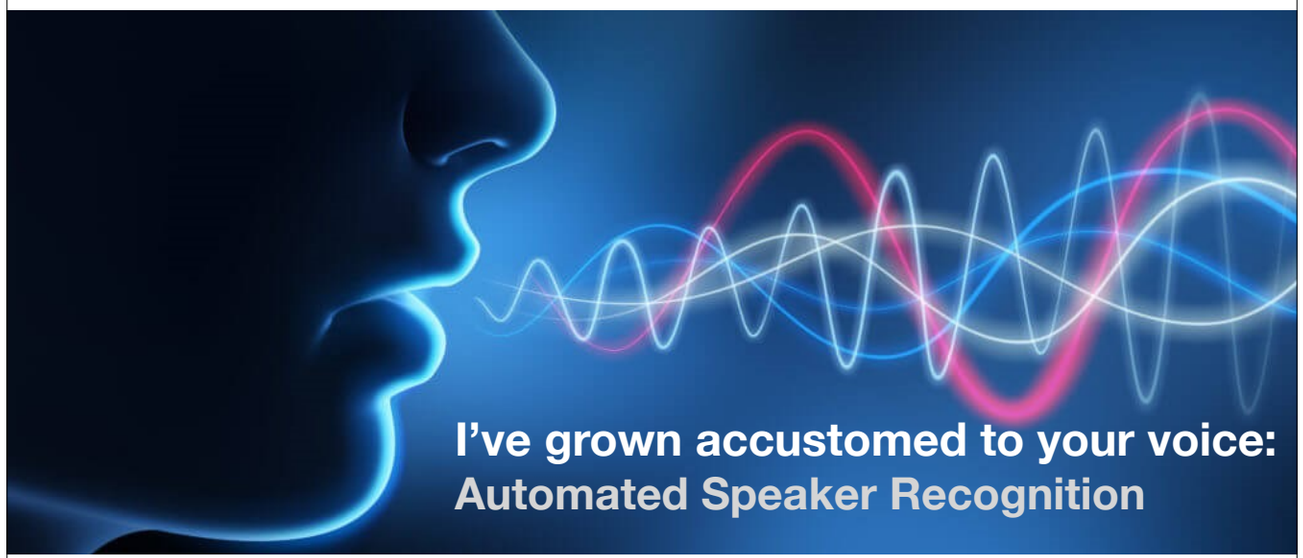


Steven Bierer, Ph.D.
Metis Data Science Bootcamp, Seattle

Speaker Recognition

- Better command interpretation
- Customized recommendations
- Authentication / security

Speaker Recognition

Objective

1. Recognize a user's voice
   - Make decision quickly
2. Function within a multi-user group
   - 1 group = 1 "household"
3. Identify unfamiliar voices

Training procedure

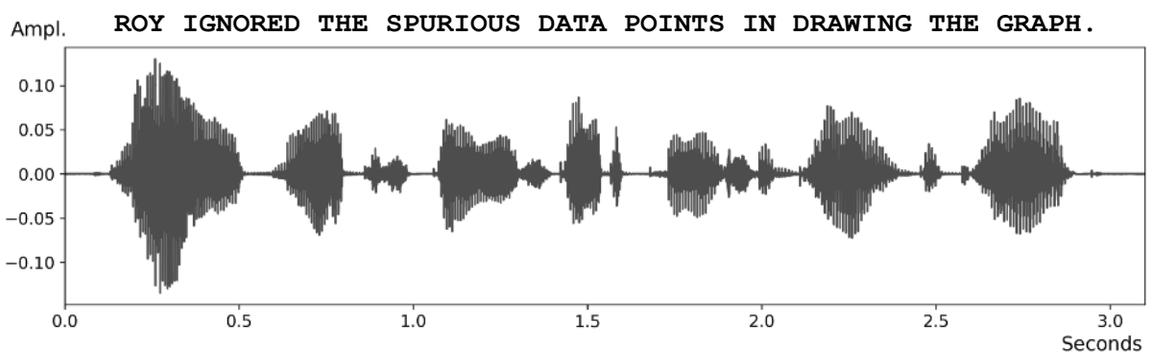
Audio Signal

TIMIT Continuous Speech Corpus

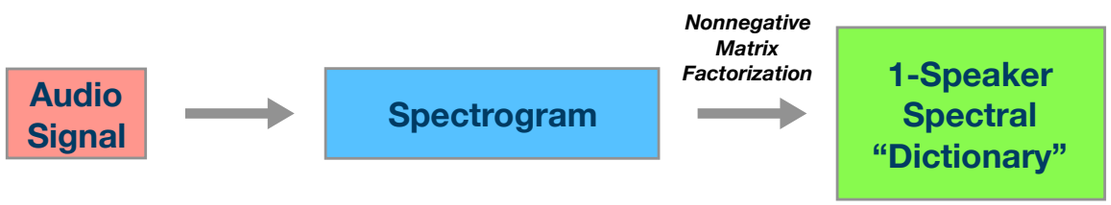
- Single-channel low-noise recordings
- 630 male + female speakers
- 7 U.S. regions

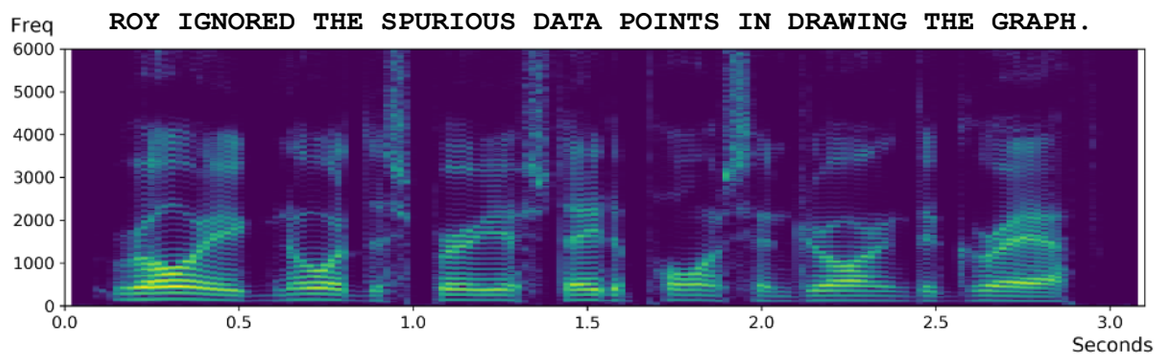
Training procedure

Audio Signal → **Spectrogram**

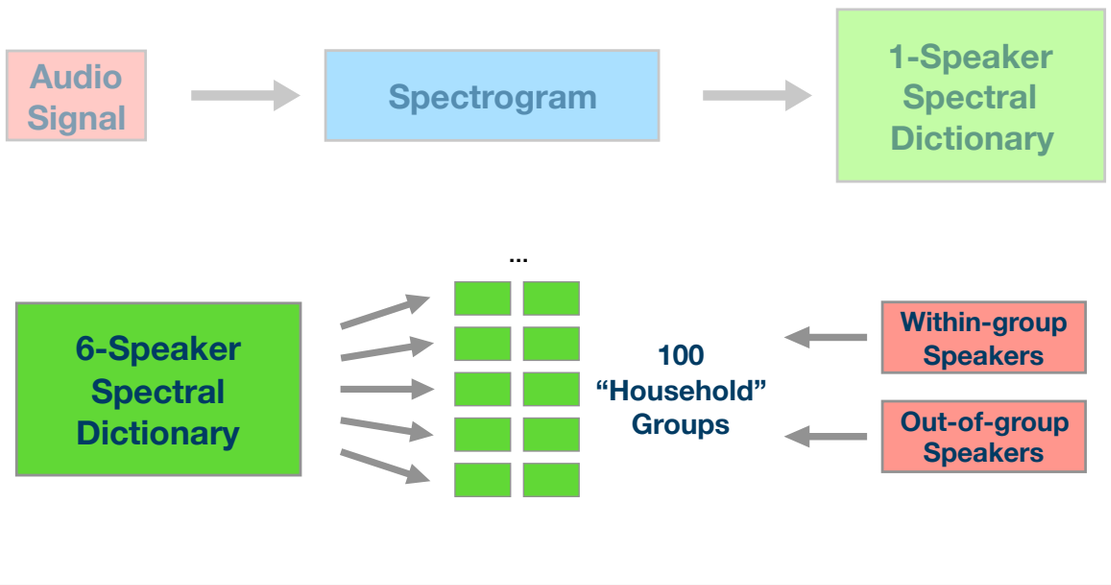


Training procedure



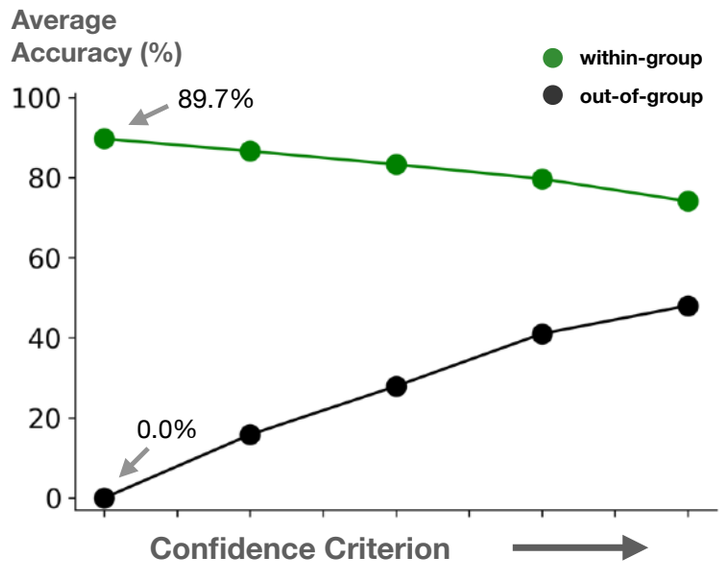


Testing procedure



Results

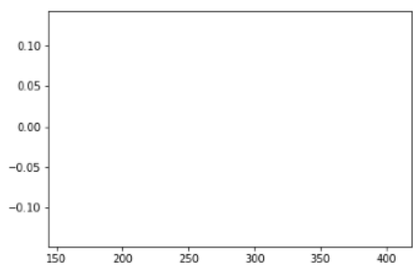
- Good within-group accuracy
  - Trained on only 7 sentences!
- Trade-off with rejection of unfamiliar speakers



Future Considerations

- Improve accuracy
  - Reduce spectral redundancy
  - Final classifier layer
- Multi-channel recordings
  - Noisy environments
  - Simultaneous speakers

- Real-time speaker identification



Thank you!

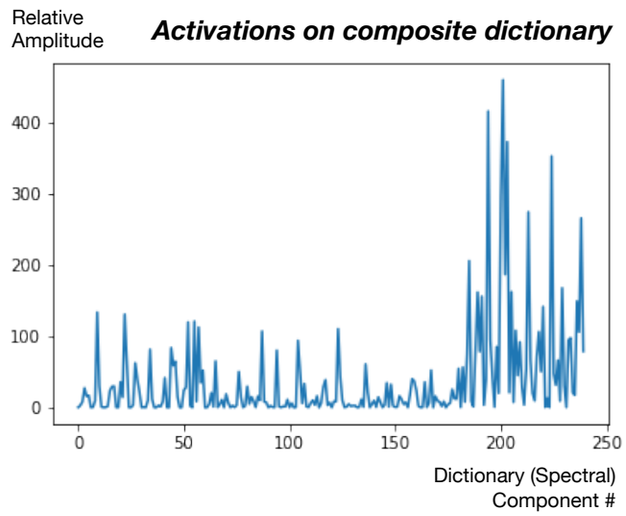
Steven Bierer

neurolaunch@gmail.com

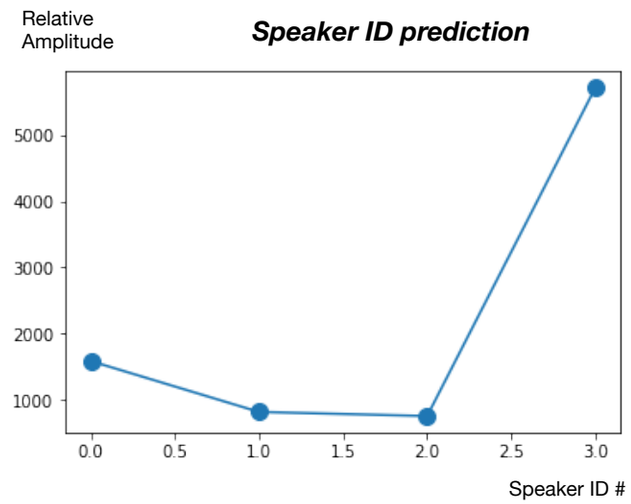
www.linkedin.com/in/steven-bierer

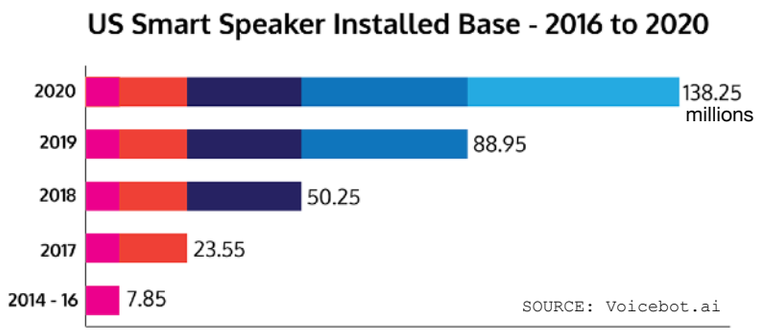
www.researchgate.net/profile/Steven_Bierer

Appendix


Relative
Amplitude
Activations on composite dictionary
400
300
200
100
0
0
50
100
150
200
250
Dictionary (Spectral)
Component #

Appendix


Relative
Amplitude
Speaker ID prediction
5000
4000
3000
2000
1000
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Speaker ID #

Amazon's Alexa started ordering dollhouses after hearing its name on TV

Interactive Speech Systems