

Project 2

Deep Learning

Students: Karina Tiurina, Mateusz Zacharecki

Plan

1. Transformer
2. ConvS2S
3. LSTM/GRU



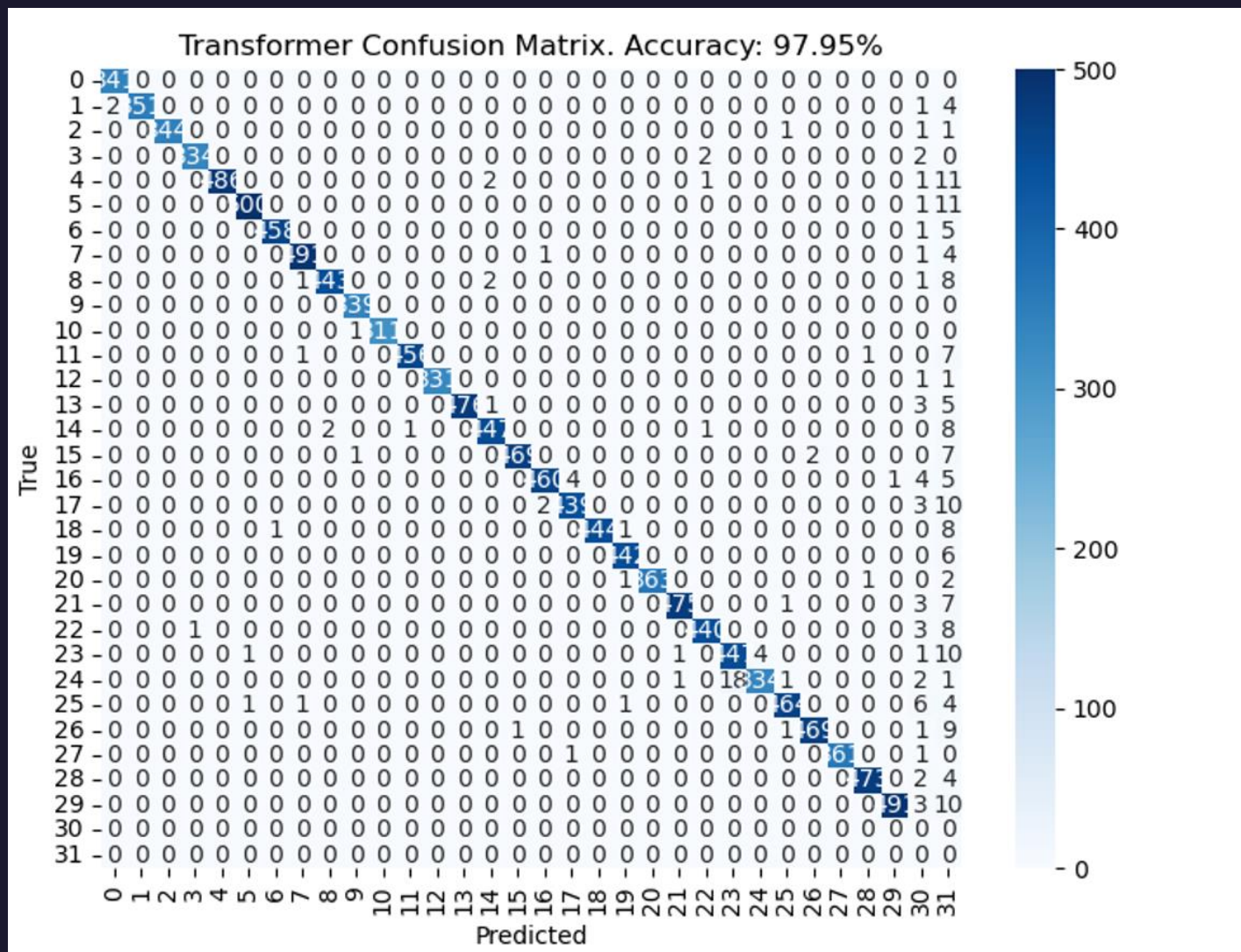
Data pre-processing approaches

1. Training on all 30 classes with artificial handling of 'unknown' and 'silence'.
 - Score is below 0.5 - 'unknown'
 - Score is below 0.3 - 'silence'
2. Training on 10 classes with 'unknown' and 'silence' as trainable classes.




81

- 81



Transformer wav2vec2

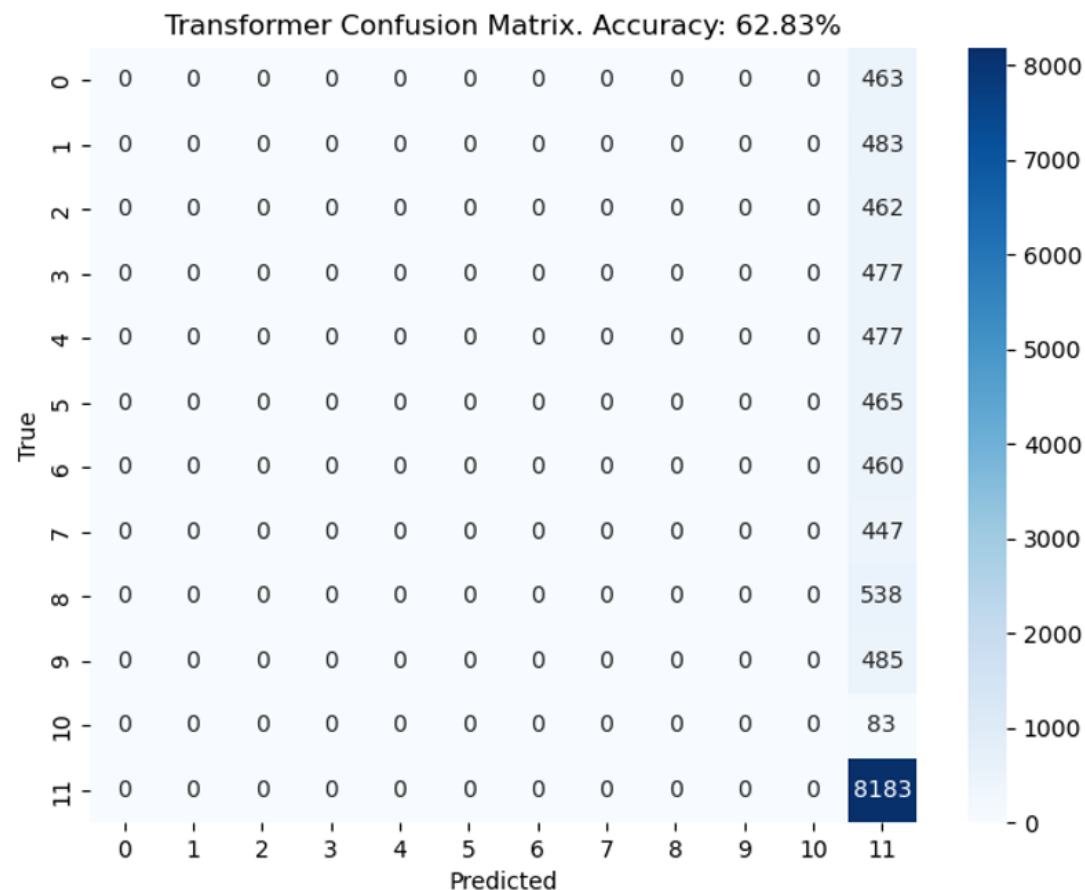
- All 30 classes;
- 'unknown' - below 0.5;
- 'silence' - below 0.3.
- Validation accuracy: 97.95%
- Kaggle accuracy: 81.09%

Submission and Description		Private Score ⓘ	Public Score ⓘ
	0_submission.csv Complete (after deadline) · 1d ago · Transformer (facebook/wav2vec2-base)	0.81099	0.80016



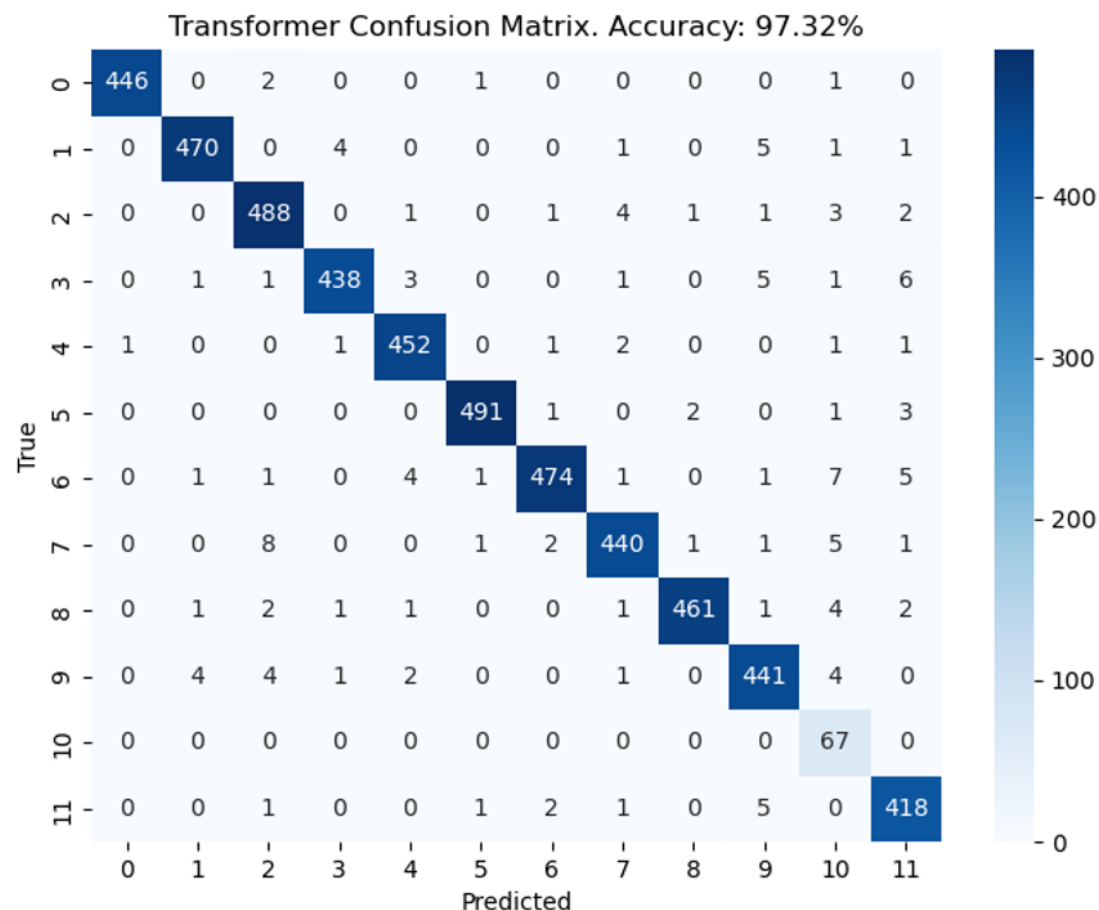
Transformer wav2vec2

- 10 classes with trainable 'unknown' and 'silence'
- Validation accuracy: 62.93%



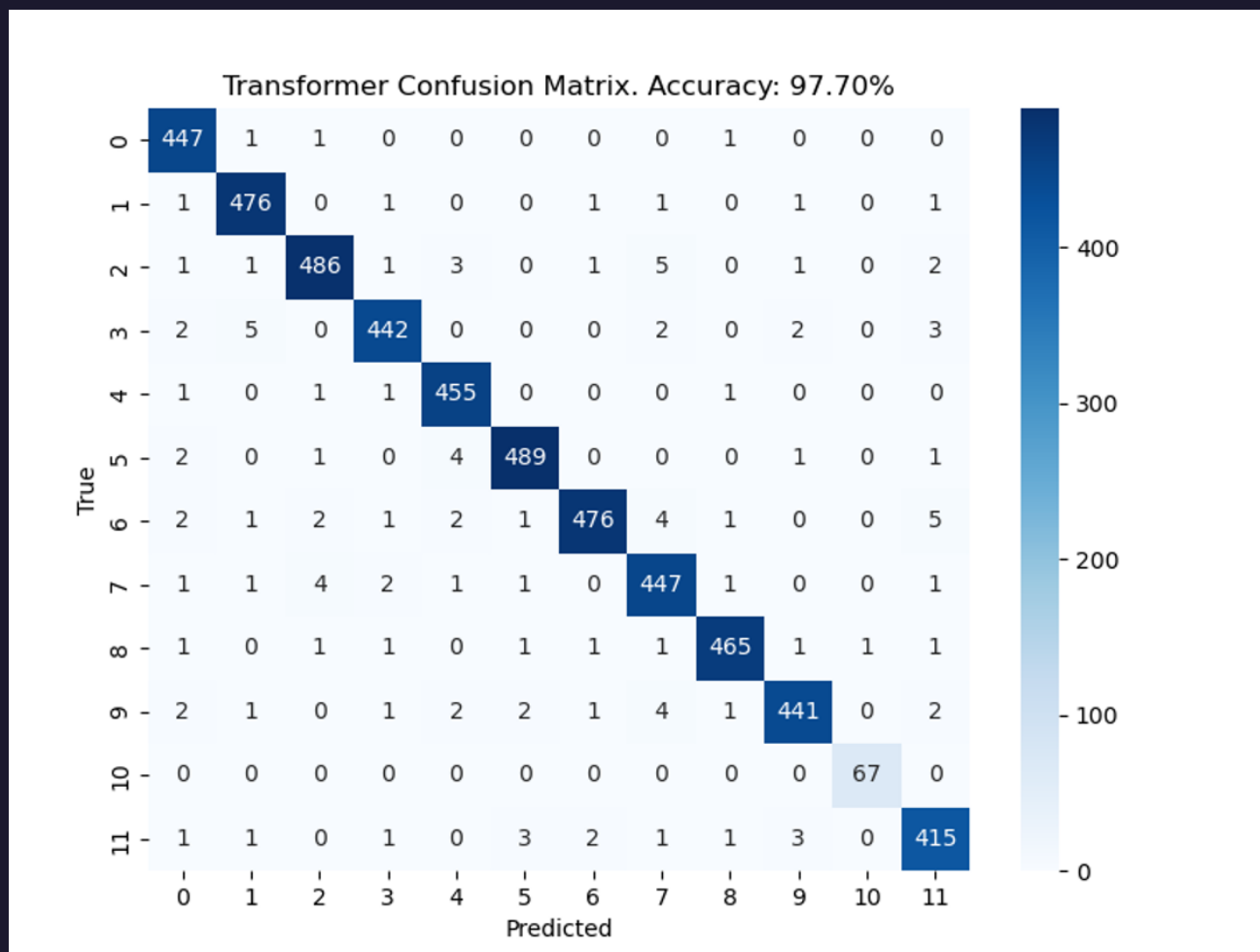
Transformer wav2vec2

- 10 classes with trainable 'unknown' and 'silence'
- Reduced the number of 'unknown' examples to ~2000
- Validation accuracy: 97.32%




Audio Spectrogram Transformer (MIT)

- 10 classes with trainable 'unknown' and 'silence'
- Reduced the number of 'unknown' examples to ~2000
- Validation accuracy: 97.70%
- Kaggle accuracy: 85.72%



Audio Spectrogram Transformer (MIT)

- 10 classes with trainable 'unknown' and 'silence'
- Validation accuracy: 97.70%
- Kaggle accuracy: 85.72%

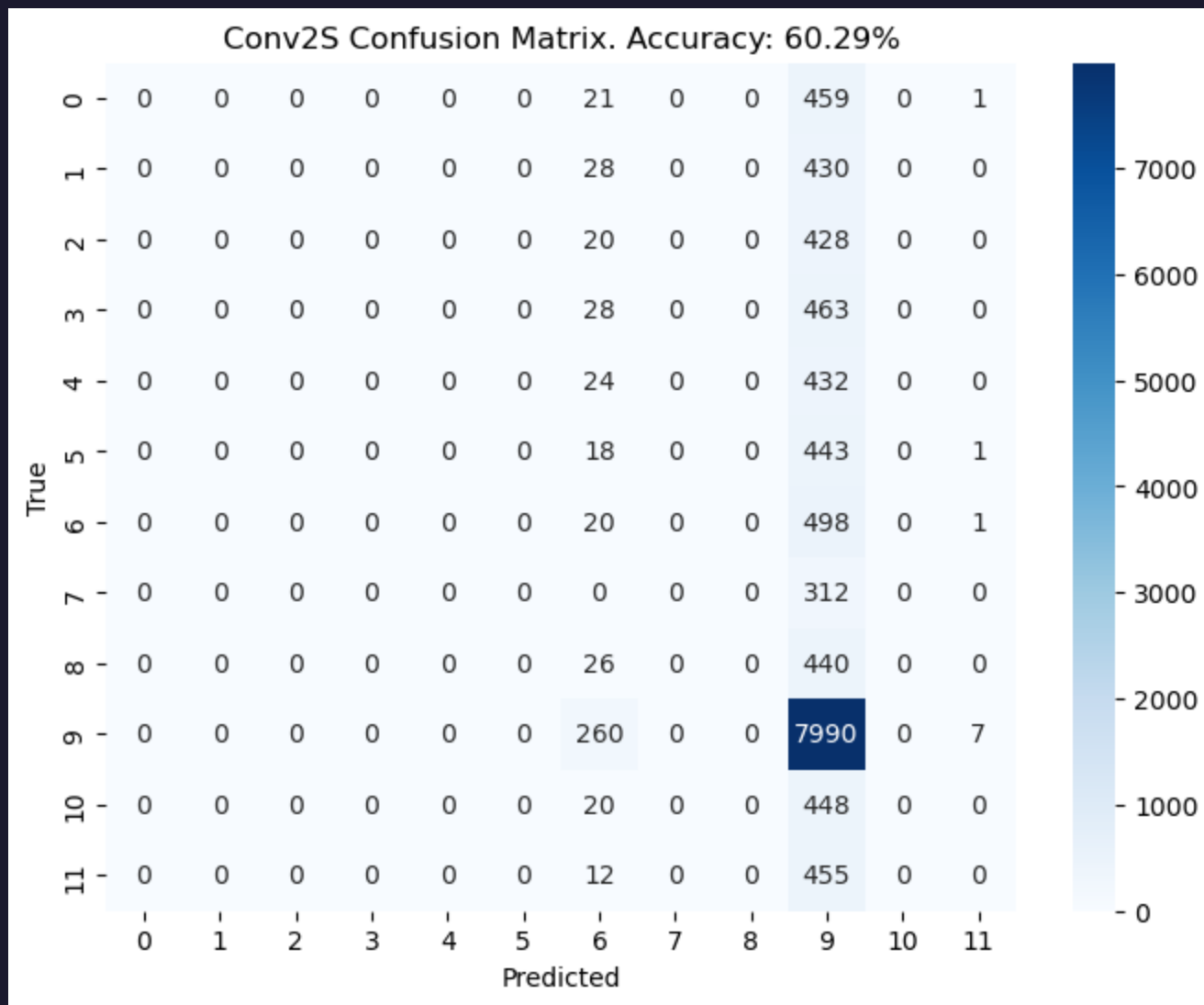
Submission and Description		Private Score ⓘ	Public Score ⓘ
	0_submission.csv Complete (after deadline) · 3h ago · MIT/ast-finetuned-speech-commands-v2	0.85727	0.84703

ConvS2S

- convs2s_1 - ConvS2S model with mse loss and adam optimizer
- convs2s_2 - ConvS2S model with mse loss and adamax optimizer
- convs2s_3 - ConvS2S model with mse loss and sgd optimizer
- convs2s_4 - ConvS2S model with categorical crossentropy loss and adam optimizer
- convs2s_5 - ConvS2S model with mse loss and RMSprop optimizer
- convs2s_6 - ConvS2S model with mse loss and RMSprop optimizer for data with reduced unknown class observations
- convs2s_7 - ConvS2S model with mse loss and RMSprop optimizer for data with reduced unknown class observations and for batch_size = 256

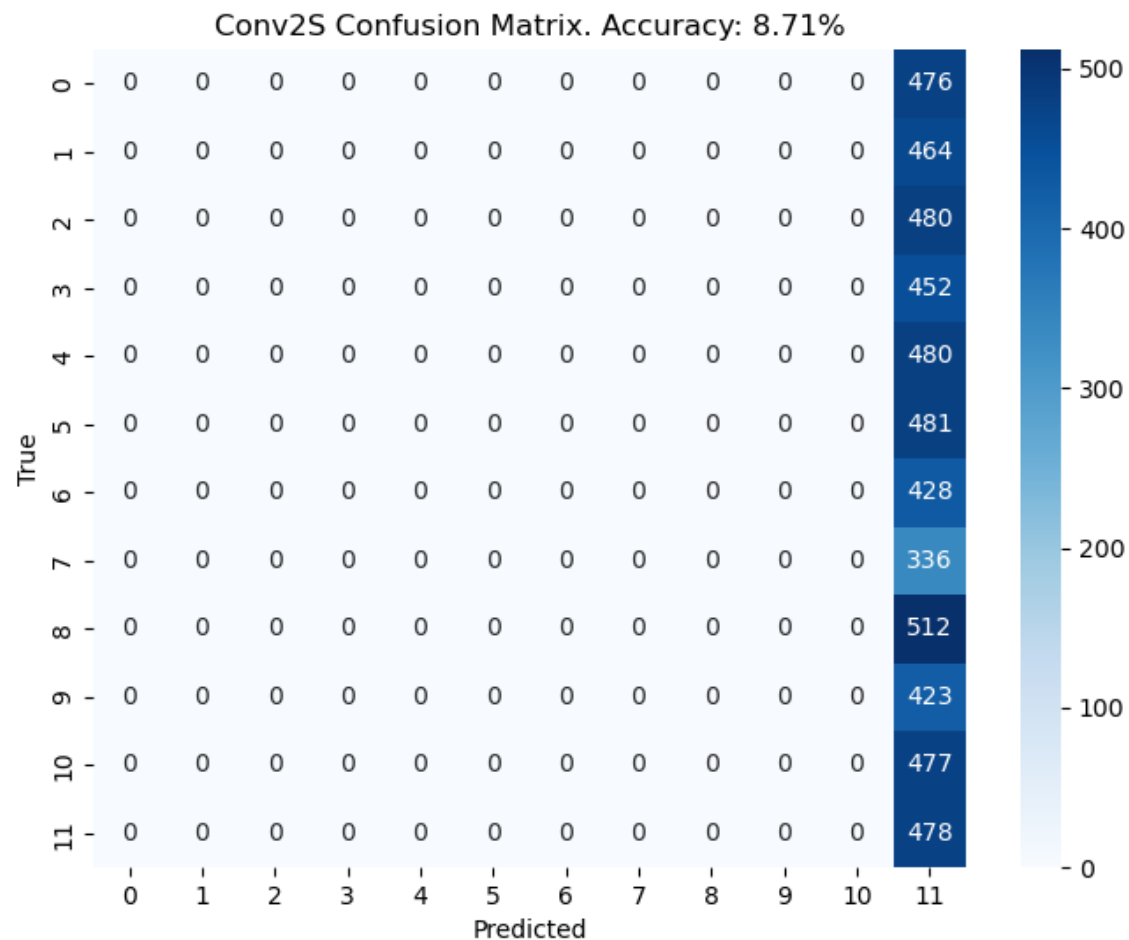
ConvS2S

The first 5 tests were conducted for observations of not considered classes assigned as unknown class and data augmentation for silence class.



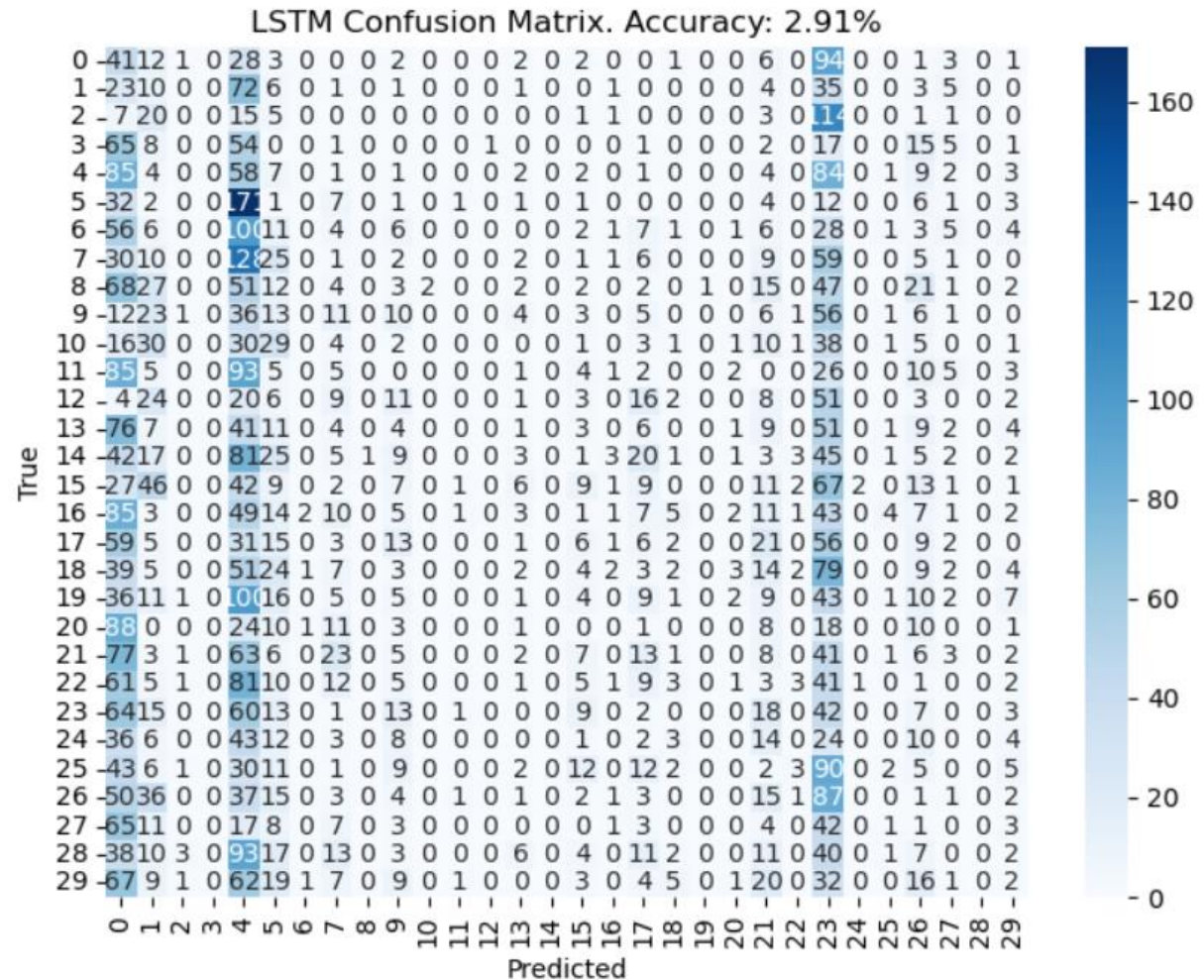
ConvS2S

The last 2 tests were conducted for additional reduction of observations from unknown class to about 2000 in order to get similar numbers of observations from each class and, as result, get better accuracy.



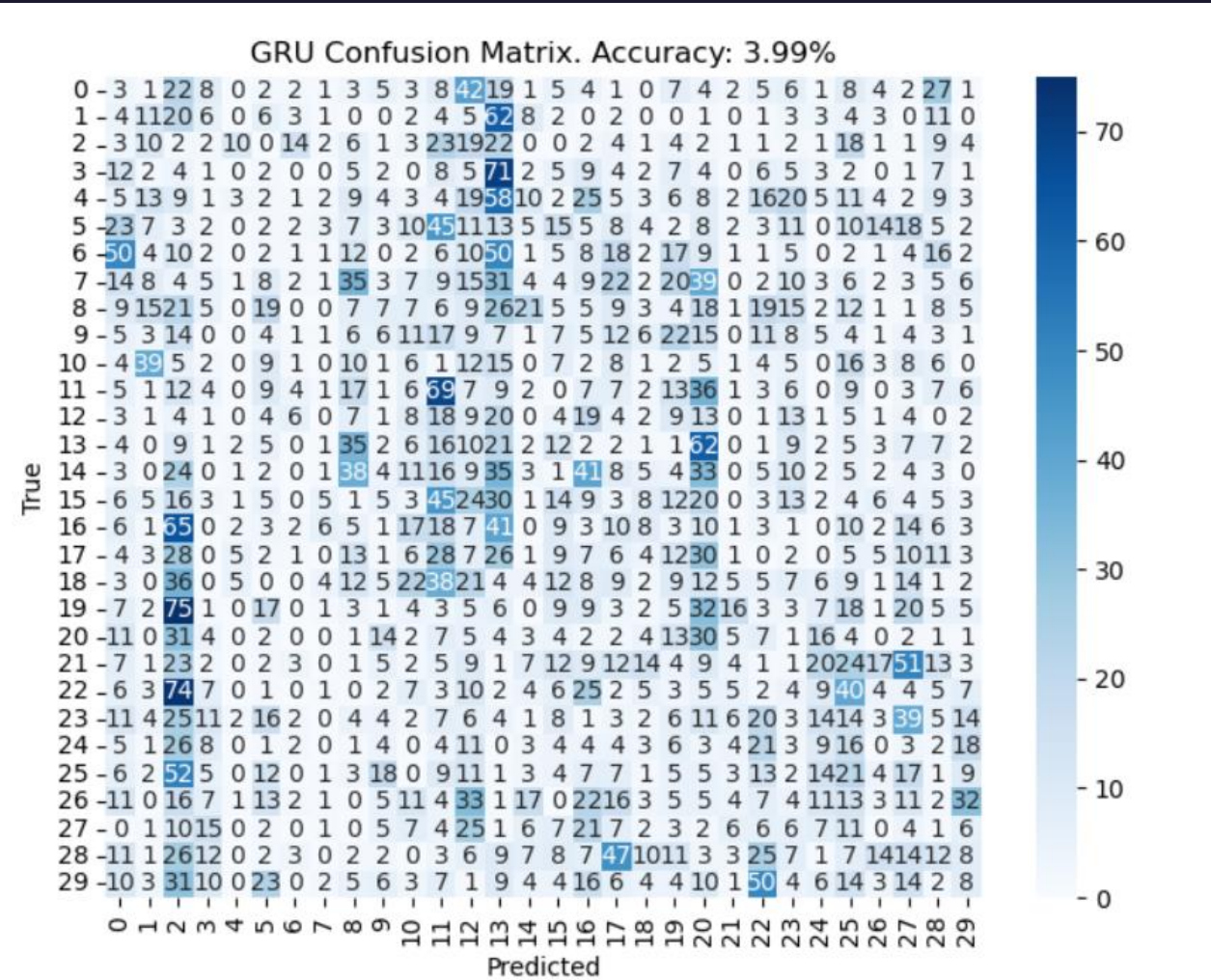
LSTM/GRU

- LSTM accuracy: 2.91%
- GRU accuracy: 3.99%



LSTM/GRU

- LSTM accuracy: 2.91%
- GRU accuracy: 3.99%



Thank you for attention!

