Research Week 2021

Extraction of Active Medications and Adherence Using Natural Language Processing for Glaucoma Patients

Wei-Chun Lin, MD, MS, Jimmy S. Chen, BA, Joel Kaluzny, MD, Aiyin Chen, MD, Michael F. Chiang, MD, Michelle R. Hribar, PhD
Email: linw@ohsu.edu
Medical Informatics & Clinical Epidemiology, Oregon Health & Science University, Portland, OR

Keywords

Natural language processing, Secondary use of EHR data, Glaucoma, Information extraction

Abstract

Purpose

Accuracy of medication data in EHR is crucial for patient care and research. Previous work has shown frequent errors in medication lists include incomplete records, duplicated prescriptions, and failed discontinuation of medications. Since medication lists are inaccurate, physicians often record medication information in progress notes, which is difficult to automatically extract since notes are written as free-text narratives. In this study, we developed and validated a named entity recognition model to identify current medication and adherence from progress notes. Also, a prototype tool for medication reconciliation using the developed model was demonstrated.

Methods

First, we sampled a dataset with 507 progress notes from office visits at the CEI in 2019, with encounter ICD10 codes associated with glaucoma. The documents were manually annotated for eight categories: Drug Name, Route, Frequency, Dosage, Strength, Duration, Adherence, and Current Medication Use. Next, we used an NLP technique called Named Entity Recognition (NER) to extract eight entities from clinical notes. The NER model was developed in Python using the spaCy library and evaluated by using F1 score, precision, and recall. Finally, the developed NER model was used to extract the patient's current medications from 150 sample progress notes for medication reconciliation.

Results

The custom NER model was trained with 381 progress note documents that were manually annotated with eight named entities and then tested on 126 progress notes. The overall performance of the NER model across all categories was F1 score = 0.955, Precision = 0.951, and Recall = 0.957. Also, we obtained accurate current medications from 150 sample progress notes (F1 score = 0.957).

Conclusions

This study shows that NLP can be used to accurately extract active glaucoma medication from free-text EHR data and help with medication reconciliation. This has implications in improving the data quality and medication reconciliation for glaucoma patients.

Table. The results of the NER model on the test dataset

Entities	**Performance on Test Data**		
	Precision	**Recall**	**F1-Score**
Drug	0.971	0.971	0.971
Frequency	0.972	0.969	0.970
Route	0.948	0.986	0.966
Dosage	0.987	0.998	0.991
Duration	1.000	0.600	0.749
Strength	0.969	0.997	0.982
Adherence	0.803	0.758	0.779
Current Medication Use	0.899	0.919	0.909
Average (micro)	0.951	0.957	0.955