

# A PERSONALIZED RECOMMENDATION ENGINE FOR PREDICTION OF DISORDERS USING BIG DATA ANALYTICS

Shobana.V<sup>1</sup>, N.Kumar<sup>2</sup>

<sup>1,2</sup>Assistant Professor, Dr. N.G.P Arts and Science College, Coimbatore, India.

<sup>1</sup>[shobana484@gmail.com](mailto:shobana484@gmail.com), <sup>2</sup>[kumar@drngpasc.ac.in](mailto:kumar@drngpasc.ac.in)

## ABSTRACT

Recommender systems has become an important research field since 1990's and its applications includes several domains. The application of recommendations in health care is a very wide area which recommends the patients about their health. Upon vast amount of data prevailing in health care sector, these data can be processed using big data tools to deliver a meaningful prediction to the patients. The predictions and recommendations will be more accurate since we are dealing with vast amount of data. Moreover it alerts the user from the occurrence of disease and takes necessary actions before it occurs. Thus recommender system along with big data will be promising solution that is prevailing in healthcare sector. The sum total of data related to the patient and their well-being constitutes the "Big Data" problem in the healthcare industry.

**KEYWORDS-** Recommender systems, big data, collaborative filtering, health care, prediction of thyroid disease.

## I. INTRODUCTION

The recent trends in healthcare support systems is the development of patient centric pervasive environments in addition to the hospital centric one. Pervasive healthcare takes steps to design, develop and evaluate computer technologies that help citizens participate more closely in their own health care on one hand and on the other to provide flexibility in patients everyday life with work, family and friends [1]. And moreover, the past several years , data has been growing immensely in all business sectors. While many industries are successful in performing big data analysis to benefit from datasets, health care sector has started to take small steps to move forward [2]. The primary goal is that prevails in health care sector are to handle large volumes of unstructured data in a secured environment. Recent trends in health care support systems are focusing on developing a patient centric pervasive environments and the use of

mobile devices and technologies in medical monitoring and healthcare systems [3]. Big data analytics involves recommendation system on prediction models which essentially concentrate upon analyzing a set of relevant data and predict a meaningful pattern. The above idea can be applied to all health care sectors ranging from prediction of diseases to predicting physical exercises to improve health.

## II. PERSONALIZED HEALTHCARE: A DATA-DRIVEN APPROACH

Healthcare is moving from a disease-centered model towards a patient-centered model. In a disease-centered model, physicians' decision making is centered around the clinical expertise, and data from medical evidence and various tests. In a patient-centered model, patients actively participate in their own care and receive services focused on individual needs and preferences, informed by advice and oversight from their healthcare

providers[2]. At the same time as this patient-centered care model is being emphasized in health care delivery, the potential for ‘personalizing’ health care from disease prevention, disease management, and therapeutics perspective is increasing. Healthcare informatics and advanced analytics (or data science) may contribute to this shift from population-based evidence for health care decision-making to the fusion of population and individual-based evidence in health care[10] [4].

### III. RECOMMENDATIONS IN HEALTH CARE

“Recommendation systems are software tools and techniques that provide suggestions for items to be of use to a user. The suggestions provided are aimed at supporting their users in various decisions”[12][5]. Patients exposed to similar risk, lifestyle and environmental factors may have similar outcomes. How to deeply leverage the “big data” resident in electronic medical records, patients’ experiences and histories, to create a personalized disease risk profile for an individual patient? Can we develop a patient-centered model for personalized care to answer questions such as: What diseases am I at risk for developing? How should I manage them? What wellness strategies may best work for me?

We approached this problem by learning from the work on ‘collaborative filtering methodology’ used in other settings by recommendation systems. Essentially, collaborative filtering is a data mining technique designed to predict a user’s opinion about an item or service based on the known preferences of a large group of users[9][10]. Most applications assume that the input is partial preference information for each user. That is, the user’s opinion or rating is known

for a few items, but usually unknown for a strong majority of the item set. The basic principle behind collaborative filtering is that users who have similar taste on some items are likely to have similar taste on other items. This information is then used to make personalized predictions on the movies one may want to watch (such as in Netflix.com) or books one may want to read (such as in Amazon.com)[4][10]

### IV. COLLABORATIVE FILTERING IN HEALTH CARE

Using collaborative filtering, we can generate predictions on other diseases based on a set of other similar patients. However, there are a number of challenges. The diseases do not have a rating system or a “preferential scoring” as in movies or books. We know if a person had a disease in the past or does not have a disease (rather not been diagnosed with the disease); there is no preferential ranking or rating provided by individuals. The challenge of absent disease cannot be ignored as that could simply imply that the patient has not been, yet, diagnosed with the disease [10][4].

Thus, we are able to incorporate a vast array of disease co morbidities, which are effectively leveraged for personalized disease prediction, management plan, and wellness for an individual patient. The work leverages the similarities and shared experiences among individuals in a healthcare system and beyond (potentially millions of individuals) on: patient history, disease timing, disease progression, prognosis, and wellness strategies. This big data is filtered to result in the personalized plan [10][4].

### V. EFFICIENT CHRONIC DISEASE DIAGNOSIS PREDICTION AND RECOMMENDATION SYSTEM AND

## **PREDICTION OF THYROID DISORDERS IN WOMEN**

Health care Recommender systems for chronic disease diagnosis (CDD) plays a major role in controlling of diseases ,by providing disease diagnosis prediction and medical recommendations. Providing an accurate real time recommendation for medical data is a tedious task and in spite of it recommendations are being made. This research applies recommender system to manage chronic disease and provide advices to patients[6].Thyroid gland produces thyroid hormone which is used to regulate body's metabolism. The abnormalities in thyroid hormone is divided into two cases. Hypothyroidism and Hyperthyroidism. Hypothyroidism is the result of less production of thyroid hormone and Hyperthyroidism is the result of excessive production of thyroid hormone. A properly functioning thyroid will maintain the right amount of hormones needed to keep the body's metabolism functioning at a satisfactory rate. As the hormones are used, the thyroid creates replacements. The quantity of thyroid hormones in the bloodstream is monitored and controlled by the pituitary gland. When the pituitary gland, which is located in the center of the skull below the brain, senses either a lack of thyroid hormones or a high level of thyroid hormones, it will adjust its own hormone (TSH) and send it to the thyroid to tell it what to do. When the thyroid produces too much hormone, the body uses energy faster than it should. This condition is called hyperthyroidism. When the thyroid does not produce enough hormone, the body uses energy slower than it should. This condition is called hypothyroidism. There are many different reasons why either of these conditions might develop. Currently, about 20 million Americans have some form of thyroid disease. People of all ages and races can get thyroid disease. However, women are five to eight times more likely than men to have thyroid problems[7].

## **VI. OVERVIEW OF MAHOUT IN RECOMMENDATION ENGINE**

This section concerns the non-distributed, non-Hadoop-based recommender engine / collaborative filtering code inside Mahout. It was formerly a separate project called "Taste" and has continued development inside Mahout alongside other Hadoop-based code. It may be viewed as a somewhat separate, more comprehensive and more mature aspect of this code, compared to current development efforts focusing on Hadoop-based distributed recommenders. This remains the best entry point into Mahout recommender engines of all kinds. A Mahout-based collaborative filtering engine takes users' preferences for items ("tastes") and returns estimated preferences for other items. For example, a site that sells books or CDs could easily use Mahout to figure out, from past purchase data, which CDs a customer might be interested in listening to. Mahout provides a rich set of components from which you can construct a customized recommender system from a selection of algorithms. Mahout is designed to be enterprise-ready; it's designed for performance, scalability and flexibility[11]. Apache Mahout is a library of Machine Learning algorithms for Hadoop. Machine learning is an area of artificial intelligence, which focuses on learning from available data to make predictions on unseen data without explicit programming. Apache Mahout recommendations module helps recommending to the users items based on his preferences[8].

## **VII. CONCLUSION AND FUTURE WORK**

Thus when the data size is large we can predict the results of particular occurrence very easily.The above study reveals the application of recommender systems in health care. It also particularly concentrates on prediction and diagnosis of thyroid disorders in women. By

making appropriate classification and clustering we can predict the disorders in thyroid hormone and recommending appropriate treatments for that. The tool mahout will be very effective one in analyzing the large data and recommending appropriate results. Our future work focuses on developing a novel recommendation engine for prediction of thyroid disorders in women.

## References

- [1].Igor Kulev, Elena Vlahu Gjorgievska,Vladimir Trajkovic,Saso Koceski "Development of a novel recommendation algorithm for collaborative health care system model" Com SiS Vol.10,No.3,June 2013.
- [2]. Weider. D. Yu, Choudhury Pratiksha, Sawant Swati, Sreenath Akhil, Medarametla sarath"A modeling approach to big data based recommendation engine in modern Healthcare environment" 39<sup>th</sup> IEEE Conference on Computer, software and application 2015.
- [3]. Igor Kulev, Elena Vlahu Gjorgievska,Vladimir Trajkovic,Saso Koceski "Collaborative Health Care system model" .
- [4] Nitesh V. Chawla, and Darcy A. Davis "Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework"
- [5] Racci .F, Rokach.L,Shapira.B, Kautor.P.B,"Recommender Systems Handbook" Springer Iedition 2011.
- [6] Asmaa S.Hussain ,Wail M.Omar, Xue.Li Modafar Ati "Efficient Chronic Disease Diagnosis prediction and Recommendation system " IEEE conference on Biomedical Engineering and Sciences 2012.
- [7]. Keles.A Ayturk Keles "ESTDD: Expert system for thyroid diseases diagnosis" Expert Systems with Applications, Science Direct,2008.
- [8].<https://mahout.apache.org/users/recommender/recommender-documentation.html>
- [9]. Davis DA, Chawla NV, Christakis NA, Barabási AL. Time to CARE: a collaborative engine for practical disease prediction. Data Mining and Knowledge Discovery. 2010;20(3):388–415.
- [10]. [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)
- [11] [www.mahout.apache.org](http://www.mahout.apache.org).
- [12]. [www.eprints.ugd.edu.mk](http://www.eprints.ugd.edu.mk).