International Journal of Multidisciplinary Research and Modern Education (IJMRME)
ISSN (Online): 2454 - 6119
(www.rdmodernresearch.org) Volume II, Issue I, 2016

DECISION ENGINE FOR HEALTH CARE QUESTIONNARIE ANALYSIS

S. Padmapriya*, M. Poornima, S. Dhivya**, N. Parkavi** & A. Rahamath Nisha****
* Associate Professor, Department of Information Technology, A.V.C College of Engineering, Mannampandal, Mayiladuthurai, Tamilnadu
** UG Scholar, Department of Information Technology, A.V.C College of Engineering, Mannampandal, Mayiladuthurai, Tamilnadu

Abstract:

In healthcare system, there is vocabulary gap between health seekers and health care experts due to occurrence of ambiguity in language processing. The ambiguity is arising at the time queries post on their pages in online health care web portals. And the results are inconsistency, complexity and ambiguity. The major challenges are data access and analytics. So implement the system to provide vocabulary gap and present a new scheme that able to return more than one answers that are well-structured by experts and various related extracted from multiple heterogeneous healthcare sources. Further, should users not be satisfied with the returned search results, our system can automatically route the unsolved questions to the professionals with relevant expertise.

Key Words: Healthcare, Local Mining, Global Learning, Vocabulary Gap & Relevant Expertise

1. Introduction:

Recent studies have exposed that patients prefer online advice rather than embracing doctor's advice passively. There is large number of health care services offer the online services. They are distributed personalized knowledge about health and connecting patients with doctors worldwide via question answering. These forums are very striking to both health seekers and professionals. For professionals, they are able to amplify their reputations among their patients and colleagues, strengthen their practical knowledge from interactions with other well-known doctors, as well as more new patients. The health care community generated content, however, may not be providing correct data due to the vocabulary gap. Users with different backgrounds do not necessarily share the equal vocabulary. Take online service as example, which is a question answering site for participants to inquire and answer health-related questions. The questions are written by patients in description language. The same question may be described in significantly different ways by two individual health seekers. On the other hand, the answers provided by the well-trained experts may have acronyms with multiple meanings, and no standardized terms. Recently, some sites have expectant experts to annotate the medical records with medical information. However, the tags used often vary passionately and medical concepts may not be health related terminologies. For example, "myocardial disorder" and "heart attack" are employed by different experts to refer to the similar medical diagnosis. It was shown that the changeability of community produced health data greatly hindered data exchange, management and integrity. Even worse, it was reported that users had encountered large challenges in reusing the archived content due to the inaptness between their search terms and those accumulated medical records. Therefore, repeatedly coding the medical records with standardized terminologies is highly desired. In this paper provide the contributions are: The first work on repeatedly coding the community generated health data, which is extra complex, ambiguous and inconsistent compared to the hospital generated health data. It proposes the concept to normalize the medical concepts locally, which naturally make a corpus-aware terminology vocabulary with the help of external knowledge. It builds improved global learning model to collaboratively enhance the local coding results. This model seamlessly integrates various heterogeneous information cues.

2. Existing Approach:

In existing there are numerous efforts implemented for automatically analyzing medical records to related terminologies. Most of these work, focused on hospital generated health data or health expert's sources by utilizing rule-based and machine learning approaches for trained data. Compared to these kinds of data, the emerging community created health data is more colloquial, in terms of inconsistency, ambiguity and complexity, which pretense challenges for data access and analytics. Further, most of the prior work simply utilizes the exterior medical dictionary to code the medical records slightly than allowing for the corpus-aware terminologies. Their reliance on the autonomous external knowledge might bring in wrong terminologies. Constructing a corpus responsive terminology vocabulary to reduce the irrelevant terminologies of specific dataset and tapered down the candidates is the tough topics we are facing. In addition, the varieties of heterogeneous cues were often not sufficiently exploited simultaneously. As a result, a robust integrated framework to draw the power from various resources and models is still predictable.

3. Local Mining:

This section provides the details about local mining approach. To accomplish this task, we set up a tri-stage framework. Specifically, specified a medical record, we initially extract the embedded noun phrases. We then identify the medical concepts from these noun phrases using specificity of their measure. Finally, we normalize the detected medical concepts to terminologies

a. Noun Phrase Extraction:

To extract every noun phrases, we initially assign part of speech as to each word in the known medical record by Stanford POS tagger. We then pull out sequences that match a fixed pattern as noun phrases. This pattern is prepared as follows:

(Adjective |Noun)*(Noun Preposition)

(Adjective |Noun)*Noun

In addition to simply pulling out the phrases, we also do several simple post processing to link the variants together, such as singularizing plural variants.

b. Medical Concept Detection:

This phase aims to differentiate the medical concepts from other general noun phrases. We assume that concepts that are relevant to medical domain occur frequently in medical domain and rarely in non-medical ones.

c. Medical Concept Normalization:

Medical concepts are distinct as medical domain specific noun phrases; we cannot ensure that they are consistent terminologies. Take "birth control" as an example, it is recognized as a medical concept by our approach, but it is not a genuine terminology. Instead, we should map it into "contraception". Therefore, it is necessary to normalize the detected medical concepts according to the external suitable standardized dictionary and this normalization is the solution to bridging the vocabulary gap.

There exist numerous authenticated vocabularies, including various NLP processes these medical and clinical terminologies were created in different times by different associations for different purposes. Local mining terminologies may suffer from various problems. The first problem is incompleteness. This is because some type of medical concepts not explicitly present in the medical records is excluded. The second one is the lower precision. This is due to some irrelevant medical concepts explicitly embedded in the medical records, and is mistakenly detected and normalized by the local approach. Another issue, which deserves further discussion here, is the terminology space. It may result in the deterioration in coding performance in terms of efficiency and Effectiveness.

4. Global Learning:

The target of this section is to study appropriate terminologies from the global terminology space to annotate each medical record q in Q. Among prior machine learning methods, graph-based learning achieves promising performance. In this work, we also discover the graph-based learning model to accomplish our terminology selection task, and expect this model is able to concurrently consider various heterogeneous cues, including the medical record content analysis, terminology-sharing networks, and the inter-expert as well as inter-terminology relationships. We will first introduce relationship identification and then we detail how to employ our proposed model to link the underlying connected medical records. Next, we present the optimal solution for our erudition model followed by the label bias estimation. Finally, we discuss the scalability of our method.

a. Relationship Identification:

The inter-terminology and inter-expert relationships be not intuitively seen or implied from medical records. We thus call them as implied relationships. This subsection aims to introduce how to discover these kinds of relationships.

b. Inter-Terminology Relationship:

The medical terminologies in SNOMED CT are prepared into acyclic taxonomic (is-a) hierarchies. For example, “viral pneumonia” is-an “infectious pneumonia” is-“pneumonia” is-a “lung disease”. Terminologies may also have multiple parents.

5. Proposed Work:

Online health seeking has altered the way of health knowledge enhance and scalability. The prior local and global health mining, however, just normally return lists of relevance documents or question answer (QA) pairs, which may overwhelm the seekers or not sufficiently meet the seekers’ expectations. As an alternative, a new system is able to return one multi-faceted answer that is well-structured and exactly provide answers from various terminology networks. Further, should the seekers not be satisfied with the returned search results, our system can automatically route the unsolved questions to the professionals with relevant expertise.

In proposed to overcome the information irrelevance, unstructured and incomplete problems, this demonstration presents a novel system. It automatically organizes all the associated healthcare knowledge into a single view for a given question. The comprehensive answer is an analytical result of heterogeneous and multilingual data sources. These sources can be broadly categorized into health provider released data, expert generated data and patient generated data.

a. Relevance Answer Selection:

The proposed work supports relevance search, so it first identifies the type of question language. It then locates the similar answer-aware question by utilizing the score approach. The expert crafted answers of the first positioned question are returned.

b. Healthcare Topic Prediction:

To better capture the semantics of healthcare data and reduce the feature dimension, proposed work exploits the mining based topic-level features for data representation instead of traditional low level n-gram features.

c. Mobile Intimation System:

Mobile based real time system implement at the time of unstructured queries. Queries are transferred to admin or relevant expertise in the from Short message service. So user can got recommended answers

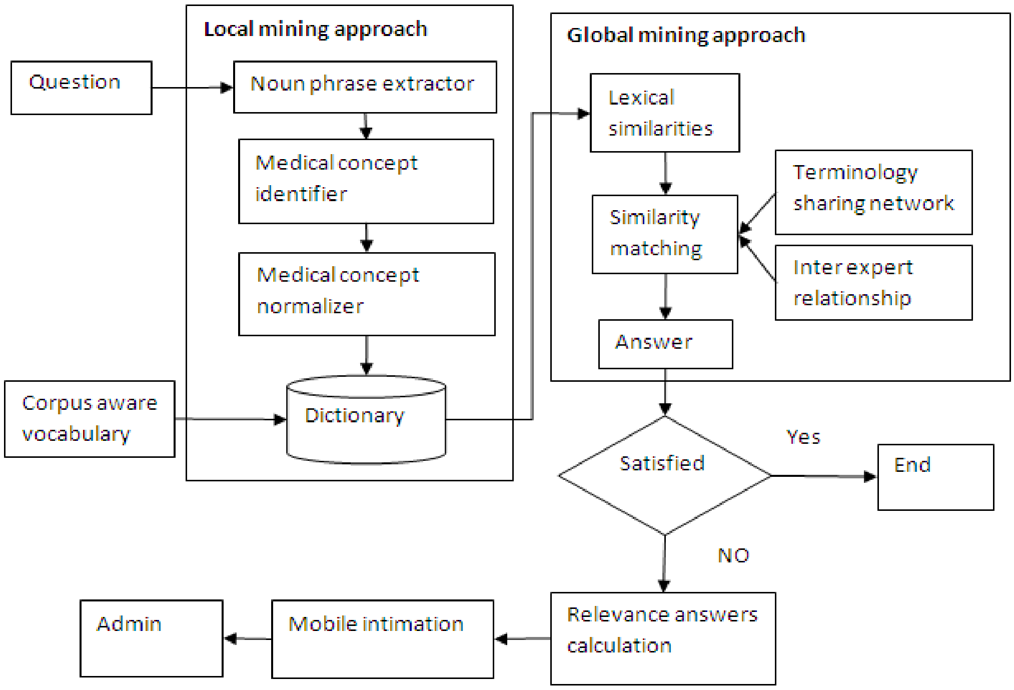
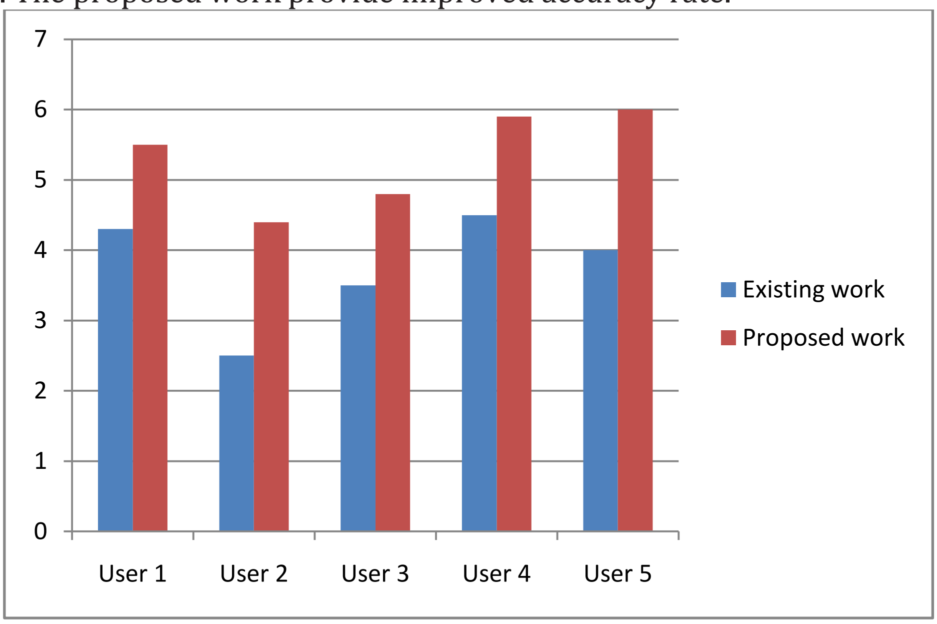


Figure1: Proposed Work

6. Experimental Results:

In experiment, we can calculate the relevance answers accuracy for each user questions. The proposed work provide improved accuracy rate.



7. Conclusion:

This paper presents a new scheme as medical terminology assignment scheme to bridge the vocabulary gap between health seekers and healthcare knowledge. It supports unidirectional connections among patients and experts, which naturally forms the tightly linked communities in terms of similar healthcare concerns, habits and practices. The latest posted questions and answers can be promptly updated and fed within each community, so users can learn key knowledge and offer advice via their personalized pages. Before adding a new question, proposed work encourages users to perform vertical search over archived question-answer pairs, which is a policy to constrain duplicate questions.

8. References:

1. L. Nie, M. Akbari, T. Li, and T.-S. Chua, "A joint local-global approach for medical terminology assignment," in Proc. Int. ACMSIGIR Conf., 2014.
2. L. Nie, T. Li, M. Akbari, and T.-S. Chua, "Wenzher: Comprehensive vertical search for healthcare domain," in Proc. Int. ACM SIGIR Conf., 2014, pp. 1245–1246.
3. AHIMA e-HIM Work Group on Computer-Assisted Coding, "Delving into computer-assisted coding," J. AHIMA, vol. 75, pp. 48A–48H, 2004.
4. G. Leroy and H. Chen, "Meeting medical terminology needs-the ontology-enhanced medical concept mapper," IEEE Trans. Inf. Technol. Biomed., vol. 5, no. 4, pp. 261–270, Dec. 2001.
5. G. Zuccon, B. Koopman, A. Nguyen, D. Vickers, and L. Butt, "Exploiting medical hierarchies for concept-based information retrieval," in Proc. Australasian Document Comput. Symp., 2012, pp. 111–114.
6. E. J. M. Laur ıa and A. D. March, "Combining Bayesian text classification and shrinkage to automate healthcare coding: A data quality analysis," J. Data Inf. Quart., vol. 2, no. 3, p. 13, 2011.
7. L. Yves A., S. Lyudmila, and F. Carol, "Automating ICD-9-cm encoding using medical language processing: A feasibility study," in Proc. AMIA Annu. Symp., 2000, p. 1072.
8. C. Dozier, R. Kondadadi, K. Al-Kofahi, M. Chaudhary, and X. Guo, "Fast tagging of medical terms in legal text," in Proc. Int. Conf. Artif. Intell. Law, 2007, pp. 253–260.
9. L. V. Lita, S. Yu, S. Niculescu, and J. Bi, "Large scale diagnostic code classification for medical patient records," in Proc. Conf. Artif. Intell. Med., 1995.
10. L. S. Larkey and W. B. Croft, "Automatic assignment of icd9 codes to discharge summaries," PhD dissertation, Dept. Comput. Sci., Univ. Massachusetts at Amherst, Amherst, MA, USA, 1995.