



A “PLUG-AND-PLAY” NATURAL LANGUAGE PROCESSING TEMPLATE FOR PHYSICIAN RESEARCHERS TO ACCURATELY DETECT CLINICAL FEATURES IN FREE-TEXT REPORTS

Fadlullah Ba'th, MD; Justin Bauzon, MD; Gustavo Romero-Velez, MD; Sofya Asfaw, MD; Benjamin Miller, MD

Objective: Natural language processing (NLP), a branch of artificial intelligence (AI), has transformed our approach to clinical research. By enabling computers to interpret and generate human language, NLP models have supported tasks that include detecting clinically relevant features from “free-text” written reports. However, a significant limitation with AI is the inability for physicians to easily modify the source code for their specific clinical questions, which may lead to inaccurate or biased outputs misaligned with their study goals. Using free-to-access software, we developed a customizable NLP template that could be adopted for various clinical use cases while still being accessible to physician researchers with limited coding experience. Our aim was to assess the utility of this AI-based template to accurately detect clinical features from two types of previously unexamined free-text reports.

Description of Technology/Method of Use: R is an open-source programming language that allows for statistical computing and data analysis. R has several “add-on” packages that can be installed to extend its capabilities. This study used the ‘spacyr’ package to create a simple NLP model template. This template was applied to two separate test cases: 1) to identify indocyanine green (ICG) use among laparoscopic cholecystectomy (LC) operative reports, and 2) to identify papillary thyroid carcinoma (PTC) from surgical pathology reports. The NLP template was designed to have four adjustable variables: report name (e.g. LC operative note), “gold standard” definition (e.g. +ICG), NLP “target terms” (e.g. ICG), and classification labels (ICG vs no ICG). Manual physician review and labeling reports for ICG and PTC pathology served as the gold standard. NLP performance was then analyzed using R version 4.4.2.

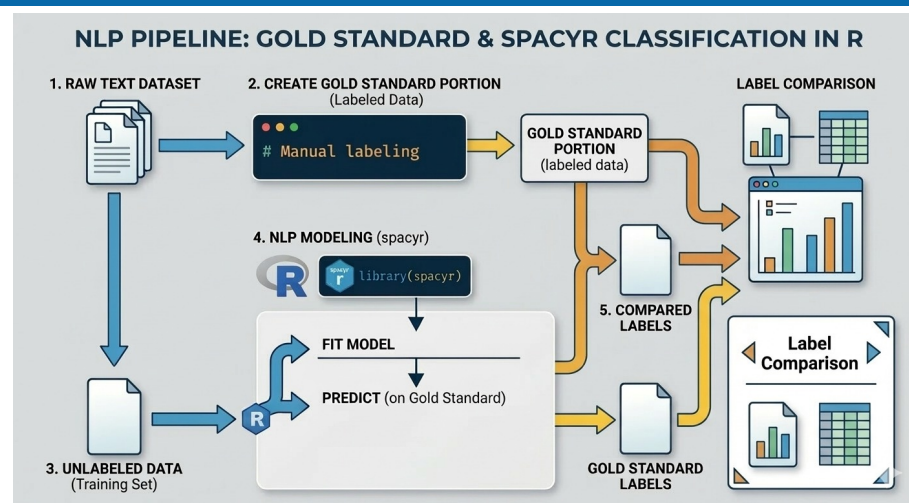


Figure 1: AI generated image representing the NLP creation process.

Preliminary Results: 674 free-text reports were analyzed: 556 LC operative notes and 118 thyroidectomy pathology reports. From the 62 LC reports where ICG was used, the NLP model correctly classified 57 (91.9%). Meanwhile, ICG was not described in 494 LC reports, which NLP identified in 493 (99.0%). From the 99 thyroid pathology reports with PTC, 93 (93.9%) were correctly classified by NLP. Meanwhile, 19 (100%) of the reports describing non-PTC pathology were accurately recognized by NLP.

Conclusions/Future Directions: A simple, customizable NLP model was able to identify clinically relevant features within unseen free-text clinician notes with high degrees of accuracy. This versatile template serves as an AI-based method for physician researchers to reliably parse through unstructured clinical data and produce binary outputs that are easily measurable. External validation with other datasets are needed.