

Results of the PoEval 2020 Shared Task 1

Post-editing and rescoring of automatic speech recognition results

Danijel Korżinek

<danijelk@pja.edu.pl>



18. czerwca 2020

Cel zadania

Celem zadania jest stworzenie rozwiązania konwertującego sekwencję wyrazów wygenerowaną za pomocą pewnego systemu rozpoznawania mowy (ASR) w inną sekwencję która lepiej odzwierciedla wymówioną wypowiedź.



Przepływ informacji

1. sygnał audio
2. rozpoznawanie mowy
 - ▶ transkrypcja ortograficzna
 - ▶ lista N-best
 - ▶ krata
3. **post-edycja**
4. finalna transkrypcja

Dane wejściowe systemu post-edycji

- ▶ 1-best
 - ▶ każde nagranie zawiera jedną najlepszą transkrypcję wypowiedzi
- ▶ n-best
 - ▶ każde nagranie zawiera do 100 najlepszych hipotez transkrypcji
- ▶ krata
 - ▶ każde nagranie zawiera listę krawędzi tworzących graf kraty
 - ▶ każda linia grafu zawiera następujące pola: węzeł początkowy, węzeł końcowy, wyraz, waga modelu języka, waga modelu akustycznego, lista stanów fonetycznych

Zbiory danych

- ▶ Treningowe
 - ▶ Korpus Clarin-PL studio ¹
 - ▶ Korpus nagrań Sejmu RP ¹
 - ▶ Korpus stenogramów sejmowych PSC ^{1 2}
 - ▶ Dane z poprzedniego podobnego konkursu ^{3 4}
- ▶ Ewaluacyjne
 - ▶ Korpus posiedzeń Parlamentu Europejskiego projektu PINC ⁵

¹ Powstał w ramach projektu Clarin-PL.

² <http://clip.ipipan.waw.pl/PSC>

³ Dzięki Markowi Kubisowi.

⁴ <https://gonito.net/gitlist/asr-corrections.git/>

⁵ <https://pincproject2020.wordpress.com/>

Ewaluacja

- ▶ Porównanie transkrypcji poprawionej przez system do transkrypcji referencyjnej sprawdzonej ręcznie

$$WER = \frac{N_{del} + N_{sub} + N_{ins}}{N_{ref}}$$

- ▶ Na podstawie pakietu NIST Sclite

Zgłoszenia

- ▶ Zgłoszono 20 systemów
- ▶ Uczestniczyło 8 różnych jednostek naukowych i badawczych
- ▶ Większość rozwiązań to nowoczesne sieci rekurencyjne i modele typu transformer

Wszystkie wyniki

Submitter	Submission	Affiliation	WER %	Changes %
Krzysztof Wróbel	KRS + spaces	UJ, AGH	25,9	3,6
Krzysztof Wróbel	KRS	UJ, AGH	26,9	1,6
Dariusz Kleczek	Polbert	https://skok.ai/	26,9	2,1
Tomasz Ziętkiewicz	BiLSTM-CRF edit-operations tagger	AMU	24,7	6,2
Kornel Jankowski	base-4g-rr	Samsung	27,7	2,0
Adam Kaczmarek	t-REx_k10	UWr	24,9	14,2
Adam Kaczmarek et al.	t-REx_k5	UWr	25,0	14,2
Adam Kaczmarek et al.	t-REx_fbs	UWr	24,0	17,2
Krzysztof Wołk	PJA_CLARIN_1k	PJAiT	33,5	9,1
Krzysztof Wołk	PJA_CLARIN_10k	PJAiT	32,0	9,6
Krzysztof Wołk	PJA_CLARIN_20k	PJAiT	31,8	9,9
Krzysztof Wołk	PJA_CLARIN_40k	PJAiT	31,8	10,3
Krzysztof Wołk	PJA_CLARIN_50k	PJAiT	31,8	10,2
Krzysztof Wołk	CLARIN_SEJM_40k	PJAiT	33,7	19,1
Krzysztof Wołk	CLARIN_SEJM_50k	PJAiT	32,5	17,7
Jim O'Regan	MLM+bert_base_polish	n/a	26,8 ⁶	2,1
Tomasz Syposz et al.	tR-Ex_xk	UWr	25,7	18,1
Tomasz Syposz et al.	tR-Ex_fbs	UWr	24,0	17,2
Tomasz Syposz et al.	tR-Ex_fx	UWr	25,0	23,3
Tomasz Syposz et al.	tR-Ex_kv2	UWr	25,5	17,1

⁶Submission missing many files.

Wyrbane najlepsze wyniki

Submitter	Submission	WER %
Adam Kaczmarek et al. (UWr)	t-REx_fbs	24,0
Tomasz Ziętkiewicz (AMU)	BiLSTM-CRF	24,7
Krzysztof Wróbel (UJ,AGH)	KRS + spaces	25,9
Dariusz Kłeczek (skok.ai)	Polbert	26,9
Kornel Jankowski(Samsung)	base-4g-rr	27,7
Krzysztof Wołk (PJATK)	CLARIN_SEJM_50k	32,5



Kontakt

danijel@pja.edu.pl