Utterance Classification for Analysis of Hospital Trauma Resuscitations
Kyle Clark, Jan Anthony Miranda, and Elizabeth Chao
kyalexclark@gmail.com, jan.anthony.m@gmail.com ec537@scarletmail.rutgers.edu
Advisors: Ivan Marsic, Yue Gu

Goal

- Create a natural language machine learning model that can predict the intent and topic of utterances spoken during hospital trauma resuscitations.
- Intents: Request, Report, Command
- Topics: Intubation, Blood Pressure, Respiratory Rate...

Motivations and Objectives

Motivations
- Interest in the field of language processing and machine learning.
- There aren't very many applications of these topics in the medical field.

Objectives
- Classify utterances in two different ways based on intent and topic.
- Achieve an accuracy greater than 70% for both types of classifications.

Methodology

1. Transcribe audio data
2. Manually tag each utterance by intent, topic (main + sub)
3. Normalize + Split corpus 80/20 (Train/Test)
4. Create Features, Select Features
5. Train and Test Classifiers
6. Measure/Compare Performance

Results

- Intent Classification
  - Best Accuracy: 75%
  - Classifier: LinearSVC

- Topic Classification
  - Accuracies left to right: 86%, 61%, 79%, 100%, 86%
  - Classifiers left to right: Ridge, Perceptron, LinearSVC, Ridge, Ridge

References

[1] TRU-IT Coding Schemes V6
[4] Language Disambiguation for Disease Analysis

Acknowledgement

We would like to thank Ivan Marsic and Yue Gu for their guidance and support throughout our project, as well as staff at RWJ for providing us with the data to analyze and making it all possible.