

Song Deep Thought Bot



Данные

Youtube (youtube-dl)

[Songs to Sing in the Shower 2017 \(Best Shower Music Playlist\)](#)

[80s Music Hits - 80s Playlist Greatest Hits \(Best 80s Songs\)](#)

[Pop Music Playlist - Timeless Pop Songs \(Updated Weekly 2021\)](#)

~700 песен, из них больше половины отсеились по наличию английских субтитров

NLP

InferSent vs SentenceTransformers (эмбеддинги)

Дополнительно использовался EMNLP dataset

message -> похожая реплика датасета -> ответ на реплику -> похожая строчка в песенном датасете

k-sampling

Синхронизация текста и аудио

- Использовали предобученный QuartzNet.
- Синхронизация: проходимся окном размера 10 символов и сравниваем по CER. Хотим найти биекцию, причём с правильной последовательностью.
- Используем механизм, аналогичный guided attention.
- Используем beam-search.

Синхронизация текста и аудио

Результат: большинство строчек синхронизируется в пределах 3-5 секунд.

Для более точной синхронизации:

- Выбираем окно +/- 3 секунды от предполагаемого отрывка
- Находим наилучший отрывок по CER

Бот @sond_deep_thought_bot

telebot

