

Plan for project 2:

Michał Iwicki

Michał Legczylin

We are planning to test two pre-trained models based on transformer encoder architecture. Wav2vec works with raw audio and second AST with mel spectrograms of observation.

We also are going to implement our own CNN-based classification model consisting of two parts — 1D convolution based on raw vector of sound and 2D convolution based on mel spectrogram and then ensemble results.

We'll be using pytorch python framework for this.

As for hyperparameters we are going to explore learning and decay rates.

We will explore denoising methods probably using given noise data. This section will be developed as a way of improving basic models.

Michał Iwicki tasks: Fine-tune pretrained models.

Michał Legczylin tasks: CNN ensemble model implementation.

Other tasks will be split during working process.