iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50588

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

AI Based Disease Predictor Model

Miss. Vaishnavi Balasaheb Gase[1], Mr. Vipul Rajendra Ande[2], Miss. Sakshi Rajendra Shirbhate[3], Miss. Vaishnavi Harishchandra Patil[4], Prof. Yogesh R. Shelokar[5]
[1, 2, 3, 4]*UG Scholar,* [5]*Professor, Computer Science & Engineering, Dr. Rajendra Gode Institute of Technology and Research, Amravati, Maharashtra, INDIA*

Abstract: ***Diseases tracing plays important role in daily life. Every one cares about their own health. According to some social study, lot of people spends their time on online searching of health related issues. By browsing they get lot of information about the medical concepts and health related issues. Normally, people use Google to search their queries and that search engine respond them with the answer but that answer is in scattered format. User does not gets the exact answer for their queries. From previous work there has been vital work on the information needs of health seekers in terms of questions and then select those that ask for possible disease of their manifested symptoms for further analytic. To resolve such issues an extensive experiments on a real-world dataset labeled by online doctor's show the significant performance. In this project, further restructuring of the question and answer has been done in order to get the exact answer of query. A tag mining framework for health seekers will be proposed; aim to identify discriminant features for each specific diseaases.***
Keywords: SVM (Support Vector Machine), sparse deep learning, Classifiers, Querying, Signature mining.

I. INTRODUCTION

In today's era, each and every human-being on earth depends on medical treatment and medicines. Every day we can hear some new diseases or new symptoms of the existing disease being discovered. But with the growing number of diseases and their symptoms, everyone cannot manage to be updated with it. So to deal with such situations, we are developing an android application "Smart Health Care" which has a list of large number of diseases, their symptoms, their treatment and medicines required to cure it. One major problem in today's world is hike in Doctor's fee. So the middle class and lower class people are unable to afford for the fee and treatment charges. The application is developed taking this fact in mind. Using this application, one can easily find what disease he/she is infected with by simply inputting the symptoms faced.

There are some other features such as inquiring about the diseases, medicines etc. [4] Recent years have seen a flourishing of community-driven question answering (cQA)Portals, which have emerged as an effective paradigm for disseminating diverse knowledge, seeking precise information, and locating outstanding expert. Around 40% of the questions in the emerging social-oriented question answering forums have at most one manually labeled tag, which is caused by incomprehensive question understanding or informal tagging behaviors. Information extraction from medical text is the basis for other higher-order analytics, such as representation and classification. However, accurately and efficiently inferring diseases is non-trivial, especially for community-based health services due to the incomplete information, correlated medical concepts, and limited high quality training samples. To solve such problems of incomplete information and correlated medical concepts, the dissertation will develop the scheme which studies the user information and health related data. It will infer a learning of the possible diseases given by the questions of health seekers [1].

The prime intention of learning comprises of two key components. The first globally mines the discriminant medical signatures from raw features. The raw features and their signatures serve as input nodes in one layer and hidden nodes in the subsequent layer, respectively.

The second learns the inter-relations between these two layers via pre-training. With incremental and alternative repeating of these two components, our scheme builds a sparsely connected pattern matching architecture with three hidden layers, In this project we are Smart Health Prediction Using Data Mining Technique M.ScFinal [Computer Science] Page 2 taking the symptoms as input and our system compares the symptoms and gives proper disease name with its related doctors. Diseases tracing plays important role in daily life. Every one cares about himself or herself health.

Generally people use Google to search their queries and that search engine respond them with the answer but that answer is in scattered format. User not gets exact answer for his / her queries. We propose a learning scheme to finding the possible diseases given the questions of health seekers.

©IJRASET: All Rights are Reserved

II. LITERATURE SURVEY

1) Here collected more than 900 popular disease concepts from EveryoneHealthy5, WebMD and Medline Plus. Also handled with a wide range of diseases, including endocrine, urinary, neurological and other aspects. Using these disease concepts as queries, we crawled more than 220 thousand community generated QA pairs from Health Tap. In order to increase the effectiveness of our proposed disease inference scheme, we compare it against three state-of-the-art techniques. Most of them can benefit from labeled data; unlabeled data supervised and unsupervised data, which ensures fair comparison. This technique mainly focused on sparse deep learning technique where each layer is incrementally added based on the user‟s need. SVM is implemented here as a classifying tool. Overall it will give a better performance in inferring a disease

2) This paper presents a medical terminology assignment scheme to bridge the vocabulary gap between health seekers and healthcare knowledge. The scheme comprises of two components, local mining and global learning. The former establishes a tri-stage framework to locally code each medical record. However, the local mining approach maIy suffer from information loss and low precision, which are caused by the absence of key medical concepts and the presence of the irrelevant medical concepts. This motivates us to propose a global learning approach to compensate for the insufficiency of local coding approach. The second component collaboratively learns and propagates terminologies among underlying connected medical records. It enables the integration of heterogeneous information. Extensive evaluations on a real-world dataset demonstrate that our scheme is able to produce promising performance as compared to the prevailing coding methods. More importantly, the whole process of our approach is unsupervised and holds potential to handle large-scale data

3) In this paper, we propose a method to enhance cancer diagnosis and classification from gene expression data using unsupervised and deep leaning methods. The proposed method, which uses PCA to address the very high dimensionality of the initial raw feature space followed by sparse feature learning techniques to construct discriminative and sparse features for the final classification step, provides the potential to overcome problems of traditional approaches with feature dimensionality as well as very limited size data sets. It does this by allowing data from different cancers and other tissue samples to be used during feature learning independently of their applicability to the final classification task. Applying this method to cancer data and comparing it to baseline algorithms, our method not only shows that it can be used to improve the accuracy in cancer classification problems, but also demonstrates that it provides a more general and scalable approach to deal with gene expression data across different cancer types..

4) In this paper we have presented a novel temporal event matrix representation and learning framework in conjunction with an in-depth validation of over 40,000 learned latent factor models. The framework has wide applicability to a variety of data and application domains that involve largescale longitudinal event data. We have demonstrated that our proposed framework is able to cope with the double sparsity problem and that the induced double sparsity constraint on the β-divergence enables automatic relevance determination for solving the optimal rank selection problem via an over-complete sparse latent factor model. Further, the framework is able to learn shift invariant high-order latent event patterns in large-scale data. We empirically showed that our stochastic optimization scheme converges to a fixed point and we have demonstrated that our framework can learn the latent event patterns within a group.

III. SYSTEM DIAGRAM

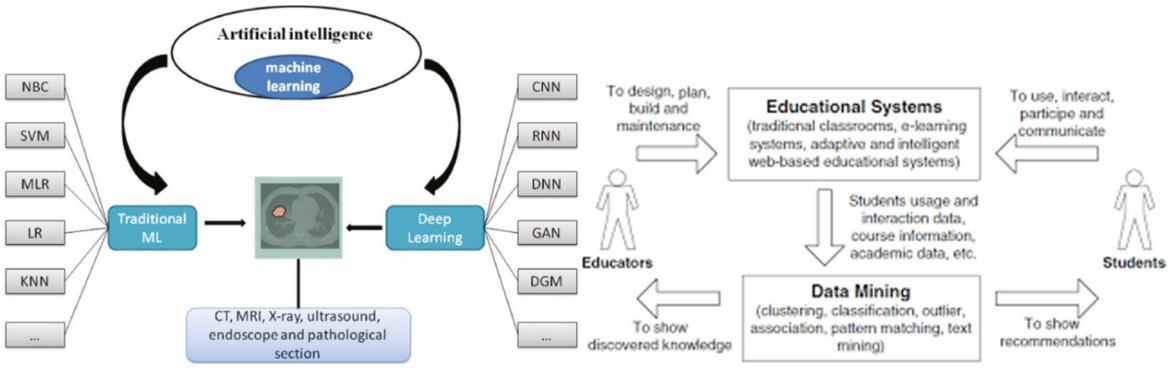


Fig: AI Based Diseases Prediction Model

IV. SCREEN SHOTS

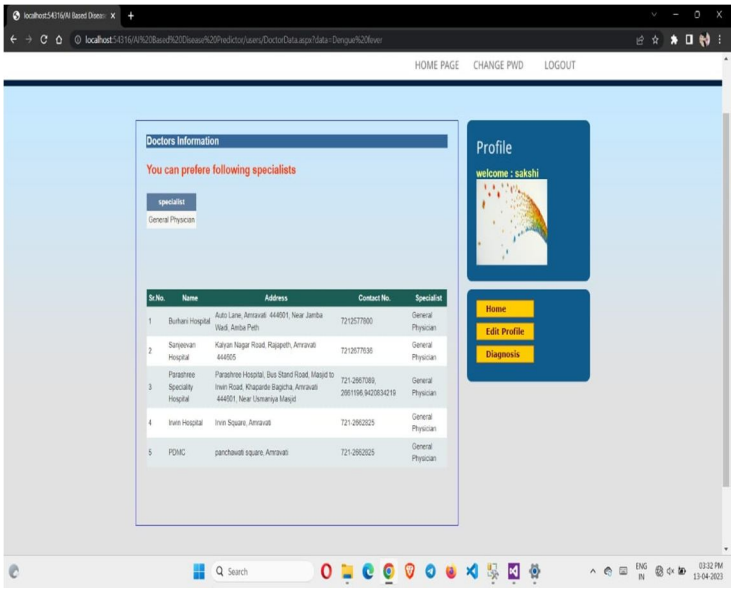


Fig: Home Page

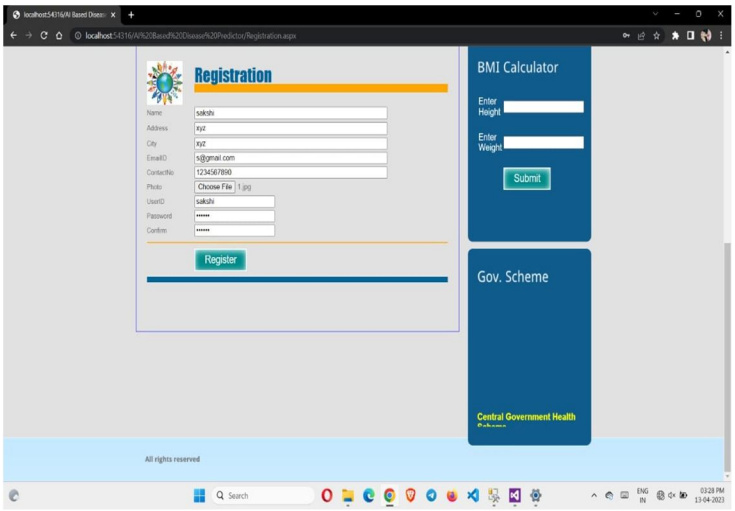


Fig: Registration Page

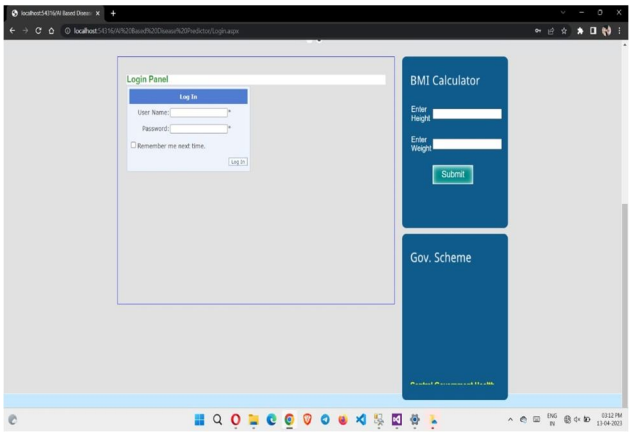


Fig: Login

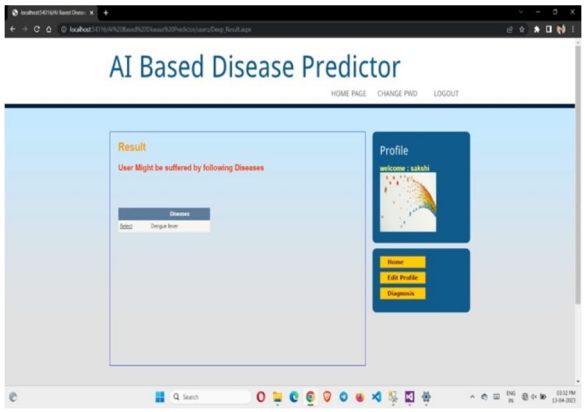


Fig: Result Page

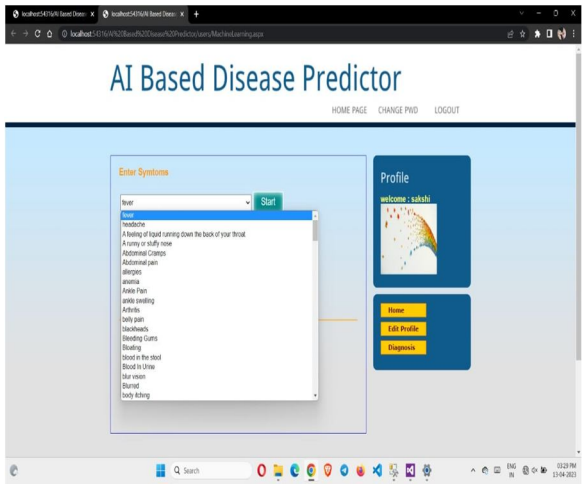


Fig: Symptoms Selection

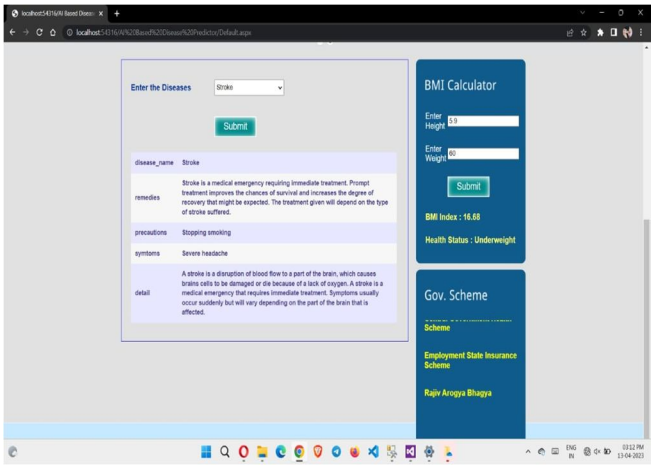


Fig: Diagnosis page

V. CONCLUSION

This paper provided implementation of diseases prediction technique. This paper established that while the current practical use of data mining in health related problems is limited, there exists a great potential for data mining techniques to improve various aspects of Clinical Predictions. Furthermore, the inevitable rise of clinical data will increase the potential for data mining techniques to improve the quality and decrease the cost of healthcare.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

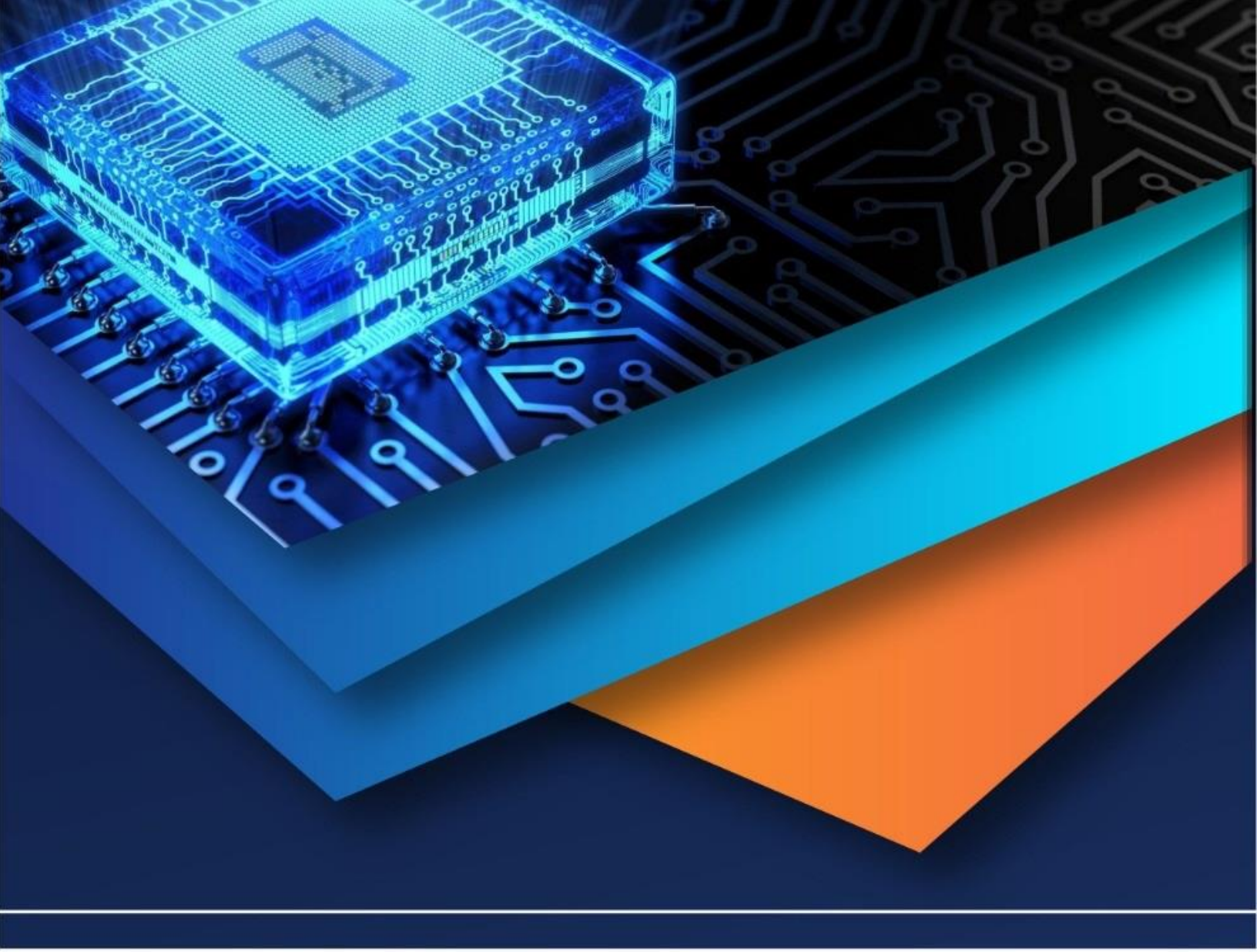
VI. ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my **Prof. Yogesh R. Shelokar** who has in the literal sense, guided and supervised me. I am indebted with a deep sense of gratitude for the constant inspiration and valuable guidance throughout the work.

REFERENCES

[1] LiqiangNie, Meng Wang, Luming Zhang, Shuicheng Yan, Member, IEEE, Bo Zhang, Senior Member, IEEE, Tat-Seng Chua, Senior Member, IEEE ”Disease Inference from Health-Related Questions via Sparse Pattern matching” IEEE Transactions on Knowledge and Data Engineering ,May 2014.
[2] L. Nie, Y.-L. Zhao, M. Akbari, J. Shen, and T.-S. Chua, “Bridging the vocabulary gap between health seekers and healthcare knowledge,” IEEE Transactions on Knowledge and Data Engineering, 2014.
[3] R. Fakoor, F. Ladhak, A. Nazi, and M. Huber, “Using pattern matching to enhance cancer diagnosis and classification,” in Proceedings of the International Conference on Machine Learning, 2013
[4] F. Wang, N. Lee, J. Hu, J. Sun, S. Ebadollah , and A. Laine, “A framework for mining signatures from event sequences and its applications in healthcare data,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013.
[5] F. Wang, N. Lee, J. Hu, J. Sun, and S. Ebadollahi, “Towards heterogeneous temporal clinical event pattern discovery: A convolution approach,” in The ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2012.
[6] David Barbella1, Sami Benzaid2, Janara Christensen3, Bret Jackson4, X. Victor Qin “Understanding Support Vector Machine Classifications via a Recommender System-Like Approach” in Proceedings of the IJSR Conference, 2013.
[7] Lejun Gong∗, Ronggen Yang, Qin Yan, and Xiao Sun, “Prioritization of Disease Susceptibility Genes Using LSM/SVD” in Proceedings of theIJSR Conference, 2011.
[8] M. Shouman, T. Turner, and R. Stocker, “Using decision tree for diagnosing heart Disease patients,” in Proceedings of the Australasian Data Mining Conference, 2011.
[9] Olumurejiwa A. Fatunde(Student Member, IEEE) 1, And Timothy W. KOTIN (Student Member, IEEE) “Refinement of the Facility-Level Medical Technology Score to Reflect Key Disease Response Capacity and Personnel Availability,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)