

“Дифтонги” в кильдинском саамском

Антон Бузанов, Илья Егоров, Василий Молчанов
2023

Экспедиция в Ловозеро летом 2022-го

<я>, <е> — /ea/, /ie/ ИЛИ /ja/, /je/

- Итконен (1958), Итконен (1971) и другие исследователи считают, что это дифтонги.
- Исследователи из России (Керт, 1971; Куруч, 1985) считают, что это монофтонги.

Сколтосаамский

- Одна носительница в России.
- Говорит на плохо описанном диалекте.
- В сколтосаамском существуют сегменты, на официальной орфографии (создана на основе финских диалектов) обозначаемые как <ie> и <eä>.
- Дифтонги ли это?

Наши задачи

- Понять, как автоматизировано отличать дифтонги от монофтонгов.
 - Посмотреть, какие способы предложены в литературе, и протестировать их
- Понять природу наших сегментов.
 - В кильдинском <я>, <е>.
 - В сколтосаамском <еӕ>, <іе>.

Rate of Change

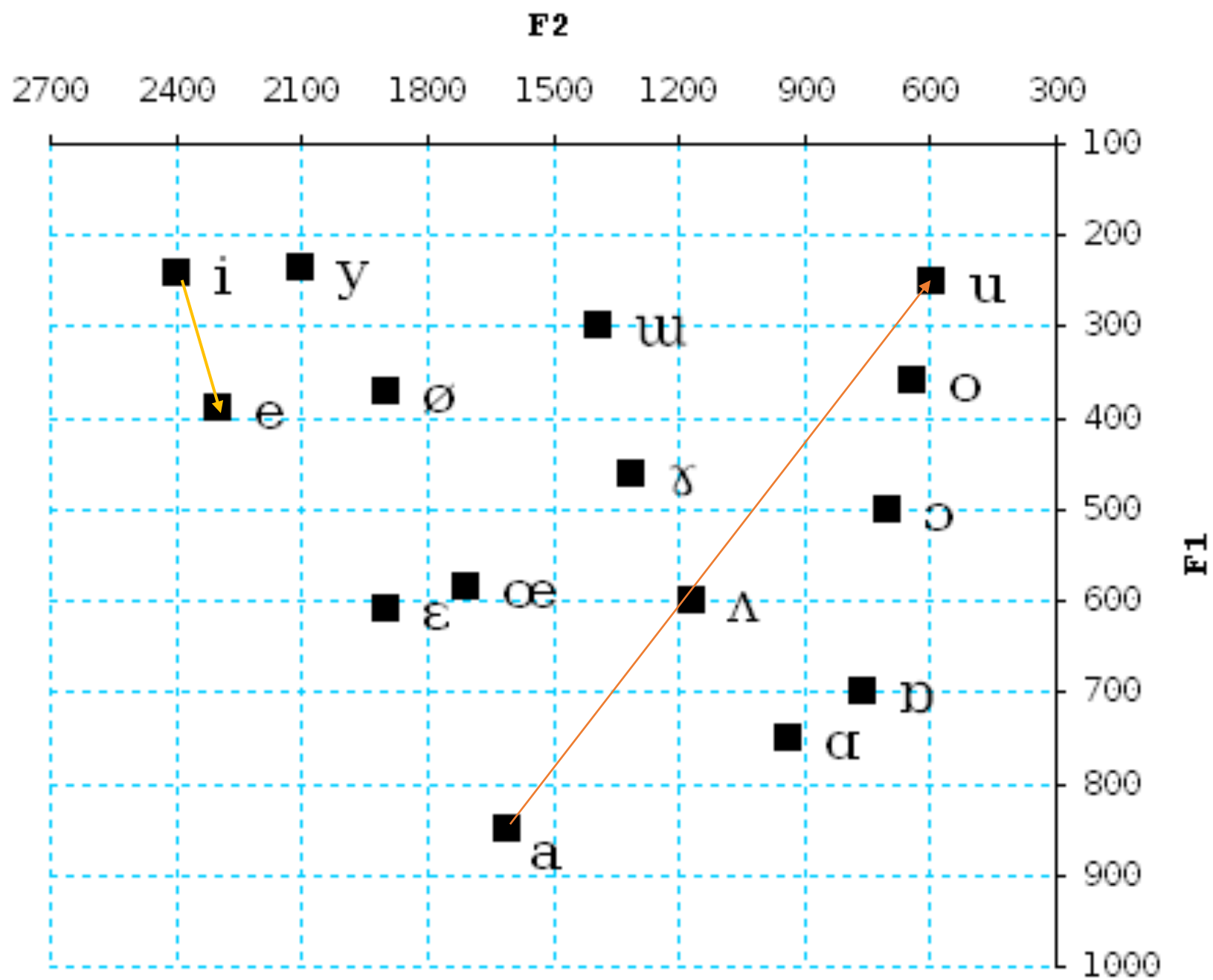
$$\frac{F1_{\text{конец}} - F1_{\text{начало}}}{\text{продолжительность в секундах}} = \text{ROC (Hz/s)}$$

- Надо устанавливать пороговое значение самостоятельно.
- Может мешать коартикуляция.

Trajectory Length

$$\sqrt{(F1_{0.2} - F1_{0.8})^2 + (F2^{0.2} - F2^{0.8})^2}$$

- Более элегантный способ: учитывает сразу две форманты.
- Та же проблема: с какого показателя начинается дифтонг?



Changepoint Detection

- Мы предложили этот метод для этой задачи.
- Семейство алгоритмов, основная суть:
алгоритм ищет точки, в которых происходят статистически значимые изменения.

У дифтонгов мы их ждём, у монофтонгов нет.

Cepstral Coefficients

- Кепстральные коэффициенты — это числа, которые используются для анализа и обработки звуковых сигналов.
- Задействуются в различных областях, включая распознавание речи, обработку звука и музыкальную аналитику.

Тест

- Берём настоящие дифтонги из языков, где они точно есть (эстонский, литовский).
- Смотрим, насколько метрики хорошо отличают их от монофтонгов из этих языков.
- Если хорошо, то с их помощью анализируем саамские сегменты.

Данные

Кильдинский	5 (женщины)
Сколто-саамский	1 (женщина)
Эстонский	2 (женщина и мужчина)
Литовский	2 (женщины)

Все сегменты, которые мы анализировали

	Монофтонги	n	Дифтонги	n
Кильдинский	i: e: a:	149	ie (ʲe:) ea (ʲa:)	120
Сколто-саамский	õ õ: â â: e e:	82	iõ eâ ie	47
Литовский	ɛ ɛ: e: i i:	531	iɛ	162
Эстонский	æ æ: æ:: e e: e:: i i: i::	439	ie æe ea	123

ROC

- Справился очень плохо.
- Анализировать саамские не берём.

данные	сегмент	ROC
эстонский	i::	0.2
(мужчина)	æe	0.2
	e::	0.3
	e:	0.6
	æ:	0.6
	i:	0.7
	æ::	0.8
	ie	0.8
	ea	1.2
	e	1.5
	i	1.5.
	æ	1.8

данные	сегмент	ROC
эстонский	æ::	0.2
(женщина)	i::	0.2
	e:	0.3
	e::	0.6
	æ:	0.6
	æe	0.7
	i:	0.8
	i	0.8
	ie	1.2
	ea	1.5
	e	1.5
	æ	1.8

данные	сегмент	ROC
литовски й	'e:	1.0
	i	1.4
	ε:	1.5
	'i	1.5
	'i:	1.7
	e:	1.7
	'iε	1.9
	'ε:	1.9
	i:	2.1
	ε	2.3
	iε	2.4

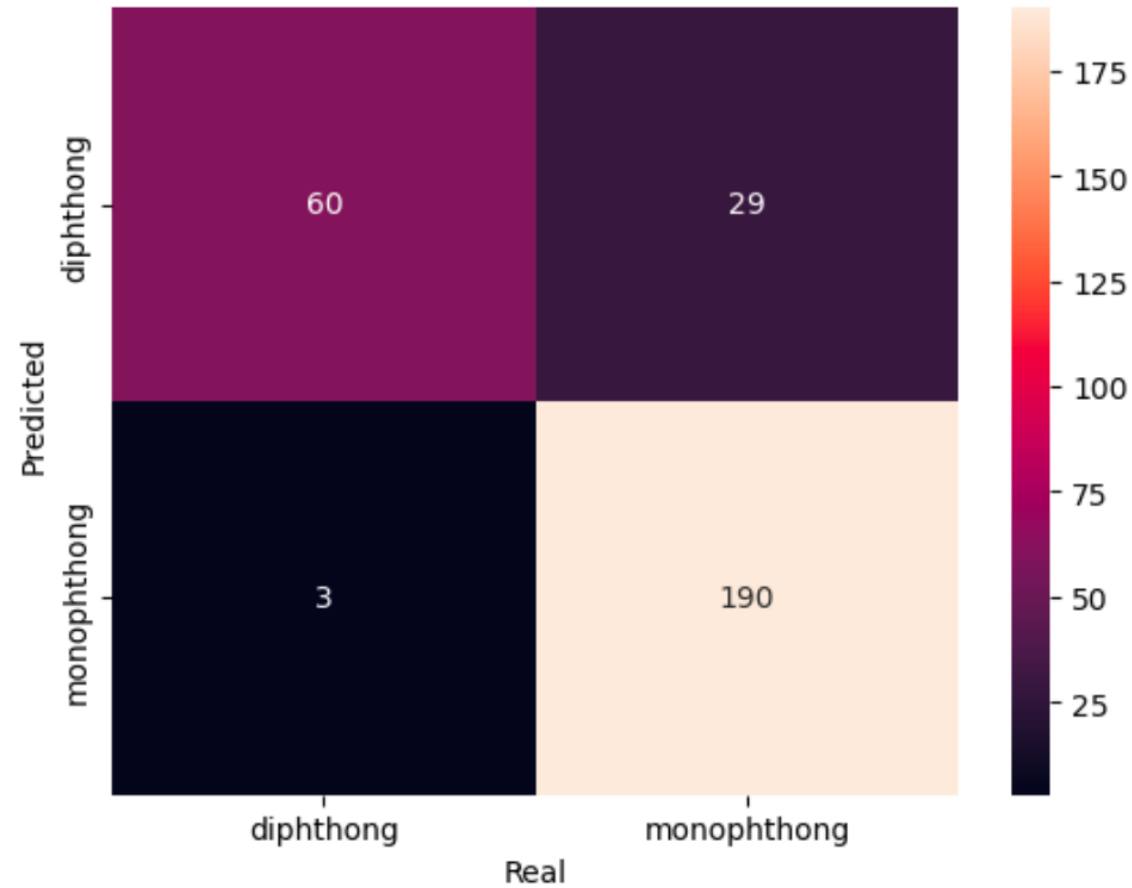
Trajectory Length

- Пороговое значение — среднее средних:

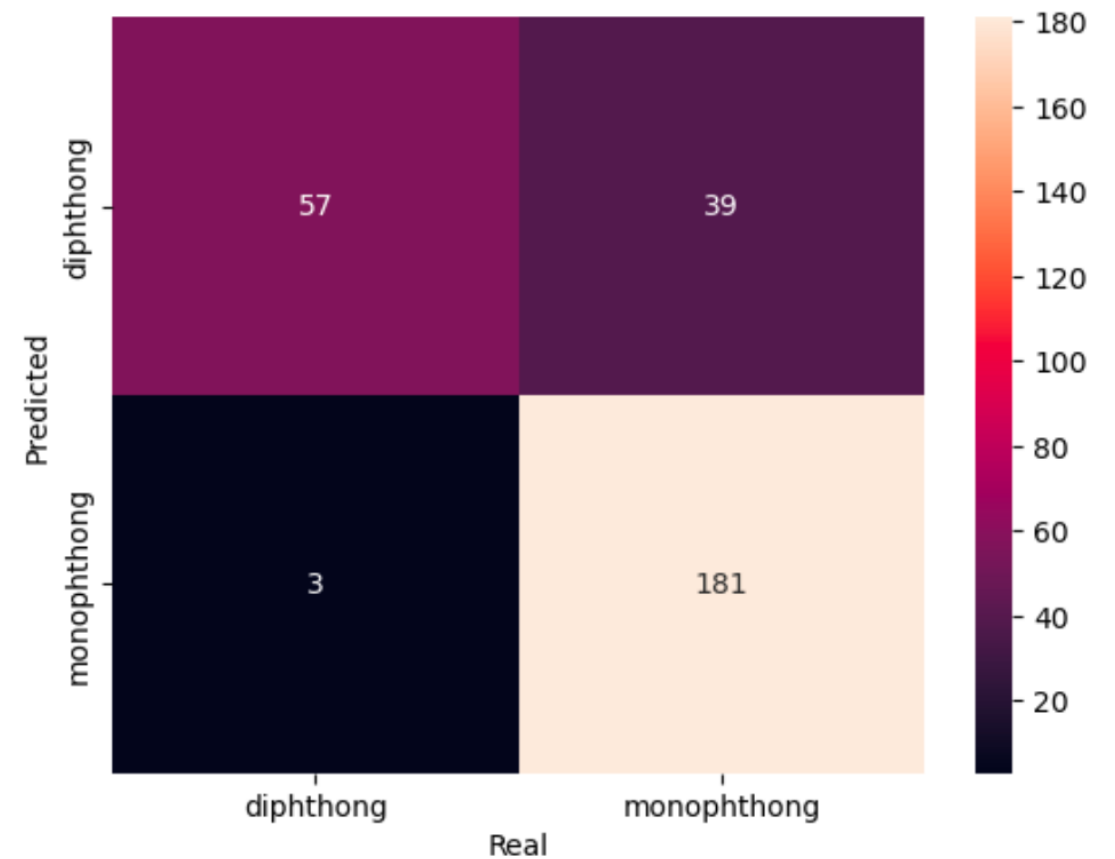
$$\frac{\text{средняя длина дифтонгов}}{\text{средняя длина монофтонгов}}$$

- Для хорошей работы метрики нужны сбалансированные данные (у нас сбалансированные).

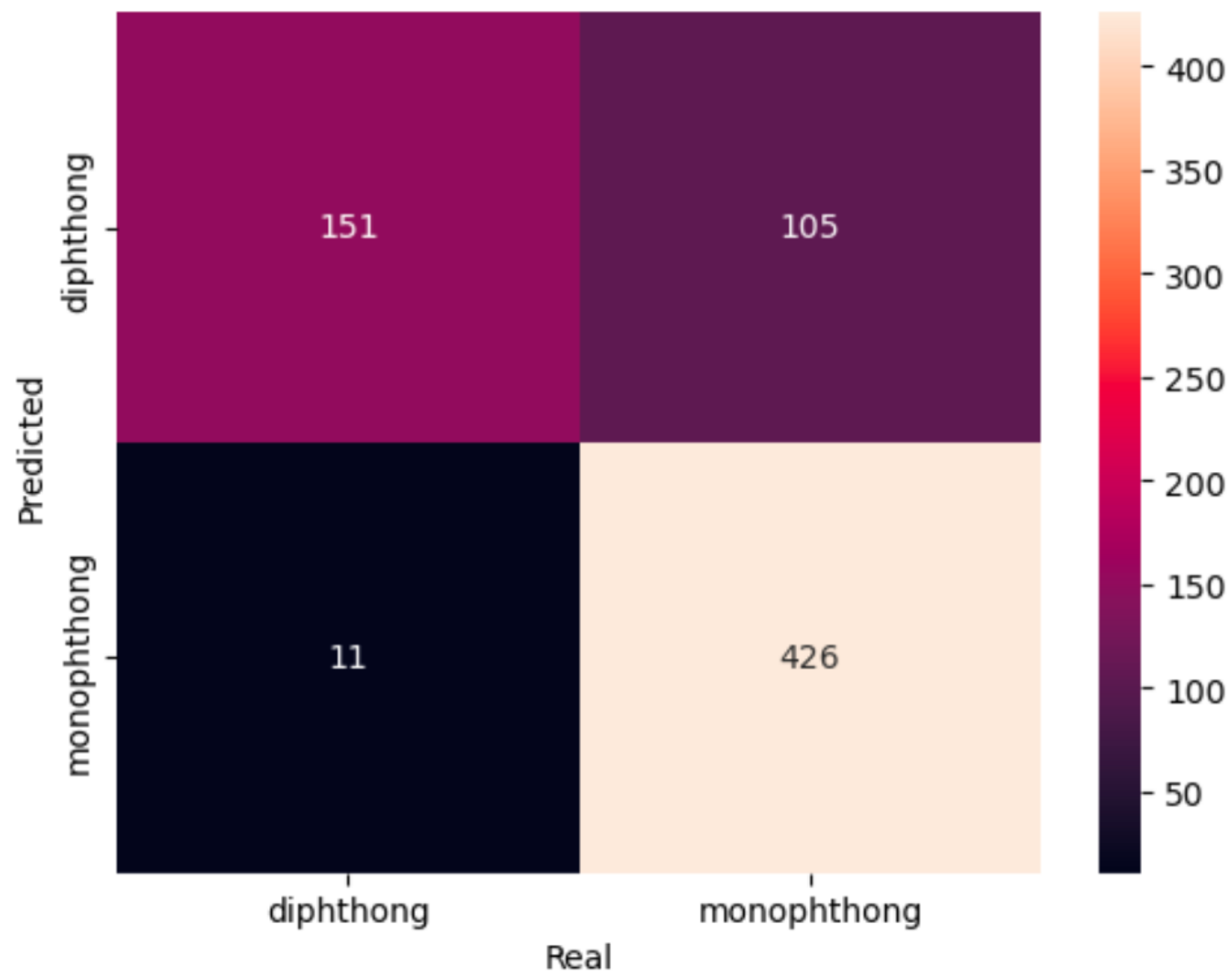
Trajectory Length



(a) Estonian (male)



(b) Estonian (female)



(c) Lithuanian

Trajectory Length

данные	сбалансированная точность
эстонский (М)	0.91
эстонский (Ж)	0.89
литовский	0.87

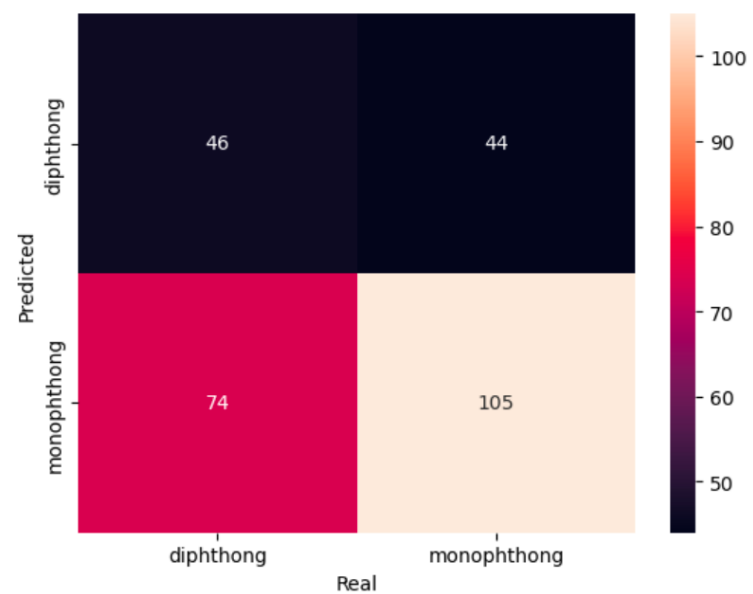
Changepoint Detection

данные	сбалансированная точность
эстонский (М)	0.76
эстонский (Ж)	0.81
литовский	0.68

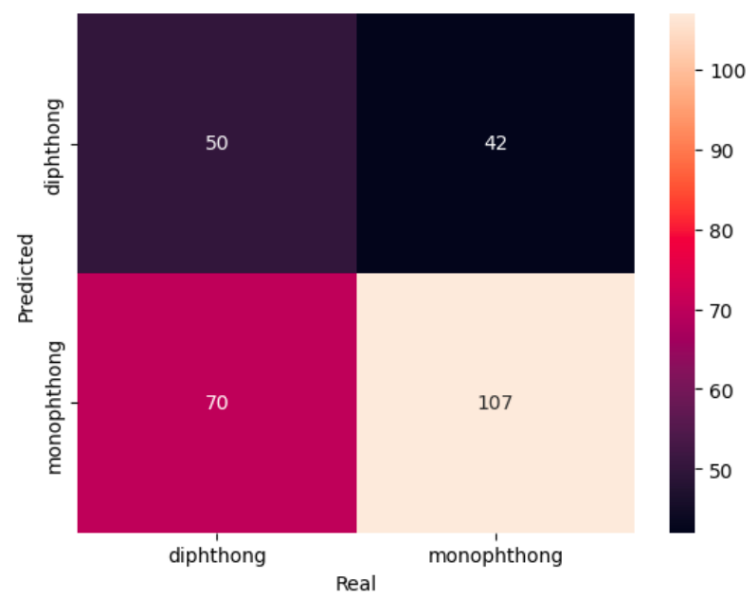
Cepstral Coefficients

данные	сбалансированная точность
эстонский (М)	0.79
эстонский (Ж)	0.89
литовский	0.81

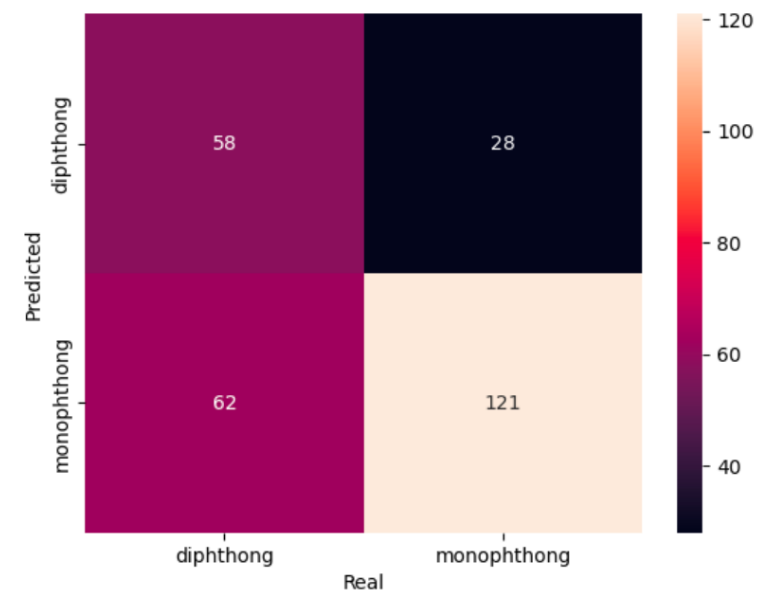
Кильдинский



(a) changepoint



(b) trajectory length

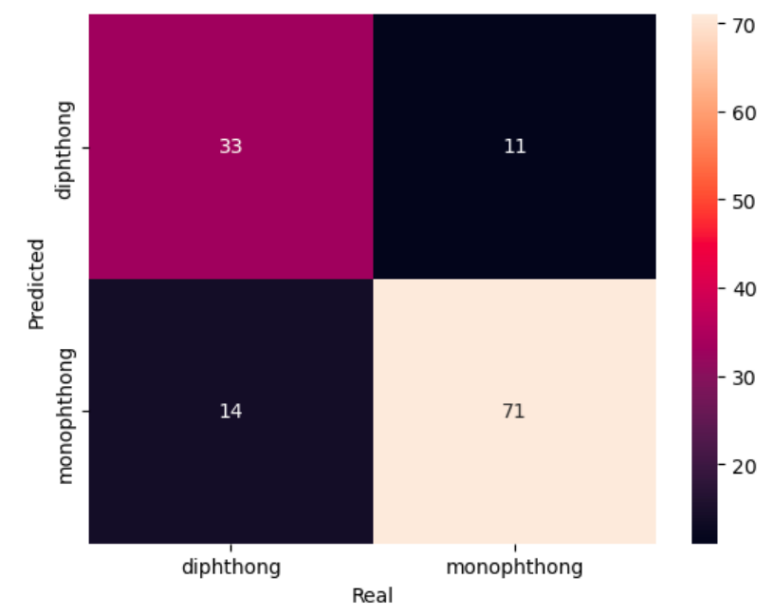


(c) cepstral coefficients

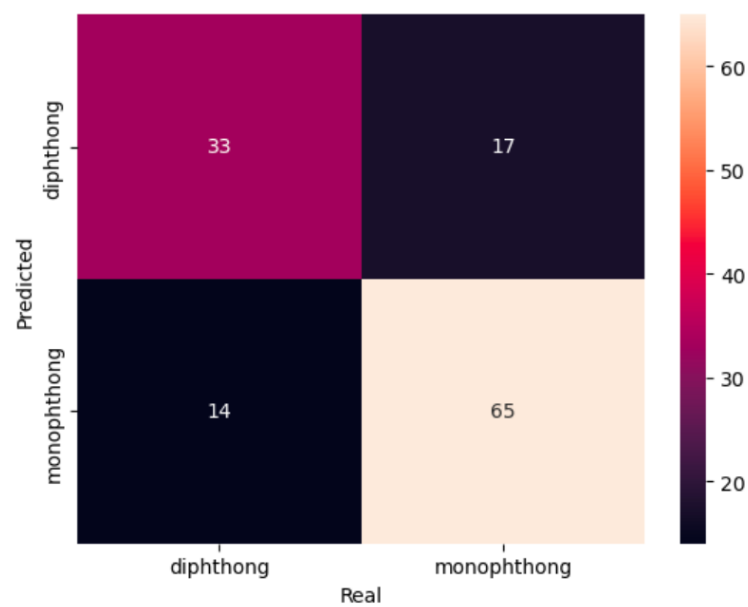
Итоговая таблица для КИЛЬДИНСКОГО

данные	сбалансированная точность
changepoint	0.54
trajectory length	0.57
cepstral coefficients	0.65

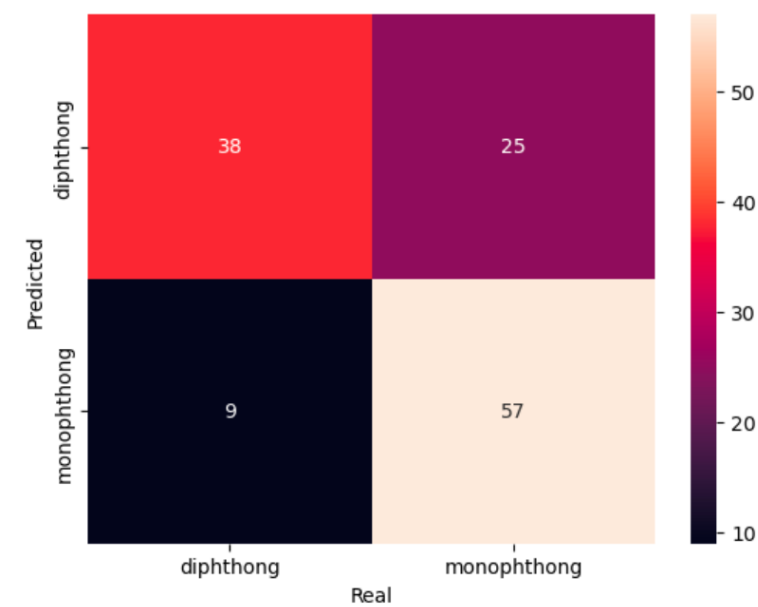
Сколто-саамский



(a) changepoint



(b) trajectory length



(c) cepstral coefficients

Итоговая таблица для сколто-с.

данные	сбалансированная точность
changepoint	0.78
trajectory length	0.75
cepstral coefficients	0.75

Выводы

- В кильдиносаамском нет дифтонгов /ie/, /ea/.
- В нотозерском диалекте сколтосаамского есть дифтонги /ie/, /eæ/

НАШ САЙТ!

<https://saamilang.github.io/online/>



Спасибо за внимание!

“Дифтонги” в кильдинском саамском

Антон Бузанов, Илья Егоров, Василий Молчанов
2023