An Interactive NLP Approach for Improving Completeness and Annotation Efficiency in Prostate Screening Reports

Jeroen Geerdink, Innovation Manager, Health & Information Technology, Hospital Group Twente
Hridya nair Suresh, MS; Jeroen Veltman, MD, PhD; Lars Bosboom, MD; Maikel Viskaal; Shenghui Wang, PhD

Background/Problem Being Solved

Radiology is an important component of healthcare, playing a vital role in disease diagnosis. Thus, the completeness of these reports is essential, as minor errors can significantly affect the diagnosis and further treatment. The mistakes or missing fields in the report can arise due to factors such as increased workload, time constraints and inexperienced radiologists. This research focuses on automating the process of checking reports and providing radiologists with suggestions for any missing information. An interactive interface is also developed where the model is deployed for the radiologist to use, and also to derive annotation from the user interactions to solve the problem of limited annotated datasets.

Intervention(s)

The prostate screening radiology reports we used for this study are Dutch semi-structured text data, thus Natural language Processing(NLP) techniques were used to extract the important information from the reports. Dutch Language models BERTje and MedRoBERTa.nl were tested for this task, but they exhibited overfitting due to a limited dataset. A hybrid Conditional Random Field model was implemented in identifying fields. The model was able to identify the majority of the fields. The lower performance for certain fields is attributed to the underrepresentation of these fields in the reports. To address the challenges of limited data and underrepresentation, we developed an interface that integrates the model into the radiologists' workflow, allowing for both the application of the model and the collection of annotations through user interactions.

Barriers/Challenges

We were unable to integrate the interactive interface into the hospital system fully. We use the non-interactive interface that shows the model results into the system, but to know the performance of the interactive one in real-time instead of at the end of reporting is not still done.

Outcome

The Hybrid Conditional Random Field model was the effective model that identified the fields except for the fields that were underrepresented. This indicates that while the CRF model is adaptable, it requires more annotated data to improve accuracy in identifying all fields consistently. For quantitative evaluation, F1 scores were used to assess the accuracy of the CRF model, which ranged from 0.94 to 0.45, with the lower scores attributed to underrepresented fields like ”aspect” and ”grootte.”

Conclusion/Statement of Impact/Lessons Learned

The model developed can aid the radiologists in checking the report's completeness and compliance. By integrating the model into an interface radiologists can use, and also by leveraging the interface for collecting annotated data, we can increase the efficiency of the annotation process.

Keywords

Artificial Intelligence/Machine Learning; Clinical Workflow & Productivity; Quality Improvement & Quality Assurance