Sina Rashidi

✉ *sinarashidi46@gmail.com* | 🌐 Homepage | 𝐢𝐧 Sina Rashidi | ⌂ sinarashidi | 🎓 Google Scholar

Education

- **Sharif University of Technology**, Tehran, Iran — *Oct 2021 - Jun 2024 (32 months)*
  M.Sc., Artificial Intelligence and Robotics, GPA: **16.73**/20, **3.6**/4.0
  Thesis: Improving Direct Speech-to-Speech Translation: A Unit-based Pipeline and Synthetic Corpus for Persian-English (Summarized Report)
  Supervisor: Prof. Hossein Sameti

- **Amirkabir University of Technology**(Tehran Polytechnic), Tehran, Iran — *Sep 2016 - Sep 2021 (60 months)*
  B.Sc., Electrical Engineering (Electronics), GPA: **15.03**/20, **3.0**/4.0 (last two years)
  Thesis: Building a Fire Detector Guard Robot for Indoor Places Using Real-time Object Detection
  Supervisor: Prof. Amir Jahanshahi, Prof. Mohsen Moezzi

Research Interests

- Speech Processing and Natural Language Processing
- Large Language Models, Multimodal Foundation Models, and Synthetic data generation
- Applications of Machine Learning and Deep Learning in Healthcare, Cognitive Science, and Social Sciences
- Explainable AI, Counterfactual Reasoning, and Causal Inference

Research Experience

- **Research Assistant, Columbia University** *(Remote)* — *Jan 2024 - Present (23 months)*
  Supervisor: Prof. Maryam Zolnoori
  - Conducted research on the application of natural language and speech processing techniques on cognitive impairment detection from interview speech project
  - Developed multi-modal Transformer-based pipelines for cognitive impairment detection from speech, text, and EHR data
  - Achieved state-of-the-art performance in Alzheimer's Disease detection from speech using Synthetic Data Generation with LLMs, Audio-LLMs, Text-to-Speech models, and Voice Conversion models
  - Fine-tuned multiple open-weight LLMs and Audio-LLMs from HuggingFace for Alzheimer's Disease detection and synthetic speech and text data generation using LoRA
  - Developed an award-winning interface for explainability of acoustic Transformer models for Alzheimer's detection from speech

- **Research Assistant, Sharif University of Technology** — *Jan 2022 - Jun 2024 (29 months)*
  Supervisor: Prof. Hossein Sameti
  Topics: Automatic Speech Recognition, Text-to-Speech Synthesis, Voice Conversion, Speaker Verification, Digital Signal Processing

- **Research Assistant, Amirkabir University of Technology** — *Feb 2021 - Sep 2021 (8 months)*
  Supervisor: Prof. Amir Jahanshahi, Prof. Mohsen Moezzi
  Topics: Real-time Object Detection, Robotics, Data Augmentation, Sensors, Actuators, Single-Board Computers & Edge Devices

Publications

- **S. Rashidi**, H. Azadmaleki, A. Zolnour, M.J. Momeni Nezhad, and M. Zolnoori. "SpeechCura: A Novel Speech Augmentation Framework to Tackle Data Scarcity in Healthcare," ***Studies in Health Technology and Informatics***, vol. 318, 2025. PMID: 40776267, doi: 10.3233/SHTI251250.

- H. Azadmaleki, Y. Haghbin, **S. Rashidi**, M. J. Momeni Nezhad, and M. Zolnoori. "SpeechCARE: Dynamic Multimodal Modeling for Cognitive Screening in Diverse Linguistic and Speech Task Contexts," ***npj Digit. Med.*** 8, 677 (2025). doi: 10.1038/s41746-025-02026-x.

- H. Azadmaleki, A. Zolnour, **S. Rashidi**, et al. "TransformerCARE: A novel speech analysis pipeline using transformer-based models and audio augmentation techniques for cognitive impairment detection," ***International Journal of Medical Informatics***, vol. 207, 2026, p. 106208. doi: 10.1016/j.ijmedinf.2025.106208.

- A. Zolnour, H. Azadmaleki, Y. Haghbin, F. Taherinezhad, M.J. Momeni Nezhad, **S. Rashidi**, et al. "LLMCARE: early detection of cognitive impairment via transformer models enhanced by LLM-generated synthetic data," ***Frontiers in Artificial Intelligence***, 2025. doi: 10.3389/frai.2025.1669896.

- **S. Rashidi**, Y. Haghbin, H. Azadmaleki, A. Zolnour, and M. Zolnoori. "Leveraging Text-to-Speech and Voice Conversion as Data Augmentation for Alzheimer's Disease Detection from Spontaneous Speech," ***ICASSP 2026***. (under review)
  S. Rashidi*, Y. Haghbin*, M. McDonald, and M. Zolnoori. "Multimodal Attention Fusion of Speech and EHR Data for Early Detection of Cognitive Decline in Home Healthcare," ***ICASSP 2026***. (under review)

- H. Azadmaleki, Y. Haghbin, **S. Rashidi**, M.J. Momeni Nezhad, M. Naserian, E. Esmaeili, A. Zolnour, and M. Zolnoori. "SpeechCARE: Harnessing Multimodal Innovation to Transform Cognitive Impairment Detection — Insights from the National Institute on Aging Alzheimer's Speech Challenge," ***Studies in Health Technology and Informatics***, vol. 318, 2025. PMID: 40776266, doi: 10.3233/SHTI251249.

- M. Zolnoori, A. Zolnour, **S. Rashidi**, I. Spens, Y. Haghbin, et al. "Detecting Mild Cognitive Impairment Using Follow-Up Call Speech and Electronic Health Record Data in Home Health Care Settings," ***Journal of Gerontological Nursing***, 2025. (accepted)

- **S. Rashidi**, H. Sameti. "Improving Direct Persian-English Speech-to-Speech Translation with Discrete Units and Synthetic Parallel Data," ***arXiv preprint*** arXiv:2511.12690 2025. Available: https://arxiv.org/abs/2511.12690.

Teaching Experience

Teaching Assistant, Sharif University of Technology
- Speech Processing (Graduate Course, Instructor: Prof. Hossein Sameti) — *Fall 2023*
- Machine Learning (Graduate Course, Instructor: Prof. Hamid Beigy) — *Spring 2024*
- Signals and Systems (Undergraduate Course, Instructor: Mina Mahmoudi) — *Fall 2024*

Work Experience

ML Engineer (Speech & NLP), Sharif Information Systems and Data Science Center (Asr Gooyesh Pardaz)   *Jan 2022 - Jan 2024 (24 months)*
- Contributed to several projects focused on speech processing and natural language processing, including large language model (LLM)-based applications, automatic speech recognition (ASR), and text-to-speech (TTS) systems
- Led the team on the development of the RAG-based custom enterprise chatbot *"Dana"*, ensuring successful integration and performance optimization
- Improved existing models and built new models for state-of-the-art ASR and TTS for Persian language by 10%-20%
- Served as an advisor to multiple groups of interns, providing guidance on projects related to speech and language processing technologies, and fostering collaboration within the team

Related Technical Projects

Improving Automatic Speech Recognition for Persian Language using Transformer and Conformer models   *2023*
- Contributed to the development of the Sharif sub-domain for Persian ASR (nevisa.sharif.edu)
- Achieved state-of-the-art performance (8%-10% Word Error Rate) with Conformer model for Persian ASR.

Joint Speech and Text Keypoint Detection from Conference Speech and Audio Books   *2023*
- Fine-tuned SpeechT5 and implemented a retrieval system to detect keypoints in speech–text multimodal data with high precision.

Sentiment Analysis of Persian Product Reviews by Fine-Tuning LLMs and Pre-Trained Transformers   *2023*
- Built the Alpaca Persian dataset and fine-tuned BERT and LLaMA-2 models for high-accuracy sentiment analysis of Persian reviews.

Honors

- Achieved the **Special Recognition Prize** and the **Explainability Prize** in the NIA Challenge Competition: Model Arena (Acoustic Track) as a member of the "SpeechCARE" team
- Full Tuition Waiver Scholarship in AI and Robotics from **Sharif University of Technology** (highest-ranked technical university in Iran), 2021
- Full Tuition Waiver Scholarship in Electrical Engineering from **Amirkabir University of Technology**, 2016
- Ranked **11**th among nearly **15,000** participants in nationwide M.Sc. computer engineering entrance exam of universities, 2021
- Ranked **339**th among nearly **200,000** participants in nationwide B.Sc. mathematics entrance exam of universities, 2016

Professional Activity

Conference Reviewer, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2026

Skills

Technical Skills
- **Programming Languages**: Python, C++, C, JavaScript, TypeScript
- **ML/DL Packages**: Pytorch, Langchain, Tensorflow, Keras, Huggingface Transformers, Numpy, Scikit-learn, Matplotlib, Pandas, FAISS, MLFlow, Gradio
- **Utilities and OS**: Linux, Docker, Git, Microsoft Word and Excel, LaTeX

Soft Skills: Team Leadership, Analytical thinking, Problem solving, Experimental design, Academic writing, Interdisciplinary teamwork, Continuous learning

Language Skills: Persian (Native), English (Fluent)
   TOEFL iBT: **111**/120 (Reading **27**/30, Listening **29**/30, Speaking **26**/30, Writing **29**/30) (November 2024)

References

- **Maryam Zolnoori, Assistant Professor**, School of Nursing, Columbia University
  Email: mz2825@cumc.columbia.edu

- **Hossein Sameti, Associate Professor**, Computer Engineering Department, Sharif University of Technology
  Email: sameti@sharif.edu