Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 04, April -2018

MEDI-Q&A: AN ONLINE MEDICAL SYSTEM BASED ON BOOTSTRAP APPROACH.

[1]Omkar Tutare, [2]Pulkit Srivastava, [3]Prof. Sumit Harale, [4]Nachiket Zadap, [5]Shubham Mishra

[1,2,3,4,5]*Dept. Computer Engineering, ICEM Indira College of Engineering and Management Pune, India*

Abstract— *Online healthcare is a 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. Liqiang Nie, Bo Zhang, were proposed user study report on the information needs of health seekers in terms of questions and then select those that ask for possible diseases of their manifested symptoms for further analytic. Next step proposed is a novel deep learning scheme to infer the possible diseases given the questions of health seekers. The proposed scheme is comprised of two key components. The first globally mines the discriminate medical signatures from raw features. The second deems the raw features and their signatures as input nodes in one layer and hidden nodes in the subsequent layer, respectively. Meanwhile, it learns the inter-relations between these two layers via pre-training with pseudo-labelled data. Following that, the hidden nodes serve as raw features for the more abstract semantic bootstrap approach. With incremental and alternative repeating of these two components. Our contribution is a proposed a question answer system for automatic disease inference. Tag have been generated from the user query to be matched with the dataset .System provides inference of having a particular disease based on tags. The proposed system identifies discriminant features for correct diagnosis of disease.*

Keywords—*Signature mining, Semantic Bootstrapping, pseudo-labelled data.*

1. INTRODUCTION

Question Answering (QA) sites have accumulated vast amount of questions and corresponding crowd-sourced answers over time. How to efficiently share the underlying information and knowledge from reliable (usually highly-reputable) answerers has become an increasingly popular research topic. A major challenge in QA tasks is the accurate matching of high-quality answers w.r.t given questions. Many of traditional approaches likely recommend corresponding answers merely depending on the content similarity between questions and answers, therefore suffer from the sparsely bottleneck of QA data. In this project, we propose a novel framework which encodes not only the contents of question-answer (Q-A) but also the social interaction cues in the community to boost the QA tasks. More specifically, our framework collaboratively utilizes the rich interaction among questions, answers and answers to learn the relative quality rank of different answers w.r.t a same question. Moreover, the information in heterogeneous social networks is comprehensively employed to enhance the quality of question-answering (QA) matching by our deep random walk learning framework. Conformity of style throughout a conference proceedings. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. Some components, such as multi-leveled equations, graphics, and tables are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.

2. LITERATURE SURVEY

The primary objective of this paper is to improve the performance of Q&A systems by actively forwarding questions to users who are capable and willing to answer the questions. To this end, we have designed and implemented Social Q&A, an online social network based Q&A system[8]. Social Q&A with security and efficiency enhancements by protecting user privacy and identifies, and retrieving answers automatically for recurrent questions we first report a user study on the information needs of health seekers in terms of questions and then select those that ask for possible diseases of their manifested symptoms for further analytic. We next propose a novel deep learning scheme to infer the possible diseases given the questions of health seekers. The proposed scheme comprises of two key components[6]. The first globally mines the discriminant medical signatures from raw features. The second deems the raw features and their signatures as input nodes in one layer and hidden nodes in the subsequent layer, respectively. Meanwhile, it learns the inter-relations between these two layers via pre-training with pseudo labelled data. Following that, the hidden nodes serve as raw features for the more abstract signature mining. With incremental and alternative repeating of these two components, our

scheme builds a sparsely connected deep architecture with three hidden layers. Overall, it well fits specific tasks with fine-tuning semantic bootstrapping can achieve superb precision while retaining good recall. Nonetheless, the working mechanism of semantic bootstrapping remains elusive. In this paper, we present a detailed analysis of semantic bootstrapping from a theoretical perspective. We show that the efficiency and effectiveness of semantic bootstrapping can be theoretically guaranteed. Our experimental evaluation results substantiate the theoretical analysis[]. We treat an SVO sentence as a three-element triple (subject, sentence pattern, object), and cast the sentence object completion problem as an element inference problem. These elements in all triples are encoded into a unified low-dimensional embedding space by our proposed TRANSFER model, which leverage's the external knowledge base to strengthen the representation learning performance. With such representations, we can provide reliable candidates for the desired missing element by a linear model. Extensive experiments on a real-world dataset have well-validated our model. Meanwhile, we have successfully applied our proposed model to factoid question answering systems for answer candidate selection, which further demonstrates the applicability of the TRANSFER model[7].We first report a user study on the information needs of health seekers in terms of questions and then select those that ask for possible diseases of their manifested symptoms for further analytic. We next propose a novel deep learning scheme to infer the possible diseases given the questions of health seekers. The proposed scheme comprises of two key components. The first globally mines the discriminant medical signatures from raw features. The second deems the raw features and their signatures as input nodes in one layer and hidden nodes in the subsequent layer, respectively. Meanwhile, it learns the inter-relations between these two layers via pre-training with pseudo labelled data. Following that, the hidden nodes serve as raw features for the more abstract signature mining. With incremental and alternative repeating of these two components, our scheme builds a sparsely connected deep architecture with three hidden layers. Overall, it well fits specific tasks with fine-tuning[1].It has been shown that semantic bootstrapping can achieve superb precision while retaining good recall. Nonetheless, the working mechanism of semantic bootstrapping remains elusive. In this paper, we present a detailed analysis of semantic bootstrapping from a theoretical perspective. We show that the efficiency and effectiveness of semantic bootstrapping can be theoretically guaranteed. Our experimental evaluation results substantiate the theoretical analysis[5].We are to highlight the problems associated with missing data in healthcare research and to demonstrate the use of several techniques for dealing with missing values. It is almost impossible to avoid at least some missing values during data collection, which in turn can threaten the validity of the study conclusions. A range of methods for reducing the impact of missing data on the validity of study findings have been developed, depending on the nature and patterns which the missing values may take[1]. iMeds iterative search advisor, which integrates medical and linguistic knowledge to help searchers improve search results iteratively. Such an iterative process is common for general Web search, and especially crucial for medical Web search, because searchers often miss desired search results due to their limited medical knowledge and the task's inherent difficulty. iMeds iterative search advisor helps the searcher in several ways. First, relevant symptoms and signs are automatically suggested based on the searcher's description of his situation. Second, instead of taking for granted the searcher's answers to the questions, iMed ranks and recommends alternative answers according to their likelihoods of being the correct answers. Third, related MeSH medical phrases are suggested to help the searcher refine his situation description[2]. A distributed Q&A system incorporating both social community intelligence and global collective intelligence, named as iASK. iASK improves the response latency and answer quality in both the social domain and global domain. It uses a neural network based friend ranking method to identify answerer candidates by considering social closeness and Q&A activities. To efficiently identify answerers in the global user base, iASK builds a virtual server tree that embeds the hierarchical structure of interests, and also maps users to the tree based on user interests. To accurately locate the cooperative experts, iASK has a fine-grained reputation system to evaluate user reputation based on their cooperativeness and expertise, and uses a reputation based reward strategy to encourage users to be cooperative. To further improve the performance of iASK, we propose a weak tie assisted social based potential answerer location algorithm and an interest coefficient based uncategorized question forwarding algorithm[3]. A probabilistic neural network (PNN) with a self-adaptive approach to perform sentiment analysis on tweets. Probabilistic Neural Network as a multi-layered feed-forward neural network is an apt choice because of its prominent features of adaptive learning, fault tolerance, parallelism and generalization which provide a superior performance. Also, the smoothing parameter of PNN plays a great role for predicting an accurate class of classifier. So a self-adaptive algorithm is used to calculate and optimize the smoothing parameter in our research. Two types of Probabilistic Neural Network models are implemented in the proposed approach. First model of PNN, also called as PNNS has single value of smoothing parameter for whole network. Second model, also called as PNNC has different values of smoothing parameter for each class. The training and testing dataset is collected from twitter using Twitter API[6].A distributed Social-based mobile Q&A System (SOS) with low overhead and system cost as well as quick response to question askers. SOS enables mobile users to forward questions to potential answerers in their friend lists in a decentralized manner for a number of hops before resorting to the server. It leverage's lightweight knowledge engineering techniques to accurately identify friends who are able to and willing to answer questions, thus reducing the search and computation costs of mobile nodes. The trace-driven simulation results show that SOS can achieve a high query precision and recall rate, a short response latency and low overhead. The feedback from the users shows that SOS can provide high-quality answers[4].

3. PROBLEMS IN THE EXISTING SYSTEM

The main challenging problem in health domain is the inter-dependent medical attributes, which is named as signature in this paper. As compared to individual raw feature, signatures are essential cues for diseases. For example, “urinary frequency”, “excessive thirst” and “blurry vision” together likely signal hints to diabetes. While lonely “blurry vision” may be the result of abnormalities present at birth such as near or far sightedness. It can also be a symptom of numerous conditions that do not directly involve the eyes, such as migraine, stroke and side effects of medications. Therefore, the medical signatures are more descriptive than raw features and will significantly reduce the dimension of feature space. However, it is difficult to extract such signatures from individual data instances, as their structures are usually implicitly distributed over a large-scale dataset.

4. ARCHITECTURE

- To build a disease inference scheme that is able to automatically infer the possible diseases of the given questions in community-based health services.
- Analyze and categorize the information needs of health seekers.
- As a byproduct, differentiate questions of this kind that require disease inference from other kinds.
- It is worth emphasizing that large-scale data often leads to explosion of feature space in the lights of n-gram representations, especially for the community generated inconsistent data.
- Distinguished from the conventional sporadic efforts that generally focus on only a single or a few diseases based on the hospital generated records with structured fields, proposed scheme benefits from the volume of unstructured community generated data and it is capable of handling various kinds of diseases effectively.
- It investigates and categorizes the information needs of health seekers in the community-based health services and mines the signatures of their generated data.
- It proposes a sparsely connected deep learning scheme to infer various kinds of diseases. This scheme is pre-trained with pseudo-labeled data and further strengthened by fine-tuning with online doctor labeled data.
- This scheme builds a novel deep learning model, comprising two components.
- The first globally mines the latent medical signatures.
- The raw features and signatures respectively serve as input nodes in one layer and hidden nodes in the subsequent layer.

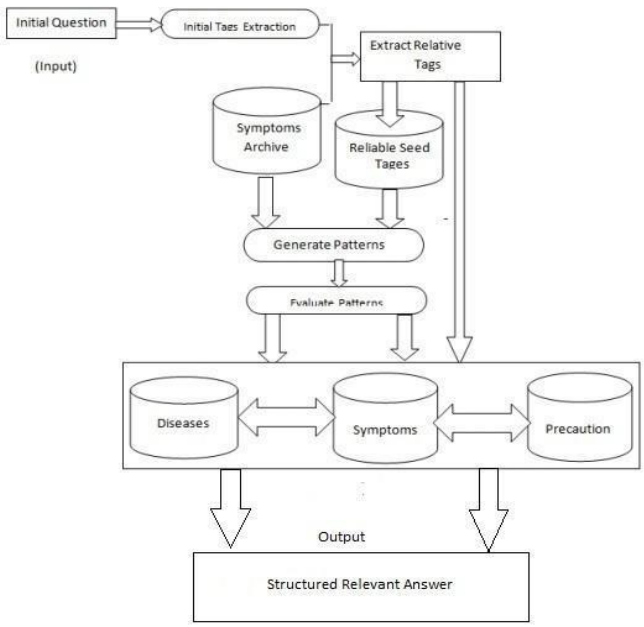

FIGURE :- ARCHITECTURE OF THE SYSTEM

5. ADVANTAGES OF THE PROPSED SYSTEM

- Different from conventional deep learning algorithms, the number of hidden nodes in each layer of proposed model is automatically determined and the connections between two adjacent layers are sparse, which make it faster.
- This model is generalizable and scalable.
- Fine-tuning with a small set of labeled disease samples fits proposed model to specific disease inference.

6. CONCLUSION

In this Era, where almost everything is digitized, the need for medical emergencies has increased to a great extent. Such emergencies can be prevented by creating fast Online QA threads where the individuals can get quick and prompt answers from Medical Experts. The Medical Experts on such threads will be genuine experts and will have all the expertise needed in the field of medicine. Their authenticity will be checked by various verification before allowing them to register as professionals on the thread. The answers provided by the experts will be then stored as patterns and then evaluated as per the query. When a different type of query is searched which is not in the database, a new pattern will be created and evaluated. After the evaluation, the exact inference of the disease will be extracted and displayed. When the new pattern is created, the system is trained for the occurrences of such patterns in the future and will display the answers accordingly.

7. FUTURE SCOPE

In the future, the system can be used for application besides medical system. Creating a thread on a large scale and will require more training. Such system can be trained for more precision with the help of various neural networks algorithms. The algorithms will run in parallel with the base pattern making and evaluation algorithm.

REFERENCE

1) Approaches for dealing with missing data in health care studies Journal paper Kay I Penny and Ian Atkinson

2) On Iterative Intelligent Medical Search Gang Luo Chunqiang Tang IBM T.J. Watson Research Center, 19 skyline drive, Hawthorne, NY 10532, USA

3) iASK: A Distributed Q&A System Incorporating Social Community and Global Collective Intelligence Guoxin Liu, Student Member, IEEE, and Haiying Shen*, Senior Member, IEEE

4) SOS: A Distributed Mobile Q&A System Based on Social Networks IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 25, NO. 4, APRIL 2014 Haiying Shen, Senior Member, IEEE, Ze Li, Student Member, IEEE, Guoxin Liu, Student Member, IEEE, and Jin Li, Fellow,
IEEE

5) Semantic Bootstrapping: A Theoretical Perspective IEEE

TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING 2016 ”-Wentao Wu, Hongsong Li, Haixun Wang, and Kenny Q. Zhu

6) .Sentiment Analysis Using Neural Network 2016 2nd International Conference on Next Generation Computing Technologies (NGCT-2016) Dehradun, India 14-16 October 2016 Akshi Kumar,Ritu Rani

7) Disease Inference from Health-Related Questions via Sparse Deep Learning IEEE Transactions on Knowledge and Data Engineering

Volume: 27, Issue: 8, Aug. 1 2016 ”Liqiang Nie, Meng Wang,
Luming Semantic Bootstrapping: A Theoretical Perspective IEEE

TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING 2016 ”-Wentao Wu, Hongsong Li, Haixun Wang, and Kenny Q. ZhuM. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

8) I Know What You Want to Express: Sentence Element Inference by Incorporating External Knowledge Base IEEE Transactions on Knowledge and Data Engineering Volume: 29, Issue: 2, Feb. 1
2017 ”Xiaochi Wei* , Heyan Huang** , Liqiang Nie, Hanwang

Zhang, Xian-Ling Mao, Tat-Seng ChuaK. Elissa, “Title of paper if known,” unpublished.

9) SocialQ&A: An Online Social Network Based Question and
Answer System IEEE , Volume: 3, Issue: 1, March 1 2017”Guoxin
Liu, Haoyu Wang, Nikhil VithlaniJ. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73

