Annotation of Discharge Summaries Based on Selection Criteria

Amber Stubbs, MA, James Pustejovsky, PhD
Brandeis University, Waltham, MA

Abstract

This paper describes an annotation framework designed to identify whether patients meet study selection criteria based on textual evidence from their medical discharge summaries. The annotation method used was light and user-friendly, rather than a comprehensive linguistic annotation of all events in the summaries, and as such did not require extensive training for the annotators.

Introduction and Background

The selection of patients for medical studies is often a time-consuming and expensive task, requiring extensive medical knowledge and reasoning. Automating or semi-automating this task would be of great benefit to researchers. One approach to this has been to codify selection criteria[1], while we approached this task by applying machine learning techniques to a corpus of discharge summaries where selection criteria have been annotated.

While other groups have annotated EHRs,[2,3] to the best of our knowledge this is the first exploration of a two stage methodology for knowledge-rich annotation tasks: 1) a light, user-friendly annotation against a coarse-grained specification followed by 2) the automated enhancement of richer specifications to the initial annotation which are then used for machine learning techniques. The first stage of our process is described here.

Methods

The construction of the annotated corpus involved four representative selection criteria: all participants must be under 55; all must be diabetic; cases must have had a cardiac event in the past 2 years; and controls must have no history of cardiac events.

Our annotation scheme specifies three tag types used for the criteria: *Selection_criterion, Modifier*, and *Modifies*. Annotation was done using MAE, a tool developed for this task at Brandeis University[4].

The annotators, two professional medical researchers from Brigham and Womens Hospital, were given 100 hospital discharge summaries and asked to annotate only the words or phrases that were directly related to the specified criteria. Once the keywords were annotated with the *Selection_criterion* tag, the annotators assigned specific attribute values to indicate which criterion the word applied to, and whether or not the criterion was met. If there was context surrounding the annotated extent determining the status of the criterion, the annotators marked this context with the *Modifier* tag, and linked the *Selection_critierion* and *Modifier* tag by means of a *Modifies* tag. "Father DMII+" is an example of a phrase that would require a modifier tag.

Results

After normalizing the annotations for formatting and criterion definitions (*cardiac event* was insufficiently defined in the guidelines, and clarification of what defines a *Modifer* was needed), by using one annotator as the gold standard we obtained precision and recall scores ranging from .57 to .79. We plan on further refining the guidelines and creating another set of annotations, which will greatly improve the agreement scores for individual tags.

More importantly, the Cohen's Kappa scores for patient classification were: recent card. event: 0.8; age under 55: 0.9; diabetic: 0.96; no history of cardiac events: 0.66. Due to repetition in the texts, the overall classifications were quite standardized.

Discussion

By identifying the relevant portions of discharge summaries pertaining to criteria, we are able to create a consistent representation of patient eligibility. This annotation paradigm can be used for representing other selection criteria, and the resulting corpora useful for training and testing semi-supervised learning algorithms.

References

1. Z Luo, SB Johnson, H Chase, R Duffy, C Weng, Semi-automatic induction of semantic classes from free-text clinical research eligibility criteria using UMLS, *Proc of AMIA 2010 Fall Symposium, 487-491.*
2. BR South, S Shen, M Jones et al. Developing a manually annotated clinical document corpus to identify phenotypic information for ibd. *BMC Bioinformatics. 2009 Sep 17;10 Suppl 9:S12.*
3. A Roberts, R Gaizauskas, M Hepple et al. Building a semantically annotated corpus of clinical texts. *J Biomed Inform. 2009 Oct; 42(5):950-66. Epub 2009 Jan 23.*
4. A Stubbs. Multi-purpose Annotation Environment cs.brandeis.edu/~astubbs/mae.html