International Journal of Engineering Science and Advanced Technology (IJESAT)

IJESAT (Enriching the Research)

Open Access Research Article

Volume: 23 Issue: 07

July, 2023

Medical Information Retrieval System

Dr Satyabrata Dash [1] J. Maneesha [2], K. Iran Kumar [3], G. Jasmine [4], J. Syam Prakash [5]

[1] Associate Professor, Department of CSE, Ramachandra College of Engineering, Eluru, A.P

[2,3,4,5] UG Students, Department of CSE, Ramachandra College of Engineering, Eluru, A.P

ABSTRACT

Medical information retrieval is becoming more crucial as a way to support decision-makingand improve access to knowledge and information for doctors and other specialists in the field. By the incorporation of medical domain knowledge for relevance evaluation, integrating the medical knowledge bases has the potential to enhance information retrieval performance. The system uses various search algorithms and techniques to provide accurate and comprehensive results to medical professionals and researchers. Medical information retrieval systems are particularly useful for healthcare providers who need to stay up-to-date with the latest medical research, guidelines, and treatment options. Due to the difficulties in using the domain information from the medical knowledge bases efficiently, this is not a simple task. In this research, we proposed a unique medical information retrieval system with a two-stage query expansion approach that can efficiently model and include the latent semantic linkages to enhance performance. There are two components to this system. The commonly used pseudo relevance feedback method was first improved using a heuristic strategy, which involved iteratively expanding the queries to increase the similarity score between the queries and the documents in order to improve query expansion. Secondly, we presented a latent semantic relevance model

based on tensor factorization to detect semantic association patterns in sparse environments, whichimproved retrieval performance with organised knowledge bases. Then, knowledge-based query expansion in medical information retrieval is triggered by these discovered patterns, which are used as inference paths. Research using the TREC CDS data set: 1) demonstrated that the performance of the proposed system is significantly better than the baseline system and the systems reported in the TREC CDS conference, and is comparable with the state-of-the-art systems, and 2) demonstrated the capability of tensor-based semantic enrichment methods for medical information retrieval tasks.

1. INTRODUCTION

The primary goal of a medical information retrieval system is to provide healthcare professionals, researchers, and patients with accurate, up-to-date, and reliable medical information. The system can help clinicians make more informed decisions about patient care, aid researchers in identifying trends and patterns in medical data, and provide patients with access to reliable health information.

Medical information retrieval systems can vary in complexity, from simple keyword-based search engines to more advanced systems that use machine learning algorithms to understand natural language

IJESAT (Enriching the Research)
Open Access Research Article
Volume: 23 Issue: 07
July, 2023

queries and provide relevant results. Some systems also incorporate data visualization tools, such as graphs and charts, to help users interpret complex medical data. Overall, a medical information retrieval system is a valuable tool for healthcare professionals, researchers, and patients looking to access and analyze medical information quickly and efficiently.

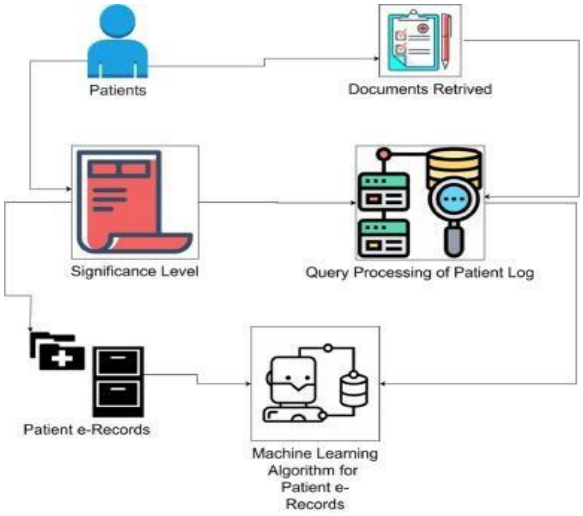


Fig 1: How the life connected with Medical IR System

Information retrieval methods can be quite challenging to use when trying to extract pertinent datafrom electronic medical records. This is because requests for information are typically represented using boolean combinations of index terms and information retrieval is typically done with huge structured data sets, sets of key words, or index terms, and stored records.

medical applications like the Clinical Decision Support System (CDSS) to provide scientific evidences to support clinical decision making, facilitate the translation of the most recent research findings into practise, and improve the quality of healthcare.

A medical information retrieval system is a software application designed to retrieve relevant medical information from various sources such as medical literature, electronic health records (EHRs), and medical databases. The system uses advanced search algorithms and natural language processing techniques to understand and interpret user queries and retrieve relevant information.

Medical information retrieval systems can be used by healthcare professionals, researchers, and patients to access accurate and up-to-date medical information. Some examples of medical information retrieval systems include PubMed, Cochrane Library, and UpToDate.

The benefits of using medical information retrieval systems include:

1. Improved accuracy: Medical information retrieval systems provide access to accurate andup-to-date medical information, which can improve patient care and treatment outcomes.
2. Time-saving: These systems can save time by quickly retrieving relevant medical information, reducing the need for manual searching and analysis.

IJESAT (Enriching the Research)
Open Access Research Article
Volume: 23 Issue: 07
July, 2023

3. Increased productivity: Medical information retrieval systems can increase productivity by providing quick access to relevant medical information, allowing healthcare professionalsand researchers to focus on their work.
4. Cost-effective: Medical information retrieval systems can be cost-effective by reducing theneed for manual searching and analysis, which can be time-consuming and costly.

A medical information retrieval system is a software tool designed to search and retrieve relevant medical information from various sources such as medical journals, clinical trials, patient records, and other medical databases. The system uses various search algorithms and techniques to provide accurate and comprehensive results to medical professionals and researchers. Medical information retrieval systems are particularly useful for healthcare providers who need to stay up-to-date with the latest medical research, guidelines, and treatment options. These systems can help healthcare professionals to find answers to specific medical questions, compare different treatment options, and access relevant patient information quickly and efficiently.

However, it is important to note that medical information retrieval systems should not replace theexpertise and judgment of healthcare professionals. The information retrieved should be evaluated and interpreted in the context of the individual patient and their unique circumstances.

2. LITERATURE SURVEY

IR research community, development of search and access technologies remains particularly challenging. A central issue in medical IR is the diversity of the users of these services. In particular, they will have varying categories of information needs, varying levels of medical knowledge, and varying language skills. These challenges can be summarized as follows: – Varying information needs: While a patient with a recently diagnosed condition will generally benefit most from simple or introductory information on the disease and its treatment, a patient living with or managing a condition over a longer term will generally be looking for more advanced information, or perhaps support groups and forums. Similarly, a general practitioner might require basic information quickly while advising a patient, but more detailed information if deciding a course of treatment, and a specialist clinician might look for an exhaustive list of similar cases or research papers relating to the condition of a patient that they are currently seeking to advise. Understanding of various types of users and their information needs is one of the cornerstones of medical IR; development of effective, potentially personalized systems that address these needs is one of the greatest challenges. – Varying medical knowledge: The different categories of users of medical IR systems have different levels of medical knowledge, and indeed the medical knowledge of different individuals within a category can also vary greatly. This affects the way in which individuals pose search queries to systems and also the level of complexity of information which should be returned to them or the type of support in understanding of retrieved material which should be provided. – Varying language skills: Given that much of medical content is written in the English language, research to date in medical information search has predominantly focused on monolingual English retrieval. However, given the large number of non-English speakers on the Internet and the lack of content in their native language, effective support for them to search English sources is highly desirable. The format, reliability, and quality of biomedical and medical information

IJESAT (Enriching the Research)
Open Access Research Article
Volume: 23 Issue: 07
July, 2023

varies greatly. A single health record can contain clinical notes, technical pathology data, images, and patient-contributed histories, and may be linked by a physician to research papers. The importance of health and medical topics and their impact on people's everyday lives makes the need for retrieval of accurate and reliable information especially important. Determining the likely reliability of available information is challenging. Finally, as with IR in general, the evaluation of medical search tools is vital and challenging. For example, there are no established or standardized baselines or evaluation metrics, and limited availability of test collections. Further discussion and progression on this topic would be beneficial to the community

3. EXISTING SYSTEM

- The model is using semantic knowledge bases to compute relatedness between terms and improve the retrieval process. The framework consists of three main components: knowledge basecreation, term similarity computation, and ranking of documents based on term similarity.
- First identifying the relevant concepts and entities related to the user's query using a combinationof syntactic and semantic analysis. These concepts are then used to generate a set of related termsusing a knowledge-based relatedness measure.

Disadvantages

- Dependence on the quality of the knowledge base
- Limited scalability: They are computationally expensive, particularly when dealing with largedatasets
- Limited flexibility and Limited evaluation

4. PROPOSED SYSTEM

The proposed system is constructing cyber We Proposed Medical IR System, which has a two-stage query expansion strategy (as shown in Figure 2) to integrate the pseudo relevance feedback and the knowledge-based query expansion to improve the performance of retrieving relevant documents for queries. The system then uses these latent factors to improve the retrieval of medical information. When a user performs a search, the system uses the latent factors to identify related concepts and medical terms. Medical Information Retrieval System based on Tensor Factorization approach offers significant advantages over traditional keyword-based search techniques, making it a valuable tool for healthcare professionals and researchers in the medical field.

Advantages of Proposed system:

- Improved accuracy: The system uses a sophisticated approach to identify latent factors in the data that can be used to improve the accuracy of information
- retrieval.
- Customizability: The system can be customized to suit the specific needs of different medical domains, allowing healthcare professionals and researchers to retrieve information.
- Improved patient care: By providing accurate and relevant medical information, the system can help healthcare professionals make better decisions and improve patient care. Efficiency and Scalability

IJESAT (Enriching the Research)
Open Access Research Article
Volume: 23 Issue: 07
July, 2023

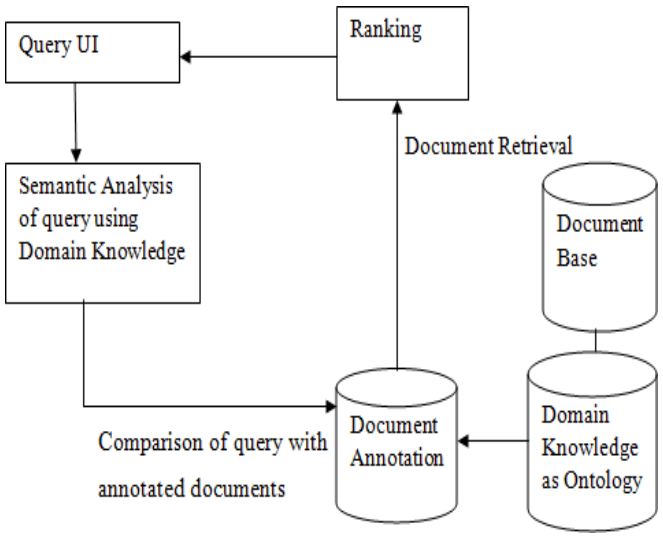
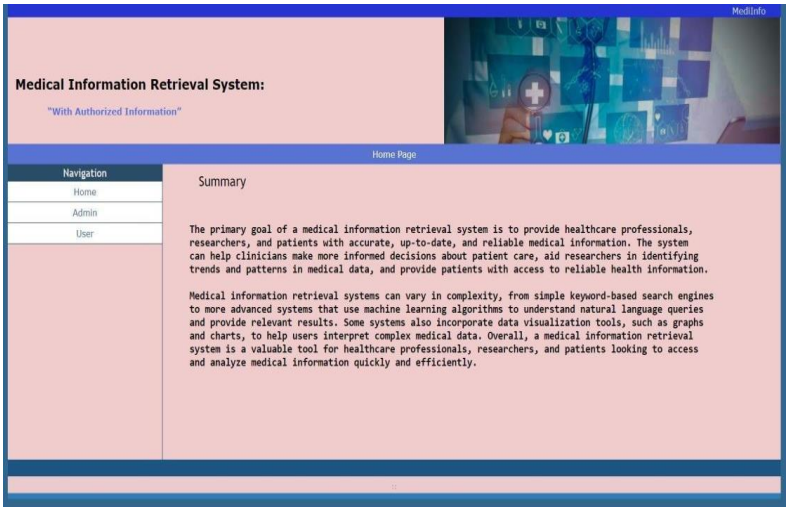


Fig 1: The framework of the Medical Information Retrieval System

DATABASE

Our database is in the format of tables. The database is named as datamining. It contains seven tables they are admin, user, document, search history, discussion form, comments and uml. Each and every table contains specific attributes. It stores the data about the user details and the search history, view the documents present in the database. The admin have all the rights to give activation of access and deactivation of user profile. These gives more security for the information and data retrieving is easy



IJESAT (Enriching the Research)

Open Access Research Article

Volume: 23 Issue: 07

July, 2023

RESULTS

Fig 2: Home Page

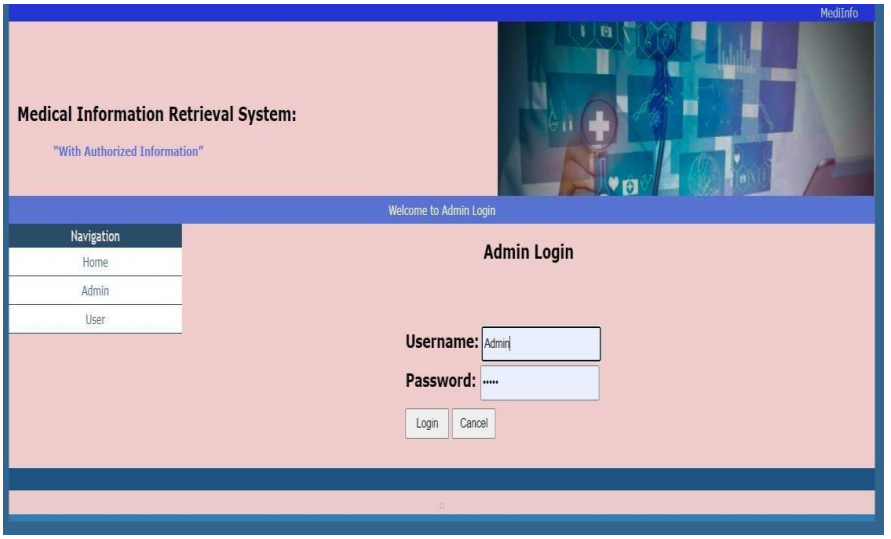


Fig 3: Admin Login Page

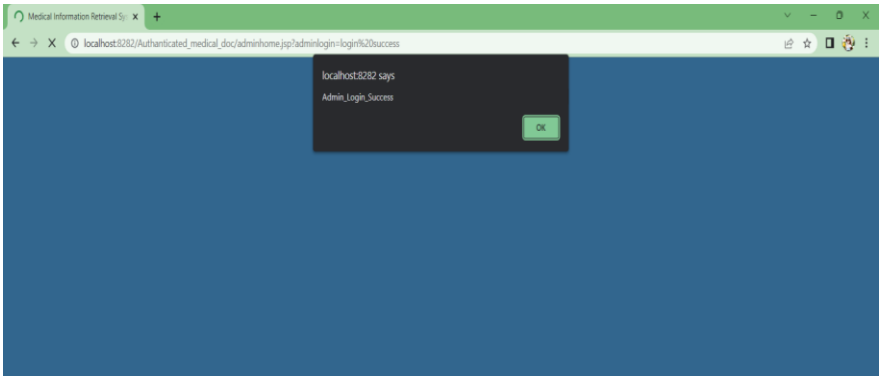
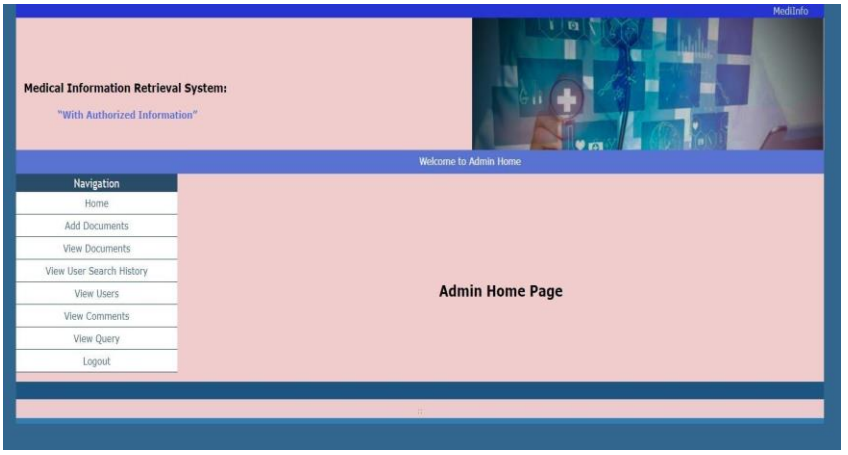


Fig 4: Login Success



Fg 5: Admin Home Page

IJESAT (Enriching the Research)
Open Access Research Article
Volume: 23 Issue: 07
July, 2023

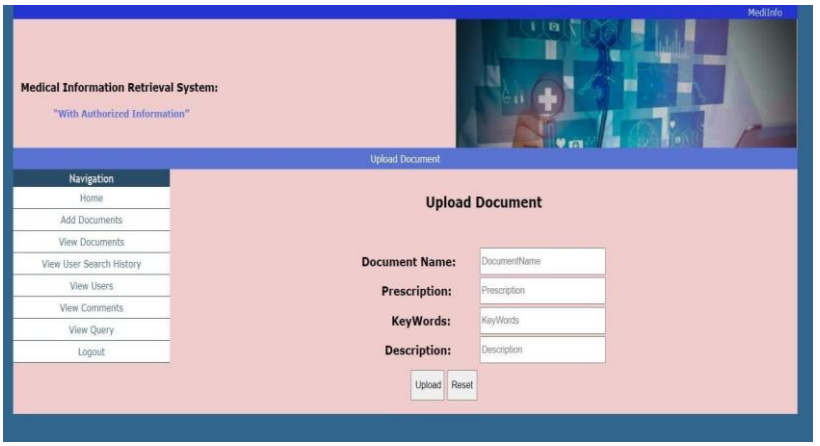


Fig 6: Admin can upload documents

Fig 7: Admin can view the user history

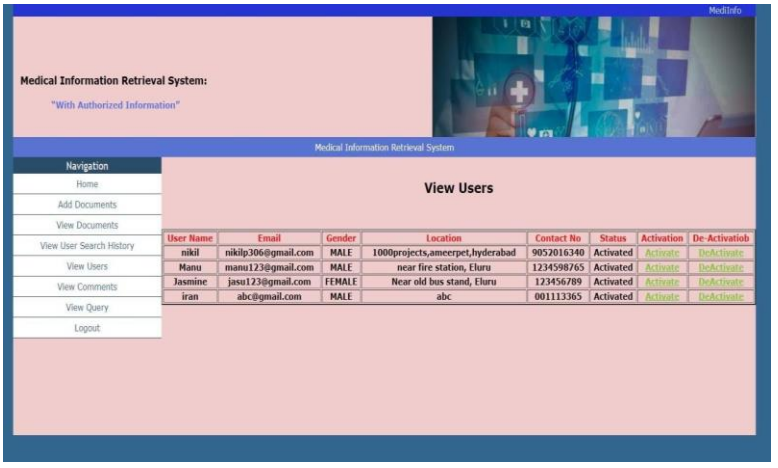


Fig 8: Admin can view no of users and give activate or deactivate status

IJESAT (Enriching the Research)
Open Access Research Article
Volume: 23 Issue: 07
July, 2023

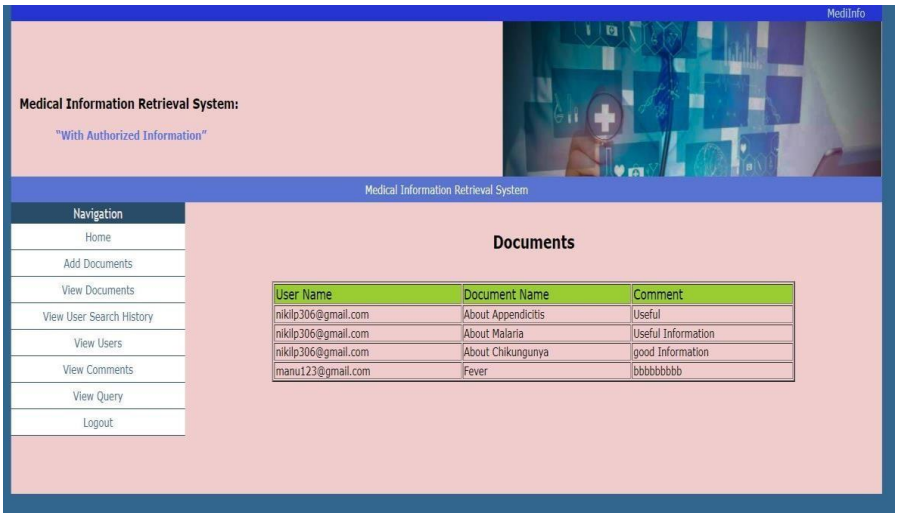


Fig 9: Admin can view comments

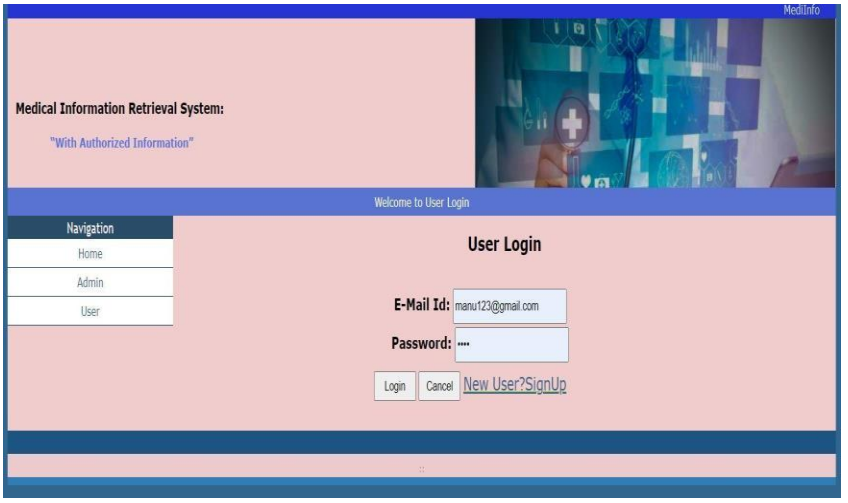


Fig 10: User Login Page

IJESAT
(Enriching the Research)

Open Access Research Article
Volume: 23 Issue: 07
July, 2023

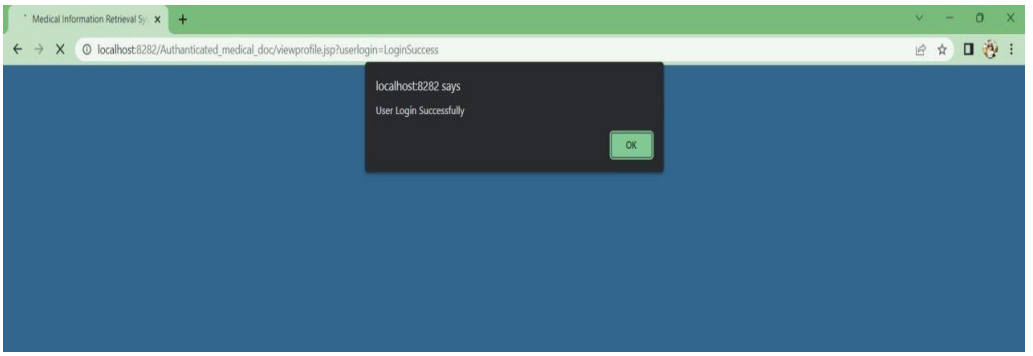


Fig 11: User login Success

Fig 12: User Home Page

Fig 13: User Profile

IJESAT (Enriching the Research)
Open Access Research Article
Volume: 23 Issue: 07
July, 2023

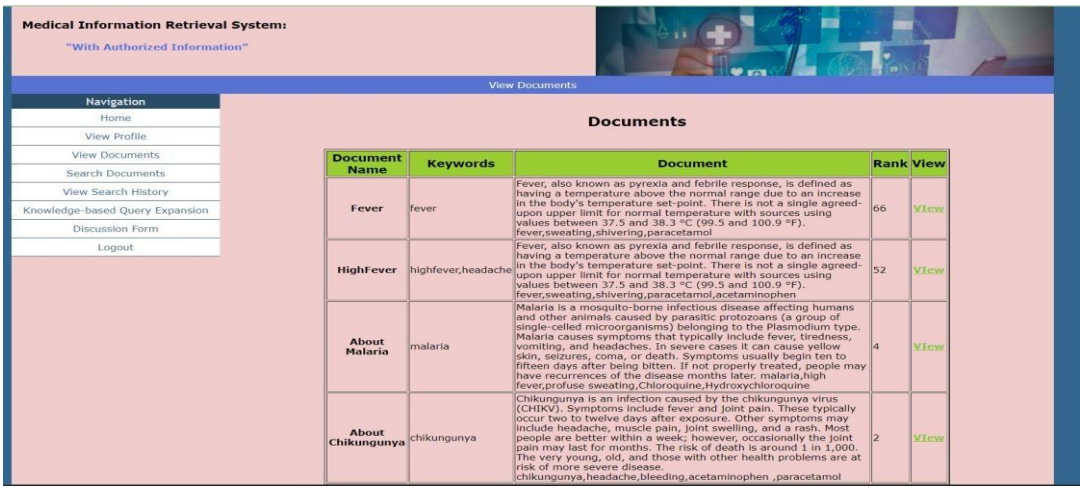


Fig 14: User can View the documents

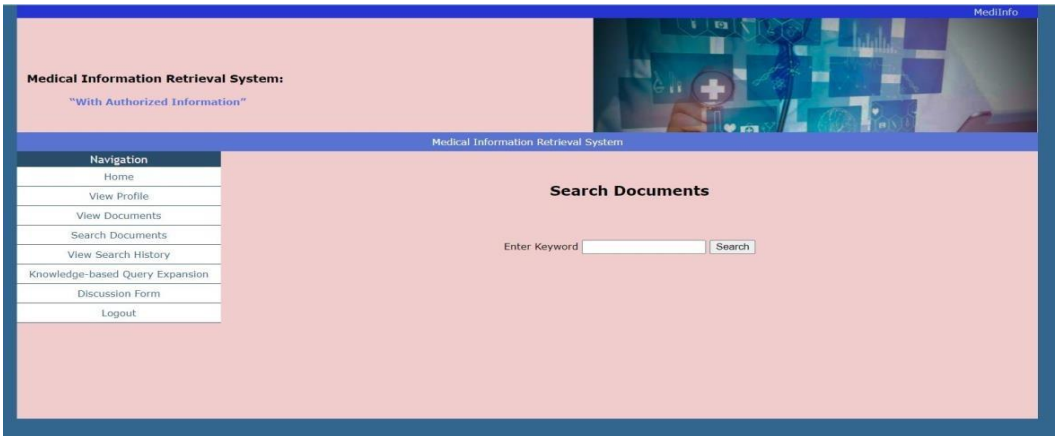


Fig 15: User can search for documents

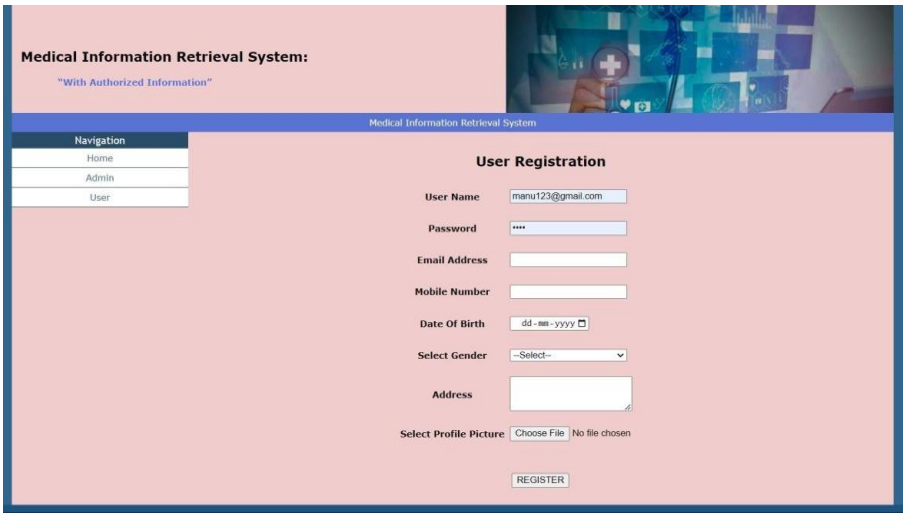


Fig 16: New User can register in this page

IJESAT (Enriching the Research)
Open Access Research Article
Volume: 23 Issue: 07
July, 2023

5. CONCLUSION

In this research, we proposed a medical IR system with a two-stage query expansion strategy, based on the incorporation of semantics from knowledge bases with tensor factorization methods. Experiments with the TREC dataset demonstrated the effectiveness of the proposed system. The proposed system has the potential to be adapted in other machine learning and medical informatics applications, like recommender systems, ontology learning, bioinformatics, etc. In our future research, we will (a) evaluate the performance of the proposed system with other medical IR datasets; (b) explore the feasibility of integrating the proposed tensor based latent semantic relevance model with the probabilistic tensor decomposition framework to further enhance the performance of the medical IR system. This approach has been shown to be effective in improving the retrieval of medical information. Factorization is a powerful approach to improving informationretrieval in the medical domain. Its ability to identify latent factors in the data and customize the retrieval process to specific medical domains makes it a valuable tool for healthcare professionalsand researchers.

REFERENCES

[1] Fei Song, W. Bruce Croft "A General Language Model for Information Retrieval" Proceedings of the eighth international conference on information and knowledge management, 1999
[2] Miguel A. Callejas P., Yue Wang, Hui Fang "Exploiting Domain Thesaurus for Medical Record Retrieval 1" NIST, 2012
[3] S. T. Bhairnallykar and V. B. Gaikwad. Content based medical image retrieval with SVMclassification and relevance feedback. International Journal of Applied Information Systems, 2013
[4] A. Castellanos, J. Benavent, X. Benavent, and A. Garc´ıa-Serrano. Using visual concept features in a multimodal retrieval system for the medical collection at ImageCLEF2012. In Working Notesof CLEF 2012, 2012
[5] L. Goeuriot, L. Kelly, G. J. Jones, H. Müller, and J. Zobel, ‘‘Report on the SIGIR 2014 workshopon medical information retrieval (MedIR),’’ ACM SIGIR Forum, vol. 48, no. 2, pp. 78–82, 2014
[6] A. Otegi, X. Arregi, O. Ansa, and E. Agirre, ‘‘Using knowledge-based relatedness for informationretrieval,’’ Knowl. Inf. Syst., vol. 44, no. 3, pp. 689–718, 2015.
[7] B. Koopman, G. Zuccon, P. Bruza, L. Sitbon, and M. Lawley, ‘‘Information retrieval as semantic inference: A graph inference model applied to medical search,’’ Inf. Retr. J., vol. 19, no. 1, pp. 6–37, 2016.
[8] M. Mørup, ‘‘Applications of tensor (multiway array) factorizations and decompositions in data mining,’’ Wiley Interdiscipl. Rev., Data Mining Knowl.
[9] Discovery, vol. 1, no. 1, pp. 24–40, 2011
[10] S. Al-Saffar, C. Joslyn, and A. Chappell, ‘‘Structure discovery in large semantic graphs using extant ontological scaling and descriptive semantics,’’ in Proc. IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol., Aug. 2011, pp. 211–218
[11] Y. Gkoufas, A. Morou, and T. Kalamboukis. Combining textual and visual information forimage retrieval in the medical domain. The Open Medical Informatics Journal, 5:50–57, 2011