

Human in the Loop NLP Interface to Accelerate Medical Coding

Medical coders across the US spend more than 100,000,000 hours every year scouring unstructured free text notes to generate billable condition codes. Imagine this: you visit a doctor with a few chief complaints which the doctor investigates. During your visit, he/she writes down a narrative description including everything you report and everything that is done while you are there. In order to get paid for this visit, the bill may need to include a list of conditions that were investigated, treated, or discussed. The person who generates that list is a medical coder. The coder's job is to read every word of that note in search of specific words/phrases that can be translated to one (or more) of the 70,000 ICD-10 codes used in medical coding. The coder's job, and the practice's revenue, depend on never missing an opportunity for revenue, and never making too strong an assumption. In the cases of missing a code, there is loss of revenue, in the case of overcoding there is audit risk and fraud.

This project presents a knowledge based approach to NLP that uses keyword extraction and naive presentation of context to streamline the seek-and-find manual elements of coding. The project consists of a backend which preprocesses the data, and a frontend which provides a streamlined interface for classification. The backend preprocessing identifies keywords/phrases from the text and stores a tuple containing the relevant excerpt from the note and the proposed ICD-10 code suggested by that excerpt. The frontend provides an interface for users to quickly iterate through those excerpts and approve or reject the proposed codes based on the presented text. Extracted phrases include both the trigger word and surrounding context, although a naive n-gram approach to context was taken which leaves room for further work. Medical coders are still required for approving or rejecting every proposed code, but they are no longer required to play the high stakes, high complexity, word search that has remained a critical part of their job for too long. Future iteration could include confidence intervals by training a classifier on true labels assigned by users.