Human-Enabled Healthcare NLP

Discovering Inaccurate Extracted Information from Unstructured Clinical Notes Using Natural Language Processing

UCSF BCHSI

Team Members: Chloe Kim, Layne Wei, Marcel Schaack
Advised by: Dr. Gundolf Schenk, Dr. Gabriel Gomes, Dr. Angela Rizk-Jackson, Eugenia Rutenberg

fi

INTRODUCTION

Despite the recent digitalization of patient information by the use of millions of **electronic health records** (EHR), some of the most important information remains hidden to machines in unstructured clinical notes and reports. **Natural language processing** (NLP) can extract un/structured information to enable precision medicine and predict epidemic trends through big data approaches. However, the rise of new conditions, such as COVID-19, frequently leads to the rise of unstandardized medical terms that are challenging to extract for the current NLP pipelines. Our NLP/Machine Learning (ML) enabled system improves the identification and extraction of medical concepts by integrating human-provided feedback into a **confidence score**. This allows users to give directed feedback on the data correctness.

Apache cTAKES

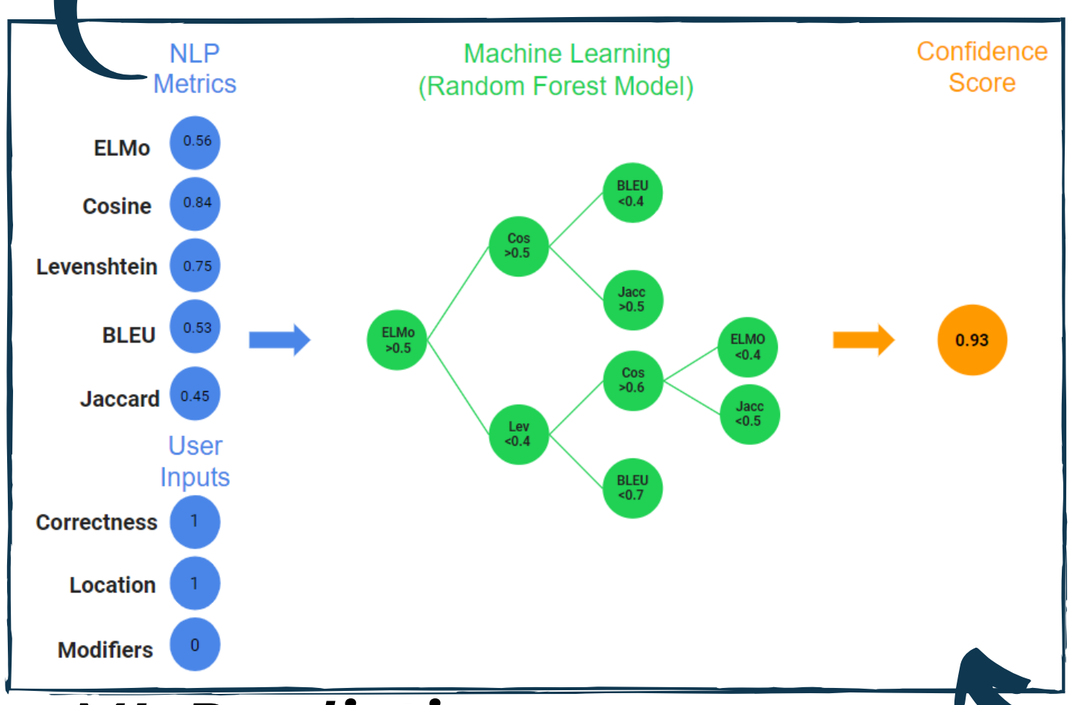
cTAKES (Clinical Text Analysis & Knowledge Extraction System) analyzes unstructured electronic health records and extracts medical concepts and health information

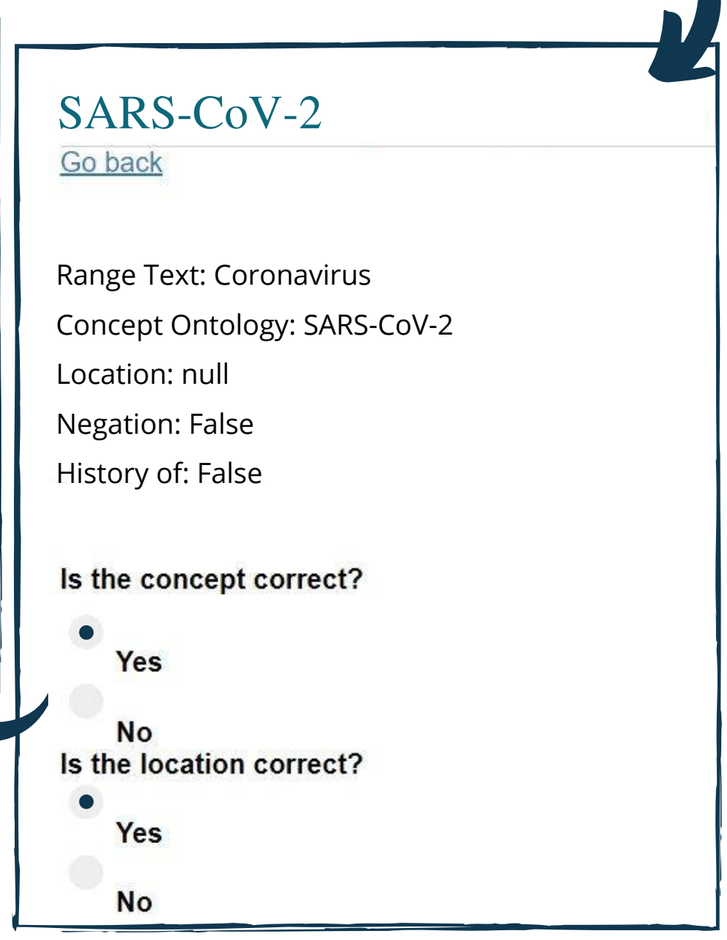
NLP Analysis

NLP analytics are used to evaluate precision of matching between concepts in the clinical notes and in cTAKES medical database

In text: Coronavirus
Concept: SARS-CoV-2

- **BLEU Score: 0**
- **Levenstein Similarity: 0.5**
- **Jaccard Similarity: 0.12**
- **Cosine Similarity: 0.76**
- **ElMO: 0.82**





ML Prediction

Trained random forest model predicts the correctness of the extracted concepts and computes reliable confidence scores which are used to improve cTAKES identification

RESULTS

*We collected **over 1500** feedback datapoints from users, and used this, together with the computed NLP metrics to build a model that identifies inaccurate extracted information at a **96% accuracy** (97% F1). Our **human-in-the-loop** system can utilize user-provided feedback to self-improve through **active learning**. It, thus, represents a superior method to collect and use feedback data and can effectively be used to increase the reliability of medical NLP pipelines.*

User Feedback

Users provide feedback ('yes' or 'no') to extracted clinical concepts and the respective modifier words through UI. Feedback is used as a label to train an ML model along with NLP the metrics

Human-Enabled Healthcare NLP: Predicting your health of tomorrow, today