Verbal Protest Recognition in Children with Autism
Jonah Casebeer; Hillol Sarker; Murtaza Dhulliawala; Nicholas Fay; Mary Pietrowicz; Amar Das
IBM Research, USA

Motivation
- Verbal protest is a commonly reported challenging behavior observed in children with autism.
- Definition: sensory overload-induced crying, screaming, shouting, and yelling.
- Detection will help us:
  - Understand frequency and context of occurrence
  - Predict and prevent
  - Reduce burden on caregiver
- Develop metric to evaluate the efficacy of therapy
- Can we detect whether a child with autism is exhibiting verbal protest from audio recordings?

Contribution
- Curation of a verbal protest audio dataset
- Development of verbal protest detection models:
  - Can work in a resource-constrained, real-time setting
  - Noise tolerant
  - Limited false positives

Dataset
- Training
  - YouTube Collection
  - Urban sound [4]
  - RML [5]
  - AudioSet [3]
- Testing
  - Verbal protest
  - Noise
  - Emotional speech
  - Music
  - Real-world data

Modeling Approach
- Augment:
  - Pre-shuffling
  - Time stretching
  - Dynamic range compression
- Training Perturbation:
  - (Urban sound dataset)
  - 35 ms window, 10 ms step
  - 40 Log Mel-Filter banks (LMB)
- Features:
  - Spectral centroid
  - Spectral contrast
  - 23 MFCC
- Feature Vector (34)

Validation
- Gaussian Mixture Model (GMM)
- \( R_{\text{Fp}} = \max \{ P(\text{Verbal Protest}) \} - P(\text{Verbal Protest}) \)
- False positive rate for real-world evidence dataset:
  - GMM-N (0.100) and CNN-N (0.367)
  - GMM-E (0.003) and CNN-E (0.007)
- Ensemble model reduces the false positive rate to 0.001

Conclusion
- GMM model is lightweight but produces more false positives
- CNN model reduces false positive rate but requires more resources
- Ensemble model can be a superior choice
- Future work is to explore reducing verbal protest episodes by utilizing biomarkers and characteristics of the child and their environment.

Acknowledgements
We would like to thank Soumya Ghosh from IBM Research for the useful comments and discussions on the methodology.

References