iStandardize: Recommendations for Healthcare Form Standardization

Joe Roussy
Deloitte
Academic Supervisor: Michael Brudno
Industrial Supervisor: Lynn Luo

Abstract

iStandardize is a natural language processing (NLP) solution that is designed to streamline the standardization of clinical order sets (i.e., forms). Currently, hospital networks use multiple versions of forms and order sets, many of them similar in nature. The lack of standardization poses challenges in integrating the data for sharing, adds additional documentation burden, and disrupts the workflow for clinicians. The solution applies current NLP and Machine Learning (ML) techniques to identify similar order sets and their elements (attributes and responses), reduce the manual work required to compare the order sets, and expedite the decision-making process for standardization.

A PoC version of the solution was enhanced by improving the quality of the clusters created and a visualization was developed to highlight the insights of the generated recommendations. In the PoC, clusters were created by applying hierarchical clustering to cosine similarity matrices of word embeddings. Cluster quality was improved by incorporating custom non-semantic contextual features into the similarity calculations. In this enhanced version, the similarity matrix used for clustering is a weighted average of the word embedding and contextual feature similarity matrices.

Over the course of the internship, iStandardize was used by a client organization in a standardization project resulting in an estimated cost savings of $750k – $1M. Next steps include further improving the quality of recommended clusters by developing additional contextual features and weighting medical terms more in similarity calculations. The solution will also be developed into an automated end-to-end processing pipeline.