Voice Annotation Tool

Release 1.8.0

Feliks Weber

Jun 15, 2022

CONTENTS

This program allows annotation of a list of audio files and exporting it in a similar format as Mozilla's CommonVoice dataset.

CONTENTS

1.1 Usage

1.1.1 Installation

The tool can be install using pip:

```
$ pip install voice-annotation-tool
```

To start the program you can use the command line script or execute the python module:

```
$ voice-annotation-tool
```

```
$ python -m voice_annotation_tool
```

1.1.2 Creating a Project

To get started, click the "Create Project" button. You will be prompted to open the project settings, where you can choose the folder with the audio samples and the location of the tsv file.

The TSV file is used to store the metadata of the audio samples, including the annotated text and the age, gender and accent of the speaker. More about the format can be found here: *The CommonVoice TSV Format*

1.2 The User Interface

1.2.1 Sample List

In this list you can select the sample you want to annotate.

You can delete the selected samples under `Edit > Delete Selected` or pressing `Delete`.

When you change the annotation of a sample it is highlighted in green. To remove this highlight select the samples and press the `Mark Unchanged`.

1.2.2 Metadata Section

In this region you can edit the metadata of all selected annotations.

You can also import the profile metadata exported from the CommonVoice account page.

1.2.3 Annotation Field

Here you can edit the annotation of the selected sample.

1.2.4 Audio Playback Controls

To play the audio of the selected sample, press the play button. There are also buttons to move to the next / previous sample. To speed up the workflow you can also assign shortcuts to these buttons: *Keyboard Shortcuts*

1.3 Keyboard Shortcuts

Keyboard shortcuts can be configured under `Settings > Keyboard Shortcuts.`

To change a shortcut, click the button and press the shortcut you want to assign.

Right-click the button to clear the shortcut.

1.4 The CommonVoice TSV Format

The TSV format can store tabular data using tab separated values.

The fields included in the TSV files of the CommonVoice dataset are explained here: https://github.com/common-voice/cv-dataset#fields

The order of the samples in the TSV file is kept, with new samples added to the end.

1.5 Import and Export

The annotations can be exported and imported in multiple formats, including Json and CSV.

These are meant for easy reading and modification by other applications, and only include the sample file name and the annotated text.

1.5.1 Json

Listing 1: An example of an exported Json file

```
[
  {
    "file": "sample1.mp3",
    "text": "Annotation One",
  },
```

```
  {
    "file": "sample2.mp3",
    "text": "Annotation Two",
  }
]
```

1.5.2 CSV

Listing 2: An example of an exported CSV file

```
file;text
sample1.mp3;Annotation One
sample2.mp3;Annotation Two
```

1.6 API

The voice-annotation-tool package also includes an API to load, modify and save CommonVoice datasets.

Listing 3: Loading a TSV file using the Project class

```python
from voice_annotation_tool.project import Project

# Create a new project.
project = Project()

# Load a TSV file.
with open('tsv_file.tsv') as file:
    project.load_tsv(file)

# Show all annotations.
print(project.annotations)
```

Listing 4: Changing the sentence and votes of an annotation

```python
from voice_annotation_tool.project import Project
from voice_annotation_tool.annotation import Annotation

project = Project()
project.load_audio_files(Path('audio'))

annotation = project.annotations[0]
annotation.sentence = 'Annotated text'
annotation.age = 'thirties'
annotation.up_votes += 1

with open('tsv_file.tsv', 'w', newline='') as file:
    project.save_annotations(file)
```