A Survey on Online Symptoms information and health related question for disease inference system

Ms. Pranita Kulkarni [1], Ms. Pradnya Phanse[2], Ms. Bhavana Patil[3], Ms. Srishti[4],
Dr. Mrs. Neeta Deshpande [5], Mrs. Suvarna Kadam[6]

[1,2,3,4] *BE Scholar, Department of Computer Engineering, D Y Patil College of Engineering, Akurdi, Pune, MH, India*

[5, 6] *Asst. Professor, Department of Computer Engineering, D Y Patil College of Engineering, Akurdi, Pune, MH, India*

Abstract

Online healthcare system is defined as web based system which satisfies health seeker needs by providing inference of the related disease this reduces communication gap between health seeker and health advisor. In healthcare accurately and efficiently inferring diseases is nontrivial especially for community-based health services due to vocabulary gap, incomplete information, correlated medical concepts, and limited high quality training samples. It is also very important to identify the discriminant features. Suppose more than one disease having some symptoms then to give the correct inference of which possible disease the health seeker may suffer from finding discriminant features is done using signature mining.

I. INTRODUCTION

The sensational growth of society, escalating cost of healthcare and fast growing computer technologies are together driving more consumer to spend longer time online to explore the heath information through the Internet in the past few years has led to a genuine need, as well as a sense of urge, for both health seeker and health advisor to bridge the communication gap over the Internet. It will provide inference instead of doctors. The reputable portals run by official sectors offers platforms where health seeker can ask health oriented questions while doctors provide knowledgeable and trustworthy answers. Disadvantages of these community based online health services is that it is very time consuming for health seeker to get their posted questions and also solved replies are conditioned on doctors expertise, experiences which may result in diagnosis conflicts among multiple doctors. The proposed system overcomes the disadvantages of existing system.

Our proposed system will provide inference instead of doctors the project is based on machine learning approaches. This project focuses on creating more secure system. The main aim is to reduce the communication gap and also save the time in which the patient can go for other health related issues. This system will provide various facilities related to health in which the user will post the queries related to symptoms and user will get two or three possible diseases from which they may suffer from.

The system is the expert system that uses expert knowledge from human expert to diagnose the diseases. Using deep learning approaches the model comprises of two parts Firstly generation of tags and globally mines the signatures. Signatures refer to interdependent medical attributes. Second these raw features along with The system takes queries related to symptoms as a input in the form of questions. Tags are generated from the given queries using bootstrap algorithm. These tags are compared with symptoms dataset accordingly, inference of disease is provided as a output.
The system provides the probability of diseases refer that can occur based on symptoms provided using signature mining and sparse deep learning methodology.. The system will work on extracting the discriminant features for the diseases having same symptoms. This will provide correct inference of the related disease to the health seeker

II. LITERATURE REVIEW

Sr. No	Author Name	Key Features	This paper refer to:
1	Nut Limsopatham, Craig Macdonald, ladh Ounis	• Focusing on the bag-of -words model it is used in document classification where the frequency of occurrence of each word is used as a feature for training a classifier • Ratio between the retrieval performance predictors computed on BoW and BoC	• Bag of words • Bag of concepts
2	Ceyhun Burak Akgul, Ahmet Ekin, Devrim Unay	• The framework helps in diagnosis of Alzheimer's disease from magnetic resonance images using visual similarity and user feedback. • It indirectly employs the distance to support vector machine decision boundary • With the experiments on open access series of imaging studies neuroimaging database establish that even with weak global visual descriptors and small training sets this framework has better diagnostic performance	• Nearest neighbour approach • SVM algorithm .
3	Shuang Hong Yang, Steven P. Crain, Hongyuan Zha	• This technique seeks to close the gap at the thematic level via topic adaption • The proposed LDA model explicitly encodes the interplay between topic and technicality hierarchies, providing an effective topic level bridge between lay and expert documents.	• LDA Model
4	Tom Chao Zhou, Michael R. Lyu, Irwin King	• To Bridge the gap a new approach has been developed to question Routing questions to participants who are providing answers. • The experimental result obtained from an evaluation and answers dataset demonstrate high feasibility of question routing . • Perform a systematically comparison on how different types of features contribute to the final results and show that question user relationship. • Find both global and local features play a crucial role in the classification task .	• Classification algorithms for eg, SVM
5	Yi Zhang, Bing Liu	• System provides semantic text classification of disease reports using supervised classification techniques. • Show both keywords and sentence semantic features are very useful for classification. The aim is to classify sentences that report possible disease outbreaks . • They have both keywords and semantic	• SVM • Naive Baise algorithm

		features for classification.	
		• This features are extracted by recognizing the named entity from sentences and finding the syntactic and semantic relationship between the words constructing the dependency trees.	
6	R. Fakoor, F. Ladhak, A. Nazi ,M. Humber	• Unsupervised feature learning can be used for cancer detection and cancer type analysis gene expression data. • Applying data from various types of cancer to automatically form features which enhance the detection and diagnosis of specific one. • They reducing the dimensionality of feature space still produces accurate classification • Developed generalized version of cancer classifier by using sparse autoencoder method to learn feature representation from unsupervised data. • By using PCA featured representation in compressed form.	• Sparse autoencoder method • PCA • SVM with Gaussian kernel • Softmax regression

III.
PROPOSED SYSTEM APPROACH

The proposed system is developed to improve the efficiency. The system is self –learning so that if any query is repeated again with same parameter then he/she will get the answer directly. The main goal of proposed system is to identify the DISCRIMINANT FEATURES of diseases.

For collecting information of discriminant features in detail proposed system will have the end to end communication of doctor and health seeker.

Our system eliminates the problem of existing system. As follows:
• The user gets the answer in very small amount of time.
• It also saves the money.
• It also identifies the discriminant feature of disease by deep learning.
• As chat application is provided by the system it is useful in bridging the vocabulary gap , incomplete infor-mation is analysed   .
• This application will also help in discriminant feature analysis.
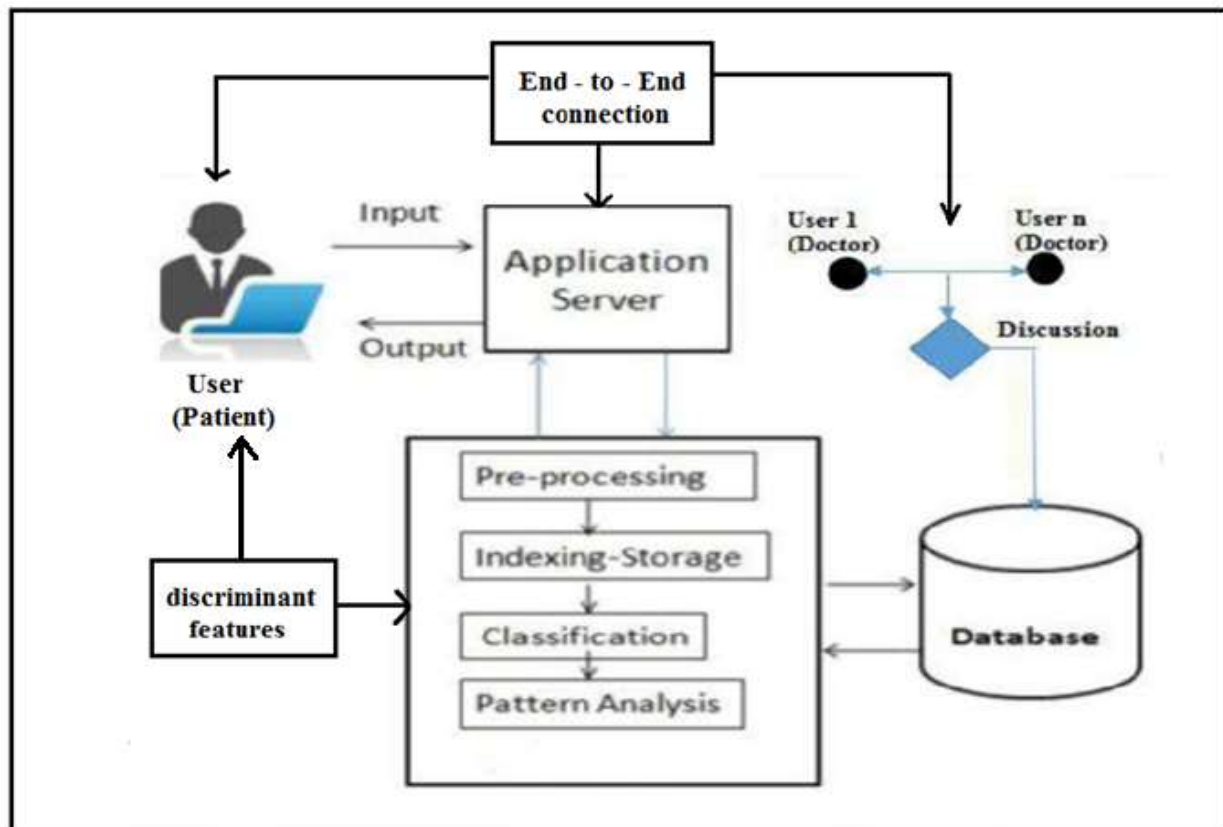
IV.   SYSTEM ARCHITECTURE

Fig.1: Disease inference system

V.   CONCLUSION

Our model is able to identify the discriminant features of disease it also reduces the communication gap between health seeker and doctor as our system provides end to end communication through text classification tags are generated which are provided as input to neural network algorithm for possible inference of disease it also identifies discriminant features of different diseases also reduces the communication gap between health seeker and health provider in this ay performance of system increases.

VI.   REFERENCES

[1]   N. Limsopatham, C. Macdonald, and I. Ounis, \Learning to combine representations for medical records search," in Proc. Int. ACM SIGIR Conf. Res. Develop. Inf.Retrieval, 2013, pp. 833

[2]   C. B. Akgul, D. Unay, and A. Ekin, \Automated diagnosis of Alzheimer's disease using image similarity and user feedback," in Proc. ACM Int. Conf. Image Video Retrieval, 2009, p. 34.

[3]   S.-h. Yang, S. P. Crain, and H. Zha,\Bridging the language gap: Topic adaptation for documents with different technicality," in Proc. 14th Int. Conf. Artif. Intell. Statist., 2011, pp. 823

[4]   T. C. Zhou, M. R. Lyu, and I. King, \A classification-based approach to question routing in community question answering," in Proc. 21st Int. World Wide Web Conf., 2012, pp. 783

[5]   Y. Zhang and B. Liu,\Semantic text classification of disease reporting," in Proc.30th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2007, pp. 747.

[6]   R. Fakoor, F. Ladhak, A. Nazi, and M. Huber, \Using deep learning to enhance cancer diagnosis and

classification," presented at the Int. Conf. Mach. Learn., Atlanta, GA, USA, 2013

VII.   PLAGIARISM REPORT