Survey of Question Answering Approaches in Health Care

Ashwini Rewatkar[#1], M. A. Potey[*2]

[#]Computer Department, Savitribai Phule Pune University, Pune, India
[*]DYPCOE, Akurdi, Pune, Pune, India

Abstract: ***Vocabulary gap between health seekers and health providers which delayed the cross-system operability and the interuser reusability. Approaches for local-global learning are discussed in this paper. Local mining extracts medical concepts from the QA pair itself and maps them to authenticated terminologies. Global learning, works towards enhancing the local mining via collaboratively discovering missing key terminologies and keeping off the irrelevant terminologies by analyzing the social neighbors. The rise of digital technologies has transformed the doctor-patient relationships. In this new era of web, when people struggle with their health concerns, most of them usually explore the Internet to research the problem before and after they see their doctors. A dictionary of vocabulary is constructed as a by-product, which is used as the terminology space for global learning.***

Keywords: Health Care, Question Answering, Local Mining, Global Learning

1. Introduction

Information technology is useful for transforming the health care information from patients to doctors via question answering. Question answering (QA) is a technique for automatically answering a question posed in natural language. The medical QA system is able to answer medical questions according to generic question taxonomy. Compared to keyword-based search systems, it immensely facilitates the communication between human and computer by naturally stating users' intention in lucid sentences. QA performance requires complex natural language processing (NLP) techniques. They take input as, a natural language question and return a short passage or precise words that provide the answers.

Question answering forums attracts the attention of both patients and doctors. The patients are provided with an instant and trusted answer for simple as well as complex health concerns. The doctors are able to increase their reputation amongst their colleagues and patient, strengthen their practical knowledge from interactions with other renowned doctors, as well as possibly attract more new patients. There is a vocabulary gap between health seekers and health provider, to bridge this gap local mining and global learning approaches are used. A tremendous number of medical records have been accumulated in their repositories, and in most circumstances, user may directly locate good answers by searching from these record archives, rather than waiting for the experts' response or browsing through a list of potentially relevant documents from the web. They connect user with provider. Few upcoming community-based healthcare services are HealthTap, HaoDF and WebMD.

Generally, the community generated content, may not be directly usable due to the vocabulary gap. Users do not share same vocabulary. For example HealthTap, is a question answering web site for users to ask health-related questions. The questions are written by any narrative language. The same question may be described in substantially different ways by two individual health seekers. The answer provided by the well-experts may contain phrase with multiple possible meanings, and non-standardized terms.

The tags used often may not be medical terminologies. For e.g., "heart attack" and "myocardial disorder" is same medical terms referred by different experts. It was reported that users had encountered big challenges in reusing the archived content due to the incompatibility between their search terms and those accumulated medical records. Automatically coding of medical records is highly desired using standardized terminologies. It facilitates the medical record retrieval via bridging the vocabulary gap between queries and archives.

Several efforts already exist on automatically mapping medical records to terminologies then they focus on hospital generated health data. Health provider released sources by utilizing either isolated or loosely coupled rule-based and machine learning approaches. The community generated heath records is more general, in terms of inconsistency, complexity and ambiguity. Previously, they simply utilized the external medical dictionary to code the medical records rather than considering the health related terminologies. Constructing a health related vocabulary dictionary to prune the irrelevant terminologies of specific dataset and narrow down the relevant information queried by candidate.

2. Literature Survey

A. Local-Global Learning Approaches

Yi Liang Zhao et al., [1] have presented their work on bridging the vocabulary gap between health seekers and providers which delayed the cross-system operability and the interuser reusability. Most of the current health providers organize and code the medical records manually. There is a growing interest to develop automated approaches for medical terminology assignment. To bridge this gap using inventory of medical terminologies local mining and global learning approaches are jointly utilized.

Local mining aims to locally code the medical records by extracting the medical concepts from individual record and then mapping them to terminologies based on the external authenticated vocabularies.

They establish a tri-stage framework, which includes noun phrase extraction, medical concept detection and medical concept normalization. Global learning complements the local medical coding in a graph based approach. It collaboratively learns missing key concepts and propagates precise terminologies among underlying connected records over a large collection. The existing techniques can be categorized into two categories: rule-based and machine learning approaches. Rule-based approaches play a principle role in medical terminology assignments. They discover and construct effective rules by making strong uses of the morphological, syntactic, semantic and realistic aspects of natural language. Machine learning approaches build inference models from medical data with known annotations and then apply the trained models to unseen data for terminology prediction. The whole process of the proposed approach is unsupervised and it holds potential to handle large-scale data.

B. Graph-Based Reranking Approach

Meng Wang et al., [2] have presented their work using community question answering (cQA) services which gained popularity over the past years. They proposed a novel scheme to answer questions using media data by leveraging textual answers in cQA. It automatically generates a query based on the QA knowledge and then performs multimedia search with the query. Query-adaptive reranking and duplicate removal are performed to obtain a set of images and videos for presentation along with the original textual answer. A graph-based learning process is then formulated based on a regularization framework. Community QA (cQA) has emerged as an extremely popular alternative to acquire information. First, information seekers are able to post their specific questions on any topic and obtain answers provided by other participants. By leveraging area efforts, they are able to get better answers than simply using search engines. Second, in assessment with automated QA systems, cQA usually receives answers with enhanced quality as they are generated based on human intelligence. Third, a tremendous number of QA pairs have been accumulated in their repositories, and it facilitates the conservation and search of answered questions.

Conventional methods usually measure the similarities based on a fixed set of features extracted from media entities, such as colour, texture, shape and bag-of-visual words. It automatically generates a query based on the QA knowledge and then performs multimedia search with the query.

C. Adaptive Probabilities Hypergraph Learning Approach

Tat-Seng Chua et al., [3] have presented their work, which used automatically annotate social questions which unravels the incomplete and biased problems of question tags. For a given question, the scheme first constructs an adaptive probabilistic hypergraph to infer the semantically similar question space. Based on this question space, a collection of probably relevant tags are roughly identified. Comprehensive information cues from users, questions, and tags are seamlessly integrated into this hypergraph. It greatly strengthens annotation by keeping off subjective, ambiguous, and generic tags. Collection of probably relevant tags is roughly identified. Comprehensive information cues from users, questions, and tags are seamlessly integrated into this hypergraph.

D. Content-Based Approach

Liqiang Nie et al., [4] have presented their work, which used content-based approach to automatically predict the search performance. They proposed a query-adaptive graph based learning approach to estimate the relevance probability of each image to a given query. The task is defined as predicting the retrieval effectiveness of a query given a search system and a collection of documents. The probabilities should be close to the ranking-based relevance probabilities. A graph is constructed based on the search results of a query, where vertices are the images and edge weights indicate the pair-wise similarities.

3. Local Mining

Medical concepts are defined as medical domain-specific noun phrases, and medical terminologies are referred to as authenticated phrases by well known organizations that are used to accurately describe the human body and associated components, circumstances and processes in a science-based method. It establishes a tri-stage framework. First to extract noun phrase then identify medical concept and finally it normalize the medical concepts to terminologies.

E. Noun Phrase Extraction

To extract all the noun phrases, it initially assigns part-of-speech tags to each word in the given medical record by Stanford POS tagger. The noun phrases should contain zero or more adjectives or nouns, followed by an optional group of a noun and a preposition, followed all over again by zero or more adjectives or nouns, followed by a single noun. To make up a noun phrase, sequences of tags are matched in a pattern. For example, the following complex sequence can be extracted as a noun phrase:“ineffective treatment of terminal lung cancer”.

F. Medical Concept Detection

In medical concept detection, words are identified using local vocabulary. By this assumption they calculated concept entropy impurity to comparatively measure the domain-relevance of a concept. The concept entropy impurity approach is used to comparatively detect and normalize the medical concepts locally, which naturally construct a corpus-aware terminology vocabulary with the help of external knowledge.

G. Medical Concept Normalization

Although medical concepts are defined as medical domain-specific noun phrases, it cannot be ensured that they are standardized terminologies. Nowadays, there exist numerous authenticated vocabulary dictionaries, for example ICD7, UMLS, and SNOMEDCT. These medical and clinical terminologies were created in different times by different associations for different purposesIn medical concept normalization; it normalizes the medical concept using SNOMED CT vocabulary.

Table I: Normalized Terminologies

Medical Concepts	*Normalized Terminologies*
Birth control	Contraception
Blood loss	Haemorrhage
Breast cancer	Malignant tumour of breast
Heart attack	Myocardial disorder

SNOMED CT vocabulary is used as it provides the core general terminologies for the electronic health record and formal logic-based hierarchical structure. For example “birth control” is mapped with “contraception”. It is essential to normalize is the key to bridge the vocabulary gap. The medical terminology hierarchy will enhance the scheme in two ways. First, it tackles the granularity mismatch problem, where the terminologies found in the medical records are very detailed and specific, while those in the query may be more general and high-level. This is achieved by rewarding the medical concept against the indexed SNOMED CT.

4. Global Learning

Global learning is to complement the local medical coding in a graph-based approach. It collaboratively collects missing key concepts and propagates precise terminologies among underlying connected records over a large collection. The graph-based learning model is used to accomplish terminology selection task and this model is able to simultaneously consider various heterogeneous cues, including the medical record content analysis, terminology-sharing networks, and the inter-expert as well as inter-terminology relationships. A novel global learning model is built to collaboratively enhance the local coding results. This model seamlessly integrates various heterogeneous information cues. [1]

The inter-terminology and inter-expert relationships are used in relationship identification. The medical terminologies in SNOMED CT are organized into acyclic taxonomic (is-a) hierarchy. For e.g., “viral pneumonia” is-an “infectious pneumonia” is “pneumonia” is-a “lung disease”. Terminologies may also have multiple parents. For e.g., “infectious pneumonia” is also a child of “infectious disease”. The graph-based learning models can be broadly categorized into simple graph-based and hypergraph based approaches. They are both built on a graph where vertices are samples, while the simple graph conveys the pair-wise relationship of vertices and overlooks the relations in higher orders.

5. QA Architecture

The main component of QA system summarized in following steps:

H. Question Analysis

The Question Analysis performance consists of classifying and analyzing the questions asked in natural language by the users can ask. This computational process is based on question classification. Question classification is assigning one of the generic patterns to each one of the questions that the user asks the system. This task starts once the user enters the question into the system. Question Analysis task firstly captures the semantics of the users question. These medical answer types can be diseases, symptoms, dose of drugs, and so on, according to the possible answers to the generic questions treated by the system.
These question keywords are directly recognized by applying a set of heuristics to the predicates and the relationships between predicates in the logic form. Like question keywords QA system identifies complex nominal’s and nouns recognized as medical expressions using Medical Named Entities Recognition including their possible adjective modifiers, the rest of the complex nominal’s and nouns including their possible adjective modifiers and the main verb in the logic form. The logic form of a sentence is derived through applying NLP rules to the dependency relationship of the words in the sentence. [13]

I. Document Retrieval

The document retrieval module can retrieve locally stored documents; its remote facility retrieves the relevant documents from medical websites using the Google search service. These medical websites are sorted from the previously defined medical website classification. This medical website classification is performed before the real-time execution of the google search engine and consists of defining the different medical website classes where the system can retrieve the medical documents. Document retrieval engine can start retrieving those relevant documents from medical websites whether there exists or not the association between the searched generic question and the medical websites. When the treated generic question has been related to at least one medical websites class then the Google search engine retrieves the relevant documents according to the question keywords in these medical websites.

J. Relevant Passage Selection

Relevant Passage Selection process consists of extracting the sentences from these medical documents that could answer questions of the user easily. These sentences are extracted by applying a technique based on comparing the question keywords in the documents and, those sentences that at least contain a question keyword are extracted from the document and are evaluated by the next Answer Extraction module that decides if the sentence correctly answers the user question.

K. Answer Extraction

Answer to any question asked by the user is extracted with the help of answer extraction module which extracts the answer by analyzing the sentences extracted by the previous relevant passage selection module. This module is performed by applying the following steps: the first one consists of inferring the logic form of the sentence and identifying the main verb in this logic form; the following step is to verify if this main verb belongs to the set of verbs that can answer the generic question; the third step is the acknowledgment of the medical entities in the logic form; the next step is of comparing if the medical entities searched as the answer is found in the logic form; and finally, the last step is the analysis of the predicates that relate the answer of the candidate, the main verb and the rest of the medical entities in the answer form. This module produces Ranking of Answers. The verb can distinctively relate two medical entities considering this feature as a direct link.

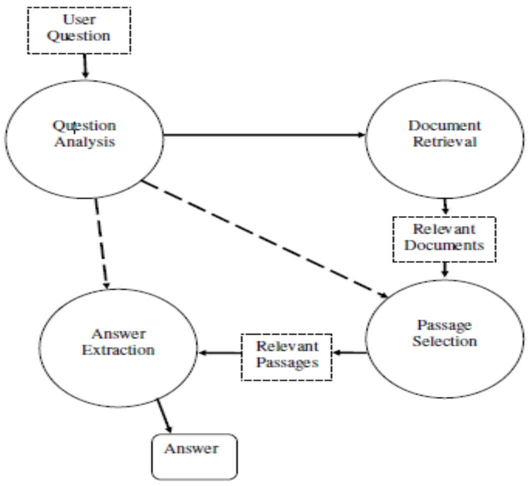

Figure 1: Medical QA System Modulate Architecture

6. Comparative Analysis

Since many questions contain multiple sentences and some of the sentences are uninformative, they are extracts the core sentence from each question. Besides question, answer can also be an important information Clue. The classification is accomplished with two steps. First, they categorize questions based on interrogatives, and second, for the rest questions, they perform a classification using a naive Bayes classifier. The answer medium selection and query selection components requires to learn classifiers based on several training data, and thus the 7,333 QA pairs are split into two parts, a training set that contains 5,866 QA pairs and a testing set of the remaining 1,467 QA pairs. The testing set consists of 800QApairs from WikiAnswers and 667 from Y!A. Classification models are trained with the whole training set, i.e., 5,866 QA pairs. [2] They are tested on the 800 QA pairs from WikiAnswers, 667 QA pairs from Yahoo! Answers, or the both.

Table II: Comparative Study between QA System

Testing set / Feature	Y!A	WikiAnswers	Both
Question-Based Classification	76.41%	80.62%	78.71%
Answer-Based Classification	59.86%	64.72%	62.51%
Integrated of Multiple Evidence	81.72%	84.97%	83.49%

7. Conclusions

A review of the research work for Question Answering in Health care is discussed in this paper. The Local-Global learning approaches are few of the most efficient approaches for reducing the vocabulary gap in health care domain. Some other approaches are Content-based approach, Adaptive Probabilities Hypergraph Learning approach, and Graph-based re-ranking approach which are also discussed in this paper, including QA architecture.

Acknowledgment

We express our thanks to publishers, researchers for making their resource available & teachers for their guidance. We also thank the college authority for providing the required infrastructure and support. Last but not the least we would like to extend a heartfelt gratitude to friends and family members for their support.

References

[1] Liqiang Nie, Yiliang Zhao, Mohammad Akbari, Jialie Shen, and T Chua. Bridging the vocabulary gap between health seekers and healthcare knowledge. 2014.
[2] Beyond text qa: Multimedia answer generation by harvesting web information. L. Nie, M. Wang, Y. Gao, Z.-J. Zha, and T.-S. Chua. IEEE Transactions on Multimedia, 36(4):8509-8522, 2013.
[3] R. Learning to recommend descriptive tags for questions in social forums. L. Nie, Y.-L. Zhao, X. Wang, J. Shen, and T.-S. Chua. In ACM Transactions on Information System, pages 669-674. ACM, 2014.
[4] Oracle in image search: A content-based approach to performance prediction. L. Nie, M. Wang, Z.-J. Zha and T.-S. Chua. ACM Transactions 2012.
[5] R. L. Cilibrasi and P. M. B. Vitanyi. The google similarity distance. IEEE Transactions on Knowledge and Data Engineering, page 389, 2007.
[6] S. Fox and M. Duggan. Survey, pew research center,. In Health online 2013, volume 2,pages 987-990. IEEE, 2013.
[7] Harvesting visual concepts for image search with complex queries. L. Nie, S. Van, M. Wang, R. Hong, and T.-S. Chua. In Proceedings of the International Conference on Multimedia,, 15(4):557-566, 2012.
[8] Dengyong Zhou, Jiayuan Huang. Learning with hyper-graphs: Clustering, classification, and embedding. In Advances in neural information processing systems, pages 1601-1608, 2006.
[9] Paola Velardi, Michele Missiko, and Roberto Basili. Identification of relevant terms to support the

construction of domain ontology's. In Proceedings of the workshop on Human Language Technology and Knowledge Management-Volume 2001, page 5.Association for Computational Linguistics, 2001.
[10] Borkur Sigurbjornsson and Roelof Van Zwol. Flickr tag recommendation based on collective knowledge. In Proceedings of the 17th international conference on World Wide Web, pages 327-336. ACM, 2008.
[11] Image retrieval via probabilistic hypergraph ranking. Y. huang, q. liu, s. zhang, and d. Metaxas. . In IEEE Conference on Computer Vision and Pattern Recognition,, pages 850-854. IEEE, 2010.
[12] Serguei VS Pakhomov, James D Buntrock, and Christopher G Chute. Automating the assignment of diagnosis codes to patient encounters using example-based and machine learning techniques. Journal of the American Medical Informatics Association, 13(5):516-525, 2006.
[13] Terol, Rafael M., Patricio Martínez-Barco, and Manuel Palomar. "A knowledge based method for the medical question answering problem." Computers in biology and medicine 37.10 (2007): 1511-1521.
[14] Niu, Yun, and Graeme Hirst. "Analysis of semantic classes in medical text for question answering.". Proceedings of the ACL 2004 Workshop on Question Answering in Restricted Domains. 2004.
[15] Rinaldi, Fabio, et al. "Answering questions in the genomics domain."Proceedings of the ACL 2004 Workshop on Question Answering in Restricted Domains. 2004.
[16] Bansal, Pratima, et al. "Bridging the research–practice gap." The Academy of Management Perspectives 26.1 (2012): 73-92.
[17] Leroy, Gondy, and Hsinchun Chen. "Meeting medical terminology needs-the ontology-enhanced Medical Concept Mapper." Information Technology in Biomedicine, IEEE Transactions on 5.4 (2001): 261-270.
[18] Lauría, Eitel JM, and Alan D. March. "Combining Bayesian text classification and shrinkage to automate healthcare coding: A data quality analysis." Journal of Data and Information Quality (JDIQ) 2.3 (2011): 13.
[19] Dozier, Christopher, et al. "Fast tagging of medical terms in legal text."Proceedings of the 11th international conference on Artificial intelligence and law. ACM, 2007.
[20] S. A. Quarteroni and S. Manandhar, "Designing an interactive open domain question answering system," J. Natural Lang. Eng., vol. 15, no. 1, pp. 73–95, 2008.
[21] H. Cui, M.-Y. Kan and T.-S. Chua, "Soft pattern matching models for definitional question answering," ACM Trans. Inf. Syst., vol. 25, no. 2, pp. 30–30, 2007.
[22] R. C. Wang, N. Schlaefer, W. W. Cohen, and E. Nyberg, "Automatic set expansion for list question answering," in Proc. Int. Conf. Empirical Methods in Natural Language Processing, 2008.
[23] L. A. Adamic, J. Zhang, E. Bakshy, and M. S. Ackerman, "Knowledge sharing and Yahoo answers: Everyone knows something," in Proc. Int. World Wide Web Conf., 2008.
[24] G. Zoltan, K. Georgia, P. Jan, and G.-M. Hector, Questioning Yahoo! Answers, Stanford Info Lab, 2007
[25] Y.-C.Wu, C.-H. Chang and Y.-S. Lee, "Cross-Language Video Question/Answering System," in Proc. IEEE Int. Symp. Multimedia Software Engineering, 2004, pp. 294–301