SpeechBrain

An Open-Source Conversational AI Toolkit

Get Started GitHub

Star 6,059 SpeechBrain 160 members

The call for Sponsors 2023 is open!

KEY FEATURES

SpeechBrain is an open-source conversational AI toolkit. We designed it to be simple, flexible, and well-documented. It achieves competitive performance in various domains.

Speech Recognition

SpeechBrain supports state-of-the-art methods for end-to-end speech recognition, including models based on CTC, CTC+attention, transducers, transformers, and neural language models relying on recurrent neural networks and transformers.

Speaker Recognition

Speaker recognition is already deployed in a wide variety of realistic applications. SpeechBrain provides different models for speaker recognition, including X-vector, ECAPA-TDNN, PLDA, contrastive learning

Speech Enhancement

Spectral masking, spectral mapping, and time-domain enhancement are different methods already available within SpeechBrain. Separation methods such as Conv-TasNet, DualPath RNN, and SepFormer are implemented as well.

Speech Processing

SpeechBrain provides efficient and GPU-friendly speech augmentation pipelines and acoustic features extraction, normalisation that can be used on-the-fly during your experiment.

Multi Microphone Processing

Combining multiple microphones is a powerful approach to achieve robustness in adverse acoustic environments. SpeechBrain provides various techniques for beamforming (e.g, delay-and-sum, MVDR, and GeV) and speaker localization.

Text-to-Speech

Text-to-Speech (TTS, also known as Speech Synthesis) allows users to generate speech signals from an input text. SpeechBrain supports popular models for TTS (e.g., Tacotron2) and Vocoders (e.g, HiFIGAN).

Other Tasks

SpeechBrain also supports Spoken Language Understanding, Language Modeling, Diarization, Speech Translation, Language Identification, Voice Activity Detection, Sound classification, Grapheme-to-Phoneme, and many others.

Research & Development

SpeechBrain is designed to speed-up research and development of speech technologies. It is modular, flexible, easy-to-customize, and contains several recipes for popular datasets. Documentation and tutorials are here to help newcomers using SpeechBrain.

HuggingFace!

SpeechBrain provides multiple pre-trained models that can easily be deployed with nicely designed interfaces. Transcribing, verifying speakers, enhancing speech, separating sources have never been that easy!

WHY SPEECHBRAIN?

EASY TO INSTALL

EASY TO USE

EASY TO CUSTOMIZE

Adapts to your needs.

SpeechBrain allows users to install either via PyPI to rapidly use the standard library or via a local install to view recipes and further explore the features of the toolkit.

Get Started Now

```
# From PyPI
pip install speechbrain

# Local installation
git clone https://github.com/speechbrain/speechbrain.git
cd speechbrain
pip install -r requirements.txt
pip install --editable .
```

SPONSORS

They are, or they sponsored SpeechBrain!
Our new call for sponsors (2023) is now open.

NAVER LABS Europe

Baidu 百度

OVHcloud

LABORATOIRE INFORMATIQUE D'AVIGNON
Avignon Université

ViaDialog
Customer Interactions Management

Mila

PREVIOUS SPONSORS

They are, or they sponsored SpeechBrain!
Our new call for sponsors (2022) is now open.

NUANCE

DOLBY

NVIDIA

SAMSUNG

COLLABORATORS

About Us

SpeechBrain isn't a company or an association. It is an open-source toolkit and a community created by Dr. Mirco Ravanelli and co-created by Dr. Titouan Parcollet. We aim at making speech technologies more accessible for the community.

Follow Us

Let us be social

Opportunities

Thanks to our sponsors, we often recruit talented candidates to continue expanding the functionalities of SpeechBrain. Feel free to contact us at: speechbrainproject@gmail.com