign key from the PATIENT table and has "one to many" relationship with corresponding tables.

PRESCRIPTION DRUGS: This table is designed to store prescription drugs. The prescription could contain one or many drugs, thereby a junction table created to put a relationship between prescriptions.

PRESCRIPTION ICD: This table designed to store prescription diagnosis codes. The prescription could contain one or many icd-10 codes, so that junction table created to put a relationship between prescriptions just as prescription drugs.

The database is designed to handle the whole transactions, procedures, and functions. Besides, all calculations and methods are stored in the database. Prescriptions, icd-10 codes, and drugs are designed to store new and previous data. Structure of our database design is very critical for our system and all data stored anonymously.

3.4.2. Application Structure

Login Page-Dashboard:

Users must log into the system to use the prescription recommendation system (see Appendix A.1). The dashboard page navigates the user to related pages.

Recommendation Page:

The prescription recommendation page is a service for recommending prescription according to the similarity rate. This page provides a variety of different information for analyzing similarity to recommend a prescription by using our algorithm. The users should enter the required areas which are respectively;

– Icd-10
– Physician branch
– Age
– Gender
– Prescription Type

After filling the required areas, the information sends to the database and waits for responding. Each feature affects the similarity percentage. Moreover, results include drugs and pharmaceutical ingredient information, gender, branch and prescription type. Additionally, the similarity percentage ordered descending order. Lack of patient information may cause mistaken results. Several factors that affect the calculations in determining the recommendation are the patient information, diagnosis, drugs, branch information. The responses relating to the percentage of similarity comparison is gathered by using these metric calculations. The results obtained from the analysis of similarity comparison are listed in descending order. Once the closest five results are listed on the user's screen, the user might be pick among these results (see Appendix A.2).

3.5. Test Methodology of Prescription Recommendation System

In this chapter, we describe the test methodology of prescription recommendation system developed in this study. Figure 6 shows the pseudo-code of how to evaluate the calculating success rate of recommendation with pseudocode at the bottom.

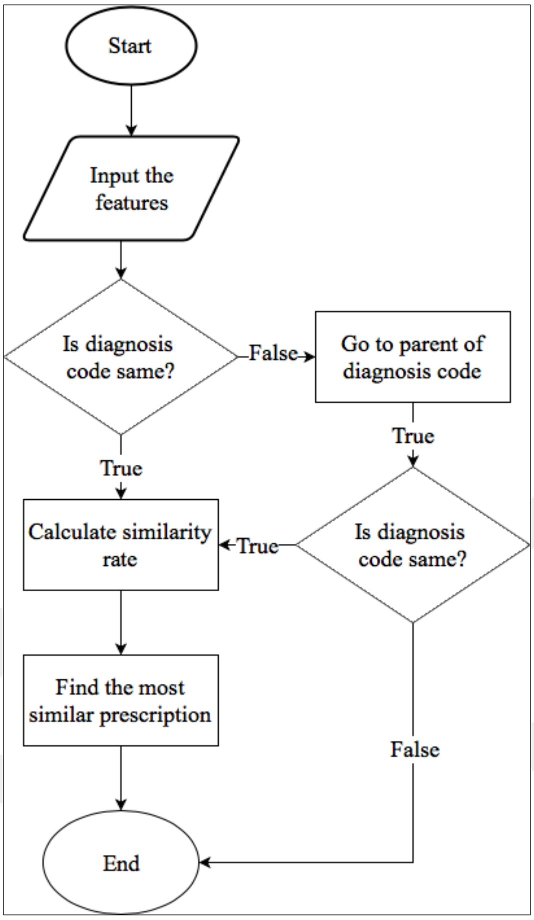

Figure 6 - *Pseudo-code of Similarity Analysis*

Two different test methodologies have been applied to the prescription recommendation system. Firstly, our recommendation system was used by the physicians for the comparison of their prescriptions. Then, it was applied to compare the previous prescriptions in the system with the latest prescriptions. Next, the comparison of the prescriptions of the physicians will be explained.

3.5.1. Calculating the Success Rate Regarding the Physicians' Prescriptions

In order to test our application, we contact with physicians. Of twenty-six physicians who were sent invitations, thirteen accepted to involve our testing process. We asked them to use the system for five different patients during the test process. By this way, the degree of success of the prescription recommendation system was evaluated.

Table 4: *Branches of physicians who participate in the testing process*

Branch	Number of Participants
Orthopedics and Traumatology	3
Gynecology	3
Cardiology	2
Practitioner	2
Urology	1
Gastroenterology	1
Family Physician	1
TOTAL	**13**

Physicians from a variety of branches were involved in the test process, namely, orthopedics and traumatology, gynecology, cardiology, urology, gastroenterology, family physician and practitioner. Table 4 presents the distribution of physicians' branches who participated in the study.

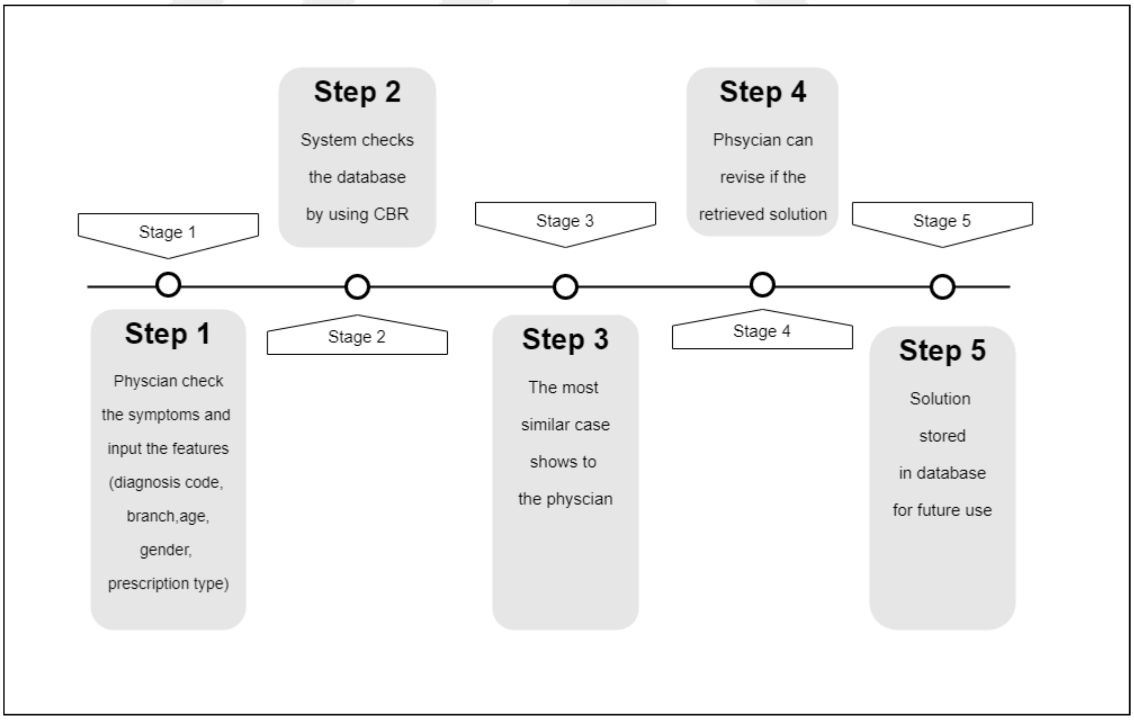

***Figure** 7 – Steps of Testing Process with the Physician*

As can be seen in Figure 7, the general pattern of the methods we used in the testing application starts with the physician's medical treatment. After the physician checks the symptoms and determines which disease is proper for the patient's symptoms, the medical diagnosis is decided. The system is ready to recommend the prescription for medical treatment. Then, we asked the physician to input the diagnosis code (in icd-10 format), brief patient information which are birth-year, gender, and prescription-type information. The application obtains the physician diagnosis, patient and prescription information using a web form.

After the required inputs and other inputs fulfilled, the system provides the most similar case by incorporating the nearest neighbor algorithm in the retrieval process in the CBR. The list appeared similarity rate from highest rate to lowest rate and similar properties highlighted on the table. The results obtained from the preliminary evaluation of similarity are shown in web application.

Table 5: *Example Comparison for CBR Recommendation System*

Case to be compared		Case in the database	
Feature	Value	Value	Weight
ICD-10	Meniscus-tear	Meniscus-tear	1
Gender	Male	Male	1
Presc. Type	White	White	1
Year	1988	1998	0.66
Branch	Orthopedics and Traumatology	Orthopedics and Traumatology	1

Table 5 can be shown as an example of calculating throughout the comparison. The value on the left side represents the data entered by the user, and on the right side, compared case in the database is displayed. Users need to select one of the listed prescriptions. If the user does not find the desired prescription on the list, they need to refine the properties. When a similar one is found on the list, the user clicks the "Select" button, and modal appears on the screen. Same drugs listed on the screen but sometimes physicians want to change drugs by some reasons so that this screen provides them an opportunity to change the drugs. The provided pieces of information save into the database for future solutions. One of the other noticeable features of the application is proving a drug reimbursement and active ingredients information to the users.

If the user detects an inappropriate recommendation from the listed prescriptions, he/she should press the out of use button. In this way, it is ensured that if the system offers an inappropriate recommendation, it will be dismissed next time. Finally, we removed the gender, age, physician branch, and prescription-type feature one by one, and calculate the recommendation regarding the remaining features.

3.5.2. Calculating the Success Rate Regarding Existing Prescriptions

In order to compare the prescriptions among them, we first determined the 50 most common diseases from our database. Then, the diagnosis is sorted in descending order concerning their number of prescriptions. Table 6 presents the number of prescriptions according to the diagnosis in descending order.

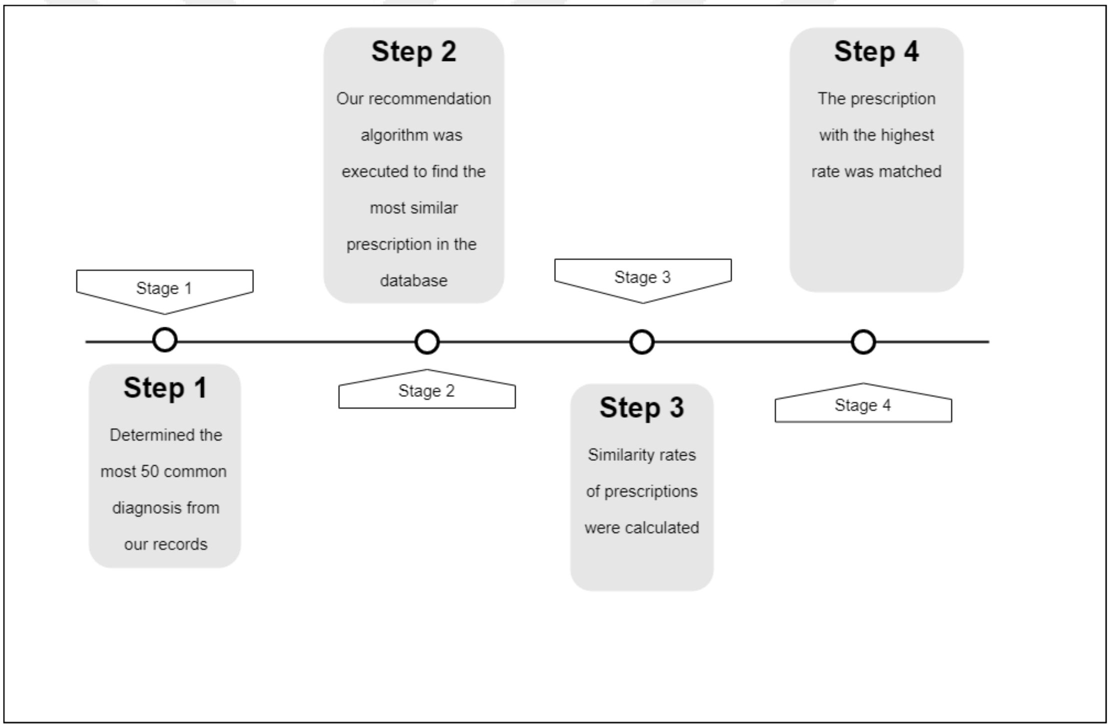

Figure 8 *– Steps of Testing Process with the Existing Prescriptions*

As can be seen in Figure 8, at first, the last prescription was taken as an input case to compare with other prescriptions. These prescriptions include the following information: patient id, prescription, physician and icd-10 information. In order to compare these prescriptions, our recommendation algorithm was executed to find the most similar prescription in the database. The system has ignored the last prescriptions to be compared in order to prevent the same prescriptions from being received as a result. For this comparison, the branch, age, gender, prescription type, and icd-10 code information of the selected prescription were used. Our recommendations system was executed for each of these prescriptions and expected to make a recommendation for each of them. The similarity rates of prescriptions were calculated, and the prescription with highest similarity rate was matched with the compared prescription as a result.

Table 6: *Top 5 Prescription Counts According to Diagnosis*

ICD-10 Code	Prescription Number
I10 (Hypertension)	250
J02 (Streptococcal pharyngitis)	198
M13 (Polyarthritis)	132
K21 (Gastro-esophageal reflux)	121
E11 (Type 2 diabetes mellitus)	112

The value on the left side represents the inputs for selected prescription among the top 50 most common diseases from our database. The right side, prescription with the highest similarity given by the system is displayed. Lastly, as similar to the previous comparison, we removed the gender, age, physician branch, and prescription-type features one by one, and calculate the recommendation regarding the remaining features.

3.5.3. Calculating the Similarity of the Drugs Between Recommended Prescriptions and Written Prescriptions by Physicians

In order to calculate the similarity of the drugs between the recommended prescription and written prescription by physicians or selected prescription by the system, we compare the prescriptions with each other. The similarity rate was calculated by comparing each drug in the recommended prescription and written prescription by physicians or selected prescription by the system as follows:

(1) Give a value of 1, if they are the same and give 0 otherwise,

(2) Sum the scores,

(3) Divide the sum by the maximum number of drugs in the prescriptions.

Table 7 presents the example calculation of the similarity of the drugs between recommended prescriptions and written prescriptions by physicians or selected prescriptions by system.

Table 7: *Example Calculation of the Drugs Similarity*

Recommended Prescription		Written/ Selected Prescription	Evaluation Score
1st Drug	APRANAX PLUS FILM TABLET 20	APRANAX PLUS FILM TABLET 20	1
2nd Drug	VOLTAREN SUPOZITUVAR 100 mg 5	VOLTAREN SUPOZITUVAR 100 mg 5	1
3rd Drug	RANTUDIL FORTE KAPSUL 60 mg 20	MUSCOFLEX CREAM %0.25 30g	0
		Similarity	**0.67**

3.6. Chapter Summary

In this chapter, the methodology used in this thesis is explained. Besides, the test environment and the test methodology detailed. Following chapter, we are going to explain the results of the application that applied to the healthcare professionals.

CHAPTER 4

RESULTS

We calculated the success rate of our CBR system concerning (1) the physicians' real-life prescriptions for a disease and the prescription recommended by the system regarding the previous prescriptions, and (2) the last prescription for a disease and the prescription recommended by the system regarding the previous prescriptions. For the first testing, our recommendation system has been tested with more than 65 patients. For the second test, we first found the most 50 common diseases and then, the last prescriptions for these diseases were found, and the system was tested against for each disease found. For each test process, we also investigated the effects on features on recommendation success rate. For example, we removed the gender, age, physician branch, and prescription-type feature one by one, and calculate the recommendation regarding the remaining features. This approach allowed us to examine the effects of each feature on our recommendation success rate. Moreover, we also investigated the similarity rate of the drugs between the prescription recommended by the system and prescription written by the physician. The test results are presented below.

4.1. The success rate of the CBR system concerning the physicians' real-life prescriptions for a disease and the prescription recommended by the system regarding the previous prescriptions

The application, which was used by the physician from 7 different branches, was run with 47 different icd-10 codes. The success rate of the CBR system (as presented in Table 7) is calculated by finding the similarity rate of the prescriptions written by physicians and the prescriptions recommended by the system.

4.1.1 Success rate of the CBR system by using all case features including age, gender, branch, and prescription-type as case features

Table 8: *The success rate of the CBR system concerning the prescriptions written by physicians and prescriptions recommended by the system*

Physician	Prescription 1 Success Rate	Prescription 2 Success Rate	Prescription 3 Success Rate	Prescription 4 Success Rate	Prescription 5 Success Rate	Average Success Rate
1	0.71	0.70	0.75	0.70	0.71	0.71
2	0.66	0.68	0.72	0.72	0.73	0.70
3	0.87	0.74	0.93	0.72	0.91	0.83
4	0.86	0.72	0.71	0.85	0.90	0.80
5	0.75	0.74	0.87	0.75	1.00	0.82
6	0.88	0.92	0.86	0.87	0.89	0.88
7	0.91	0.96	0.87	0.95	0.95	0.92
8	0.88	0.81	0.89	0.93	0.86	0.87
9	0.75	0.75	0.69	0.63	0.74	0.71
10	0.70	0.93	0.48	0.43	0.48	0.60
11	0.69	0.71	0.67	0.63	0.86	0.71
12	0.85	0.87	0.82	0.74	0.70	0.79
13	0.69	0.94	0.68	0.90	0.88	0.81
					Average	**0.78**

The average similarity rate of our recommendation system is 0.78. This result also shows the success rate of our CBR system. A total of 114 drugs are prescribed for 65 prescriptions. Of the prescriptions, 31 of them are above average. The remaining 34 prescriptions are matched below the average recommendation rate. The most frequent rates are founded as 0.72, 0.87 and 0.75 in our result set, respectively. One complete paired prescription is found through the testing process.

The comparison results of prescriptions written by physicians and prescriptions recommended by the system obtained from the application are presented in Table 8. Table 9 presents the results of the average success rates of prescriptions by grouping them according to their branches.

Table 9: *Comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system grouped by physician's branches*

Physician Branch	Total Average
Gynecology	0.90
Family Physician	0.82
Urology	0.82
Orthopedics and Traumatology	0.78
Cardiology	0.76
Gastroenterology	0.71
Practitioner	0.66

The results, as shown in Table 9, indicates that similarity percentage in decimal format. These results show the similarity averages of the branches obtained by physicians. According to the comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system, the most successful branch is Gynecology. The similar prescription is recommended at 0.90. A family physician with 0.82 follows this similarity rate. The other two branches are above average, while the other two branches were below average. The lowest rate of similarity is 0.66 in the practitioner.

4.1.2. Success rate of the CBR system by removing age and using gender, branch, and prescription-type as case features

Table 10 indicates that comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system after removing the age feature from the calculations.

Table 10: *Comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system after removing the age feature*

Physician	Prescription 1 Success Rate	Prescription 2 Success Rate	Prescription 3 Success Rate	Prescription 4 Success Rate	Prescription 5 Success Rate	Average Success Rate
1	0,66	0,66	0,66	0,66	0,66	0,66
2	0,66	0,66	0,66	0,66	0,66	0,66
3	1,00	0,66	1,00	0,66	1,00	0,86
4	1,00	0,66	0,66	1,00	1,00	0,86
5	0,66	0,66	1,00	0,66	1,00	0,78
6	1,00	1,00	1,00	1,00	1,00	1,00
7	1,00	1,00	1,00	1,00	1,00	1,00
8	1,00	1,00	1,00	1,00	1,00	1,00
9	0,66	0,66	0,66	0,66	0,66	0,66
10	0,66	1,00	0,33	0,33	0,33	0,53
11	0,66	0,66	0,66	0,66	1,00	0,72
12	1,00	1,00	1,00	0,66	0,66	0,86
13	0,66	0,66	0,66	1,00	1,00	0,86
					Average	**0,80**

Without age feature, the average similarity rate of our recommendation system is 0.80. This result also shows the success rate of our CBR system without age feature. The most frequent rate is found to be 0.66 (occurred 32 times out of 65). The results also demonstrate that there are 29 one-to-one matches. Lastly, 0.33 appears four times in the results. As can be seen from the data in Table 10, the success rate is increased to 0.80.

4.1.3. Success rate of the CBR system by removing gender and using age, branch, and prescription-type as case features

Table 11 indicates that comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system after removing the gender feature from the calculations.

***Table 11:** Comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system after removing the gender feature*

Physician	Prescription 1 Success Rate	Prescription 2 Success Rate	Prescription 3 Success Rate	Prescription 4 Success Rate	Prescription 5 Success Rate	Average Success Rate
1	0,96	0,61	1,00	0,94	0,96	0,82
2	0,89	0,91	0,97	0,96	0,98	0,94
3	0,84	0,99	0,91	0,96	0,89	0,91
4	0,82	0,97	0,95	0,81	0,87	0,88
5	0,67	0,66	0,84	0,67	1,00	0,76
6	0,85	0,90	0,82	0,83	0,86	0,85
7	0,88	0,95	0,83	0,94	0,94	0,90
8	0,85	0,76	0,86	0,91	0,81	0,83
9	0,67	0,67	0,60	0,51	0,86	0,62
10	0,61	0,91	0,64	0,58	0,65	0,67
11	0,60	0,96	0,90	0,85	0,82	0,82
12	0,81	0,83	0,77	0,99	0,94	0,86
13	0,59	0,59	0,91	0,87	0,85	0,76
					Average	**0,82**

After the gender feature removed, the average similarity rate of our recommendation system is 0.82. This result also shows the success rate of our CBR system without gender feature. The most frequent rate is found to be 0.96 (occurred 5 times out of 65). The results show that there are two one-to-one matches. Of 23 results are matched below the average while 42 of them are above. As can be seen from the data in Table 11, the success rate is increased to 0.82 percent concerning the original dataset.

4.1.4. Success rate of the CBR system by removing the branch and using age, gender, and prescription-type as case features

Table 12 indicates that comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system after removing the branch feature from the calculations.

Table 12: *Comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system after removing the branch feature*

Physician	Prescription 1 Success Rate	Prescription 2 Success Rate	Prescription 3 Success Rate	Prescription 4 Success Rate	Prescription 5 Success Rate	Average Success Rate
1	0,62	0,94	0,67	0,60	0,62	0,69
2	0,55	0,57	0,64	0,63	0,65	0,60
3	0,84	0,66	0,91	0,63	0,89	0,78
4	0,82	0,63	0,62	0,81	0,87	0,75
5	1,00	0,99	0,84	1,00	1,00	0,96
6	0,85	0,90	0,82	0,83	0,86	0,85
7	0,88	0,95	0,83	0,94	0,94	0,90
8	0,85	0,76	0,86	0,91	0,81	0,83
9	1,00	1,00	0,93	0,84	0,99	0,95
10	0,94	0,91	0,64	0,58	0,65	0,74
11	0,93	0,62	0,57	0,52	0,82	0,69
12	0,81	0,83	0,77	0,66	0,61	0,73
13	0,59	0,92	0,58	0,87	0,85	0,76
					Average	**0,79**

Table 12 presents the comparison results obtained by after removing the branch feature from our dataset. The average similarity rate of our recommendation system is 0.79. The most frequent rates are found to be as 1.00, 0.94, 0.62 (occurred 4 times out of 65), respectively. The results indicate that there are five one-to-one matches. Of 20 results are below the average and, the rest of them are above. The overall average decreased to 0.79.

4.1.5. Success rate of the CBR system by removing prescription type and using age, gender, and branch as case features

Table 13 indicates that comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system after removing the prescription type feature from the calculations.

Table 13: *Comparison results of prescriptions prescribed by physicians and prescriptions recommended by the system after removing the prescription type feature*

Physician	Prescription 1 Success Rate	Prescription 2 Success Rate	Prescription 3 Success Rate	Prescription 4 Success Rate	Prescription 5 Success Rate	Average Success Rate
1	0,62	0,61	0,67	0,60	0,62	0,62
2	0,55	0,57	0,64	0,63	0,65	0,60
3	0,84	0,66	0,91	0,63	0,89	0,78
4	0,82	0,63	0,62	0,81	0,87	0,75
5	0,67	0,66	0,84	0,67	1,00	0,76
6	0,85	0,90	0,82	0,83	0,86	0,85
7	0,88	0,95	0,83	0,94	0,94	0,90
8	0,85	0,76	0,86	0,91	0,81	0,83
9	0,67	0,67	0,60	0,51	0,66	0,62
10	0,61	0,91	0,31	0,25	0,32	0,48
11	0,60	0,62	0,57	0,52	0,82	0,62
12	0,81	0,83	0,77	0,66	0,61	0,73
13	0,93	0,59	0,58	0,87	0,85	0,76
					Average	**0,72**

Table 13 presents the results obtained by after removing the prescription type feature from our dataset. As can be seen from Table 13, the average similarity rate of our recommendation system is 0.72. The most frequent rate is found to be 0.67 (occurred 5 times out of 65). The results also demonstrate that there are one one-to-one matches. From the result set, of 34 results are below the average and, the rest of them are above the average. The overall average decreased to 0.72.

4.1.6. Similarity of the drugs between recommended prescriptions and written prescriptions by the physicians

Table 14 presents the comparison results of the drugs in the prescriptions written by physicians and prescriptions recommended by the system.

Table 14: Comparison results of prescriptions drugs prescribed by physicians and prescriptions recommended by the system

Physician	Prescription 1 Drugs Comparison Rate	Prescription 2 Drugs Comparison Rate	Prescription 3 Drugs Comparison Rate	Prescription 4 Drugs Comparison Rate	Prescription 5 Drugs Comparison Rate	Average Comparison Rate
1	0,33	0,67	1,00	0,50	0,67	0,63
2	1,00	0,50	0,67	0,67	1,00	0,77
3	0,33	0,33	1,00	0,50	0,75	0,58
4	0,67	0,50	0,50	0,33	1,00	0,60
5	0,50	0,33	0,67	1,00	1,00	0,70
6	0,67	0,50	0,50	0,33	1,00	0,60
7	0,75	0,33	0,33	0,67	1,00	0,62
8	1,00	1,00	0,50	0,33	0,33	0,63
9	0,50	0,33	0,50	1,00	0,67	0,60
10	0,67	0,33	1,00	0,50	1,00	0,70
11	1,00	0,50	1,00	0,50	0,67	0,73
12	0,75	0,50	0,33	0,67	0,33	0,52
13	1,00	0,33	0,33	1,00	0,50	0,63
					Average	**0,63**

Table 14 presents the results obtained by comparing the drugs in the prescriptions in our dataset. As can be seen from Table 14, the average similarity rate of drugs recommended by system is 0.63. The most frequent rates are found to be 1 (occurred 18 times out of 65), 0.33 and 0.50 (occurred 16 times out of 65). From the result set, of 33 results are below the average and, the rest of them are above the average. The overall average calculated as 0.63.

4.2. The success rate of the CBR system concerning the most recent prescription for a disease and the prescription recommended by the system regarding the previous prescriptions

The comparison results of the last prescription for a disease and the prescription recommended by the system presents us a different perspective. In this process, firstly, we find out the most 50 frequent icd-10 codes. Then, the last prescriptions for each diagnosis codes are found. The system is tested against each diagnosis code that we found in the previous step. Next, for each diagnosis code, we compare the features of the l prescription. From the set of possible prescriptions given by the system, the one with the highest similarity is chosen among the previous prescriptions.

4.2.1. Success rate of the CBR system by using all case features including age, gender, branch, and prescription-type as case features

Table 15: *Comparison results of the most recent prescription for a disease and the prescription recommended by the system*

ICD-10	Average Success Rate
R52, J45, K59.0, D51, R07.0, N39.0, J39.9, F41.1, H10, A08, M79.7, M62, J02, J30.2, J01, R42, R11, I10, M65.0, I25.1, F33, J20, M79.1, F32	0.95 – 1.00
I83, L30, M06, D64, G43, K21, F41, K21.9, R51, F41.9, K25, G20, D50, M54.5, T20–T, Z25.1	0.90 – 0.95
M13, I25, E04, I84, K60.0	0.85 – 0.90
J30, E11	0.80 – 0.85
L50	0.75 – 0.80
M19, E56	0.70 – 0.75
Average	**0.91**

The average similarity rate of our recommendation system is calculated as 0.91. According to the results obtained from the system, seven one-to-one matches found. The one-to-one matched codes are found to be as follows; R52, J45, K59, D51, R07, N39, J39. The most frequent rates are found to be 0.98 (occurred 8 times out of 50), 1.0 (occurred 7 times out of 100) and 0.92 (occurred 5 times out of 100). According to the results, 11 prescriptions selected by the system is below average. The least frequent rate is found to be 0.88. The results obtained from this experiment, the success rate of the system is calculated as 0.91.

4.2.2. Success rate of the CBR system by removing age and using gender, branch, and prescription-type as case features

Table 16 indicates that comparison results of the most recent prescription for a disease and the prescription recommended by the system after removing the age feature from the calculations.

***Table 16:** Comparison results of the most recent prescription for a disease and the prescription recommended by the system after removing age feature*

ICD-10	Average Success Rate
L30, M79.7, Z25.1, J30–J, F41, R07.0, M65.0, T20–T, A08, K21.9, R11, K25, H10, K60.0, F32, K59.0, M79.1, M62, I83, R52, F33, F41.1, G43, J20, M13, I25.1, K21, J02, M06, N39.0, J01, J30.2, D64, R51, J39.9, R42, D50, D51, E04, I10, I84, I25, G20, J45, F41.9, E11, M54.5	1
E56, M19, L50	0,66
Average	0.95

Table 16 presents the results obtained by after removing the age feature from our dataset. As can be seen from Table 16, the average similarity rate of our recommendation system is 0.95 after removing the age feature from our dataset. The most frequent rate is found to be 1.00 (occurred 47 times out of 50). The other similarity rate is found to be 0.66 (occurred three times out of 50).

4.2.3. Success rate of the CBR system by removing gender and using age, branch, and prescription-type as case features

Table 17 presents the comparison results of the most recent prescription for a disease and the prescription recommended by the system after removing the gender feature from the calculations.

***Table 17:** Comparison results of the most recent prescription for a disease and the prescription recommended by the system after removing gender feature*

ICD-10	Average Success Rate
M79.7, R07.0, M65.0, A08, R11, H10, F32, K59.0, M79.1, M62, R52, F33, F41.1, I25.1, J02, N39.0, J01, J30.2, J39.9, R42, D51, I10, J45	0.95 – 1.00
L30, F41, K21.9, I83, G43, J20, K21, M06, D64, R51, F41.9	0.90 – 0.95
Z25.1, T20–T, K25, M13, D50, E04, I25, G20, M54.5	0.85 – 0.90
K60.0, I84	0.80 – 0.85
J30, E11	0.70 – 0.80
M19, L50	0.60 – 0.70
E56	0.50 – 0.60
TOTAL	**0.91**

Table 17 presents the results obtained by after removing the gender feature from our dataset. The average similarity rate of our recommendation system is founded to be 0.91. The most frequent rates are found to be as 1.00 and 0.95. Of 23 results are above the average and, the rest of them are below. Overall, the overall average decreased to 0.91.

4.2.4. Success rate of the CBR system by removing the branch and using age, gender, and prescription-type as case features

Table 18 provides the comparison results of the most recent prescription for a disease and the prescription recommended by the system after removing the branch feature from the calculations.

Table 18: *Comparison results of the most recent prescription for a disease and the prescription recommended by the system after removed branch feature*

ICD-10	Average Success Rate
M79.7, R07.0, M65.0, A08, R11, H10, F32, K59.0, M79.1, M62, R52, F33, F41.1, I25.1, J02, N39.0, J01, J30.2, J39.9, R42, D51, I10, M19, J45, L50	0.95 – 1.00
L30, F41, K21.9, I83, G43, J20, K21, M06, D64, R51, F41.9	0.90 – 0.95
Z25.1, T20, K25, K60.0, M13, E56, D50, E04, I84, I25, G20, M54.5	0.85 – 0.90
K60.0, E56, I84	0.80 – 0.85
J30, E11	0.70 – 0.80
Average	**0.93**

Table 18 presents the results obtained by after removing the branch feature from our dataset. The average similarity rate of our recommendation system is 0.93. The most frequent rate is found to be range between 1.00 and 0.95. Of 25 results are below the average while 25 of them are above the average.

4.2.5. Success rate of the CBR system by removing prescription type and using age, gender, and branch as case features

Table 19 gives the comparison results of the most recent prescription for a disease and the prescription recommended by the system after removing the branch feature from the calculations.

Table 19: *Comparison results of the most recent prescription for a disease and the prescription recommended by the system after removing prescription type feature*

ICD-10	Average Success Rate
M79.7, R07.0, M65.0, A08, R11, H10, F32, K59.0, M79.1, M62, R52, F33, F41.1, I25.1, J02, N39.0, J01, J30.2, J39.9, R42, D51, I10, J45	0.95 – 1.00
L30, F41, K21.9, I83, G43, J20, K21, M06, D64, R51, F41.9	0.90 – 0.95
Z25.1, T20–T, K25, M13, D50, E04, I25, G20, M54.5	0.85 – 0.90
K60.0, I84	0.80 – 0.85
J30, E11	0.70 – 0.80
M19, L50	0.60 – 0.70
E56	0.50 – 0.60
Average	**0.91**

Table 19 presents the results obtained by after removing the prescription type from our dataset. The average similarity rate of our recommendation system is 0.91. The most frequent rates are found to be 1.00 - 0.95 and 0.95 - 0.90. Of 34 results are above the average and, 16 of them are below the average. Overall, the overall average decreased to 0.91.

4.2.6. Similarity of the drugs between recommended prescriptions and the selected prescriptions by the system

Table 20 gives the similarity comparison results of the drugs in the prescriptions between the most recent prescription for a disease and the prescription recommended by the system.

Table 20: *Comparison results of the most recent prescription for a disease and the prescription recommended by the system comparing with their drugs*

ICD-10	Average Comparison Rate
M79.7, R07.0, M65.0, A08, R11, H10, I83, G43, F32, K59.0, M79.1, M62, R52, F33, F41.1, I25.1, J02, N39.0, J01, J30.2, J39.9, R42, D51, I10, J45	0.75 – 1.00
L30, F41, K21.9, J20, K21, M06, D64, R51, F41.9	0.65– 0.75
Z25.1, T20–T, K25, M13, D50, E04, I25, G20, M54.5, J30, E11	0.5 – 0.65
K60.0, I84, M19, L50, E56	0.30 – 0.5
Average	**0.71**

Table 20 presents the results obtained by comparing the drugs in prescriptions in our dataset. The average similarity rate when comparing the drugs in the most recent prescription for a disease and the prescription recommended by the system is 0.71. The most frequent rate is found to be 1.00 - 0.75. Of 34 results are above the average and, 16 of them are below the average.

4.3. The success rate of the CBR system concerning the physicians' opinions about the recommended prescriptions based on a survey question

The application tested by the 13 physicians from 7 different branches. To calculate the performance evaluation of the system regarding the physicians' opinions about the recommended prescriptions, we ask them to grade the recommendations on a scale of 0-10 (0 indicates the lowest grade, and 10 indicates the highest grade). The results of the physicians' evaluations of the recommendations are presented in Table 21.

Table 21: *The Success rate of the CBR system concerning the physicians' opinions about the recommended prescriptions*

Physician	Evaluation score
1	8
2	7
3	9
4	8
5	8
6	8
7	9
8	8
9	7
10	9
11	8
12	8
13	7
Average	**8**

The average evaluation score of our recommendation system concerning physicians' opinions about the recommended prescriptions is positive that is 8 out of 10. The most frequent score that is given by seven physicians out of 13 is 8. The highest score that is given by three physicians out of 13 is 9. Ten responses are rated above or equal to the average while only 3 of them are below.

CHAPTER 5

CONCLUSIONS

The primary purpose of this study is developing a case-based recommendation system for prescription prediction. Through this application, we aimed to help health-care professionals while writing prescriptions. This thesis has presented the findings of applying the CBR system in the health domain. CBR is accepted as an artificial intelligence technique which helps to solve a problem by retrieving the most similar previous solution to the current problem. In real life, more than one factor can affect diagnosing disease and writing a prescription for that disease. However, one physician may not quantify every factor and may not formalize these factors efficiently. Our recommendation system may help the physicians by automatically consider the different factors during the prescription writing process. The results of this study indicate that health-care professionals can benefit from the recommendation system developed in this study.

The success rate of the CBR system concerning the physicians' real-life prescriptions for a disease and the prescription recommended by the system was 0.78. Moreover, the success rate of the CBR system concerning the last prescription for a disease and the prescription recommended by the system was 0.91. The difference between the results of the tests might emerge from (1) the test users choose only the first five preferences, (2) the limited number of sample records for the relevant diagnosis code in the database. This result showed that the average success rate of the system concerning the physicians' real-life prescriptions for a disease was low compared to the results concerning the last prescription for a disease. This result could be because of the physicians' preference was not always the first recommended prescription. One of the more significant findings to emerge from this study is that the prescriptions written by gynecology branch physicians have been identified as the most accurate recommendations. The reason why Gynecology has identified as the highest similarity rate is that it is only interested in women. One of our features in our recommendation algorithm, gender, was only the women in Gynecology. Therefore, gender has always been the same for this branch, and this might affect the results positively.

One of the other findings from the results, the age feature also affects the results. It was observed that the results were increased after the age feature was removed from the dataset the average result increased to 0.80 (from 0.78) and 0.95 (from 0.93). The results of physicians' real-life prescriptions for a disease were increased after gender feature was removed from the dataset the average result increased to 0.79 (from 0.78). On the other hand, it had not any positive effect on the success rate of the CBR system concerning the last prescription for a disease. The reason for this is that in the second comparison, there were more cases during the calculation. The presence of a few prescriptions in the same diagnostic code caused the gender data to have a negative effect in the first comparison. Removing branch affected the success rate of the CBR system concerning the physicians' real-life prescriptions for a disease positively. When the branch feature was removed from the dataset, the average result increased to 0.79 (from 0.78). The use of different branches prescriptions reduced the average. For example, for any diagnose code, the family physician may use the practitioner prescription. It was observed that the results of physicians' real-life prescriptions for a disease were increased after prescription type feature was removed from the dataset the average result decreased to 0.72 (from 0.78). On the other hand, it had not any effect on the success rate of the CBR system concerning the last prescription for a disease. As the prescription type feature differed only in specific diagnose code and drugs, it was found that it did not affect the second comparison. Finally, physicians in different branches have the opportunity to benefit from similar prescriptions in the diagnosis of other branches. In general, the use of CBR may help to prevent incorrect prescription writing and to help prescribe more accurate prescriptions.

When we compared our results with the previous related studies' results, we observed the consistency between our results and previous studies' results. For example, Rocha et al. [53] stated that the combination of Natural Language Processing and Case-Based Reasoning techniques create an opportunity to benefit from previous experiments. They The similarity rate of the system calculated by extracted information in text form of problems and solutions adopted by distributed software projects. Based on their experiment, the similarity rate is calculated at approximately 0.9. The success rates of similarity between the two studies are consistent. Although the cases in our study are quite small, the similarity rates are similar to the results of their study. Our success rates of similarity depending on our different comparison scenarios were between 0.72 – 0.91.

Cobb [54] studied Evolutionary Microelectromechanical System Design using Case-Based Reasoning. The primary purpose of this system is to revise the information in the database according to the existing problems. The success rate of the similarity retrieval method obtained from resonators test cases is 0.82. Also, they state that the initial success rate of similarity retrieval method is 0.3. This situation is similar to our study. The proliferation of cases increases the performance rate.

Janssen et. Al [55] conducted an experiment using Case-Based Reasoning for predicting the success of therapy. The authors state that the most suitable and effective methodology is Case-Based Reasoning for that study. The results show the effect of the multiplicity of cases and features. Moreover, for that research, the nearest neighbors' algorithm is used for calculating similarities between cases. The results are positively correlated with our results. The success rates are observed satisfactory in both experiments. They defined the CBR as a useful advisor.

Lamy et al. [56] define that "Case-Based Reasoning is a form of analogical reasoning in which solution for a new case is determined using previous cases with their solutions." They conducted a study in which they proposed a visual and explainable CBR system and the best classification accuracy obtained was

0.8. Also, they stated that the cases are very limited in this experiment and the increase of cases have a positive effect on the results.

Tabatabaee et al. [57] point out that the impact of the studies increases with the integration of the data series. Another research, conducted by Kiragu, states that The CBR is more accurate when the number of cases increases [58]. In our study, we also observed that increasing the quantity and quality of cases have a positive effect on the results. Furthermore, in medical cases, a large number of features make adaptation and generalization difficult in CBR use [59]. Thus, reliability cannot be guaranteed in medical CBR systems, even if the study has yielded accurate results in the used diagnosis [60].

Overall, by using the previous experience of the health-care professionals, our recommendation system can help to write correct prescriptions. At the same time, with our recommendation system, health-care professionals are supported to make faster and more accurate decisions during the prescription writing process.

CHAPTER 6

FUTURE WORKS

The research that has been undertaken for this thesis has contained a large scale of diagnosis code. Many different diagnosis codes have been tested by physicians and the system for calculating the success rate of our recommendation system. Since our CBR recommendation system has only been tested on a limited feature set, it will be worth testing with more features. In our study, we used five different features to recommend the most similar prescription. For the prescription recommendation system with using CBR, it is expected that the correct recommendations will be made by increasing these features and multiplying the data [55].

Further research would be useful to work on a single diagnosis code instead of general diagnosis codes. Because our recommendation system was worked on the general diagnosis codes, the features for specific diagnosis have been generalized. This specification would be useful to distinguish in very different diagnosis which is needed to make different examinations, and these specific features may need to be known for each diagnosis.

Finally, the recommendation system that we designed here can be enlarged with more diagnosis and prescriptions from other possible scenarios. When the cases getting enlarged, one needs to consider that these cases need to be fixed and maintenance for the recommendation.

REFERENCES

[1] R. C. Schank, Dynamic Memory: A Theory of Reminding and Learning in Computers and People. New York, NY, USA: Cambridge University Press, 1983.

[2] J. L. Kolodner, “An introduction to case-based reasoning,” Artif. Intell. Rev., vol. 6, no. 1, pp. 3–34, Mar. 1992.

[3] R. L. De Mantaras et al., “Retrieval, reuse, revision and retention in case-based reasoning,” Knowl. Eng. Rev., vol. 20, no. 3, pp. 215–240, 2005.

[4] D. B. Leake, “Problem Solving and Reasoning: Case-based,” in International Encyclopedia of the Social and Behavioral Sciences, N. J. Smelser and P. B. Baltes, Eds. Oxford: Pergamon, 2001, pp. 12117–12120.

[5] A. Aamodt and E. Plaza, “Case-based Reasoning: Foundational Issues, Methodological Variations, and System Approaches,” AI Commun., vol. 7, no. 1, pp. 39–59, 1994.

[6] T. Roth-Berghofer and I. Iglezakis, “Six Steps in Case-Based Reasoning: Towards a maintenance methodology for case-based reasoning systems,” in In: Professionelles Wissensmanagement: Erfahrungen und Visionen (includes the Proceedings of the 9th German Workshop on Case-Based Reasoning (GWCBR, 2001, pp. 198–208.

[7] A. Kyrilov, “Using Case-based Reasoning to Improve the Quality of Feedback Generated by Automated Grading Systems,” in Proceedings of the Tenth Annual Conference on International Computing Education Research, 2014, pp. 157–158.

[8] A. Kyrilov, “Using case-based reasoning to improve the quality of feedback generated by automated grading systems,” ICER 2014 - Proc. 10th Annual Int. Conf. Int. Comput. Educ. Res., 2014.

[9] Montaner, M. 2003. Collaborative recommender agents based on case-based reasoning and trust. Ph.D. thesis, Universitat de Girona, Girona, Spain.

[10] X. Blanco, S. Rodríguez, J. M. Corchado, and C. Zato, “Case-Based Reasoning Applied to Medical Diagnosis and Treatment,” in Distributed Computing and Artificial Intelligence, 2013, pp. 137–146.

[11] T. C. S. G. Kurumu, “Sosyal Güvenlik Kurumu 2018 Yılı Faaliyet Raporu”, 24 April 2019. [Online]. Available: http://www.sgk.gov.tr/wps/portal/sgk/tr/kurumsal/merkez-teskilati/danisma_birimleri/strateji_gelistirme_baskanligi/anasayfa_duyurular/duyuru_20190424_25. [Accessed 26 April 2019].

[12] W. H. Organization, “World Health Statistics 2017: Monitoring health for the SDGs”, 1 March 2019. [Online]. Available: https://www.who.int/gho/publications/world_health_statistics/2017/en/. [Accessed 26 April 2019].

[13] J. Kolodner, Case-based Reasoning. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993.

[14] D. B. Leake and J. Kendall-Morwick. Towards case-based sup- port for e-science workflow generation by mining provenance. In Advances in Case-Based Reasoning, ECCBR 2008, LNAI 5239, pp. 269–283. Springer, 2008

[15] R. C. Schank, J. L. Kolodner, and G. DeJong, “Conceptual Information Retrieval,” in Proceedings of the 3rd Annual ACM Conference on Research and Development in Information Retrieval, 1981, pp. 94–116.

[16] R. C. Schank and R. P. Abelson, “Scripts, Plans, and Knowledge”, in Proceedings of the 4th International Joint Conference on Artificial Intelligence - Volume 1, 1975, pp. 151–157.

[17] E. Plaza and L. McGinty, “Distributed Case-based Reasoning”, Knowl. Eng. Rev., vol. 20, no. 3, pp. 261–265, 2005.

[18] K. D. Ashley, “Case-based Reasoning and Its Implications for Legal Expert Systems”, Artif. Intell. Law, vol. 1, no. 2–3, pp. 113–208, 1992.

[19] P. Koton, “Using experience in learning and problem solving”, Massachusetts Institute of Technology, Cambridge, MA, USA, 1988.

[20] M. Richter and S. Wess, “Similarity, Uncertainty and Case-Based Reasoning in PATDEX”, vol. 1, pp. 261–265, 1999.

[21] R. L. de Mántaras and E. Plaza, Eds., “Machine Learning: (ECML) 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain, May 31 - June 2, 2000, Proceedings”, 2000, vol. 1810.

[22] C. Zhang, X. Wang, Y. Yu, L. Liu, and Y. Zhou, “Fault diagnosis method for complex equipment using cased-based reasoning”, in the 26th Chinese Control and Decision Conference (2014 CCDC), 2014, pp. 810–815.

[23] A. Lawanna and J. Wongwuttiwat, “Problem-based model for the improvement of case-based reasoning system”, in 2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2017, pp. 349–352.

[24] C. Feng, N. Yang, and T. Lei, “Research on Method of Emergency Cases Retrieval Based on Case-Based Reasoning”, in 2017 10th International Symposium on Computational Intelligence and Design (ISCID), 2017, vol. 1, pp. 83–86.

[25] A. Gatzioura and M. Sànchez-Marrè, “A case-based reasoning framework for music playlist recommendations”, in 2017 4th International Conference on Control, Decision and Information Technologies (CoDIT), 2017, pp. 242–247.

[26] C. Tsatsoulis, Q. Cheng, and H.-Y. Wei, “Integrating Case-Based Reasoning and Decision Theory”, IEEE Expert Intell. Syst. Their Appl., vol. 12, no. 4, pp. 46–55, 1997.

[27] Q. Zhang, G. Zhang, J. Lu, and D. Wu, “A Framework of Hybrid Recommender System for Personalized Clinical Prescription”, in 2015 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), 2015, pp. 189–195.

[28] H. Jin, E. Jang, and S. Lee, “Clinical network for case-based reasoning methodology”, in 2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW), 2011, pp. 956–957.

[29] C. Hsiao, C. Chuang, and J. Jeng, “A hybrid case base reasoning system for forecasting”, in 2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 2017, pp. 558–560.

[30] J. Gao, W. Xiao, Y. Xie, F. Gu, and B. Yao, “An intelligent fault diagnosis approach integrating cloud model and CBR”, in 2017 6th International Conference on Computer Science and Network Technology (ICCSNT), 2017, pp. 294–298.

[31] O. Chergui, A. Begdouri, and D. Groux-Leclet, “CBR approach for knowledge reuse in a community of practice for university students”, in 2016 4th IEEE International Colloquium on Information Science and Technology (CiSt), 2016, pp. 553–558.

[32] M. El Ajjouri, S. Benhadou, and H. Medromi, “New model of intrusion detection based on multi agent systems and CBR paradigm”, in 2016 4th IEEE International Colloquium on Information Science and Technology, 2016, pp. 133–138.

[33] X. P. Yu, Q. L. BeiHang, and X. Hu, “Aircraft fault diagnosis system research based on the combination of CBR and FTA”, in 2015 First International Conference on Reliability Systems Engineering (ICRSE), 2015, pp. 1–6.

[34] X. Deng, R. Luo, and J. Li, “Similarity matching algorithm of equipment fault diagnosis based on CBR”, in 2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS), 2015, pp. 998–1002.

[35] J. Ashraf and P. S. Mohod, “Literature review on CBR approaches for matching composite sketches to facial photographs”, in 2015 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC), 2015, pp. 189–194.

[36] S. Y. Sharma, D. P. Roy, and B. Chakraborty, “Health Services Support System (HESS): An Ontology supported CBR system for cognitive diagnosis

applied on HIV/AIDS detection", in 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), 2015, pp. 284–290.

[37] S. El-Sappagh, S. El-Masri, M. Elmogy, and A. M. Riad, "A diabetes diagnostic domain ontology for CBR system from the conceptual model of SNOMED CT", in 2014 International Conference on Engineering and Technology (ICET), 2014, pp. 1–7.

[38] J. Karim, M. Aufaure, R. Chiky, and M. Manceny, "A generic ontology for CBR-based recommender systems", in 2014 International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM), 2014, pp. 1–5.

[39] D. Xiaopeng and L. Donghui, "Research on Rice Pests and Diseases Warning Based on CBR", in 2013 International Conference on Computational and Information Sciences, 2013, pp. 1740–1742.

[40] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, and M. Folke, "Case-Based Reasoning Systems in the Health Sciences: A Survey of Recent Trends and Developments", Trans. Sys. Man Cyber Part C, vol. 41, no. 4, pp. 421–434, 2011.
[41] W. H. Organization, "ICD-10 Codes." 2016.

[42] Richter, M.M., Weber, R.O.: Case-Based Reasoning - A Textbook. Springer (2013).

[43] K.-D. Althoff, "Evaluating Case-Based Reasoning Systems: The INRECA Case Study", 1997.

[44] S. Wess, "Fallbasiertes Problemlösen in wissensbasierten Systemen zur Entscheidung-sunterstützung und Diagnostik", 1995.

[45] C. von Wangenheim and V. Wangenheim, "Case-Based Reasoning – A Short Introduction", p. 5, 2019.

[46] D. W. Aha and R. L. Goldstone, "Concept Learning and Flexible Weighting", in In Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society, 1992, pp. 534–539.

[47] Z. Zhang and Q. Yang, “Feature Weight Maintenance in Case Bases Using Introspective Learning”, J. Intell. Inf. Syst., vol. 16, pp. 95–116, 2001.

[48] A. K. Çınar, 2009, “INTEGRATING CASE BASED REASONING AND GEOGRAPHIC INFORMATION SYSTEMS IN A PLANNING SUPPORT SYSTEM: ÇESME PENINSULA STUDY”, Ph.D. thesis, Izmir Institute of Technology, Izmir, Turkey.

[49] A. V. A. N. HOFF, “THE CASE FOR JAVA AS A PROGRAMMING LANGUAGE”, pp. 51-56, 1997.

[50] Y. Q. Hamid Mcheick, “DEPENDENCY OF COMPONENTS IN MVC DISTRIBUTED ARCHITECTURE”, vol. 691, 2011.

[51] S. L. Eva Harb Paul Kapellari, “Responsive web design”, pp. 1–38, 2011.

[52] K. V Natda, “Responsive web design”, vol. 1, pp. 1–14, 2011.

[53] R. G. C. Rocha, R. R. Azevedo, Y. C. Sousa, E. de A. Tavares, and S. Meira, “A Case-based Reasoning System to Support the Global Software Development”, Procedia Comput. Sci., vol. 35, pp. 194–202, 2014.

[54] C. Cobb and A. Agogino, “Case-Based Reasoning for Evolutionary MEMS Design”, J. Comput. Inf. Sci. Eng., vol. 10, 2010.

[55] R. Janssen, P. Spronck, and A. Arntz, “Case-based reasoning for predicting the success of therapy”, Expert Syst., vol. 32, no. 2, pp. 165–177, 2015.

[56] J.-B. Lamy, B. Sekar, G. Guezennec, J. Bouaud, and B. Séroussi, “Explainable artificial intelligence for breast cancer: A visual case-based reasoning approach”, Artif. Intell. Med., vol. 94, pp. 42–53, 2019.

[57] H. Tabatabaee, H. Fadaeiyan, A. Alipour, and M. R. Baghaeipour, “Using Case-Based Reasoning for Diagnosis in Medical Field”, 2016.

[58] M. K. Kiragu and P. W. Waiganjo, “Case based Reasoning for Treatment and Management of Diabetes”, 2016.

[59] M. Nilsson and M. Sollenborn, “Advancements and Trends in Medical Case-Based Reasoning: An Overview of Systems and System Development.”, 2004, pp. 178–183.

[60] N. Choudhury and S. A. Begum, “A Survey on Case-based Reasoning in Medicine”, Int. J. Adv. Comput. Sci. Appl., vol. 7, no. 8, 2016.

APPENDICES

Appendix A. Screenshots of the Software

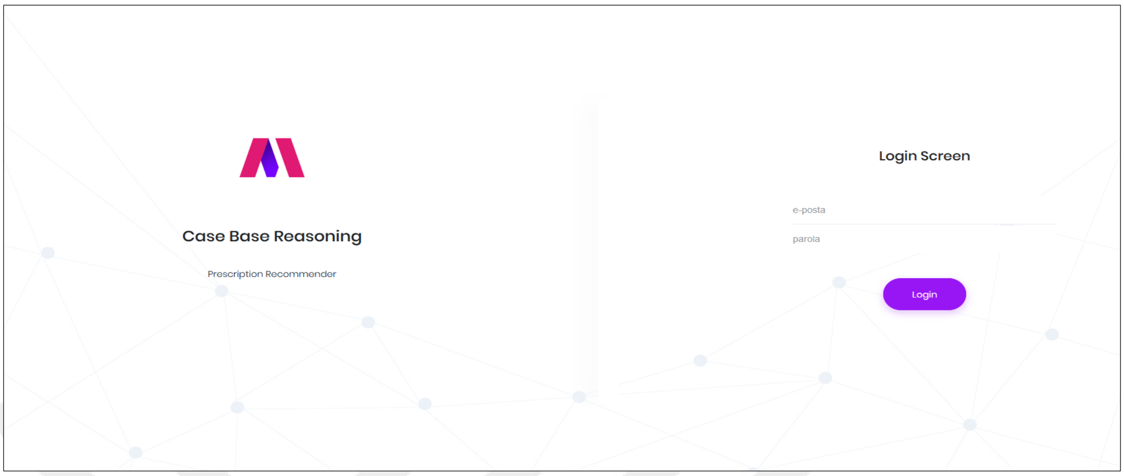

Figure A.1 Login Page-Dashboard

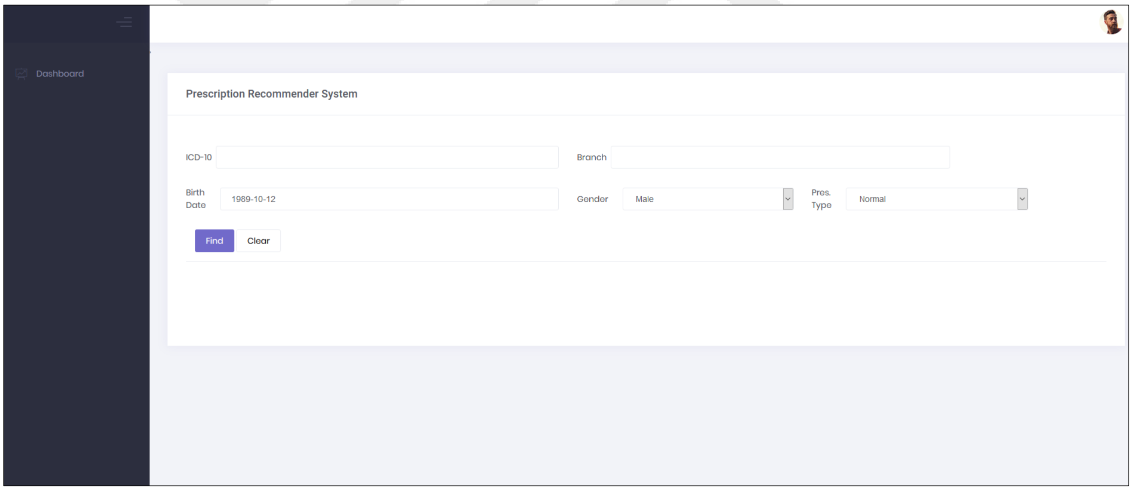

Figure A.2 Recommendation Page