Transcribing EMS calls using Deep learning for Clinical Decision Support

Pramath Shukla, Computer Science Major
Mentor: Dr. Nakul Gopalan, Assistant Professor
Ira A. Fulton Schools of Engineering

Introduction and Motivation

Current EMS audio calls are noisy, unstructured, and often incomplete, making them difficult to analyze automatically. Accurate speech-to-text transcription is a crucial first step before extracting clinically relevant information.
This ongoing work focuses on fine-tuning OpenAI's Whisper model specifically for EMS audio, enabling reliable text generation even in challenging acoustic conditions.
The goal is to create a seamless pipeline from raw audio → accurate transcript → machine learning analysis for patient triage and decision support.

Procedure

EMS call recordings were segmented into smaller clips and formatted with timestamped labels for supervised training.
Used Parameter-Efficient Fine-Tuning (PEFT)
With LoRA adapters to adapt the base Whisper-small model to EMS-specific language and acoustic patterns.
Leveraged mixed-precision training for computational efficiency on GPUs.
Employed spec-augmentation and dynamic batching to enhance robustness against environmental noise.
Implemented both batch and streaming inference modes for real-time deployment.
Streaming mode allows near-instant transcription of incoming EMS audio while maintaining high accuracy.

Results

- The fine-tuned model demonstrated a 20% reduction in Word Error Rate compared to the base Whisper model on EMS-specific data.
- Performance improvements were particularly strong for overlapping dialogues and background-noisy scenarios, such as ambulance sirens and radio interference.
- Manual evaluation confirmed improved recognition of clinically critical terms (e.g., "unconscious," "pulse," "blood pressure"), maintaining higher semantic accuracy.

Fig : Left section shows transcriptions for base whisper model, and the right section shows fine-tuned whisper model.

DC Fire Ambulance Florida Hospital Zero Two Children's National. DC Four? Yeah, it's getting to DC Fire Ambulance Zero Four. I have a three to four miniatur facility with an eight year old male. A two breast-reader stress secondary to asthma. This one patient is a little more oriented. He is surely receiving a duodenab with improvement. Patient is tachycardic at 110.0. He is currently 98% on the nebulizer. Bestiary rate is 32.0. And title 34. Blood pressure is 138 over 60. Thank you for the rest of that heavy copy. Copy, loud and clear. Any changes? Let us know, otherwise we'll see you upon arrival.

1 DC Fire Ambulance Florida Hospital, Zero-Two Children's National. DC-4. Yes, your name is DC Fire Ambulance Zero-Four. I have a three to four minute ETA facility with an eight year old male. A two-rest for distress, secondary to asthma. This patient is a little no-oriented. He is surely receiving a duodenab with improvement. Patient is tachycardic at 1.10. So the time of fact, no dawn for monitor. He is currently 98% on the nebulizer. Best surgery is 32. And the title is 34. Blood pressure is 138 over 60. Thank you for the rest of that heavy copy. Copy, loud and clear. Any changes? Let us know, otherwise we'll see you upon arrival.

Medical 14, hospital two. Medical 14. That's a 70-year-old male. History asthma. The tree of multiple breed of treatments in the mailer is from the mother. I also gave him one breed of treatment by the skinner. I said, second one more. We're going to do it throughout up to 96% after the breathing treatment. So, I think you're going to be able to feel it on the shore, Brad. We'll be taking, like I said, 10. Heart rate's 110. Ball 5, like I said, is 96%. Working on the pressure as well. Had he gotten a 2 and a half or just a 2 to a half? We did a 2 and a half. Now we're going to be all. Copy, see you in 10 minutes. If you have any changes, please let us know. Yes.

1 Medical 14, hospital 2. Medical 14. That's a 70-year-old male. History of asthma. The treatment of multiple breathing treatments in the mailer's from the mother also given one breathing treatment by the skin. I just have a second one more. We've been through it all up to 96% after the breathing treatment. So, I think you're on the shore, Brad. We're thinking like I said, 10. Heart rate's one to... Heart rate's 110. Balls rise, like I said, 96% working on the pressure as well. Have you gotten a do a nab or just a utero? We didn't do a nab. Now we're going to be all. Copy. See you in 10 minutes. If you have any changes, please let us know. Yes.

Conclusion

- The results demonstrate that domain-specific fine-tuning improves Whisper's performance in noisy, real-world emergency environments.
- The model successfully handles overlapping speech, background noise, and critical medical terminology with near-perfect accuracy.
- Results confirm that PEFT-based fine-tuning makes Whisper highly effective for real-world EMS transcription.

Future Work

- Evaluate the model on larger and unseen EMS datasets to confirm that the observed performance generalizes beyond the fine-tuning set.
- Combine transcription and NLP-based classification models (gender, age, condition) into a single automated system.

References

[1] A. Radford, J. Kim, T. Xu, et al., "Robust Speech Recognition via Large-Scale Weak Supervision," *OpenAI Whisper Technical Report*, 2022.

[2] S. H. Chiu, **"ChineseTaiwaneseWhisper: Fine-tuned Whisper for Mandarin and Taiwanese Speech,"** *GitHub Repository*, 2023. [Online].
Available: https://github.com/sandy1990418/ChineseTaiwaneseWhisper

FURI

Acknowledgements
I would like to acknowledge Dr. Nakul Gopalan for his invaluable mentorship throughout this research project. I am also grateful to the FURI Research Program for providing me with the opportunity and resources to explore and contribute to impactful work in the field of NLP

Ira A. Fulton Schools of Engineering
Arizona State University