# Disease Inference System from Health Related Question

Monali Agale[1], Shraddha Birari [2], Sheetal Diwate[3], Priti Lahane[4]

[1] *Student, Information Technology, MET BKC IOE, Maharashtra, India*
[2] *Student, Information Technology, MET BKC IOE, Maharashtra, India*
[3] *Student, Information Technology, MET BKC IOE, Maharashtra, India*
[4] *Assistant Prof., Information Technology, MET BKC IOE, Maharashtra, India*

## ABSTRACT

*Disease inference system from health related questions is able to automatically predict the possible diseases for the given questions asked by health seeker. Medical terminologies are collected in Dataset. Proposed system builds a deep learning architecture which has layered architecture. The deep learning model has L layers. The first layer contains the input raw features and the L layer denotes the output disease types. Hidden layers are unseen from the data. Because of the deep architectures, it repeatedly learns the more abstract patterns layer by layer. This enables the system to extract the underlying connections among medical attributes.*

**Keyword: -** *QA, RDQA, Deep Learning, Disease Inference*

## 1. INTRODUCTION

Medical care and research are the most important part of science for humans. Many of us are surfing Internet to get any disease related information but still they did not get the appropriate information they require so for them proposed system will give accurate information. Disease Inference system which will give the disease information which he/she is suffering on the basis of health related questions. In a less amount of time health seeker will get to know what he/she is facing and that to by sitting at the home. Using the proposed system health seeker will get immediate response. Proposed system's approach is distinctly different in that we are trying to build a general predictive system.

The proposed scheme includes three key components. First, collection of dataset. The second, finds the raw features and their signatures as input nodes in one layer and hidden nodes in the subsequent layer, respectively. Third, it learns the inter-relations between these two layers via pre-training [7] with pseudo-labeled data. Our scheme builds deep learning architecture [1] .

## 2. Literature Survey

Question Answering (QA) is a technique for automatically answer a question asked in natural language. It defined as the task of searching and extracting the text that contains the answer for a specific question, stated in natural language, from a collection of text documents.
QA systems can be classified in two types: Open-Domain Question Answering (ODQA) and Restricted-Domain Question Answering (RDQA) systems. While the first type is concerned with a wide variety of questions (e.g. who

is the president of India?).The second type is properly for a particular area (e.g. what antibiotic was needed for person suffering from fever and cold?).

The current online health resources categories:

- Official portals.

  E.g.WebMD, MedlinePlus

- Community-based          health

  services E.g. HealthTap

## 2.1 WebMD

WebMD is an official portal of online health system which provides the information related to health terms and disease concepts. It is only use question-answer format, there is no disease inference techniques used [4].



**Fig -1**: The illustration of QA example from Official Portals.   (WebMD).

## 2.2 Healthtap

Healthtap offers interactive platforms, where health seekers can anonymously ask health oriented questions while doctors provide the knowledgeable and trustworthy answers.Taking HealthTap as an example, as of January 2014, it had gathered 50 thousand doctors and accumulated more than 1 billion answers, i.e., on average each doctor has online replied approximately 23 thousand times since its foundation in 2010. Fig 2 illustrates one question answer (QA) example [4].



**Fig -2:The illustration of QA example from community-based health services.(HealthTap)**

Disease Inference is different from tags assignment to short questions where tags are direct summarization of given data instances and may explicitly appear in system [3].

## 3. PROPOSED SYSTEM

"Disease Inference System" will infer the possible diseases for the given question asked by health seekers/user.

Deep architectures are composed of multiple levels of non-linear operations, such as in neural networks with many hidden layers.[6]

Deep learning model has L layers with dl(1<=l<=L) nodes in each layer[1]. To be more specific, the first layer contains the input raw features and the $L^{th}$ layers denotes the output disease types, while the intermediate layers are hidden layers, which are unseen from the data. Then they are pre-trained sequentially [7].Supervised fine tuning is step to fit our model to specific tasks.

Deep learning has several advantages. First, it is able to learn Representative and scalable features from other disease types [5]. Take the lung cancer inference learning as an example. When building its classifier, the training data can be liver cancer or other disease samples rather than strictly constrained to lung cancer. Second, reduces the need of feature engineering, one of the most time consuming parts of machine learning practices. Third, deep learning can be adapted to new problems relatively easy.
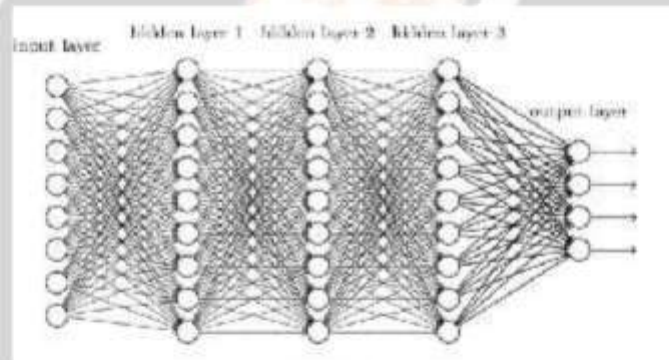


Fig- 2:  Deep Learning Architecture

## 4. CONCLUSION

"Disease Inference system from health related questions" is automatically able to infer the possible disease by doing the analysis of the question asked by health seeker.

## 5. REFERENCES

[1] Liqiang Nie, Meng Wang, Luming Zhang, Shuicheng Yan,Bo Zhang,Tat-Seng Chua,"Disease Inference from Health-Related Questions via Sparse Deep Learning", IEEE transaction on knowledge and data engineering, vol. 27, NO. 8, August  2015.

[2] L. Nie, M. Wang, Z. Zha, G. Li, and T.-S. Chua, "Multimedia answering: Enriching text qa with media information", in Proc. Int. ACM SIGIR Conf. Res. Develop. Inf.Retrieval, 2011, 695704.

[3] L. Nie, Y.-L. Zhao, X. Wang, J. Shen, and T.-S. Chua, "Learning to recommend descriptive tags for questions in social forums", ACM Trans. Inf. Syst., vol. 32, no.1, p. 5, 2014.

[4] Liqiang Nie, Mohammad Akbari ,Tao Li Zhejiang,Tat-Seng Chua," A Joint Local-Global Approach for Medical Terminology Assignment" ,MedIR 2014, July 11, 2014, Gold Coast, Australia.

[5] Y. Bengio, A. Courville, and P. Vincent, "Representation learning:A review and new perspectives," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

[6] Y. Bengio, "Learning deep architectures for ai", Found. Trends Mach. Learn., vol. 2,no. 2, pp. 1127, 2009.

[7] D. Erhan, Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent, and S. Bengio, "Why does unsupervised help deep learning?" J. Mach. Learn. Res., vol. 11, pp. 625–660, 2010.