

Classifying by Rhythm:

Oct 20

Nina Wang

Datasets

- Research papers dataset – 220 documents related to artificial intelligence
 - More genres
- Gutenberg Dataset - 3,036 English books written by 142 authors
 - Includes mostly creative authors (Dickens, Poe, Austen), but also non-creative like Einstein, Lincoln)
- Shakespeare plays dataset - Kaggle

This Week

- Figuring out all the functionalities of Praat
- Working on script to automate audio analysis using Praat
- Doing more research on intonation & rhythm

Next Steps

- Continue working with larger sets of data
- Write script for processing the data