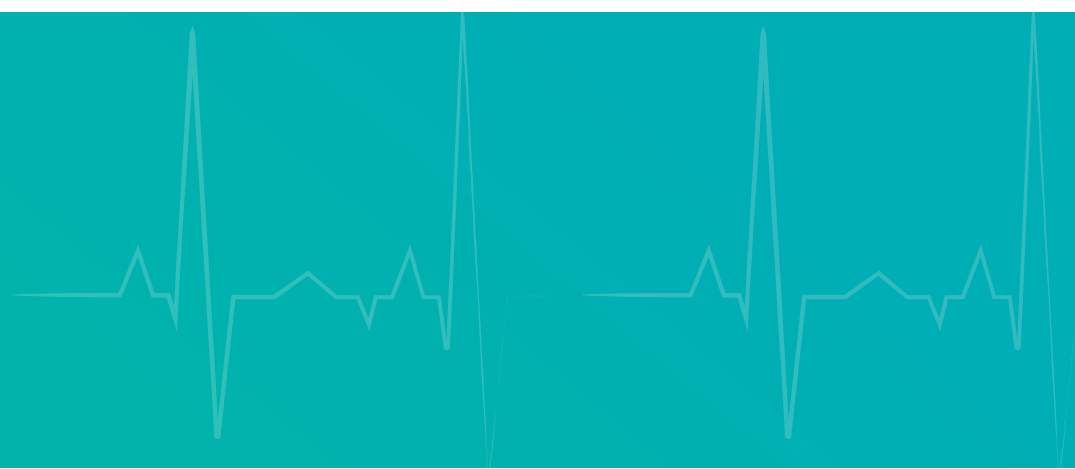


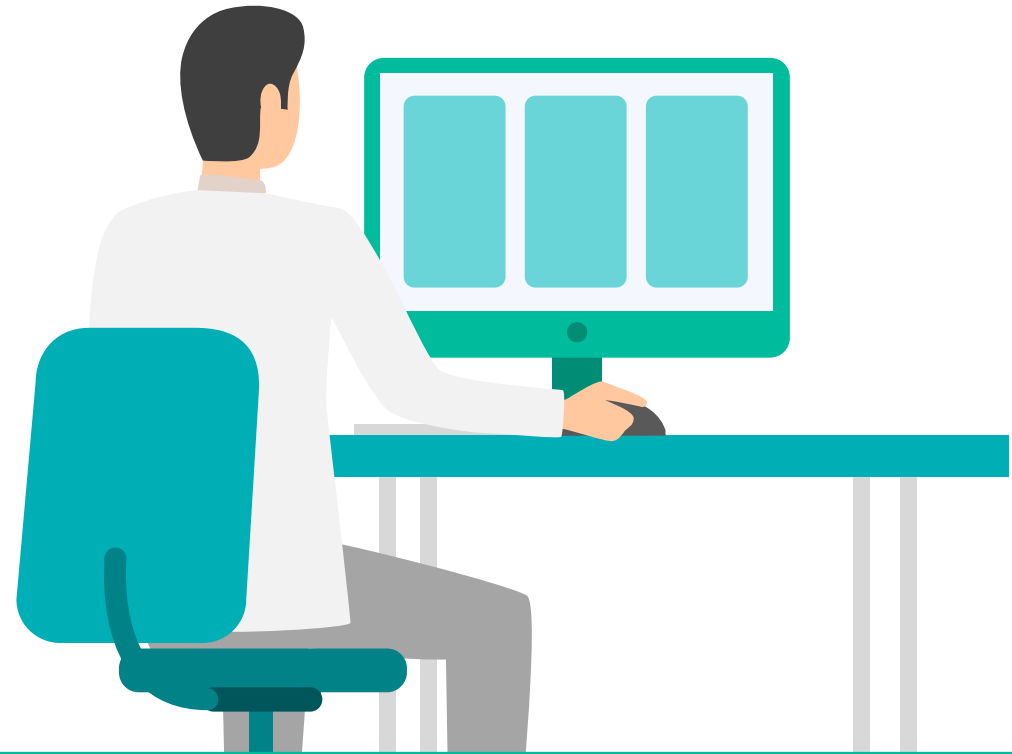
NLP Assisted Coding

Amanda Potter



Introduction





Medical coders translate physician notes into a set of diagnosis and procedure codes. These codes are used to send a claim to the payer for reimbursement



Physician Notes

Recorded by physician.



Transcription

Transcriptionist or voice recognition.



Claim Coded

Trained coders.



Coding Errors are Expensive

Errors in coding cost hospitals millions

- Claims Denials
- Incorrect base rate used for reimbursement



Coding Errors are Expensive

Denied claims

- Average \$25 to reprocess
- 50-65% of denied claims are never worked



The Data

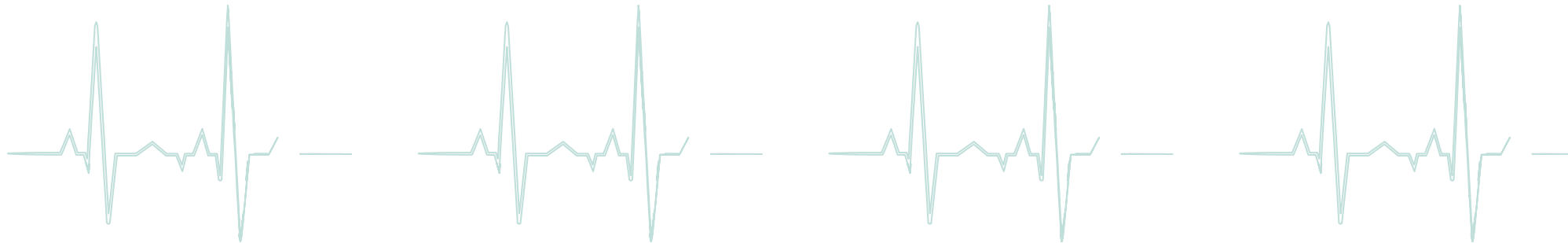


The Data

Data set from MIT

MIMIC-III and

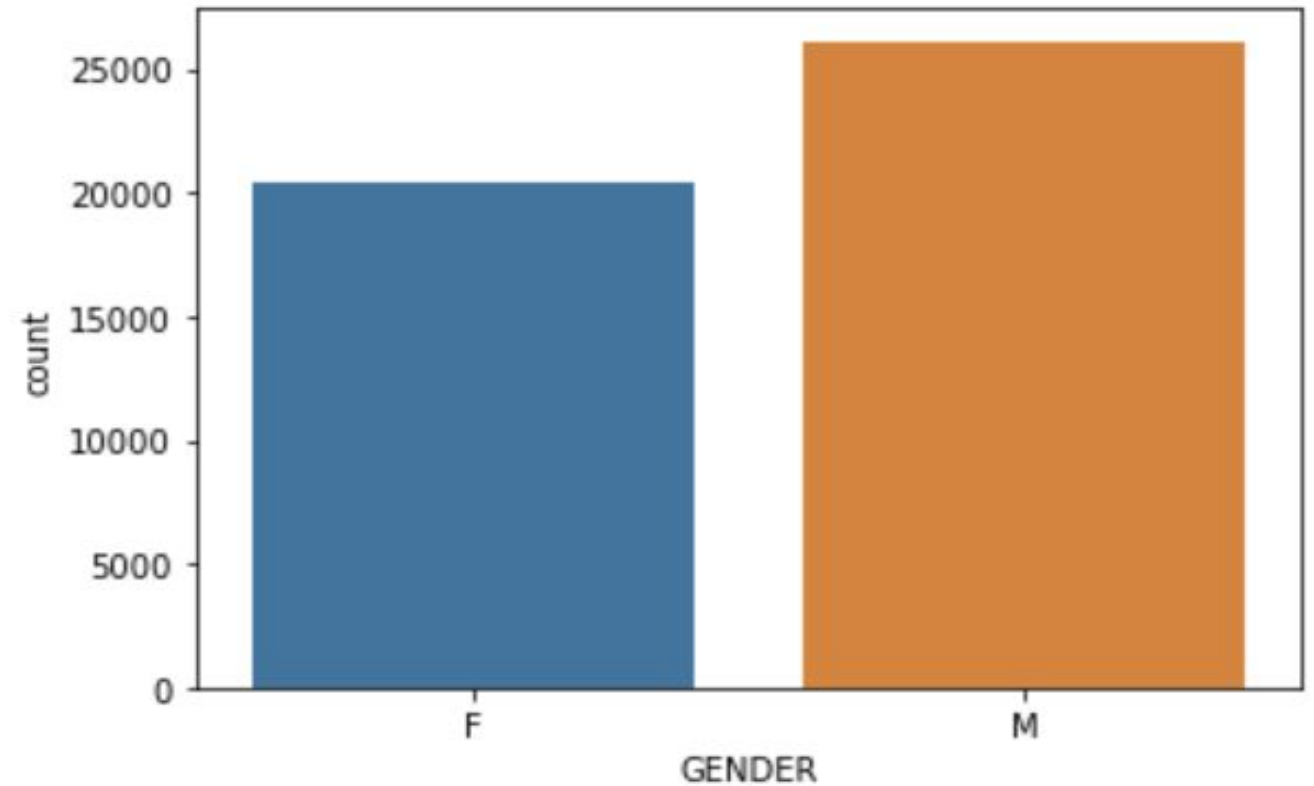
ICU patients at Beth Israel
Deaconess Medical Center
2001-2012



Patient Data

Over 40K
individual patients

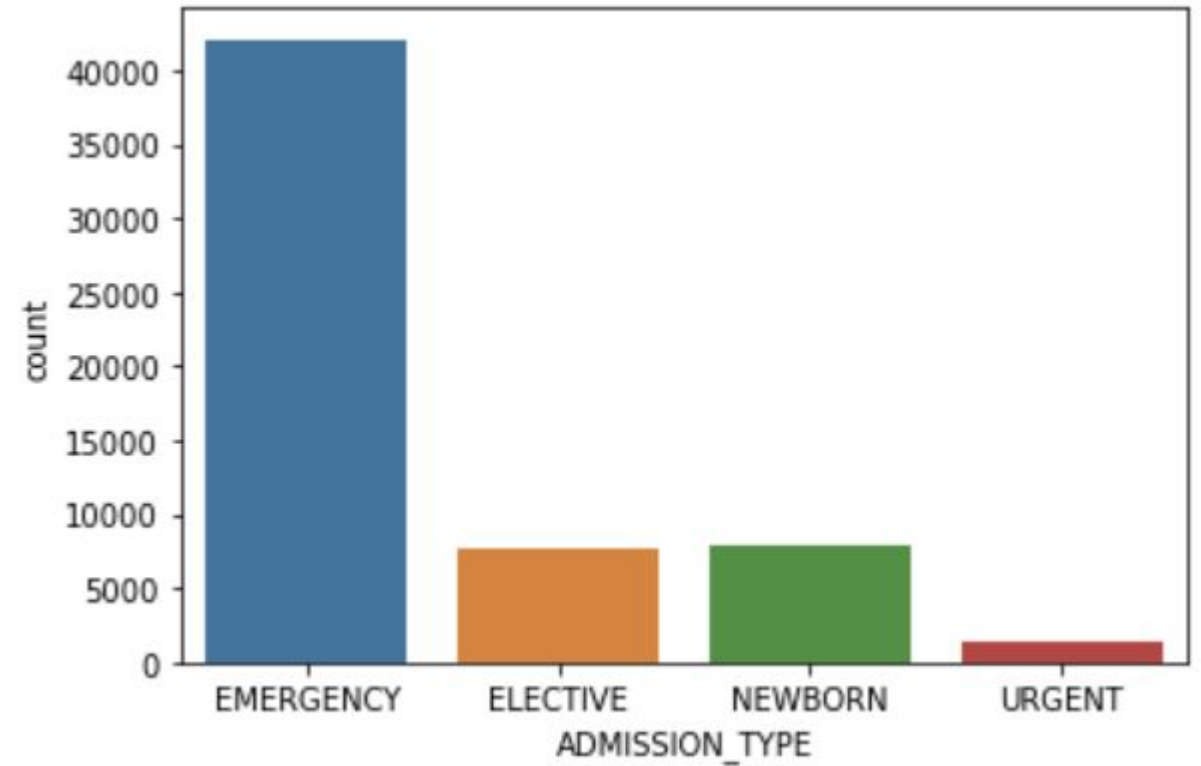
Count of Male/Female patients



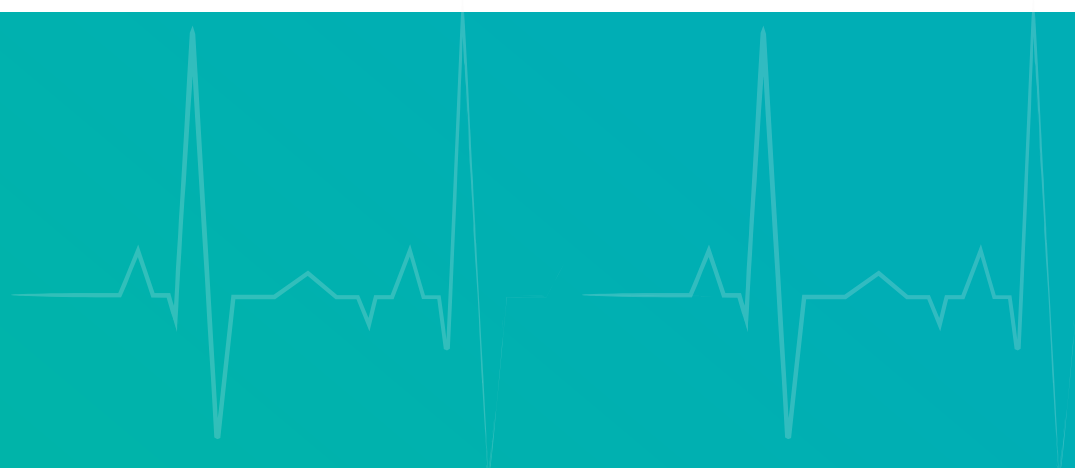
Patient Data

Majority admitted through the ED

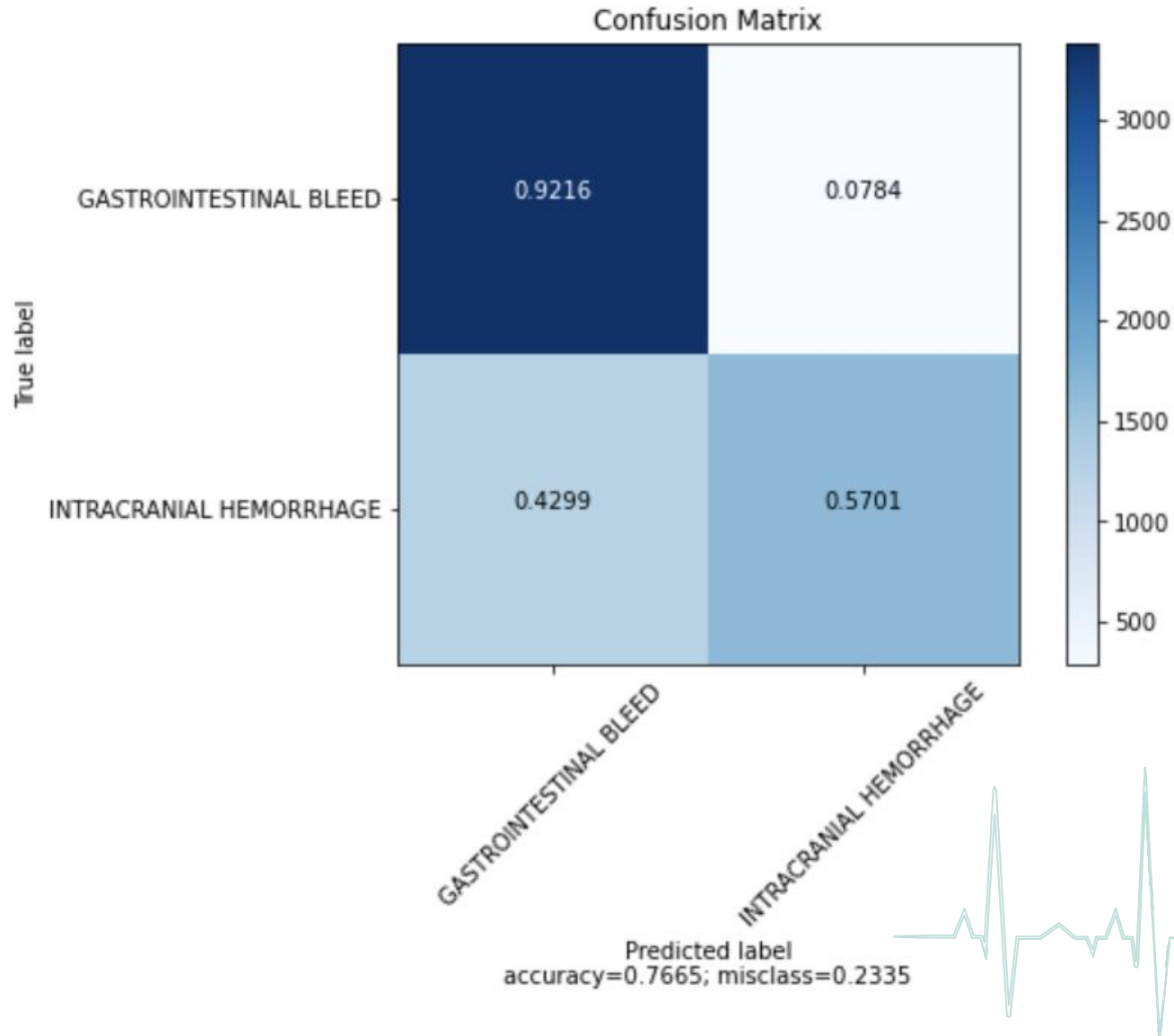
Count of patients by Admission Type



The Model

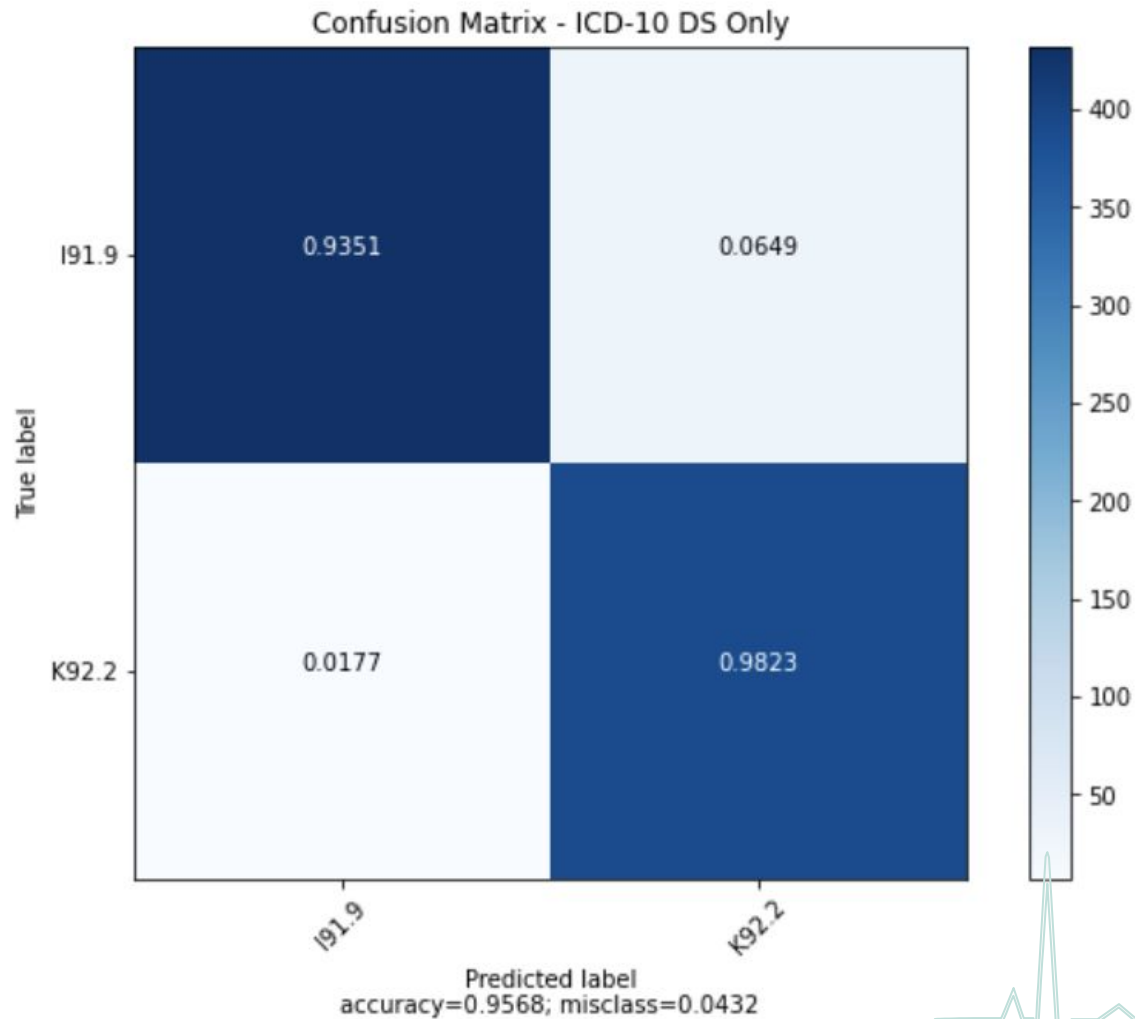


First Simple Model



- Chose 2 diagnoses
- Count vectorization
- Lemmatized
- 500 features
- Accuracy = .77

Final MVP Model



- Chose 2 ICD-10 codes
- Count vectorization
- Lemmatized
- 500 features
- Accuracy = .96

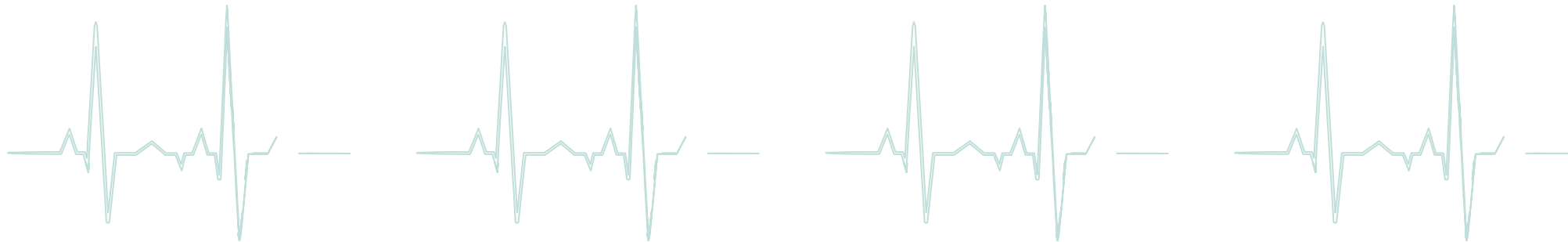


Evaluation and Next Steps

A teal-colored horizontal band at the bottom of the slide. It features a white ECG (heart rate) line graphic that starts on the left, has a small peak, then a very tall, sharp peak, followed by several smaller peaks and valleys. The line continues across the band towards the right.

Recommendations/Next Steps

- Continue to improve model
- Add additional diagnoses
- Add in DRG



Amanda Potter

amandaspotter@gmail.com

Contact

