

Использование обучения с подкреплением для решения задачи распознавания диктора в интерактивном режиме

Головин Вячеслав Сергеевич

2023

Содержание

Введение	1
1 Распознавание диктора в интерактивном режиме	2
1.1 Задача распознавание диктора	2
1.2 Интерактивный режим	2
2 Детали реализации и результаты	2
2.1 Данные для обучения и извлечение эмбедингов	2
2.2 Обучение <code>Guesser</code>	2
2.3 Обучение <code>Enquirer</code>	2
2.4 Эвристическая модель выбора слов	2
3 Модификации метода	2
3.1 От идентификации к верификации	2
3.2 <code>CodebookEnquirer</code>	3
3.3 Добавление шума	3
3.4 Альтернативные эмбединги	3
Заключение	3

Введение

бла-бла-бла, распознавание диктора это важно и полезно...

1 Распознавание диктора в интерактивном режиме

1.1 Задача распознавание диктора

...

1.2 Интерактивный режим

...

2 Детали реализации и результаты

2.1 Данные для обучения и извлечение эмбеддингов

Аналогично [1] для обучения и тестирования моделей мы использовали датасет TIMIT[2]. Он составлен из аудиозаписей речи 630 дикторов, говорящих на 8 основных диалектах американского английского языка. Эти дикторы

2.2 Обучение Guesser

...

2.3 Обучение Enquirer

...

2.4 Эвристическая модель выбора слов

...

3 Модификации метода

3.1 От идентификации к верификации

Enquirer менять вообще не нужно, Guesser — совсем немного.

3.2 CodebookEnquirer

вроде работает

3.3 Добавление шума

обучается норм, результаты такие же

3.4 Альтернативные эмбединги

внезапно эмбединги из 2018 года оказались не очень

Заключение

все работает, но хотелось бы большего

Список литературы

- [1] M. Seurin, F. Strub, P. Preux и O. Pietquin, *A machine of few words – interactive speaker recognition with reinforcement learning*, 2020. arXiv: 2008.03127 [eess.AS].
- [2] Garofolo, John S. и др., *TIMIT acoustic-phonetic continuous speech corpus*, 1993. DOI: 10.35111/17GK-BN40. url: <https://catalog.ldc.upenn.edu/LDC93S1>.