ANITI
Porté par l'Université de Toulouse

POSTDOC PROPOSAL IN ARTIFICIAL INTELLIGENCE

Automatic speech recognition for an in-car voice assistant

Advisor (s): Thomas PELLEGRINI - thomas.pellegrini@irit.fr

https://www.irit.fr/~Thomas.Pellegrini/

Net salary: according to experience

Duration: 24 months

Location: Computer Science Research Institute of Toulouse (IRIT), Toulouse, France, https://www.irit.fr/en/

DESCRIPTION

This PostDoc position is proposed in the framework of the Audio Mobility 2030 (AM2030) project, which started in April 2023. AM2030 aims at enabling car manufacturers to have their own in-car audio application, regardless of the operating system. They will be able to deploy a global audio experience and offer the best content and proactive services to drivers. It is positioned as a true road companion that will help consumers adopt eco-responsible behaviors: vehicle self-diagnosis and maintenance reports, advice on driving and the use of on-board equipment.

Project partners: ETX Studio (Lead), Continental Automotive FRANCE SAS, **ANITI**, École Polytechnique de Paris

ANITI's role in the project is related to working on human-computer interactions, in particular on natural language understanding. The role of the hired PostDoc researcher will be to work more specifically on automatic speech (ASR, Speech-To-Text) in a noisy environment (the interior of a car). Two lines of research are envisaged: 1) adapting state-of-the-art open-source ASR models and self-supervised speech representation models (Wav2Vec2) to the noisy context of vehicles (presence of music/radio in the background, engine noise, wind noise, rain, etc.), 2) working on the language models that constrain end2end systems. Depending on the candidate research profile, one of these research lines will be chosen, This research will be conducted in connection with the two other aspects treated by ANITI: 1) the study of the conversational structures between the driver and the assistant and their semantic interpretation, 2) the detection of emotions and states of mind based on speech and transcription cues.

The hired PostDoc will be based at the Computer Science Research Institute of Toulouse (IRIT, https://www.irit.fr/en/), located in the campus of the Toulouse III Paul Sabatier University. They will be integrated in the Samova team, composed of about twenty permanent staff, PhD students and PostDocs whose research is related to various aspects of AI applied to speech and audio processing (https://www.irit.fr/SAMOVA/site/).

References

Baevski, A., Zhou, H., Mohamed, A., and Auli, M. wav2vec 2.0: A framework for self-supervised learning of speech representations. arXiv preprint arXiv:2006.11477, 2020

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision. *arXiv preprint arXiv:2212.04356*.

L. Gelin, M. Daniel, J. Pinquier, T. Pellegrini, 2021. End-to-end acoustic modelling for phone recognition of young readers. Speech Communication, 134, pp. 71-84.

REQUIRED SKILLS

Applicants should have a PhD in machine learning, ideally in speech/natural language processing. Good programming and English communication skills are also required.

APPLICATION PROCEDURE

Formal applications should include detailed cv, a motivation letter and reference letters.
Samples of published research by the candidate will be a plus.
> applications should be sent by email to: advisor email
More information: https://aniti.univ-toulouse.fr/