Q-Map: Clinical Concept Mining from Clinical Documents

Authors : Sheikh Shams Azam, Manoj Raju, Venkatesh Pagidimarri, Vamsi Kasivajjala

Abstract : Over the past decade, there has been a steep rise in the data-driven analysis in major areas of medicine, such as clinical decision support system, survival analysis, patient similarity analysis, image analytics etc. Most of the data in the field are well-structured and available in numerical or categorical formats which can be used for experiments directly. But on the opposite end of the spectrum, there exists a wide expanse of data that is intractable for direct analysis owing to its unstructured nature which can be found in the form of discharge summaries, clinical notes, procedural notes which are in human written narrative format and neither have any relational model nor any standard grammatical structure. An important step in the utilization of these texts for such studies is to transform and process the data to retrieve structured information from the haystack of irrelevant data using information retrieval and data mining techniques. To address this problem, the authors present Q-Map in this paper, which is a simple yet robust system that can sift through massive datasets with unregulated formats to retrieve structured information aggressively and efficiently. It is backed by an effective mining technique which is based on a string matching algorithm that is indexed on curated knowledge sources, that is both fast and configurable. The authors also briefly examine its comparative performance with MetaMap, one of the most reputed tools for medical concepts retrieval and present the advantages the former displays over the latter.

Keywords : information retrieval, unified medical language system, syntax based analysis, natural language processing, medical informatics

Conference Title : ICSWNLP 2018 : International Conference on Semantic Web and Natural Language Processing
Conference Location : New York, United States
Conference Dates : October 08-09, 2018