Introducing a natural language processing (NLP) approach to Entrustable Professional Activities (EPAs) text analytics classification and scoring

Anhphan Ly[1], Florante Garcia[1], Beth Baron[1], Michelle Chen[2], Akhil Punia[2], Dazun Sun[2], Henry Park[1]

1. Columbia University Vagelos College of Physicians and Surgeons 2. Columbia University Data Science Institute

BACKGROUND

In May 2014, the Association of American Medical Colleges (AAMC) published new guidelines called Entrustable Professional Activities (EPAs) consisting of 13 activities that all medical students should be able to perform upon entering residency. The Columbia University Vagelos College of Physicians and Surgeons was chosen as one of the 10 pilot schools tasked with testing the implementation of this new framework. Faculty at the medical school currently provide feedback and ratings to medical students based on their performance across 8 different clerkships.

DESIGN

Faculty narrative data was collected from 541 medical students from Fall 2016 to Spring 2018. This data consisted of students' first, second, third, and fourth year performance evaluation throughout their normal course curriculum. Data was collected and de-identified by the Center for Education Research and Evaluation and assigned a random ID.

Faculty comments were individually tagged with EPAs at a sentence-level granularity by human-labelers to establish a ground truth to base our model predictions on. Faculty comments were then selected and processed using an NLP pipeline to delimit comments by sentence and extract features based on the occurrence of specific words and linguistic patterns within the sentence-level comments. Features were created using either a 'Bag of Words' approach, TF-IDF, or n-grams.

We trained a multiclass classifier (One versus Many) using 80% of the data, with a portion of this data used for model selection purposes during the cross-validation. Various combinations of NLP feature extraction methods and machine learning models were experimented with to see which would yield the best result. The final model uses a Random Forest Classifier coupled with CountVectorizer ('Bag of Words') for feature extraction. The following schema gives a detailed breakdown of the preprocessing, machine learning, and NLP phases and how each of these feeds into the final entrustment dashboard.
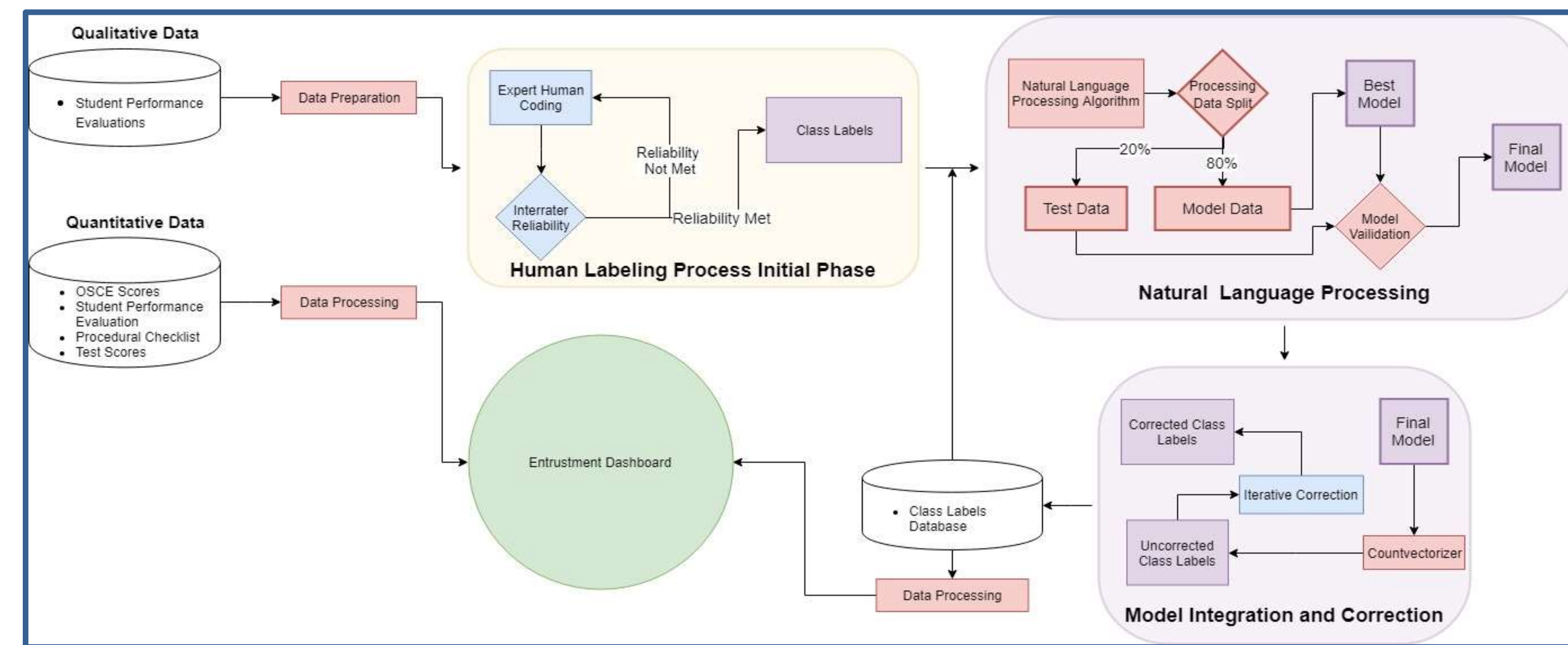
SENTIMENT ANALYSIS

Upon analyzing the comment data, we found that most of the feedback were generally very positive. Using StanfordNLP and Textblob Python packages, we assigned sentiment ratings to comments on a scale from -2 to +2, with positive and negative signs denoting the polarity of the comment and magnitude of the values representing the weight of the comment.
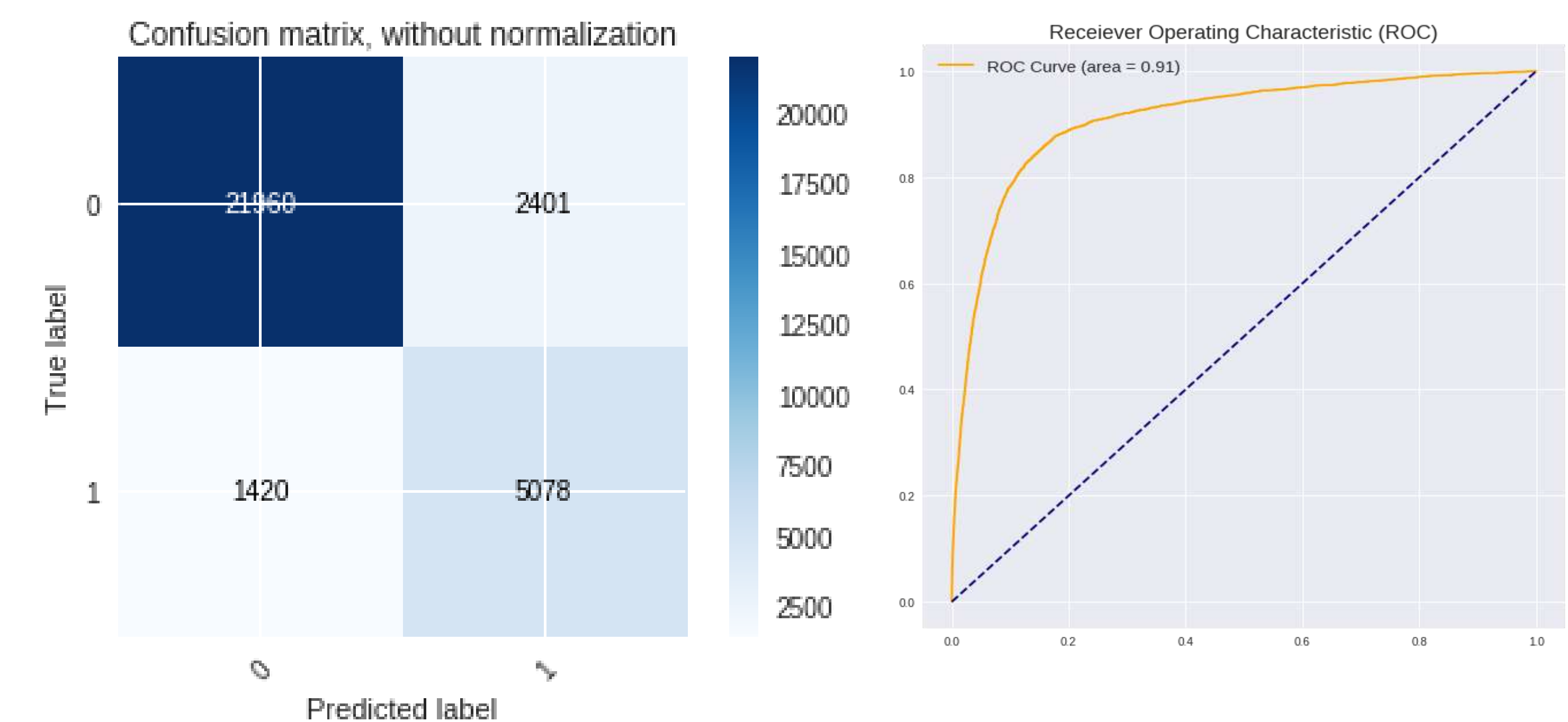
ENTRUSTMENT DASHBOARD
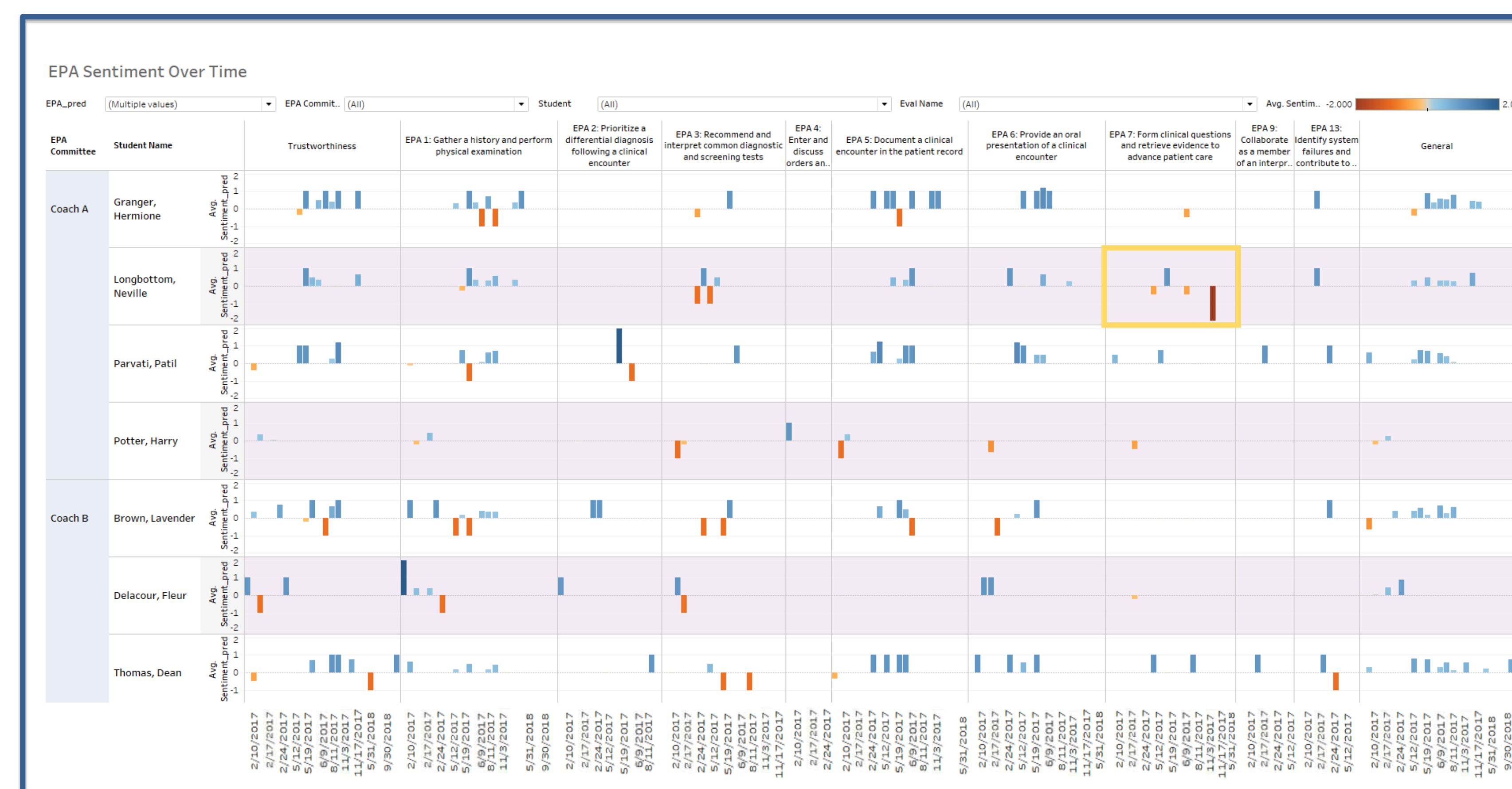
ENTRUSTMENT DASHBOARD BREAKOUT

MODELING RESULTS

STRENGTHS AND LIMITATIONS

Machine learning and NLP techniques provide immense classification power for our particular task. Limitations within the training data such as large variation in review-length (detailed versus succinct) and even writing style may make it difficult for our model fully capture the linguistic patterns of human-written text as the amount of training data increases.

NEXT STEPS

- Incorporate more EPA standards
- Incorporate additional data and environments from other institutions
- Encourage user feedback through flagging incorrect results to reduce misclassifications

USER-INTERFACE APPLICATION

We created a user-friendly web interface hosted on Google Cloud where organizations can upload their review data and our model will return the data classified into its respective EPAs. We hope this application will help our model learn nuanced regional linguistic patterns, but more importantly allow other institutions to reap the benefits of time and cost-savings through efficient classification.

REFERENCES

1. Olle ten Cate (2013) Nuts and Bolts of Entrustable Professional Activities. Journal of Graduate Medical Education: March 2013, Vol. 5, No. 1, pp. 157-158. DOI: https://doi.org/10.4300/JGME-D-12-00380.110.3109/0142159X.2010.513719
2. Hanneke Mulder, Olle Ten Cate, Rieneke Daalder & Josephine Berkvens (2010) Building a competency-based workplace curriculum around entrustable professional activities: The case of physician assistant training, Medical Teacher, 32:10, e453-e459, DOI:
3. Chirstioher Watling, Erik Driessen, Cees PM van der Vleuten, Meredith Vanstone & Lorelei Lingard (2013) Beyond individualism: professional culture and its influence on feedback. Medical education: 2013: 47: 585-594. DOI: 10.1111/medu.12150