

Patient Record Maintenance using Clustering

K. Karthick¹, Dr. N. Rajkumar², Dr. N. Suguna³

¹PG Scholar, ²Associate Professor, ³Professor

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

ABSTRACT

Article Info

Volume 7 Issue 6

Page Number: 40-44

Publication Issue :

November-December-2020

A lot of software's automating several tasks is coming live each and every day. A variety of improvements have been peeping out in almost every domain that we witness day in and day out. By accentuating the present medical system, a lot of technical enhancements have not been brought into action. For instance, consider a typical patient's life who undergoes treatments in regular intervals and waits for the proper results to be out. In spite of a having a hard day by treating so many individuals in hospitals, a doctor has to find time to check the results and submit the report back on time. If the patient count is more in a hospital, the validation process will literally eat up more time which in the end turns out to be a huge complication. Now if a software that could automate the resulting process comes into play, it brings in two major differences, that is, the patient need not wait for the results to be out for a long span and the doctor need not find time to check and explain the results. Moreover, the bias of being partial will also be broken and the patient will be awarded with the results for what he/she had in their body. This aims at creating a digitalized platform to release the medical reports and intimate to the patients, leading to the end of paper pen culture for results. By doing so, a lot of time spent on check and result explanation can be cut down which in the end saves an ample amount of time.

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

Article History

Accepted : 01 Nov 2020

Published : 10 Nov 2020

I. INTRODUCTION

Patient 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 better 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 project aims at creating a digitalized platform to publishing patient health condition leading to the end to the paper pen culture. Patient Record Maintenance is a digital platform that would enable us to evaluate health condition, validate the results which again is carried out as an online process saving a lot of time. 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.

IV. CONCLUSION

Since Patient Record Maintenance, system has gain a greater attraction in the field of Medical, this project proposes an approach to effectively upload the results and can be access anywhere from country using clustering. System is designed to overcome the issues in existing similar systems. Core part of the project was achieved using two methodologies which are keyword extraction and comparison of similarity. Keyword extraction of both document and records was done in three steps which were removing adverbs, supporting adverbs, making Unique keywords from each answer, intersecting of two sets of keyword.

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.

VI. REFERENCES

- [1]. Medical Records Clustering: A Survey, Mangesh Mali, Dr Parag Kulkarni, ProfVirendra Bagade M.EStudent, Department of Computer Engineering, Pune Institute of Computer Technology, Pune, India Chief Scientist, Research Department, iknowlation Research Labs, Pune, India AsstProfessor, 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.PNg, S.HOng, K.W.CFoong, P.SGoh5, W.LNowinski .
- [4]. Stephan Bloehdorn, Philipp Cimiano, and ndreas Hotho2006Learning ontologies to improve text clustering and classificationIn From data and information analysis to knowledge engineeringSpringer, 334–341.
- [5]. Carsten G`org, Hannah Tipney, Karin Verspoor, William A Baumgartner Jr, K Bretonnel Cohen, John Stasko, and Lawrence E Hunter2010Visualization and language processing for supporting analysis across the biomedical literatureIn International Conference on Knowledge-Based and Intelligent Information and Engineering SystemsSpringer, 420–429.
- [6]. Jun Gu,Wei Feng, Jia Zeng, Hiroshi Mamitsuka, and Shanfeng Zhu2013Efficient semisupervised MEDLINE document clustering with MeSH-semantic and global content constraintsIEEE transactions on cybernetics 43, 4 (2013), 1265–1276.
- [7]. Rasmus Knappe, Henrik Bulskov, and Troels Andreasen2007Perspectives on ontology-based queryingInternational Journal of Intelligent Systems 22, 7 (2007),739–761.
- [8]. Teuvo Kohonen1998 self-organizing mapNeurocomputing 21, 1 (1998), 1–6.
- [9]. Teuvo Kohonen, Samuel Kaski, Krista Lagus, Jarkko Salojarvi, Jukka Honkela, Vesa Paatero, and An.i Saarela2000Self-organization of a massive document collectionIEEE transactions on neural networks 11, 3 (2000), 574–585.
- [10]. S Logeswari and K Premalatha2013 Biomedical document clustering using ontology based concept weightIn Computer Communication and Informatics (ICCCI), 2013 International Conference onIEEE, 1–4.
- [11]. Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and JeDean2013Distributed representations of words and phrases and their

compositionality In Advances in neural information processing systems 3111–3119.

- [12]. SPFGH Moen and Tapio Salakoski2 Sophia Ananiadou2013Distributonal semantics resources for biomedical text processing(2013)

Cite this article as :

K. Karthick, Dr. N. Rajkumar, Dr. N. Suguna, " Patient Record Maintenance using Clustering, International Journal of Scientific Research in Science, Engineering and Technology(IJSRSET), Print ISSN : 2395-1990, Online ISSN : 2394-4099, Volume 7, Issue 6, pp.40-44, November-December-2020. Journal URL : <http://ijsrset.com/IJSRSET20768>