

Software Development Project



About Us : Infographic



Skills - Python, JS ,C
Experience -



Skills-Python, Linguistic
Experience – Database Creation



Skills – Scala , Python
Experience -



Skills – Java , Python

Project Planning

- Obtain multilingual dataset (build corpus if needed)
- Preprocess dataset to acquire useful representation for the task
- Design an appropriate nn architecture
- Research most suitable loss functions
- Reusing pretrained embeddings from the chosen languages
- Implement a speech synthesis tool uttering inputted text along with the annotations needed for the correct pronunciation,
- Fine-tune hyperparameters
- Build a simple GUI for the end-user to communicate with the application conveniently

Technology Stack

- Python
- PyTorch
- Web Development (GUI)
- Processing and Playing audio

Objective

- Main goal: a web interface for TTS
- Multilingual TTS - English and French
- Input: text
- Output: speech

Datasets

English: LJS - female voice (already preprocessed)

<https://keithito.com/LJ-Speech-Dataset/>

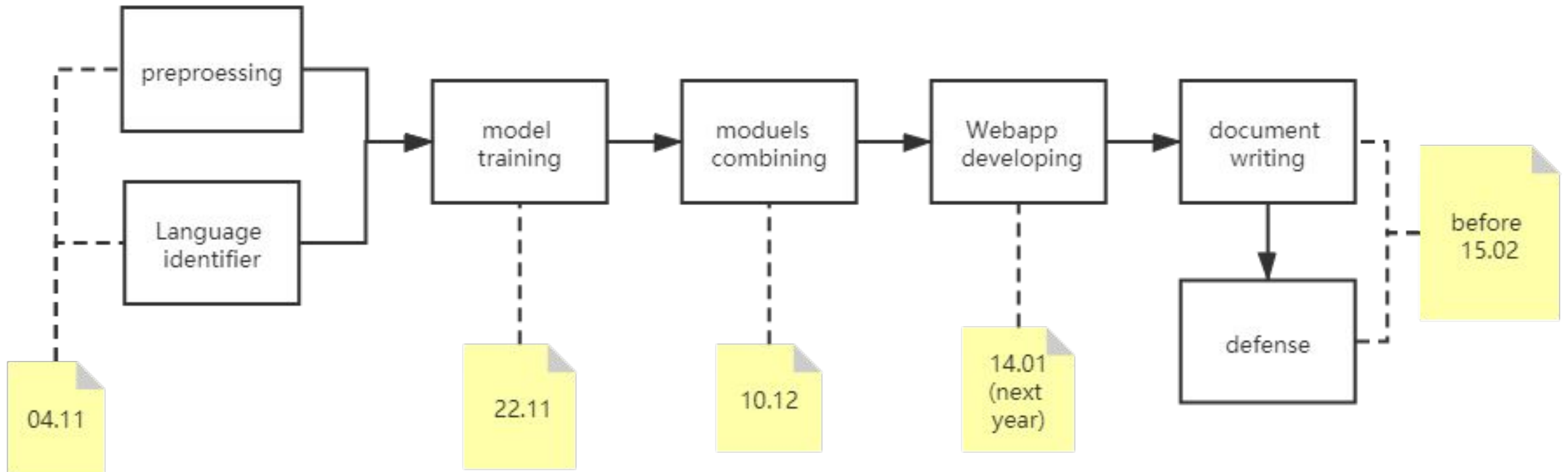
- 13100 utterances (total length ~24h)

French: SIWIS - female voice (**to be preprocessed**)

<https://datashare.ed.ac.uk/handle/10283/2353>

- 9750 utterances (~10+h)

Work planning



Strategy

- Access to Grid 5k
- Work in pairs first and split when possible
- Read paper (Grad-TTS)
- Possibility of collaboration with the other group for pre-processing, deployment, ...
- PyTorch

Model : Grad TTS

GradTTS to synthesize the speech from the preprocessed data
<https://github.com/huawei-noah/Speech-Backbones/tree/main/Grad-TTS>

- should already be pre-trained on English
- need to train with French

Given a sequence of phonemes, generate acoustic features as a Mel spectrogram that can be turned into speech waveform

Current Work

- A repo has been created
- Preprocessing for French
 - (1) perl script
 - (2) generate the labels & extract main phonemes
- Language Identifier
 - (1) LSTM classifier
 - (2) French & English monolingual data