International Journal of Advanced Culture Technology Vol.8 No.4 167-176 (2020)
DOI https://doi.org/10.17703/IJACT.2020.8.4.167

IJACT 20-12-22

K-Means Clustering with Content Based Doctor Recommendation for Cancer

Rethina kumar * 1, Gopinath Ganapathy ** 2, Jeong-Jin Kang *** 3

1 Assistant Professor, Dept of Information and Communication, Dong Seoul University, Korea.*
Research Scholar, Bharathidasan University, India.
*2** Registrar, Bharathidasan University, India.*
*3*** Professor, Dept of Information and Communication, Dong Seoul University, Seongnam, Korea.*

Abstract

Recommendation Systems is the top requirements for many people and researchers for the need required by them with the proper suggestion with their personal indeed, sorting and suggesting doctor to the patient. Most of the rating prediction in recommendation systems are based on patient's feedback with their information regarding their treatment. Patient's preferences will be based on the historical behaviour of similar patients. The similarity between the patients is generally measured by the patient's feedback with the information about the doctor with the treatment methods with their success rate. This paper presents a new method of predicting Top Ranked Doctor's in recommendation systems. The proposed Recommendation system starts by identifying the similar doctor based on the patients' health requirements and cluster them using K-Means Efficient Clustering. Our proposed K-Means Clustering with Content Based Doctor Recommendation for Cancer (KMC-CBD) helps users to find an optimal solution. The core component of KMC-CBD Recommended system suggests patients with top recommended doctors similar to the other patients who already treated with that doctor and supports the choice of the doctor and the hospital for the patient requirements and their health condition. The recommendation System first computes K-Means Clustering is an unsupervised learning among Doctors according to their profile and list the Doctors according to their Medical profile. Then the Content based doctor recommendation System generates a Top rated list of doctors for the given patient profile by exploiting health data shared by the crowd internet community. Patients can find the most similar patients, so that they can analyze how they are treated for the similar diseases, and they can send and receive suggestions to solve their health issues. In order to the improve Recommendation system efficiency, the patient can express their health information by a natural-language sentence. The Recommendation system analyze and identifies the most relevant medical area for that specific case and uses this information for the recommendation task. Provided by users as well as the recommended system to suggest the right doctors for a specific health problem. Our proposed system is implemented in Python with necessary functions and dataset.

Keywords: Recommended system, K-Means Clustering, Content Based, E-health, Health social network

Manuscript received: November 27, 2020 / revised: December 17, 2020 / accepted : December 23, 2020
Corresponding Author: brkumar76@yahoo.com
Author's affiliation
Assistant Professor, Dept of Information and Communication, Dong Seoul University, Korea

1. Introduction

The health is one of the major aspects which mainly to be consider in person's life [1]. When we need to find a doctor who is able to properly diagnose and treat a medical condition and family or friends are the main source to get the very first information to identify the right doctor. E-health, defined as the healthcare practice supported by the electronic process and communication [2]. A recent survey demonstrated that 72% of U.S. Internet users looked online for health information within the past years.1 young people between 18 and 35 uses the Web as a source of health information.2 This new phenomenon can be viewed as the evolution of the process that a patient plays out in order to find a solution to her conditions, i.e.to ask the family or friends who had the same problem how they treated it. Indeed, in the same aforementioned survey, 60% of U.S. adults got information or support from friends and family when they have a health problem and 24% of adults got information or support from others who have the same health condition. The general thread shared by all talks was the exploitation of E-health tools for empowering individuals to be actively engaged in the management of their health. Indeed, the sharing of information generates a more informed and empowered patient by reconfiguring the patient/care team relationship towards a patient-centered medicine. In this paper we present a Content Based Recommended system to help patient to find an optimal solution to their health issues by providing top list of doctors. The choice of a doctor or a health facility is a typical problem of information asymmetry where the patient has too little information to make an informed choice and thus needs to be somehow supported. Accordingly, Crowd sources helps patients to share knowledge, to find other similar patients, in order to share their experiences. KMC-CBD implements: (1) K-Means Clustering Doctors according to their profile, (2) Content Based recommended system that allows to identify the similarity between Doctors and Patient to exploit the data coming from the crowd source and community of patients to identify the top ranked doctors with respect to the patient's health condition.

2. Related Research

2.1 K-Means Clustering

K-means clustering is a type of unsupervised learning, which is used when you have unlabeled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K. The algorithm works iteratively to assign each data point to one of K groups based on the features that are provided. Data points are clustered based on feature similarity. The results of the K-means clustering algorithm are: The centroids of the K clusters, which can be used to label new data , Labels for the training data (each data point is assigned to a single cluster).Rather than defining groups before looking at the data, clustering allows you to find and analyze the groups that have formed organically. The "Choosing K" section below describes how the number of groups can be determined. Each centroid of a cluster is a collection of feature values which define the resulting groups. Examining the centroid feature weights can be used to qualitatively interpret what kind of group each cluster represents.

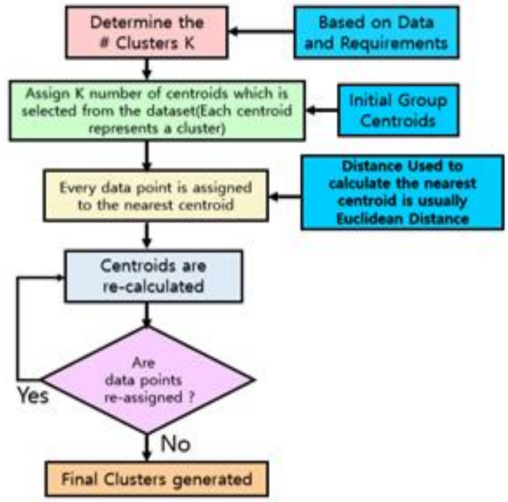


Figure 1. K-Means Algorithm

The mathematics of clustering in very simple terms involves minimizing the sum of square of distances between

$$Minimize \sum_{j=1}^{k} \sum_{i=1}^{n} (x_{ij} - c_j)^2$$

the cluster centroid and its associated data points:

- K = number of clusters
- N= number of data points
- C=centroid of cluster j
- $(X_{ij} — c_j)$– Distance between data point and centroid to which it is assigned.

Deciding on the optimum number of clusters 'K'

The main input for k-means clustering is the number of clusters. This is derived using the concept of ***minimizing within cluster sum of square (WCSS)***. A scree plot is created which plots the number of clusters in the X axis and the WCSS for each cluster number in the y-axis.

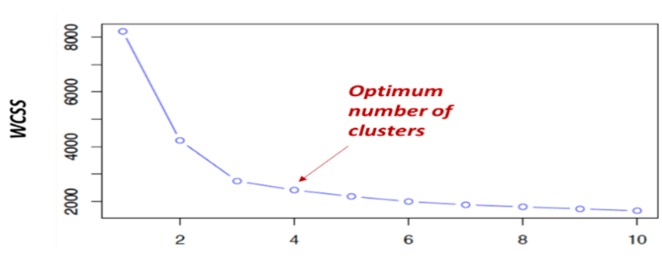


Figure 2. Elbow method to determine optimum number of clusters

As the number of clusters increase, the WCSS keeps decreasing. The decrease of WCSS is initially steep and then the rate of decrease slows down resulting in an elbow plot. The number of clusters at the elbow formation usually gives an indication on the optimum number of clusters. This combined with specific knowledge of the business requirement should be used to decide on the optimum number of clusters.

2.2 Recommender System

A recommender system provides required information needed by a particular patient so that the patent can choose the best. List of successful software tools [3] are providing successful recommendations to the patient, the objective is to give users content that is more personalized [4]. The individualized content minimizes the information load for users and provides refined content that users want. The recommender system's classification has various types of filtering, such as hybridization and collaborative filtering, filtering on the basis of content, and collaborative filtering.

$$TF_{ij} = \frac{f_{i,j}}{\max_{z} f_{z,j}}$$

2.3 Content Based Recommendation Algorithms

Content-based recommendation algorithm make recommendations by matching up the attributes of user profile with the attributes of items. The user profile, in our model, the matching is realized by calculating the cosine similarity between the attributes of user profile and the attributes of medicines Content-based methods can solve the problems of cold start and high sparsity of the data, which are suffered by Collaborative filtering methods. From a conceptual perspective, a Content Based Recommender is a recommender system and it does not necessary work with clustering strategies, instead of this, it could implement any strategy. A Content Based Recommender could apply classifications, prediction, clustering or merge all these strategies to provide a recommendation for something we call as a Decision Support System. K-means is a strategy that use the attributes of a dataset as vectors and based on Euclidean distance between the items, it measures a given k number of clusters of each item on the dataset belongs. A Content Based Recommender could use k-means as part of a strategy to provider a recommendation to a Decision Support System.

TF-IDF (Term Frequency –Inverse Document Frequency) measure

TF-IDF - One of the most known and most used CBF techniques is the TF-IDF (Term Frequency –Inverse Document Frequency) measure (Salton, 1989). TF-IDF measure is defined as follows: For total number of documents N that can be recommended to user's keyword k_i appears in n_i of them. If $f_{i,j}$ is the number of times that k_j appears in document d_j then $TF_{i,j}$ is the term frequency of keyword k_i in document d_j and it is defined as

Keywords may appear in many documents, so with TF, keywords are not useful in distinguishing the relevance between documents. For that reason the inverse document frequency (IDFi) is used in combination with the simple term frequency ($TF_{i,j}$). The IDF_i is defined as

Finally the TD-IDF weight for each keyword k_j in each particular document d_j is defined as

And the content for each document d_j is defined as

Cosine similarity

$$Content(d_j) = (w_{1j}, \dots \dots . w_{kj})$$

TF-IDF vectorizer will also take into account how frequent these combinations are among all, assigning a higher score to those that appear the least. Similarity between vectors -The next step will be to find similar TF-IDF vectors. Recall that we've *encoded* each genre into a TF-IDF representation, now we want to define a *proximity measure.* A commonly used measure is the Cosine similarity this similarity measure owns its

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}},$$

name to the fact that it equals to the cosine of the angle between the two vectors being compared. The lower the angle between two vectors, the higher the cosine will be, hence yielding a higher similarity factor. It is expressed as follows:

Where, since the inner product can be expressed as the product of the magnitudes times the cosine of the angle

$$\mathbf{A} \cdot \mathbf{B} = \|\mathbf{A}\| \, \|\mathbf{B}\| \cos\theta$$

between the two vectors, it becomes clear that the above can be expressed as the cosine:

$$\mathbf{A} \cdot \mathbf{B} = \|\mathbf{A}\| \, \|\mathbf{B}\| \cos\theta$$

So the measure is obtained by taking the inner product between both vectors, and normalizing by their respective magnitudes, resulting in the *cosine* between both vectors. To compute the cosine similarities between all TF-IDF vectors, we can again use scikit-learn Metrics. Pairwise contains many pairwise distance metrics, among them, cosine similarity which will compute the cosine similarities between all the input rows, in this case TF-IDF vectors.

3. Proposed System

The Figure 3 shows the general recommendation process. First Initiates the process then it gets the input information from the patient according to their medical requirements .Then the general recommendation system recommends the list of doctors imprecisely. After that the system records the patient’s choice of doctors and feedbacks about the doctor for the future to improve the recommendation system performance. Then the general recommendation system check for any other active patient, if any other patient is active then the general recommendation system re initiates the process again or else it will stop the process.

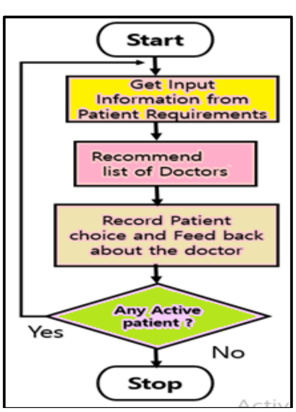


Figure 3. The Recommendation System.

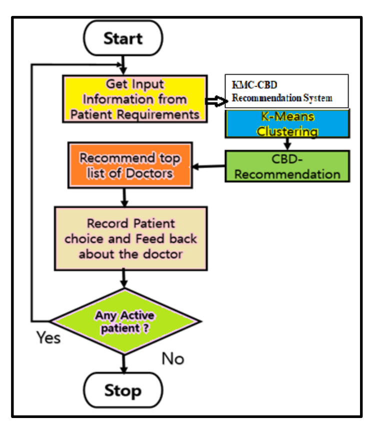


Figure 4. The KMC-CBD Recommendation System

The Figure 4 shows the KMC-CBD process. First initiates the process then it gets the input information from the patient according to their medical requirements. Then the Proposed KMC-CBD recommendation system gets activated and then the proposed system list the top recommended doctors precisely. After that the system records the patient’s choice of doctors and feedbacks about the doctor for the future to improve the recommendation system performance. Then the general recommendation system check for any other active patient, if any other patient is active then the general recommendation system re initiates the process again or else it will stop the process.

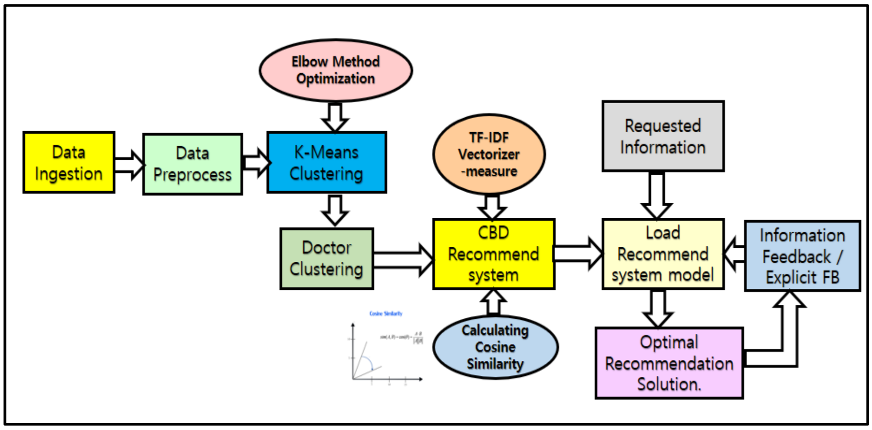


Figure 5: Proposed KMC-CBD Recommended System Frame Work.

The Figure 5 explains about the Proposed KMC-CBD Recommended System Frame Work. In which the process starts with Data ingestion and then the system preprocess the dataset accordingly. The dataset then proceeds by K-means clustering algorithm. The goal of this algorithm is to cluster the doctors according to their medical profile from the dataset. The number of clusters are derived by the concept of minimizing within cluster sum of square (WCSS) and represent with the elbow plot. The content based recommendation system is then activated by clustering of doctor's dataset with TF-IDF (Term Frequency –Inverse Document Frequency) representation.in order to define a proximity measure we used Cosine similarity. Since we have the patient requested information for the recommendation, the system loads the recommended system model which provides optimal recommended solution for the patient. The Proposed KMC-CBD recommendation system lists the top recommended doctors precisely according to the patient medical requirements. The information and feedback are recorded for the future to improve the recommendation system performance.

4. Experimental Evaluation

Our proposed system is implemented in Python with necessary functions and dataset.

4.1 Experiment Environment

We experimented our KMC-CBD Recommended system with Python 3.8.2 / Python 3.8.2 shell.
Dataset (csv Database file) Explicit dataset /From Crowd sources. OS- Window 10 32bit

4.2 Dataset and Evaluation Measures

Building Dataset

Through feedback information techniques, the recommender systems need to collect information about users' profiles. This process is the basis for these systems to be able to provide valid and interesting information to users [6].Commonly, these feedback techniques are categorized in explicit and implicit feedback techniques.

• Explicit feedback: It is the mechanism that allows a user to unequivocally express her interest in an object or set of objects. Typically, users assign a score to these objects through a survey process, such as the 5 star rating system or like/dislike rating system, [8]. As discussed in [7], recommender systems usually collect users' preferences using some of the rating systems cited above. For example, social networks such as Facebook, Twitter, Instagram, LinkedIn or YouTube use the like/dislike rating system as a mechanism for users to be able to rate contents explicitly. To the best of our knowledge, there are no freely available datasets in the health domain for research purposes. Therefore, the first step we performed was to build such a dataset for carrying out our experimental evaluation. We crawled the posts, there are few dedicated healthcare websites giving the doctor and hospital review facilities in India, there is hardly any neutral or impartial website which may really help the patients. Below are the list

1. Practo.com : Practo provides the review of doctors who are enlisted in their platform. If anyone books appointment through practo, he/she can give the review of the doctor.
2. Sehat.com : Sehat has review facility for doctors as well as hospitals. No matter you book with their platform or not (I guess they have a booking system as well), you can write your view with a simple registration through their site.
3. Credihealth.com : This site is almost in the same line of Practo. They also give you the chance to write review, only if you are booking a doctor through their system.
4. lybrate.com : The idea is same as sehat. They are also in the field of online booking of doctors, diagnostic center, online doctor consultation etc.
5 .https://www.credihealth.com/doctor/

4.3 Evaluation Measures

Table 1: List of Doctors for K-Means Clustering.

DID	DOCTOR_NAME	description	RATING	AGE
1	Dr. Bellarmine V L	HEART SPECILALIST - MUMBAI	5	55
2	Dr. Mohammed IM	BREAST CANCER - TRICHY	5	56
3	Dr. Raj G	LUNG CANCER - TRICHY	4	57
4	Dr. Veda PPS	BREAST CANCER - CHENNAI	3	58
5	Dr. Sathish SK	BREAST CANCER - CHENNAI	5	59
6	Dr. G Amarnath	BREAST CANCER - TRICHY	4	60
7	Dr. J Balaji	BREAST CANCER - TRICHY	3	61
8	Dr. Sujay	BREAST CANCER - CHENNAI	5	62
9	Dr.KR Prasanna	BREAST CANCER - CHENNAI	3	63
10	Dr. JACK	BREAST CANCER - CHENNAI	2	64
11	Dr. Rajni	LUNG CANCER -TRICHY	5	40
12	Dr. MM	LIVER CANCER - CHENNAI	2	42
13	Dr. B.ANIS	BREAST CANCER -TRICHY	1	44
14	Dr. Rajni Gupta	BREAST CANCER -TRICHY	4	46
15	Dr. Vincent Lawrence	HEART SPECILALIST - MUMBAI	5	48
16	Dr. Lawrence	HEART SPECILALIST - MUMBAI	5	50

Results 1 - Doctors for K-Means clustering.

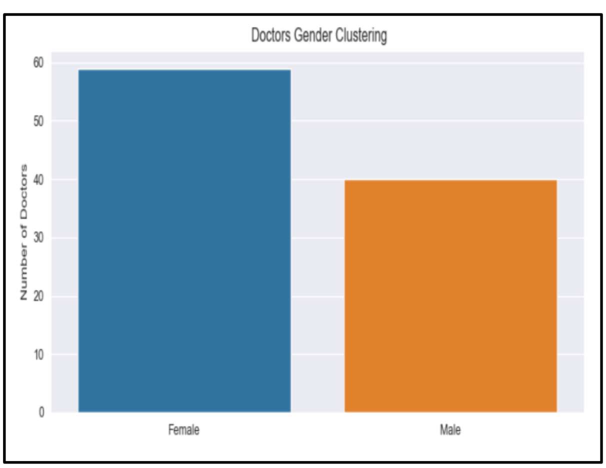


Figure 6. Doctors gender clustering.

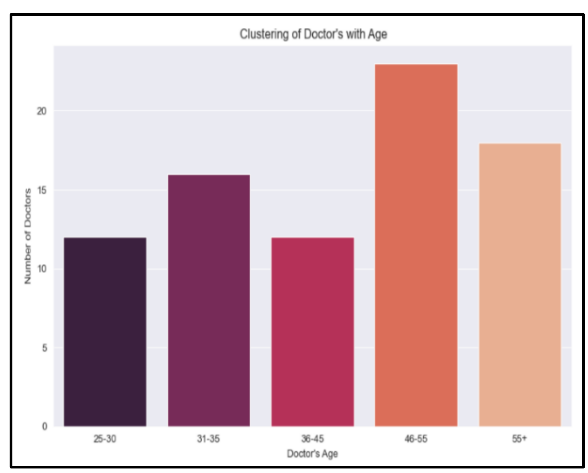


Figure 7. Clustering of Doctor's with Age.

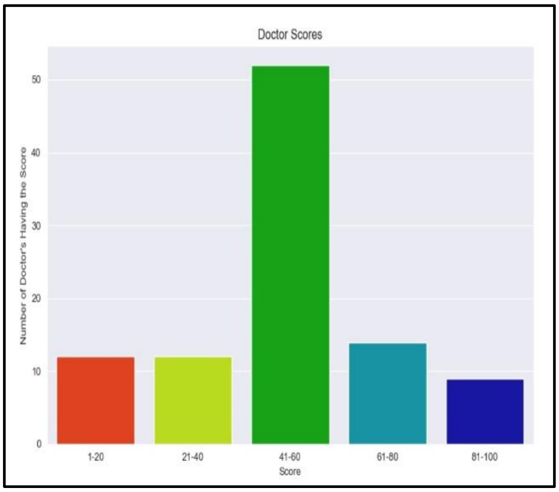


Figure 8. Doctor's scores.

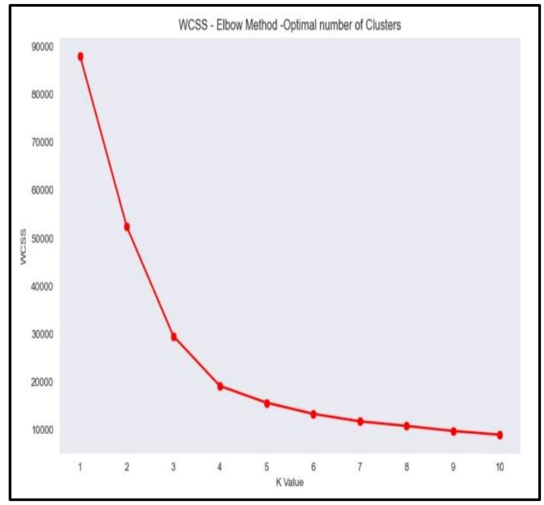


Figure 9. WCSS- Elbow Method Optimal number of clusters.

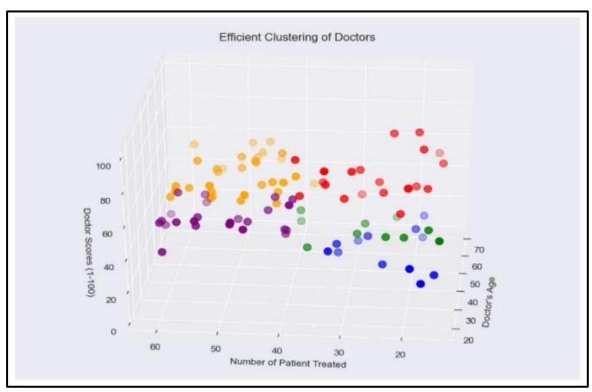


Figure 10. Efficient clustering of doctors.

In our experiments, Table 1 shows the List of Doctors for K-Means Clustering which contains the doctor's information with their medical profile. From the Results 1 the Figure 6. Shows K-Means for Doctors Gender Clustering information. The Figure 7. Shows K-Means of Doctor's with Age Clustering. Figure 8. Shows K-Means of Doctor's clustering with their Scores. Figure 9.Shows WCSS- Elbow Method to represent Optimal Number of Clusters finally Figure 10.Shows the Efficient Clustering of Doctors for the given doctor dataset.

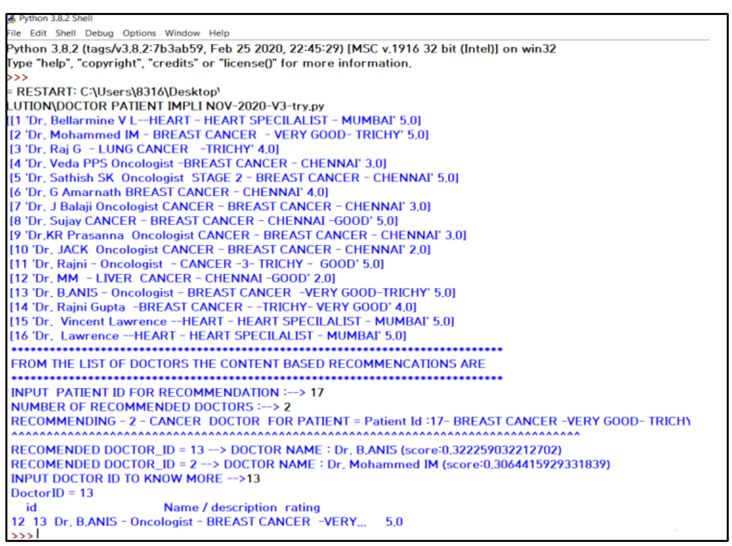
Table 2: List of Doctors for CB Recommendation.

id	Name / description	rating
1	Dr. Bellarmine V L--HEART - HEART SPECILALIST - MUMBAI	5
2	Dr. Mohammed IM - BREAST CANCER - VERY GOOD- TRICHY	5
3	Dr. Raj G - LUNG CANCER -TRICHY	4
4	Dr. Veda PPS Oncologist -BREAST CANCER - CHENNAI	3
5	Dr. Sathish SK Oncologist -BREAST CANCER - CHENNAI	5
6	Dr. G Amarnath Oncologist -BREAST CANCER - TRICHY	4
7	Dr. J Balaji Oncologist - BREAST CANCER -VERY GOOD-TRICHY	3
8	Dr. Sujay CANCER - BREAST CANCER - CHENNAI -GOOD	5
9	Dr.KR Prasanna Oncologist CANCER - BREAST CANCER - CHENN	3
10	Dr. JACK Oncologist CANCER - BREAST CANCER - CHENNAI	2
11	Dr. Rajni - Oncologist - CANCER -3- TRICHY - GOOD	5
12	Dr. MM - LIVER CANCER - CHENNAI -GOOD	2
13	Dr. B.ANIS - Oncologist - BREAST CANCER -VERY GOOD-TRICH	5
14	Dr. Rajni Gupta -BREAST CANCER - -VERY GOOD-TRICHY	4
15	Dr. Vincent Lawrence --HEART - HEART SPECILALIST - MUMBA	5
16	Dr. Lawrence --HEART - HEART SPECILALIST - MUMBAI	5
17	Dr. JACK Oncologist CANCER - BREAST CANCER - DELHI	3
18	Dr. Ram - Oncologist - CANCER -3- TRICHY - GOOD -DELHI	4
19	Dr. Manu - LIVER CANCER - GOOD-HIMACHAL P	5
20	Dr. B.ANI - Oncologist - BREAST CANCER - Bangalore	3
21	Dr. Raj -BREAST CANCER - -Hyderabad	3
22	Dr. Vince --HEART SPECILALIST - Kolkata	4
23	Dr. Law --HEART - Bangalore	5

Table 3: Patient Information for CB Recommendation.

PID	PATIENT INFO
1	Patient Id :17- Mrs.xyz- BREAST CANCER VERY GOOD. TRICHY
2	Patient Id :18- Mrs.xyzabc- BREAST CANCER GOOD. CHENNAI
3	Patient Id :19- Mrs.asdf HEART SPECILALIST. VERY GOODCHENNAI
4	Patient Id :20- Mrs.mno HEART SPECILALIST. GOODMUMBAI
5	Patient Id :24- BREAST CANCER Oncologist VERY GOOD- TRICHY
6	Patient Id :25- BREAST CANCER Oncologist GOOD- CHENNAI
7	Patient Id :26- LUNG CANCER -TRICHY
8	Patient Id :27- LIVER CANCER - MUMBAI

Result 2 - KMC-CBD Recommended System



```
Python 3.8.2 Shell
File Edit Shell Debug Options Window Help
Python 3.8.2 (tags/v3.8.2:7b3ab59, Feb 25 2020, 22:45:29) [MSC v.1916 32 bit (Intel)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:\Users\8316\Desktop\
LUTION\DOCTOR PATIENT IMPLI NOV-2020-V3-try.py
[[1 'Dr. Bellarmine V L--HEART - HEART SPECILALIST - MUMBAI' 5.0]
 [2 'Dr. Mohammed IM - BREAST CANCER  - VERY GOOD- TRICHY' 5.0]
 [3 'Dr. Raj G  - LUNG CANCER   -TRICHY' 4.0]
 [4 'Dr. Veda PPS Oncologist -BREAST CANCER - CHENNAI' 3.0]
 [5 'Dr. Sathish SK  Oncologist  STAGE 2 - BREAST CANCER - CHENNAI' 5.0]
 [6 'Dr. G Amarnath BREAST CANCER - CHENNAI' 4.0]
 [7 'Dr. J Balaji Oncologist CANCER - BREAST CANCER - CHENNAI' 3.0]
 [8 'Dr. Sujay CANCER - BREAST CANCER - CHENNAI -GOOD' 5.0]
 [9 'Dr.KR Prasanna  Oncologist CANCER - BREAST CANCER - CHENNAI' 3.0]
 [10 'Dr. JACK  Oncologist CANCER - BREAST CANCER - CHENNAI' 2.0]
 [11 'Dr. Rajni - Oncologist  - CANCER -3- TRICHY -  GOOD' 5.0]
 [12 'Dr. MM  - LIVER  CANCER - CHENNAI -GOOD' 2.0]
 [13 'Dr. B.ANIS - Oncologist - BREAST CANCER  -VERY GOOD-TRICHY' 5.0]
 [14 'Dr. Rajni Gupta  -BREAST CANCER - -TRICHY- VERY GOOD' 4.0]
 [15 'Dr.  Vincent Lawrence --HEART - HEART SPECILALIST - MUMBAI' 5.0]
 [16 'Dr.  Lawrence --HEART - HEART SPECILALIST - MUMBAI' 5.0]
 ************************************************************************
 FROM THE LIST OF DOCTORS THE CONTENT BASED RECOMMENCATIONS ARE
 ************************************************************************
 INPUT  PATIENT ID FOR RECOMMENDATION :--> 17
 NUMBER OF RECOMMENDED DOCTORS :--> 2
 RECOMMENDING - 2 - CANCER  DOCTOR  FOR PATIENT = Patient Id :17- BREAST CANCER -VERY GOOD- TRICHY
 ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
 RECOMENDED DOCTOR_ID = 13 --> DOCTOR NAME : Dr. B.ANIS (score:0.322259032212702)
 RECOMENDED DOCTOR_ID = 2 --> DOCTOR NAME : Dr. Mohammed IM (score:0.3064415929331839)
 INPUT DOCTOR ID TO KNOW MORE -->13
 DoctorID = 13
    id                         Name / description  rating
 12  13  Dr. B.ANIS - Oncologist - BREAST CANCER  -VERY...     5.0
>>>
```

Figure 11 - KMC-CBD Recommended System

In our experiment, Table 2 shows the List of Doctors for the content based recommendation which contains the doctor's information with their medical profile. Table 2 shows the List of Patients with their information which is the key content for the content based recommendation. From the Result 2 the Figure 11. Shows the cluster of information for Doctors dataset. The patients are requested to input their id by inputting the patient id from the Table 3: (Patient Information.) they are also asked to input the required number of top doctors with their score generated by the content based recommendation. Finally KMC-CBD Recommended System Recommends the top doctors from the list of doctors (Table 2) according to the patient medical profile. If the patient is willing to know more information and medical profile of the particular doctor then it is also possible with KMC-CBD Recommended System.

5.Conclusions and Future work

In this paper, we presented KMC-CBD Recommended System, The main objective is to recommend top doctors and health facilities, which provides optimal match by given patient profile. Accordingly, the core component of KMC-CBD Recommended System will be able to find doctor's based on the requirements, conditions and treatments required by the patient and to recommend top score doctors with a similar medical profile who already treated patient successfully. The recommended system will be able to assign the most relevant medical area to the health status described by a patient using explicit feedback natural language. We carried out experiments for testing the KMC-CBD Recommended System, and the results are good and satisfactory. The health status description provided by patients is enough to effectively identify the most relevant medical area for that problem, and the K-Means Clustering similar Doctors in a specific medical area drastically improves the recommendation accuracy and the Content Based Recommendation provide the patient with optimal top-scored doctors according to their health requirements. As a future work, we are planning to assess the performance of the recommended system to improve the recommended systems by also taking into account the structured information for representing symptoms, conditions and treatments.

References

[1] T.R. Hoens, M. Blanton, A. Steele, N.V. Chawla, Reliable medical recommendation systems with patient privacy, ACM Transactions on Intelligent Systems and Technology (TIST) 4 (4) (2013) 67.

[2] V. Della Mea, What is e-health (2): The death of telemedicine? JMIR 3 (2) (2001) e22, doi:10.2196/jmir.3.2.e22.

[3]. Adomavicius, G.; Tuzhilin, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. 2005, 17, 734–749.

[4]. Amatriain, X. Mining Large Streams of User Data for Personalized Recommendations. ACM SIGKDD Explor .Newsl. 2013, 14, 37–48.

[5]. An adaptive doctor-recommender system ISSN: 0144-929X (Print) 1362-3001 (Online) Journal homepage: https://www.tandfonline.com/loi/tbit20- Muhammad Waqar, Nadeem Majeed, Hassan Dawood, Ali Daud & Naif Radi Aljohani

[6] F. Narducci, C. Musto, M. Polignano, M. de Gemmis, P. Lops, G. Semeraro, Recommending doctors and health facilities in the healthnet social network, in:Proceedings of the First International Workshop on Social Media for Personalization And Search (SoMePeAS), CEUR, Aberdeen, Scotland, 2017, pp. 1–7.

[7] L. Duan, W.N. Street, E. Xu, Healthcare information systems: data mining methods in the creation of a clinical recommender system, Enterp. Inf. Syst. 5 (2) (2011) 169–181.

[8] L. Fernandez-Luque, R. Karlsen, L.K. Vognild, Challenges and opportunities of using recommender systems for personalized health education. MIE, 2009,pp. 903–907.

[9] A Rating-Ranking Method for Crowdsourced Top-k Computation,Kaiyu Li, Xiaohang Zhang, Guoliang Li SIGMOD/PODS '18, June 10–15, 2018, Houston, TX, USA.

[10] Disagreement between Hospital Rating Systems: Measuring the Correlation of Multiple Benchmarks and Developing a Quality Composite Rank, American Journal of Medical Quality1–9, 2019.