Overview on Information Extraction.

Bexultan Yelaman.

I read 3 good articles that describe good uses and techniques for using NLP. As an example, it was interesting to read the article A Biomedical Information Extraction Primer for NLP Researchers how to use NLP in medicine, this article describes well the problems and solutions for extracting information from various scientific articles, biological literature and electronic patient records that doctors can use to make information in their fields, this system is called BioIE. BioIE is used to help create database research. This database can have extracted data from drugs and illnesses as well as treatment methods and human tests. In this article, I got to know the NER system, this system helps to extract diseases, medical tests and treatment taken from patient records. They use a semi-standard random field (semi-CRF) to display labels for all tokens in a sentence. The article also uses some of the concept mapping features using existing annotation tools, as well as Brown clustering to form 128 clusters on top of unmarked data. Each improvement of these systems leads to an improvement in the quality of data extraction from scientific research that is on the Internet and has a good rating and reviews.

The second article that I read is closely related to the first, relates to medicine, in particular the use of NLP methods in clinical tasks. The article focuses on the use of neural networks, in particular two implementations of the BERT models. The authors claim that their language models have surpassed past models in solving clinical problems. The article describes the problems of extracting information from electronic medical supplies, where the manual extraction of patient data, drugs, analyzes and other important data related to patient research is very expensive and time-consuming, and for this support of clinical research and applications is strongly recommended to use processing technologies natural language for extracting patient information from clinical notes. Recently, the use of machine learning methods such as RNN and SNN have been used to extract clinical relationships and have shown good results. A very good result was shown by the BERT method, which is pre-trained in a large building without supervision, and then can be precisely configured for the subsequent task with a simple layer on top of the architecture. The BERT method has demonstrated its advantages in extracting information, answering questions, classifying text, and generating text.

The third article was devoted to the use of NLP methods in social networks. The article describes the Lithium Natural Language Processing (NLP) system. Lithium NLP extracts a rich set of information from text, including entities, topics, hashtags, and moods. Current people using social networks use different styles of context, volume and heterogeneous information like (text, emoticons, hashtags, and tags). modern NLP methods can integrate and study information from various sources well. But the problem arises, the problem is that in social networks people will not get a formal style of speech, slangs, phrases and jargons, etc. that are not related to simple grammar. Therefore, according to the authors, modern NLP methods should be developed for such content. For example, multilingual content, modern NLP techniques should be multilingual and support many languages. And also have large-scale data sets – modern NLP systems must be able to work with large-scale data sets, such as social media data, in particular millions of documents. and have a rich set of information extracted from social networks. Lithium NLP is capable of extracting a rich set of information, including entities, topics, hashtags and moods as well as marks. Lithium NLP currently supports several languages, including Arabic, English, French, German, Italian, and Spanish. According to research, Lithium NLP supports large-scale data from several social networks, such as Twitter, Facebook, Linkedin, etc. And all these studies and analyzes as well as the development of NLP methods are designed to process and extract data from social networks.