Cross Lingual Speaker Adaptation for TTS Applications

Software Project-2021/2022

Rasul, Anna, Claesia, Sharmila

Outline

- ❖ Four model versions
- ❖ Finished tasks
  - ➢ New data
  - ➢ Multi-gpu
  - ➢ Training
  - ➢ Inference
- ❖ Demo
- ❖ Plans
  - ➢ Evaluation
  - ➢ Conferences
- ❖ Timeline

Four model versions

	Language embedding	Speaker embedding
Version 1	ID	ID
Version 2	ID	Embedding Network
Version 3	Embedding Network	ID
Version 4	Embedding Network	Embedding Network

Finished tasks

1) New data

- Tundra
- Synpaflex (extracted 5 speakers)

Standard preprocessing

Data Ratio: 12 speakers (french 6 and english 6 speakers / 2 languages) 26478 training, 414 test, 101 validation data

2) Multi-gpu support

Code modifications

Finished tasks

3) Start training

Modifications of the model. With one script run any version:

```
python train.py --speaker id --language embedding
```

4) Inference

Inference takes 3 parameters:

```
python inference.py - -file <text_file> - -v <1,2,3,4
model, int for each model> --timesteps <diffusion rate 1 to
1000>
```
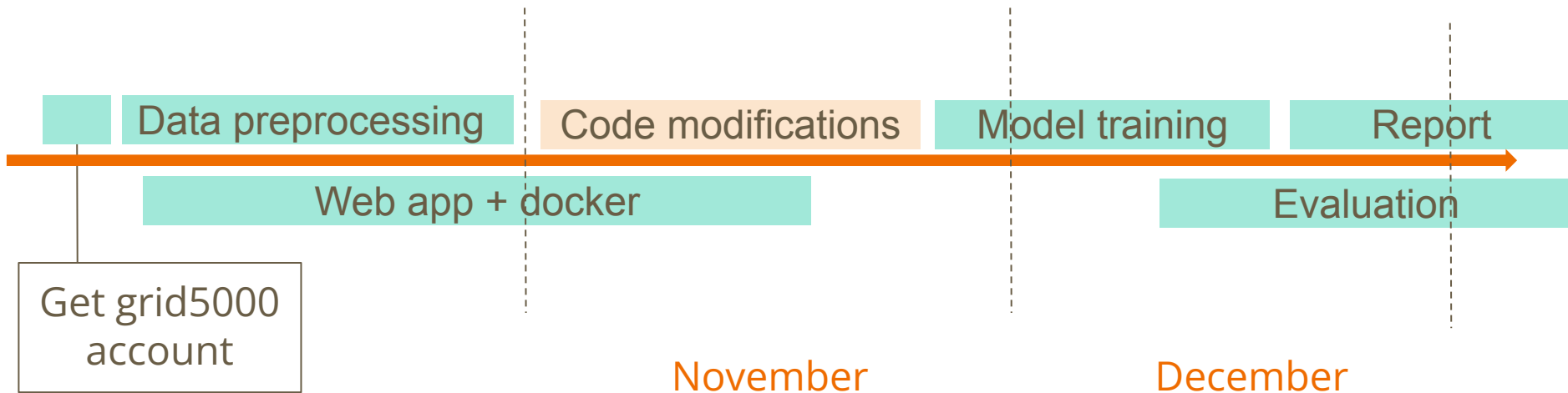
Demo time!

Plans

1) Evaluation

Write script for evaluation.

2) Conferences

- European Signal Processing Conference: https://2022.eusipco.org/
- Interspeech: https://interspeech2022.org/

Submit a paper based on our report.

Timeline

Thank you! Any questions?