Matthew Perez

https://matthewkperez.com/

TECHNICAL PROFILE

Expert in leveraging advanced machine learning models for practical speech and language applications with experience in fields such as speaker diarization, automatic speech recognition, speech synthesis, and emotional AI. Skilled in data-driven analysis for health, behavior, and mood insights, with a solid background in multimodal systems, transformer-based models, and end-to-end machine learning pipeline development.

EDUCATION

Ph.D. in Computer Science and Engineering, University of Michigan, 2024
M.S. in Computer Science and Engineering, University of Michigan, 2019
B.S. in Computer Science, University of Notre Dame, 2017, *Cum Laude*

RESEARCH EXPERIENCE

Computational Human Artificial Intelligence Lab *(University of Michigan)* 2024

Engineered robust speech-based machine learning methodologies for improving ASR systems and natural language understanding in low-resource and challenging acoustic environments. Conducted comprehensive research on error analysis and anomaly detection in speech patterns, leading to enhanced model interpretability and reliability. Played a central role in interdisciplinary teams, merging technical machine learning development with practical insights from healthcare professionals to innovate solutions across various domains.

PROFESSIONAL EXPERIENCE

TeachFX (Machine Learning Research Scientist) Present

Investigated transformer-based E2E as well as cascaded models for speaker diarization in classroom settings. Helped improve performance by 18% through the use of LLM re-scoring. Led machine learning efforts from initial design to deploying model endpoints.

Google *(Speech Synthesis Team)* 2022

Developed and trained a seq2seq model for grapheme-to-phoneme conversion, integrating language identification to leverage multilingual datasets. This led to performance improvements in all tested languages and up to 30% improvement in low-resource settings. This work led to a publication accepted at Interspeech 2023.

Cogito Corporation *(Speech/Signals Team)* 2021

Researched novel approaches for multimodal speech emotion recognition. I showed that modeling paralinguistic events such as silence within BERT framework improved emotion recognition performance by up 14.6%.

PUBLICATIONS

- **Matthew Perez**, Aneesha Sampath, Minxue Niu, Emily Mower Provost, "Beyond Binary: Multiclass Paraphasia Detection with Genertive Pretrained Transformers and E2E Models". Interspeech 2024
- Yang Yu, **Matthew Perez**, Ankur Bapna, Fadi Haik, Siamak Tazari, Yu Zhang. "PronScribe: Highly accurate multimodal phonemic transcription from speech and text". Interspeech 2023
- **Matthew Perez**, Mimansa Jaiswal, Minxue Niu, Cristina Gorrostieta, Reza Lotfian, John Kane, Emily Mower Provost. "Mind the gap: On the value of silence representations to lexical-based speech emotion recognition". Interspeech 2022.
- **Matthew Perez**, Amrit Romana, Noelle Carlozzi, Praveen Dayalu, Jennifer Ann Miner, Angela Roberts, and Emily Mower Provost. "Articulatory Coordination for Speech Motor Tracking in Huntington Disease" Interspeech 2021.
- Zakaria Aldeneh, **Matthew Perez**, and Emily Mower Provost. "Learning Paralinguistic Features from Audiobooks through Style Voice Conversion" NAACL 2021.
- Amrit Romana, John Bandon, **Matthew Perez**, Emily Mower Provost. "Automatically Detecting Errors and Disfluencies in Read Speech to Predict Cognitive Impairment in People with Parkinson's Disease". Interspeech 2021
- **Matthew Perez**, Zakaria Aldeneh, and Emily Mower Provost. "Aphasic Speech Recognition using a Mixture of Speech Intelligibility Experts" Interspeech 2020.
- **Matthew Perez**, Wenyu Jin, Duc Le, Noelle Carlozzi, Praveen Dayalu, Angela Roberts, and Emily Mower Provost. "Classification of Huntington's Disease Using Acoustic and Lexical Features." Interspeech 2018.
- Louis Daudet, Nikhil Yadav, **Matthew Perez**, Christian Poellabauer, Sandra Schneider, Alan Huebner. "Portable mTBI Assessment Using Temporal and Frequency Analysis of Speech." IEEE Journal of Biomedical and Health Informatics 2017.