Multipurpose Intelligent Medical System Using Machine Learning Approach.

- **ShreyasGhare, ChinmayDiwan, SagarKardile, SanketWaghole,Prof M.K Kodmelwar**

college details : TSSM Bhivrabai sawant College of Enginering and Research

Abstract:

For Storage, analysis and sharing about health issues, Health Information Technology (HIT) is used. Most of the healthcare providers rely on the Health Information Technology (HIT) to provide health related information about disease to public. Now-a-days patients need instant answers about health issues. So, Question answering forum is attracted by both health seekers and healthcare providers. When health seekers ask their question in their language. Expert provides answers using medical terminology. So, there is wide vocabulary gap prevailing between patients and provider. For example, “kidney failure” and “Renal failure” used by different experts to refer same disorders. In such cases patients don't understand the medical terms. So, to avoid this vocabulary gap between health seeker and health provider, this paper presents a system in which user will post their queries in free text format and then system will provide efficient answers to the health seeker.

Keywords:

J.3 LIFE AND MEDICAL SCIENCES
- Biology and genetics
- Health
- Medical information systems

I.2.7 Natural Language Processing
- Discourse
- Language generation
- Language models
- Language parsing and understanding
- Machine translation
- Speech recognition and synthesis
- Text analysis

Introduction

The field of study that focuses on the interactions between human language and computers is called Natural Language Processing, or NLP for short. It sits at the intersection of computer science, artificial intelligence, and computational linguistics. Natural Language Processing is a field that covers computer understanding and manipulation of human language. Machine learning explores the study and construction of algorithms that can learn from and make predictions on data. Machine learning is closely related to computational statistics, which also focuses on prediction-making through the use of computers.Many of the healthcare providers rely on the Health Information Technology (HIT) to provide health information about disease to public. One of the advancement of HIT is EHR used to store patient health information for long time in a digital format. Electronic Health Record (EHR) contains

healthinformation about health seekers medical condition, laboratory test reports, patient history, treatment description-rays and scan report which is securely shared among other department like laboratories, pharmacies, specialists.

Problem Definition:

When patient consults a doctor, sometimes they don't know the meaning of medical term. This leads to vocabulary gap between health seeker and health care provider. To avoid this, we have presented a system in which patient will submit his query and the health provider will return efficient answer.

Objectives

1- To simplify the medical terminologies for better understanding of an end user.
2- To reduce communication gap between healthcare system and users with no expertise in medical field.
3- We present techniques capable of extracting semantic locations from GPS data. We capture the relationships between locations and users with a graph

System Architecture

This system is multi-purpose system. Our intelligent system has been divided into three modules. System provides functionality to contact the ambulance near your location. Also, the system provides meaning of medical terms by referring dataset. Last module of the system is for prediction of disease based on the symptoms.

1. Find meaning of medical term:
   User will have to enter medical term. System will look out that term into dataset. If found then system will fetch its meaning available in dataset. And that meaning will be displayed as output of this module.
2. Prediction of disease:
   For prediction we have provided 20 questions to the user related to their health and any other symptoms. User will answer those questions in yes or no. Then by using naïve bayes algorithm and arff files the system will predict the disease. Also, system will provide recommendation of medicines for that disease and do's and don'ts of that disease.
3. Emergency contacts:
   User will select his/her location. Based on this location system will fetch the ambulance details available in that area.

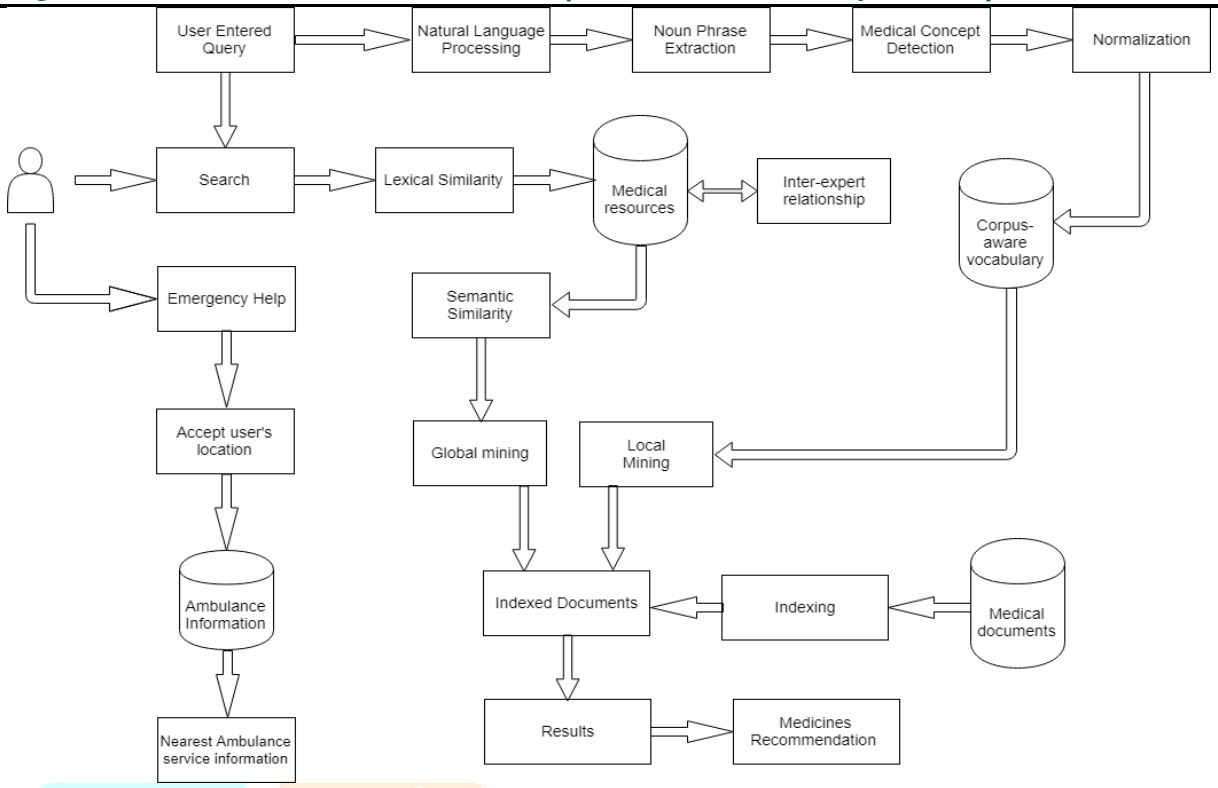

Fig 1: System Overview

Related Works

Author	Paper Title	Description	Limitation
Arshdeep Bahga, and Vijay K. Madisetti, Fellow	A Cloud-based Approach for Interoperable EHRs	This paper proposes an EHR system - Cloud Health Information Systems Technology Architecture (CHISTAR), that achieves semantic interoperability through the use of a generic design methodology which uses a reference model that defines a general-purpose set of data structures and an archetype model that defines the clinical data attributes.	Development of a cloud-based Information Integration and Informatics (III) framework for healthcare applications. This framework will allow development of smart and connected healthcare applications backed by massive scale healthcare data integrated from heterogeneous and distributed health- care systems within a scalable cloud

			infrastructure. Problem was to de velop a data thinning and progressive sampling approach within the CHISTAR infrastructure that will further improve the querying efficiency and accuracy.
Jabez Christopher H. Khanna Nehemiah A. Kannan	A Swarm Optimization Approach for Clinical Knowledge Mining	Rule sets that are extracted with data mining tools and techniques are optimized using heuristic or meta-heuristic approaches in order to improve the quality of the rule bases. In this work, a meta-heuristic approach called Wind-driven Swarm Optimization (WSO) is used. The uniqueness of this work lies in the biological inspiration that underlies the algorithm.	Parameter tuning is still a broad area of research. The optimizer can be tested over many real world and synthetic datasets in order to gain a better perspective to design and develop efficient Rule-based Clinical Decision Support Systems.
Danushka Bollegala, Yutaka Matsuo, and Mitsuru Ishizuka	A Web Search Engine-Based Approach to Measure Semantic Similarity between	To identify the numerous semantic relations that exist between two given words, we propose a novel pattern extraction algorithm and a pattern clustering algorithm. The optimal combination of page counts-based co-	First, a snippet can be a fragmented sentence, second, a search engine might produce a snippet by selecting

	Words	occurrence measures and lexical pattern clusters is learned using support vector machines. The proposed method outperforms various baselines and previously proposed web-based semantic similarity measures on three benchmark data sets showing a high correlation with human ratings. Moreover, the proposed method significantly improves the accuracy in a community mining task.	multiple text fragments from different portions in a document. Because most syntactic or dependency parsers assume complete sentences as the input, deep parsing of snippets produces incorrect results.
Koby Crammer, Mark Dredze, Kuzman Ganchev, Partha Pratim Talukdar	Automatic Code Assignment to Medical Text	The paper presents a system for the assignment of clinical codes to free text radiology reports. Our system assigns a code configuration, predicting one or more codes for each document. We combine three coding systems into a single learning system for higher accuracy.	This approach is likely to help on our task as well but we were unable to test this since the limited number of codes removes any hierarchy. Other approaches have used a variety of NLP techniques.
Liqiang NIE, Yiling ZHAO,Jialie SHEN	Bridging the Vocabulary Gap between Health Seekers and Healthcare Knowledge	The vocabulary gap between health seekers and providers has hindered the cross-system operability and the inter-user reusability. To bridge this gap, this paper presents a novel scheme to code the medical records by jointly utilizing local mining and	To Investigate how to flexibly organize the unstructured medical content into user needs-aware ontology by leveraging the recommended

		global learning approaches, which are tightly linked and mutually reinforced. Local mining attempts to code the individual medical record by independently extracting the medical concepts from the medical record itself and then mapping them to authenticated terminologies.	medical terminologies.
Cem Tekin, OnurAtan, Mihaela van der Schaar.	Discover the Expert: Context-Adaptive Expert Selection for Medical Diagnosis	In this paper we propose an expert selection system that learns online the best expert to assign to each patient depending on the context of the patient. In general, the context can include an enormous number and variety of information related to the patient's health condition, age, gender, previous drug doses, etc., but the most relevant information is embedded in only a few contexts. If these most relevant contexts were known in advance, learning would be relatively simple but they are not. Moreover, the relevant contexts may be different for different health conditions.	It is difficult to track the changes in a clinician's performance by exploiting a recent time window of patient histories, and adapting LEX to discover expertise in other settings.
Gondy Leroy, Hsinchun Chen	Meeting Medical Terminology	Our system is valuable for patients whose knowledge of medical	WordNet is not yet ready to bridge the gap between plain

	Needs—The Ontology-Enhanced Medical Concept Mapper	vocabularies is inadequate to find the desired information, and for medical experts who search for information outside their field of expertise. The Medical Concept Mapper maps synonyms and semantically Related concepts to a user's query. The system is unique because it integrates our natural language processing tool	English and specific medical terminologies.
Mamta Puppala, Tiancheng He, Shenyi Chen, Richard Ogunti	METEOR: An Enterprise Health Informatics. Environment to Support Evidence-Based Medicine"	The aim of this paper is to propose the design and implementation of next generation enterprise analytics platform developed at the Houston Methodist Hospital (HMH) System to meet the market and regulatory needs of the healthcare industry.	Due to the big data clinical reports in the METEOR warehouse, it is a challenge to screen and extract all patients' clinical information from the large number of reports without automated methods.
GokceBanu Laleci, Mustafa Yuksel, and Asuman Dogac	Providing Semantic Interoperability Between Clinical Care and Clinical Research Domains	This paper describes an initial implementation of the Semantic Framework developed within the scope of SALUS project to achieve interoperability between the clinical research and the clinical care domains. In our Semantic Framework, the core ontology developed for semantic mediation is based on the shared conceptual model	It is a difficult to extend the SALUS core ontology and semantic mediation mechanisms, based on CDISC SHARE results, as soon as they are available.

		of both of these domains provided by the Biomedical Research Integrated Domain Group (BRIDG) initiative.	
Liqiang Niey, Tao Lix, Mohammad Akbariy, Jialie Shen	WenZher: Comprehensive Vertical Search for Healthcare Domain	Instead, our multilingual system is able to return one multi-faceted answer that is well-structured and precisely extracted from multiple heterogeneous healthcare sources. 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.	Information over-loading, unstructured and incomplete problems occurs.

Limitation of Study:

In this paper we have used dataset to find the meaning of medical term. If user will enter term that is not available in dataset then system won't be able to fetch the meaning of the term.

Design ofthe Study

- Input: Medical Term, User Location, Symptoms, Historical Report.
- Output: Meaning of medical term, Ambulance details according to users location, disease prediction and medicine recommendation.
- Classifier: Naïve Bayes.
- Natural Language Processing techniques.

Tools Used

- **Software Requirement:**
  - Operating System : windows 8 and above.

- Application Server : Tomcat5.0/6.X
- Language : Java
- Front End : HTML, JSP
- Database : MySQL

- **Hardware Requirement:**
  - Processor - Pentium –III
  - RAM - 1 GB (min)
  - Hard Disk - 20 GB

Statistical Technique Used

We have developed Login and Registration which manages the user profiles, so that the users can post the Medical term and get its meaning. Database stores the information of all users, doctors, reports, also ambulance details are stored in the database.

Algorithm

This paper use Naïve Bayes algorithmsfor prediction we have provided 20 questions to the user related to their health and any other symptoms. User will answer those questions in yes or no. Then by using Naïve Bayes algorithm and ARFF files the system will predict the disease. Also, system will provide recommendation of medicines for that disease and do's and don'ts of that disease.

Our Approach:

This system is multi-purpose system. Our intelligent system has been divided into three modules. System provides functionality to contact the ambulance near your location. Also, the system provides meaning of medical terms by referring dataset. Last module of the system is for prediction of disease based on the symptoms.

1. Find meaning of medical term:
   User will have to enter medical term. System will look out that term into dataset. If found then system will fetch its meaning available in dataset. And that meaning will be displayed as output of this module.
2. Prediction of disease:
   For prediction we have provided 20 questions to the user related to their health and any other symptoms. User will answer those questions in yes or no. Then by using Naïve Bayes algorithm and ARFF files the system will predict the disease. Also, system will provide recommendation of medicines for that disease and do's and don'ts of that disease.
3. Emergency contacts:
   User will select his/her location. Based on this location system will fetch the ambulance details available in that area.

Experiment Result:

This system will return meaning of medical terms, ambulance details of users location, prediction of disease, medicine recommendation of predicted disease, do’s and don’ts for predicted disease.

Future scope:

This project is used for non-expert user. it is difficult to understand the meaning of medical terminologies. So, we proposed a system to minimize the communication gap between health seeker and health provider. Effective answers will be provided to the queries raised by user. In future, we are trying to develop the process by uploading images and videos.

Acknowledgment: (optional)

Conclusion:

The system provides way to overcome communication gap between healthcare provider and heath seeker. The efficient answers will be provided to user by using machine learning approach. System provides results with accuracy around 90% as the question is passes to classifier, Naive Bayes.

Reference:

- A Web Search Engine-Based Approach to Measure Semantic Similarity between Words.
http://ieeexplore.ieee.org/document/5582093/
- Automatic Code Assignment to Medical Text
http://portal.acm.org.sci-hub.cc/citation.cfm?doid=1572392.1572416
http://talukdar.net/papers/cmc_bionlp07.pdf
- Bridging the Vocabulary Gap between Health Seekers and Healthcare Knowledge
http://ieeexplore.ieee.org/document/6843980/
- Discover the Expert: Context-Adaptive Expert Selection for Medical Diagnosis
http://medianetlab.ee.ucla.edu/papers/Tekin_TETC2015.pdf
- Providing semantic interoperability between clinical care and clinical research domains.
https://www.ncbi.nlm.nih.gov/pubmed/23008263