

Medical Concept Extraction: A Comparison of Statistical and Semantic Methods

Nyein Pyae Pyae Khin

nyeinyapyaekhin@gmail.com

University of Computer Studies, Mandalay, UCSM

Khin Thida Lynn

khinthidalinn@ucsm.edu.mm

University of Computer Studies, Mandalay, UCSM

The goal of medical concept extraction is to identify phrases that refer to medical concepts of interest such as problems, treatments and tests from medical documents. In this study, three types of medical concept extraction models are developed and then compared them. The first concept extraction task is mainly based upon semantic features obtained from a domain-knowledge based method using MetaMap, and the other two are machine-learning methods with using sequential classifier Conditional Random Fields (CRF) for both with and without MetaMap outputs as features. Among the three concept extraction models, the combined approach of CRF with MetaMap features obtained the best results.