

Ćwiczenia z tokenizacji

Monday, 27 September 2021 10:20 (30)

Notatnik: using-tokenizer.ipynb

Modele: Roberta (polska i angielska), BERT, itp.

- podstawy działania Google Collab, ładowanie bibliotek, ściąganie modeli
- ładowanie modeli tokenizacji
- rozumienie outputu tokenizatorów
- porównanie działania różnych modeli

Presenter(s) : PRZEWLOCKI, Pawel (NCBJ)