



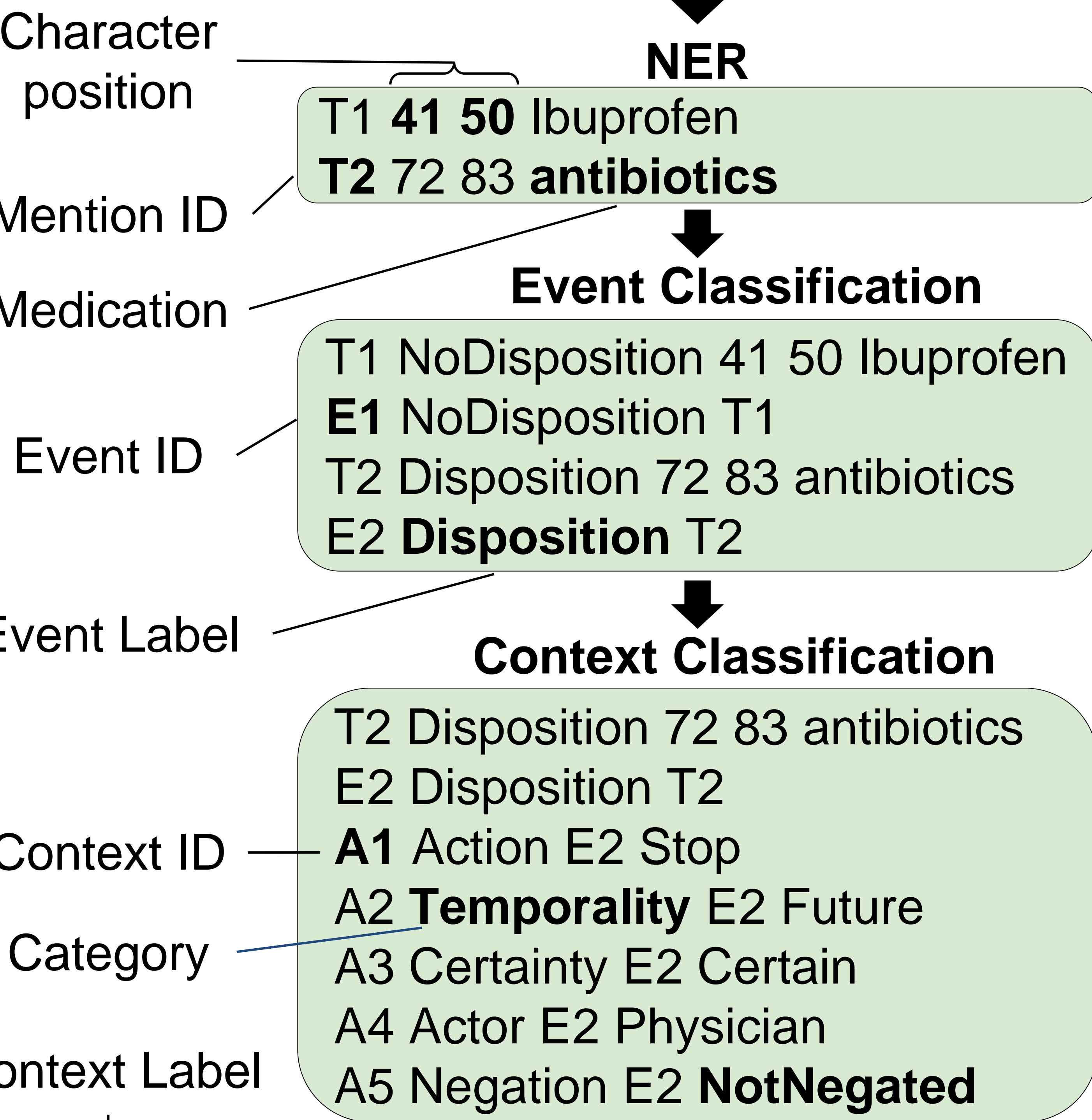
Abstract

This project details our participation in the 2022 n2c2 challenge. The challenge's goal is to identify useful context categories from unstructured medical notes. Our solution involves several steps: (1) identify all mentions of medications within a given medical note, (2) for each medication, determine whether the medication mention involved disposition, and (3) for each disposition, classify the disposition across five different context dimensions. To address this challenge, we implement a series of token and sequence classification models.

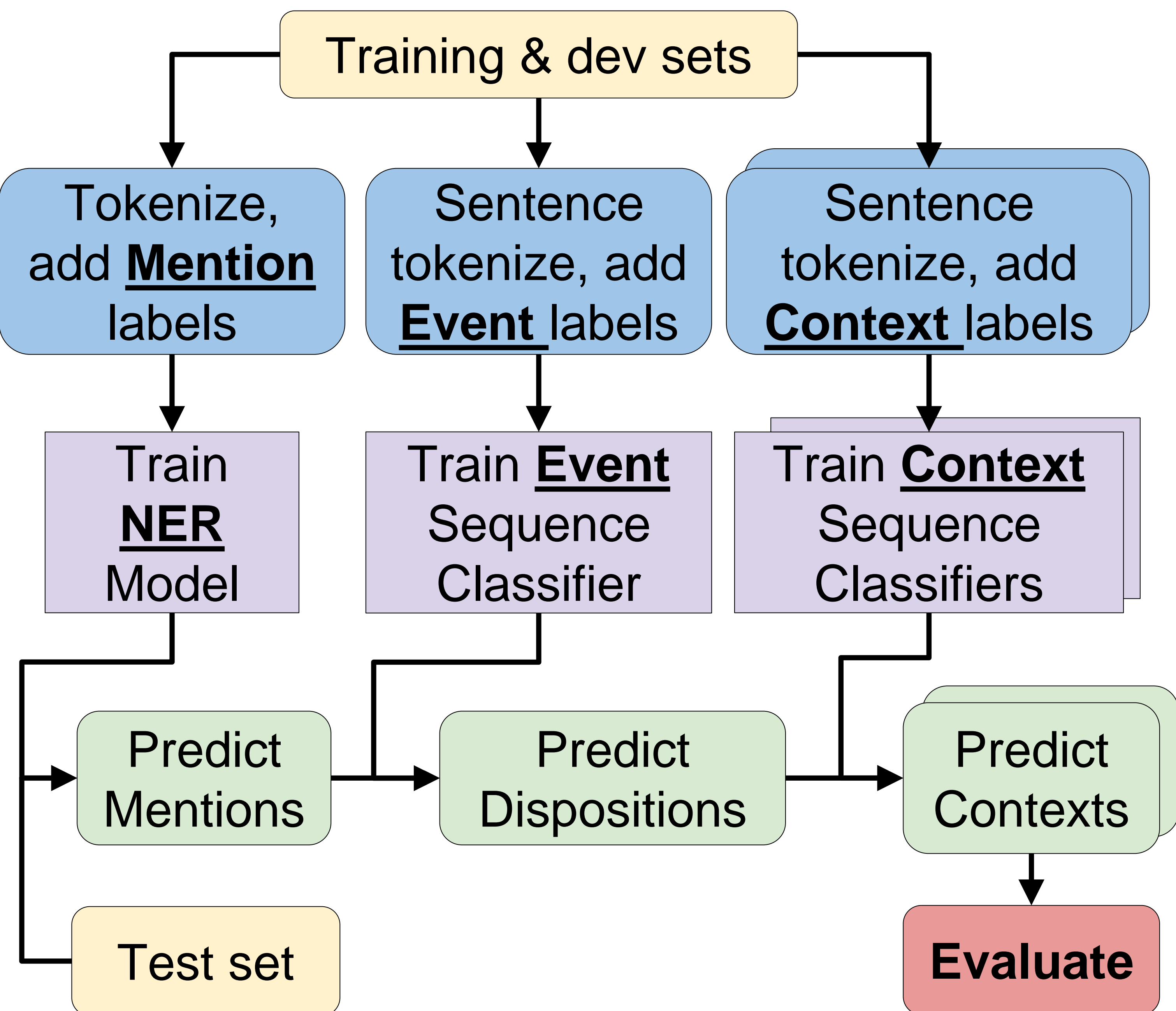
Annotation Process

Example Medical Note

The patient is feeling better. They take Ibuprofen in the morning. Stop antibiotics after 2 weeks.



Model Flow



Methods

Official training data included 400 clinical notes. After holding out 10 percent for testing, we used 315 training notes, 45 dev notes, and 40 test notes in our analysis.

For the NER task, we trained a token classification model using the BIO framework. For the Event task and Context tasks, we trained six separate sequence classification models (one for Event, five for Context) using three sentences around each medication mention as context. For each subtask, we used the Bio_ClinicalBERT model, a BERT model pre-trained specifically for clinical text. The Bio_ClinicalBERT model was pre-trained on all MIMIC III notes, and initialized from BioBERT.

Results

Task	Subtask	F1 (Ours)	F1 (Mahajan)
NER	Med. mention	0.97	
	Disposition	0.82	
Event	No Disposition	0.91	
	Undetermined	0.56	
	Overall (micro)	0.87	0.88
	Action	0.54	0.75
Context	Temporality	0.64	0.83
	Certainty	0.69	0.90
	Actor	0.79	0.93
	Negation	0.78	
	Overall (micro)	0.69	
	Combined	0.42	

Conclusion

While our results are promising, there is still room for improvement. Further error analysis will help pinpoint how to solve issues where multiple medications are mentioned in the same sentence but with different contexts and other common errors. Also, other methods can be considered for the different classifications, such as the general BERT model or LSTM.

References

- Alsentzer, E., Murphy, J. R., Boag, W., Weng, W.-H., Jin, D., Naumann, T., & McDermott, M. B. A. (2019). *Publicly Available Clinical BERT Embeddings*.
- Mahajan, D., Liang, J. J., & Tsou, C.-H. (2020). *Toward Understanding Clinical Context of Medication Change Events in Clinical Narratives*.