

Semantic Annotation on Medical Records

Sayantani Ghosh and Samir Kumar Bandyopadhyay*

¹Research Institute Pasteur, Russia

²Agrophysical Research Institute, Russia

³All-Russian Institute for Agricultural Microbiology, Russia

Submission: February 04, 2017; **Published:** March 24, 2017

***Corresponding author:** Samir Kumar Bandyopadhyay, Senior Member IEEE, Department of Computer Science and Engineering, University of Calcutta, India, Email: 1954samir@gmail.com

Abstract

Statistical and machine learning approaches are used for recognizing various types of information included within natural language texts. Such applications prepared domain data for training and testing but it is difficult to get these data due to law. For searching unstructured information in patient medical records semantic annotated text is incorporated in patient data. In this paper an outline of method on semantic annotation on medical records is proposed.

Keywords: Electronic medical record (EMR); Semantic annotation; NLP and IE System

Introduction

The term “annotation” implies to attach data to some other piece of data. Semantic annotation is a technique to enrich target’s information by attaching additional information to various concepts such as people, things, places, organizations etc. in a given text or any other content. The fast growth of shared digital image and video collections require new effective methods to search and retrieve in these collections. It is necessary to understand the semantic content of images and to build linguistic descriptions of their content in an automatic way as an interface between humans and computers. Annotation can be categorized into:

Manual annotation

Where an object is manually annotated with explicative keywords and is shared with the world. Like a web page containing the image of flower vase of rose if annotated with the keyword “ROSE” and shared online will provide a better understandability of the image [1]. Although manual annotation yields maximum accuracy, yet it quite a laborious and time consuming task as the keywords are usually generated based on human imagination. In addition, it may generate revivification error.

Semi-automatic annotation

In this case an image is semantically annotated by keywords selected automatically from a pre-defined set. It is a two-

step process. In the first step annotator annotates the object (especially Web Resources) manually. In the second step the tagged object is scrutinized by the annotation tools based on the set of annotation rules and generates candidate annotation. This technique aims at gaining a combinative benefit of the accuracy of manual annotation technique and the efficiency of the automatic annotation technique.

Automatic annotation

The entire process of annotation is carried out by the annotation system. Based on the specified criteria by the user, the annotation tool defines the anchor and the shape of the annotation and then assigns specific property to it.

Review Works

Active Semantic Electronic Medical Record is used in health care which aims to reduce medical errors, improve physician efficiency, improve patient safety and satisfaction in medical practice, improve quality of billing records leading be better payment, and make it easier to capture and analyze health outcome measures. Two main approaches to the task of annotating text data are manual annotation and manual correction of automatically assigned labels. Manual annotation is a time-consuming and expensive process and it is error-prone. These data are very hard to extend and modify and every change imposes extra effort for checking the consistency of the result.

Some researchers use cancer patient records annotated with information related to their important clinical records [2]. About clinical relations, entities, and temporal information. Semantic manual annotation is described in [2-4].

Outline of Method

Many doctors are now investing in electronic medical records (EMR) systems which allow them to have all patient data at their fingertips. Due to complexity of today's healthcare, medical practitioners face challenges for managing their patient's record for improving the quality of care needed by the patient. Natural language processing (NLP) systems can help to monitor patient care by automated processing of medical records and extraction of quality-of-care indicators. NLP systems is based on machine learning training system that uses annotated documents from the intended medical domain. Both rule-based and machine learning-based NLP tools require rigorous testing before deployment within a new medical field.

In contrast to many annotated text corpora which were built by manually assigning labels to appropriate text fragments, we decided to adopt an existing IE system [1] for the task.. However, IE systems do not contain all the information needed. To solve the problem, the idea of combining two extraction grammars was introduced. On the basis of the existing grammar a simplified version, consisting of a subset of the original rules, was created. The final information associating text fragments with semantic labels is the effect of a comparison of the results

of these correlated IE grammars. The limits of text fragments representing attribute values are recognized in the simplified grammar, while their correctness is justified by more complex grammar rules which describe the contexts in which a particular phrase has a desired meaning.

Conclusion

The healthcare industry is changing towards the use of electronic medical records systems to manage the increasingly large amount of patient data for reducing medical errors. In addition to patient data there is a large amount of data describing procedures, treatments, diagnoses, drugs, insurance plans, coverage, formularies and the relationships between these data sets. This paper proposes an outline of method of semantic annotation on medical records.

References

1. Mykowiecka A, Marciniak M, Kupść A (2009) Rule-based information extraction from patients' clinical data. *Journal of Biomedical Informatics* 42(5): 923-936.
2. Roberts A, Gaizauskas R, Hepple M, Demetriou G, Guo Y, et al. (2009) Building a semantically annotated corpus of clinical texts. *J Biomed Inform* 42(5): 950-966.
3. South BR, Jones M, Garvin J, Samore MH, Chapman WW (2009) Developing a manually annotated clinical document corpus to identify phenotypic information for inflammatory bowel disease. *BMC Bioinformatics* 10 (Suppl 9): S12.
4. Marciniak M, Mykowiecka A (2011) Construction of a medical corpus based on information extraction results. *Control and Cybernetics* 40(2): 337-360.



This work is licensed under Creative Commons Attribution 4.0 License
DOI: 10.19080/AIBM.2017.02.555600

Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats
(Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission

<https://juniperpublishers.com/online-submission.php>