Abstract

Natural Language Processing (NLP) and Machine Learning concepts are gaining rapid importance in the era of digitalization of data. The value of data keeps changing over time and makes it important to harness that value for performing in depth research in various domains. Extracting information from clinical text helps in automated terminology management, data mining, de-identification of clinical text, research subject identification and studying effect of research on them, predicting the onset and progress of various chronic diseases, disease-treatment-side effect analysis etc. Methods based on NLP and Machine Learning tends to perform better in this area but more experience is required to analyse clinical text than the biomedical literature. The issues and opportunities in information extraction from the clinical
text need to be intensively reviewed to find new avenues in this domain of research.

**References**

- Peter B. Jensen, Lars J. Jensen & Soren Brunak, "Mining EHRs towards better research applications and clinical care", Nature Reviews Genetics, June 2012.
- Sweeney L, "Replacing personally-identifying information in medical records, the Scrub system, Proceedings of AMIA Annual Fall Symposium 1996.
Information Extraction from Clinical Text using NLP and Machine Learning: Issues and Opportunities

- Chapman W, Chu D, Dowing JN, "ConText: An algorithm for identifying contextual features from clinical text," BioNLP 2007: Biological, translational, and clinical language processing, Prague, CZ.

Index Terms

Computer Science

Information Sciences

Keywords

Natural Language Processing  Machine Learning  Clinical Text  Information Extraction

Electronic Health Records

Viterbi Algorithm.