REAL-TIME AI-BASED SUPPORT SYSTEM FOR IDENTIFYING EMOTIONAL DISTRESS IN FAMILY MEMBERS OF SUBSTANCE ADDICTS

THIS STUDY DEVELOPS AN AI-DRIVEN SYSTEM TO ASSIST CALL CENTER OPERATORS IN IDENTIFYING AND RESPONDING TO EMOTIONAL DISTRESS. USING SPEECH EMOTION RECOGNITION (SER) AND NLP, TRANSCRIBES SPEECH, ANALYZES EMOTIONS, AND PROVIDES REAL-TIME INTERVENTION SUGGESTIONS.

Authors
Beyza Gökkaya[1], Christina Todorova[2]

Affiliations
[1]Hogeschool Utrecht, Utrecht, the Netherlands | [2]CeADAR, Dublin, Ireland

INTRODUCTION

Call center operators assisting families of individuals with addiction often lack psychological training, making it difficult to recognize and respond to distress. Without AI support, they rely on subjective judgment, increasing the risk of misinterpretation and ineffective intervention. Despite advances in Speech Emotion Recognition (SER) and Natural Language Processing (NLP), real-time emotional distress detection in call centers remains underdeveloped. Current AI solutions focus on text-based sentiment analysis, but recognizing emotions from speech in live conversations presents challenges, including privacy concerns, transparency, and AI reliability.

PRELIMINARY CONSIDERATIONS

AI's ability to detect emotional distress in real-time conversations needs to be evaluated for accuracy and effectiveness.
The most effective technical approach for SER, comparing ASR, speech characteristic descriptors, and audio encoders.
AI-driven SER can be leveraged to provide useful intervention suggestions for call center operators.
Enhancing AI suggestions with past caller interactions and guidelines may improve response effectiveness.

LITERATURE REVIEW

Speech Emotion Recognition (SER) identifies emotions in speech but struggles with differences in human and machine perception, requiring integration across multiple fields (Wani et al., 2021). While promising for mental health, it remains underexplored. SVM achieved 74% accuracy, highlighting SER's potential (Madanian et al., 2022). Multimodal Emotion Recognition (MER) improves accuracy by combining audio, video, and text, but challenges remain in data fusion and model performance (A.V. et al., 2024).

METHODOLOGY

1. Define evaluation framework for comparing results from various models (manual labeling of ground truth data)
2. Evaluate various technical approaches:
-Speech to text transcription with Whisper,
-Augmenting with additional natural language descriptions of speaker characteristics such as pitch, volume, and speed, and
-Directly transforming latent representations of audio input into recognized emotions
3. Experiment with various models and prompting techniques to produce useful real-time suggestions based on real-time SER.

CONCLUSION

This research shows how AI can collaborate with humans to improve emotional support in a measurable way. By evaluating different technical approaches, the goal is to maximize model performance and speed. Future research could explore video integration and expand the system to other mental health support applications.
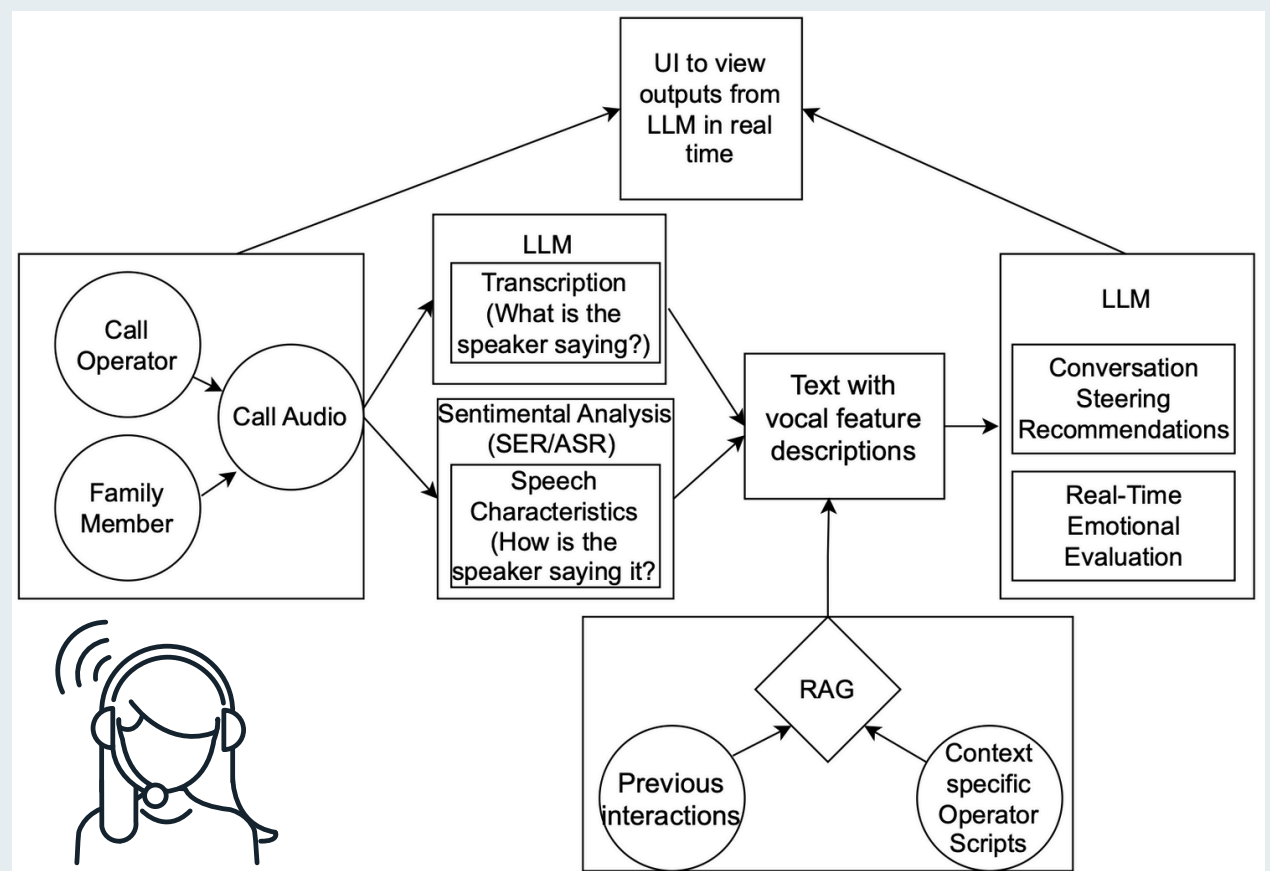
REFERENCES

A.V., G., T., M., D., P., & E., U. (2024). Multimodal Emotion Recognition with Deep Learning: Advancements, challenges, and future directions. Information Fusion, 105, 102218. https://doi.org/10.1016/j.inffus.2023.102218
Bu, K., Liu, Y., & Ju, X. (2024). Efficient utilization of pre-trained models: A review of sentiment analysis via prompt learning. Knowledge-Based Systems, 283, 111148. https://doi.org/10.1016/j.knosys.2023.111148
Madanian, S., Parry, D., Adeleye, O., Poellabauer, C., Mirza, F., Mathew, S., & Schneider, S. (2022). Automatic Speech Emotion Recognition Using Machine Learning: Digital Transformation of Mental Health.
Wani, T. M., Gunawan, T. S., Qadri, S. A. A., Kartiwi, M., & Ambikairajah, E. (2021). A Comprehensive Review of Speech Emotion Recognition Systems. IEEE Access, 9, 47795–47814. https://doi.org/10.1109/ACCESS.2021.3068045
Wu, Z., Gong, Z., Ai, L., Shi, P., Donbekci, K., & Hirschberg, J. (2024). Beyond Silent Letters: Amplifying LLMs in Emotion Recognition with Vocal Nuances (No. arXiv:2407.21315). arXiv. https://doi.org/10.48550/arXiv.2407.21315

ACKNOWLEDGEMENTS