**Mini Project Report on**



**RECOMMENDATION SYSTEM FOR**

**HEALTHCARE**



**Submitted in partial fulfillment of the requirement for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted by:**

**Student Name**  **University Roll No.**

**Anudeepa Giri 2018696**

***Under the Mentorship of***

**Dr. Surender Singh Samant**

**Associate Professor**



**Department of Computer Science and Engineering**

**Graphic Era (Deemed to be University)**

**Dehradun, Uttarakhand**

**January-2024**



**CANDIDATE’S DECLARATION**

I hereby certify that the work which is being presented in the project report entitled **“Recommendation System for Healthcare”** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineeringof the Graphic Era (Deemed to be University), Dehradun shall be carried out by the under the mentorship of **Dr. Surender Singh Samant, Associate Professor**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Anudeepa Giri 2018696

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No.** |
| Chapter 1 | Introduction | **4-6** |
| Chapter 2 | Literature Survey | **7-8** |
| Chapter 3 | Methodology | **9-11** |
| Chapter 4 | Result and Discussion | **12-13** |
| Chapter 5 | Conclusion and Future Work | **14** |
|  | References | **15** |

**Chapter 1**

**Introduction**

* 1. **Introduction**

Healthcare is the term used to describe the maintenance of health through disease and injury prevention, diagnosis, and treatment, as well as various mental and physical impairments in people provided by professionals in the health areas [1]. Massive patient data must be analyzed in order to provide insights and support illness prediction in all primary, secondary, and tertiary care as well as public health activities that are part of healthcare systems. One strong and expanding platform for healthcare services is healthcare recommender system (HRS).

Many clinical applications have been proposed to address hospital patient load. Physicians seeking physiological data, naturally become confused by growing number of diabetics and their high-risk variables. Elderly people with chronic illnesses are also unable to see doctors on a regular basis, which is unnecessary and expensive.

New approaches to putting these designed treatments into practice are emerging as a result of technological advancements, and researchers and policymakers require easy-to-use instruments to assess the quality and applicability of proposed designs. One of these novel methods for computer-based health interventions is recommender systems (RSs). RSs are information retrieval software tools that employ machine learning (ML) to forecast an item's relevance for a particular user [2]. Retrieving information from users' past allows RS to select, modify, and deliver health messages that are relevant to them.

Viral disorders are among major areas in which HRSs are applied, in addition to chronic illnesses.[1]Applications of HRSs can help limit infection and virus transmission by improving people's knowledge and awareness, providing preventive activities, rapid screening, early diagnosis, and quarantine ideas, as demonstrated by the experiences of global epidemic, such as COVID-19.Access to health information is necessary so that data analysts can use analytical techniques to evaluate health data and help doctors improve treatment outcomes, predict illness, and cut expenses.

In order to acquire insights from a vast amount of patient data (lab reports, treatment outcomes, and medical plans), researchers and experts in recommendation systems and healthcare services engaged in relevant studies. This allowed them to gain knowledge, aid in illness predictions, and improve the accuracy of HRSs.

Medical professionals are interested in identifying strategies for early disease detection, lowering healthcare costs, recommending medications based on patient histories and profiles, improving patient care, and creating new health services.

* 1. **Problem Definition**

Reliable and accurate predictions can be made for physicians if patient privacy is preserved while contributing to the recommender system and several hospitals or health centers with healthcare data collaborate in a secure manner. Patients can upload their symptoms to the healthcare recommender system, which uses the symptoms to determine the disease and suggest doctors to the patient.

With wireless Body Sensor Nodes (BSN), which can send patient data to a smartphone via Wireless Body Area Network (WBAN), patients' symptoms, such as blood pressure, blood sugar, heart rate, and temperature, can be recorded. Additionally, smartphone uses Wi-Fi or a mobile network to send this data to a healthcare recommender system. Following a diagnosis, recommender system suggests the doctor to the patient.

* 1. **Objectives**

A major transformation in the healthcare sector has been brought about by tremendous increase in information technology. Patients utilize Internet to research illnesses, diagnoses, therapies, doctors, and hospitals. Selecting a qualified doctor to treat their illness is one of the patients' top worries.

This methodology has two advantages: first, it is transparent because model outcomes mimic logical decision processes based on the hierarchy of significant physiological parameters; second, it is safer than conventional deep learning methods against adversarial attacks because it significantly reduces number of parameters that need to be trained.

However, to explore methods to protect the patient's identity and guarantee data privacy we implemented a methodology on a variable-by-variable basis by fitting a sequence of regression models and drawing synthetic values from the corresponding predictive distributions using linear regressions and norm rank.

* 1. **Scope of the project**

Using healthcare recommender systems is one way to reduce information overload because it can be challenging to extract relevant, individualized information from massive amounts of data. Though it's unlikely that friends and relatives have same medical history as the patient, recommendations they provide are reliable. Through the internet, patients can access reviews or other information about doctor.

Patients use smart gadgets such as laptops, tablets, and mobile phones to send health data to the healthcare system. Smart gadgets and healthcare facilities must communicate securely since patient health data contains sensitive information. Trusted Authority (TA), which creates the master key and private keys for healthcare facilities, handles the key establishment and key sharing mechanism for patient registration in the healthcare system. The medical facilities create their own private keys by registering with TA, using those keys to create private keys for their patients.

While some HRSs are currently in use, there is still a long way to go before they are generally used in domains connected to health. Potential of these systems could represent the cause. Pursuing efficacy, robustness, speed, and precision must be the main goals of RSs.

**Chapter 2**

**Literature Survey**

* 1. **Related Work**

Numerous HRSs, their applications, and their difficulties have all been investigated by researchers. This section offers a systematic review that undertakes a thorough investigation of HRS subjects, summarizes the findings, and assesses and illustrates the classification of HRSs. Unlike the unstructured review process, systematic literature review (SLR) procedure minimizes bias while defining open challenges and directions by adhering to the correct steps in the scientific sequence.[2]

Sub-classes of techniques include context-based, knowledge-based, content-based, collaborative-based, and hybrid approaches. Empowering patient-centric decision-making within medical realms has the potential to significantly optimize effectiveness of current healthcare recommender systems. This can be achieved by ensuring seamless collection, efficient mining, and thorough analysis of data dispersed across various geographical regions.

Various sites using Arbitrary Distributed Data (ADD) of healthcare services at different nodes might work together to generate consumer preference, which would benefit both parties and solve the problem of insufficient ratings for different medical services. However, some parties choose not to share their private information owing to concerns about privacy, finances, and legality. The parties may consent to productive cooperation if they are guaranteed confidentiality of their data. [4]

Collaborative filtering (CF), is the most widely used and traditional method for producing recommendations. In order to provide predictions for the present user,user-based CF approach finds previous users whose rating patterns resemble current user's. [4]

Compared to user-based CF techniques, these methods are more scalable and yield higher-quality recommendations. In model-based CF, all data is utilized to develop a model offline, which is subsequently used to make predictions. [3]In memory-based CF, all data is used directly to generate predictions. Because they require less online computing, model-based CF techniques are more scalable than memory-based CF techniques. Model-based recommender systems are preferred over memory-based ones because a high number of ratings are used in the prediction generating process. Usage of best-suited therapy keys by doctors is intended to lower patient management expenses. Enhance management of disease diagnosis by assisting medical professionals in finding more keys that may enhance patient care.

We have implemented a collaborative filtering strategy in our recommender system. Intention is to create a model based on the historical choices made by a number of doctors on treatment keys that correspond to the patient's diagnosis keys. As a result, a list of potential treatment keys with the best possible therapy is presented to clinician when a new patient with a particular ICD joins the system.[3] Thus, we apply single-concept extraction in electronic medical records to extract organized medical concepts, such as diseases and treatment protocols.

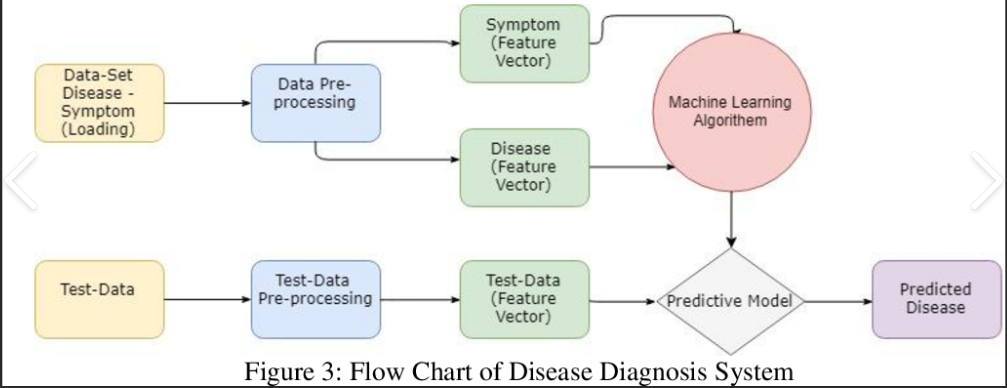
Sending patients required information at the appropriate moment while maintaining integrity, reliability, and privacy of their medical records is HRS's top priority .Additionally, these technologies should reduce the time and effort costs associated with making decisions pertaining to healthcare.

**Chapter 3**

**Methodology**

Following their treatments, patients provide healthcare facilities with their opinions or evaluations of the doctors they were treated by. S1=(P1,P2,....Ph2)..Sn= (P1,P2,....Phn) is a collection of patients registered with healthcare recommender systems H1,H2,.......Hn respectively, so that S1^ S2....Sn **≠** 0. It is believed that H1,H2,.......Hn represent n healthcare recommender systems. The distribution of patient ratings throughout various healthcare facilities can be done in three ways: horizontally, vertically, or arbitrarily.

The deployed recommender system generates a recommendation based on the desired events after analyzing the frequency of medical occurrences in the EHR. As a result, our system selects items based on the rating history of the patients, much like collaborative filtering.

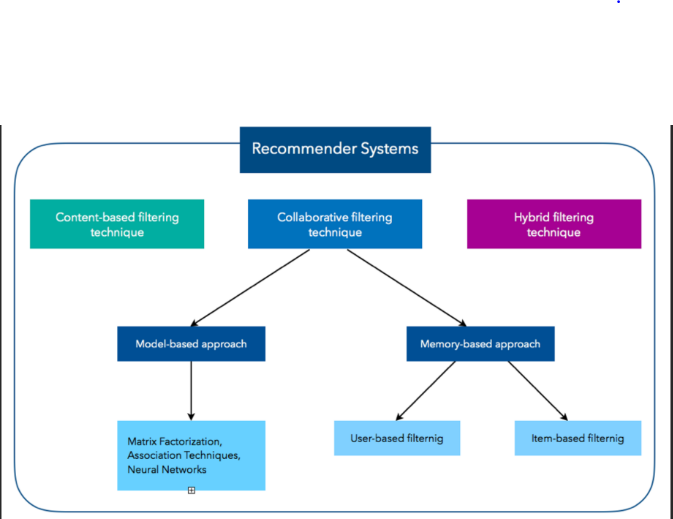
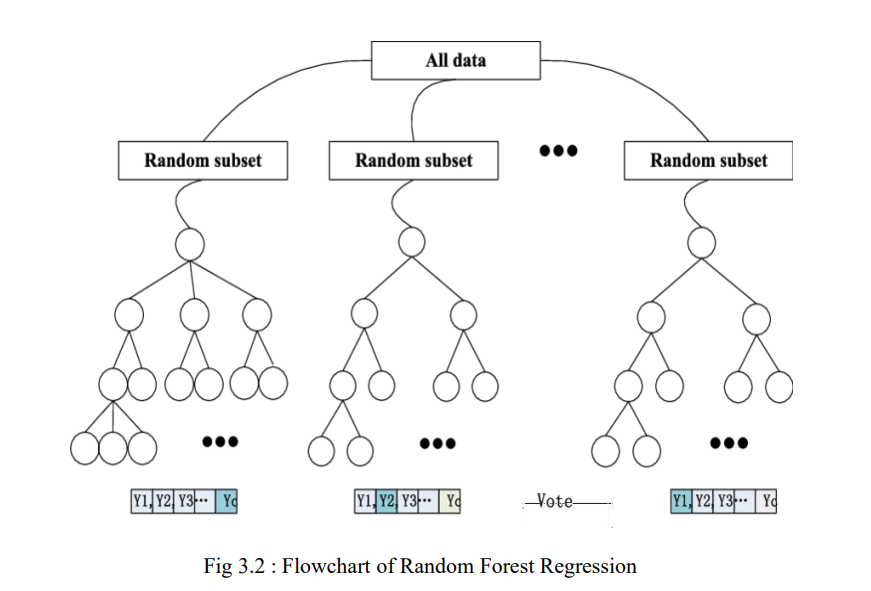


Using the terms "health recommender systems," "medicine recommender systems," "recommender systems in the wellness domain," and "e-Health systems," we first gathered a collection of studies about HRS. We used the following extra keywords to find references in order to take a closer look at recommendation scenarios in the healthcare industry: "food recommendation," "nutrition recommendation," "drug recommendation," "heath status prediction," "healthcare service recommendation," "physical activity recommendation," and "doctor recommendation."

* 1. **Algorithms Used**

**3.1.1 Random Forest Regression**

Random Forest is a tree-based bootstrapping algorithm based on that tree that includes a certain number of decision trees to build a powerful predictive model. The final prediction may be the function of all predictions made by each learner. In sales prediction, random forest classifier is used because it has decision tree like hyperparameters. It hows the relation between decision trees and random forest. To solve regression tasks of prediction by virtue of random forest, the sklearn.ensemble library’s random forest regressor class is used**.**



**Fig 3.2 Recommender Systems**

* + 1. **Decision Trees**

A decision tree is a type of tree structure that resembles a flowchart, with internal nodes representing features, branches standing for rules, and leaf nodes for algorithmic results. It is a supervised machine-learning approach that can be applied to regression and classification issues. This method assesses the degree of impurity or randomness in the subsets using metrics like entropy or Gini impurity, and during training it chooses the best characteristic to split the data depending on. The objective is to identify the characteristic that optimizes the information gain or impurity reduction following the split.

* + 1. **Recommendation Techniques**

Collaborative Filtering (CF)

According to this method, "patients would have similar treatments/health-care services if they share similar disease profiles/health conditions."[2]

Content-based Filtering (CB)

This method looks for products that fit the user profile and are comparable to ones the user has previously liked. This method in HRS recommends medical services that are appropriate for the patient's health and that resemble previous treatments that were given to them.[2]

Knowledge-based Recommendation (KB)

In domains where the user wants to clearly specify what they need, this method is used. With the use of explicit user preferences, item knowledge, and a set of constraints outlining the relationships between users' preferences and item properties, this method generates recommendations. [2]

Hybrid Recommendation (HyR)

This method aims to address the shortcomings of a previous approach while leveraging the benefits of a combination of recommendation techniques. Since CB's prediction for new things is typically based on descriptions of these items that are currently accessible, it can handle when a new item is added to the system and has no user ratings. [2]

**Chapter 4**

**Discussion**

Healthcare professionals as well as end consumers are two user types that HRS may handle. End users may be patients or healthy users. The system must store a user profile detailing each end-user's health status. After registering in this system, a patient must provide a predetermined set of data, including age, blood pressure, BMI, triceps thickness, insulin, and glucose. [1]These data are then evaluated to generate customized suggestions. Patients whose qualities most closely match those of the active patient are identified using a user-based CF. A cardiovascular patient's profile, for example, might have following details: name, date of birth, height, weight, cardiovascular type, and blood pressure reading.[3] HRS uses this information to determine which prescription drugs are right for the patient. Physicians, nurses, clinicians, doctors, and pharmacists are examples of healthcare professionals. Additionally, policy makers and medical researchers can gain from HRS.

An increase in amount of drug information available has created challenges for finding pertinent medications and drug-disease interactions. In this regard, drug recommender systems have been created to help healthcare providers and end users find the right drugs for a given illness.

**Result**

We compared performance of random forest, support vector machine (SVM), and decision tree classifiers in cancer, diabetes and cardiovascular diseases prediction. The decision tree outperformed the other two methods, with an accuracy of 99%, compared to 92% and 85% for random forest and SVM, respectively. Further analysis suggests that the decision tree's ability to handle non-linear relationships in the data might have contributed to its superior performance. We examined the accuracy, MAE, and MSE of several machine learning algorithms, such as random forests, decision trees, and SVM classifiers, in order to predict cancer, diabetes, and heart disease from a dataset.

Health recommender systems (HRSs) have ability to influence and involve users in changing their behavior while offering better options and useful information based on behavior that has been observed. By providing individualized, technology-assisted advice, the HRS seeks to enable people to monitor and improve their health.[2]

It will take a lot of future work to create a treatment recommender system that offers solid therapy suggestions based on clinical facts and ethical design. This issue could be resolved by making efforts in data gathering and cleaning, as well as by increasing data sharing throughout care institutions.

Real-time functioning of such a system in a therapeutic context will require more investigation, and testing and optimization of the reward functions and exploration-exploitation ratios that systems use. Design of a system that functions well in health care treatment recommendations should be pursued due to the potentially immense benefits it could have on individuals and society.

**Chapter 5**

**Conclusion**

The purpose of this study was to evaluate previous HRS studies about benefits claims. The primary contribution of this study was the presentation of a technical taxonomy for categorizing HRSs. This taxonomy can identify methods and resources that can be used for quick disease diagnosis, care cost reduction, medication prescription based on patient profiles and histories, improving healthcare, and developing health services.

Health recommender systems are instrumental in aiding patients and healthcare professionals in making informed decisions. Our project explores various recommendation scenarios, including medicine recommendations, health status predictions for conditions like cancer and chronic diseases, and healthcare professional recommendations involving addition of hospitals based on specific diseases, along with budget calculations. Diverse algorithms, spanning recommendation and machine learning techniques, are employed to enhance these scenarios.

**Future Work**

Future study and possible technological extensions of this system include customized health advice for every patient and the ability to add recommendations for additional medical conditions and treatments through the use of APIs, combining medical data from several sources. Along with suggesting tests to be performed and predicting therapies based on symptoms, it can also include functionality to provide a secure system for storing private patient and physician data along with smooth UI experience for users. [5]

It goes a step further by recommending essential tests for accurate assessments. It can also ensure the utmost protection for confidential data of patients and doctors globally. It's not just a leap forward in healthcare; it's a revolution poised to transform way we approach well-being, bringing personalized, predictive, and secure healthcare solutions.

**References**

[1] https://www.sciencedirect.com/science/article/abs/pii/S0957417422018413 **(Example : Article)**

[2] Recommender systems in the healthcare domain: state-of-the-art and research issues Vol57, Pages 171-201(2022) [Thi Ngoc Trang Tran](https://link.springer.com/article/10.1007/s10844-020-00633-6#auth-Thi_Ngoc_Trang-Tran-Aff1), [Alexander Felfernig](https://link.springer.com/article/10.1007/s10844-020-00633-6#auth-Alexander-Felfernig-Aff1),  [Christoph Trattner](https://link.springer.com/article/10.1007/s10844-020-00633-6#auth-Christoph-Trattner-Aff2) & [Andreas Holzinger](https://link.springer.com/article/10.1007/s10844-020-00633-6#auth-Andreas-Holzinger-Aff3)  **(Example : Conference papers)**

# [3] Health Recommender Systems: Systematic Review Monitoring Editor: Gunther Eysenbach Reviewed by André Calero Valdez and Jinglu Jiang [Robin De Croon](https://pubmed.ncbi.nlm.nih.gov/?term=De%20Croon%20R%5BAuthor%5D), PhD,corresponding author1 [Leen Van Houdt](https://pubmed.ncbi.nlm.nih.gov/?term=Van%20Houdt%20L%5BAuthor%5D), MSc,1 [Nyi Nyi Htun](https://pubmed.ncbi.nlm.nih.gov/?term=Htun%20NN%5BAuthor%5D), PhD,1 [Gregor Štiglic](https://pubmed.ncbi.nlm.nih.gov/?term=%C5%A0tiglic%20G%5BAuthor%5D), PhD,2 [Vero Vanden Abeele](https://pubmed.ncbi.nlm.nih.gov/?term=Vanden%20Abeele%20V%5BAuthor%5D), PhD,1 and [Katrien Verbert](https://pubmed.ncbi.nlm.nih.gov/?term=Verbert%20K%5BAuthor%5D), PhD11 Department of Computer Science, KU Leuven, Leuven, Belgium2 Faculty of Health Sciences, University of Maribor, Maribor, Slovenia **(Example : Research paper)**

[4]Ethical Design of a Health Care Treatment Recommender System Lam, Peterson, Schnall Harvard University Professor Barbara Grosz December 12, 2016

**(Example: Journal Paper)**

[5] <https://www.sciencedirect.com/science/article/abs/pii/S0167739X17327012>

**(Example : Article)**