Patient Record Maintenance using Clustering

K. Karthick[1], Dr. N. Rajkumar[2], Dr. N. Suguna[3]

[1]PG Scholar, [2]Associate Professor, [3]Professor

Department of Computer Science and Engineering, Akshaya College of Engineering and Technology, Coimbatore, Tamil Nadu, India

ABSTRACT

A ton of programming's mechanizing a few undertakings is coming live every single day. An assortment of enhancements has been peeping out in pretty much every area that we witness throughout each and every day. In the field of clinical office, a ton of specialized upgrades have been brought right into it as medicines yet not in keep up understanding records. Consider an ordinary patient's life who goes through medicines in normal stretches and trusts that the legitimate outcomes will be out. Regardless of a having a hard day by treating such countless people in clinics, a specialist needs to figure out how to check the outcomes and present the report back on schedule. On the off chance that the patient include is more in an emergency clinic, the approval interaction will in a real sense gobble up additional time which in the end ends up being a gigantic difficulty. Presently if a product that could computerize the subsequent interaction becomes an integral factor, it gets two significant contrasts, that is, the patient need not trust that the outcomes will be out for a long range and the specialist need not figure out how to check and clarify the outcomes. Additionally, the inclination of being incomplete will likewise be broken and the patient will be granted with the outcomes for what he/she had in their body. This target making a digitalized stage to deliver the clinical reports and private to the patients, prompting the finish of paper pen culture for results. Thusly, a great deal of time spent on check and result clarification can be chopped down which in the end saves a plentiful measure of time.

Keywords - Clustering, Medical Reports, Documents, Patient Details.

I. INTRODUCTION

Patient Health Record Maintenance is a digital platform that would enable us to release the patient result in a webpage which is secured, authorized and authenticated based on document clustering.

Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups.

Document clustering is the application of cluster analysis to textual documents. It has applications in automatic document organization, topic extraction and fast information retrieval or filtering

If a patient has consulted with various doctor for their various health problems, it is hard to maintain the record and documents from the origin. The hospital can upload the documents and details of a patient into the specific portal by using their credentials. It enables the patient to check their health condition by using their specific and their registered mobile number.

Our aim is to provide the transparent conditions of a patient to their relatives and family members and it reduces the time of verifying the medical certificates of patient in critical condition and consulting the results with the respective doctors.

The documents of a patient are classified based on their health problems and it classifies the health records effectively without any classification routines

II. LITERATURE SURVEY

"Clinical Documents Clustering Based on Medication/Symptom Names using Multi-View Non-Negative Matrix Factorization", Clinical documents are rich free-text data sources containing valuable medication and symptom information, which have a great potential to improve health care. In this paper, we build an integrating system for extracting medication names and symptom names from clinical notes. Then we can apply Non-Negative Matrix Factorization (NMF) and multi-view NMF to cluster clinical notes into meaningful clusters based on sample-feature matrices. Our experimental results show that multi-view NMF is a preferable method for clinical document clustering. Moreover, we find that using extracted medication/symptom names to cluster clinical documents out performs just using word.

"Medical Image Segmentation using K-Means Clustering and Improved Watershed Algorithm", a methodology that incorporates k-means and improved watershed segmentation algorithm for medical image segmentation. The use of the conventional watershed algorithm for medical image analysis is widespread because of its advantages, such as always being able to produce a complete division of the image. However, its drawbacks include over-segmentation and sensitivity to false edges. We address the drawbacks of the conventional watershed algorithm when it is applied to medical images by using K-means clustering to produce a primary segmentation of the image before we apply our improved watershed segmentation algorithm to it. The K-means clustering is an unsupervised learning algorithm, while the improved watershed segmentation algorithm makes use of automated thresholding on the gradient magnitude map and post-segmentation merging on the initial partitions to reduce the number of false edges and over-segmentation. By comparing the number of partitions in the segmentation maps of 50 images, we showed that our proposed methodology produced segmentation maps which have 92% fewer partitions than the segmentation maps produced by the conventional watershed algorithm.

"Biomedical Document Clustering and Visualization based on the Concepts of Diseases", Document clustering is a text mining technique used to provide be.er document search and browsing in digital libraries or online corpora. A lot of research has been done on biomedical document clustering that is based on using existing ontology. But, associations and co-occurrences of the medical concepts are not well represented by using ontology. In this research, a vector representation of concepts of diseases and similarity measurement between concepts are proposed identify the closest concepts of diseases in the context of a corpus. Each document is represented by using the vector space model. A

weight scheme is proposed to consider both local content and associations between concepts. A Self-Organizing Map is used as document clustering algorithm. Vector projection and visualization features of SOM enable visualization and analysis of the clusters distributions and relationships on the two dimensional space. experimental results show that the proposed document clustering framework generates meaningful clusters and facilitate visualization of the clusters based on the concepts of diseases.

"Medical Records Clustering Based on the Text Fetched from Records", describes how the rich available data from patient's medical records can be clustered and hidden information can be retrieved out of it. We first collect the 49 patient's medical records, use annotators to extract the text based on symptom occurred and medical drug name. The fetched text are clustered and stored in a file. When a combination of medical terms taken from medical documents is given as a query through the search engine shows the clustered documents. We use MetaMap and Medex as annotators for extracting the symptom names and the pharmaceutical names. For clustering the fetched data, we are using the multi view NMF, which is a clustering technique.

"BMC Medical Informatics and Decision Making", Multiplication of data sources within heterogeneous healthcare information systems always results in redundant information, split among multiple databases. Our objective is to detect exact and approximate duplicates within identity records, in order to attain a better quality of information and to permit cross-linkage among stand-alone and clustered databases. Furthermore, we need to assist human decision making, by computing a value reflecting identity proximity. The proposed method is in three steps. The first step is to standardise and to index elementary identity fields, using blocking variables, in order to speed up information analysis. The second is to match similar pair records, relying

on a global similarity value taken from the Porter-Jaro-Winkler algorithm and the third is to create clusters of coherent related records, using graph drawing, agglomerative clustering methods and partitioning methods.

"Medical Records Clustering: A Survey", Retrieving similar medical cases from the medical case repository for user search case, the similarity measure and good clustering is useful. While to finding similarity between cases several methods have been proposed, but measuring the similarity between patient cases is a challenging problem. In that survey we focus on different similarity measures and clustering techniques. We are working on the data of medical records. Data is high dimensional, that much of features not gives much accuracy so we extract features from the medical records and build case library. We compare the result of different clustering algorithms using clustering validation.

III. PROPOSED SYSTEM

The venture targets making a digitalized stage to distributing patient ailment prompting the finish to the paper pen culture. Patient Record Maintenance is a computerized stage that would empower us to assess medical issue, approve the outcomes which again is done as an online cycle saving a ton of time. Bunching is the errand of separating the populace or information focuses into various gatherings with the end goal that information focuses in similar gatherings are more like other information focuses in similar gathering than those in different gatherings. Record grouping is the utilization of bunch investigation to literary reports. It has applications in programmed report association, point extraction and quick data recovery or sifting If a patient has talked with different specialist for their different medical issues, it is difficult to keep up the record and archives from the cause. The medical clinic can transfer the reports and subtleties of a patient into the particular gateway by utilizing their

certifications. It empowers the patient to check their ailment by utilizing their particular and their enlisted portable number. Our point is to give the straightforward states of a patient to their family members and relatives and it decreases the hour of confirming the clinical declarations of patient in basic condition and counselling the outcomes with the individual specialists. The archives of a patient are arranged dependent on their medical issues and it characterizes the wellbeing records viably with no characterization schedules.

IV. CONCLUSION

Since Patient Record Maintenance, framework has acquired a more prominent fascination in the field of Medical, this task proposes a way to deal with adequately transfer the outcomes and can be access anyplace from country utilizing grouping. Framework is intended to beat the issues in existing comparative frameworks. Centre piece of the task was accomplished utilizing two strategies which are watchword extraction and examination of likeness. Catchphrase extraction of both archive and records was done in three stages which were eliminating intensifiers, supporting verb modifiers, making Unique watchwords from each answer, crossing of two arrangements of watchword.

V. FUTURE WORK

In the Proposed System, the patient documents are not analyzed and the enhancement will be delivered in analyzing their documents and automatically extract the health condition will update in database.

V. REFERENCES

[1]. Medical Records Clustering: A Survey, Mangesh Mali, Dr Parag Kulkarni, Prof. Virendra Bagade M.E. Student, Department of Computer Engineering, Pune Institute of Computer Technology, Pune, India Chief Scientist, Research Department, iknowlation Research Labs, Pune, India Asst. Professor, Department of Computer Engineering, Pune Institute of Computer Technology, Pune, India.

[2]. Clinical Documents Clustering Based on Medication/Symptom Names using Multi-View Nonnegative Matrix Factorization Yuan Ling, Xuelian Pan, Guangrong Li*, Xiaohua Hu, Member, IEEE

[3]. Medical Image Segmentation using K-Means Clustering and Improved Watershed Algorithm H.P. Ng, S.H. Ong, K.W.C. Foong, P.S. Goh5, W.L. Nowinski .

[4]. Stephan Bloehdorn, Philipp Cimiano, and ndreas Hotho. 2006. Learning ontologies to improve text clustering and classi€cation. In From data and information analysis to knowledge engineering. Springer, 334–341.

[5]. Carsten G¨org, Hannah Tipney, Karin Verspoor, William A Baumgartner Jr, K Bretonnel Cohen, John Stasko, and Lawrence E Hunter. 2010. Visualization and language processing for supporting analysis across the biomedical literature. In International Conference on Knowledge-Based and Intelligent Information and Engineering Systems. Springer, 420–429.

[6]. Jun Gu,Wei Feng, Jia Zeng, Hiroshi Mamitsuka, and Shanfeng Zhu. 2013. Efficient semisupervised MEDLINE document clustering with MeSH-semantic and global content constraints. IEEE transactions on cybernetics 43, 4 (2013), 1265–1276.

[7]. Rasmus Knappe, Henrik Bulskov, and Troels Andreasen. 2007. Perspectives on ontology-based querying. International Journal of Intelligent Systems 22, 7 (2007),739–761.

[8]. Teuvo Kohonen. 1998 self-organizing map. Neurocomputing 21, 1 (1998), 1–6.

[9]. Teuvo Kohonen, Samuel Kaski, Krista Lagus, Jarkko Salojarvi, Jukka Honkela, Vesa Paatero, and An.i Saarela. 2000. Self-organization of a massive document collection. IEEE transactions on neural networks 11, 3 (2000), 574–585.

[10]. S Logeswari and K Premalatha. 2013. Biomedical document clustering using ontology based concept weight. In Computer Communication and Informatics (ICCCI), 2013 International Conference on. IEEE, 1–4.

[11]. Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Je. Dean. 2013. Distributed representations of words and phrases and their compositionality. In Advances in neural information processing systems. 3111–3119.

[12]. SPFGH Moen and Tapio Salakoski2 Sophia Ananiadou. 2013. Distributional semantics resources for biomedical text processing. (2013)

Cite this article as :