

Zastosowanie sieci Gpt-2 w wybranym problemie NLP

autor: Michał Dubiel

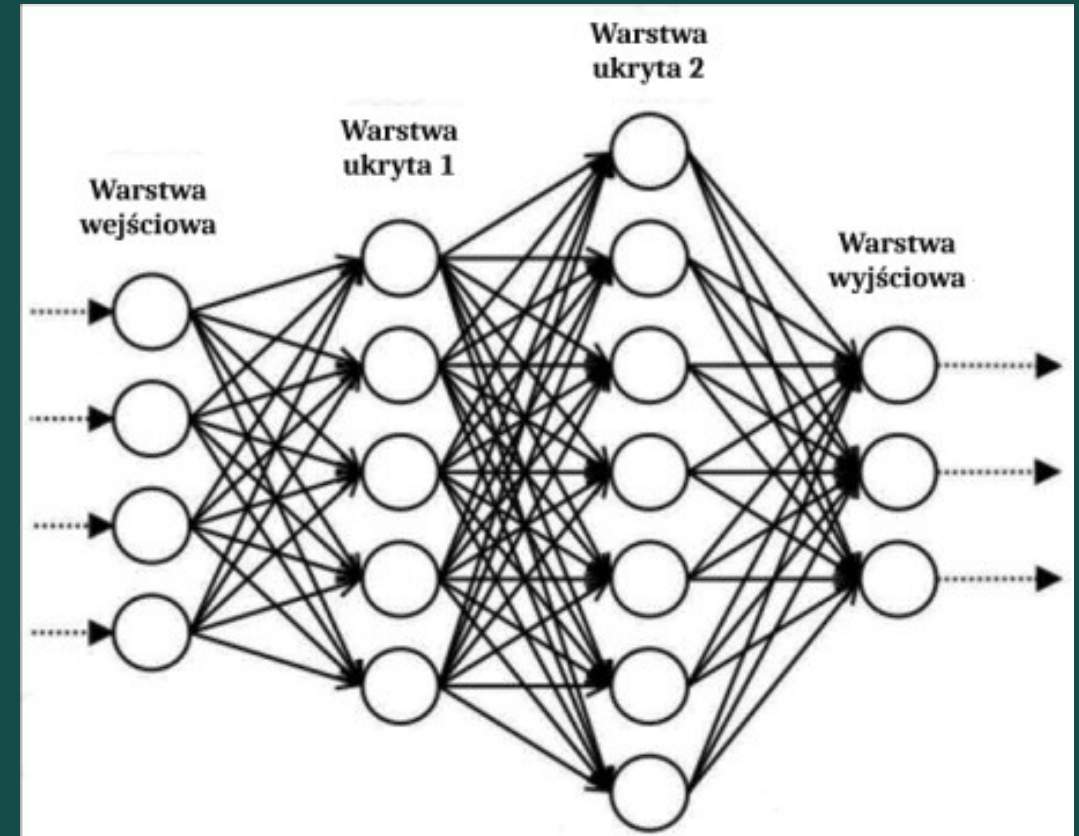
promotor: mgr. inż. Kazimierz Kiełkiewicz

Przetwarzanie języka Naturalnego

- Definicja języka
- Problemy z językiem naturalnym
- Zadania dziedziny
- Rozwiązania

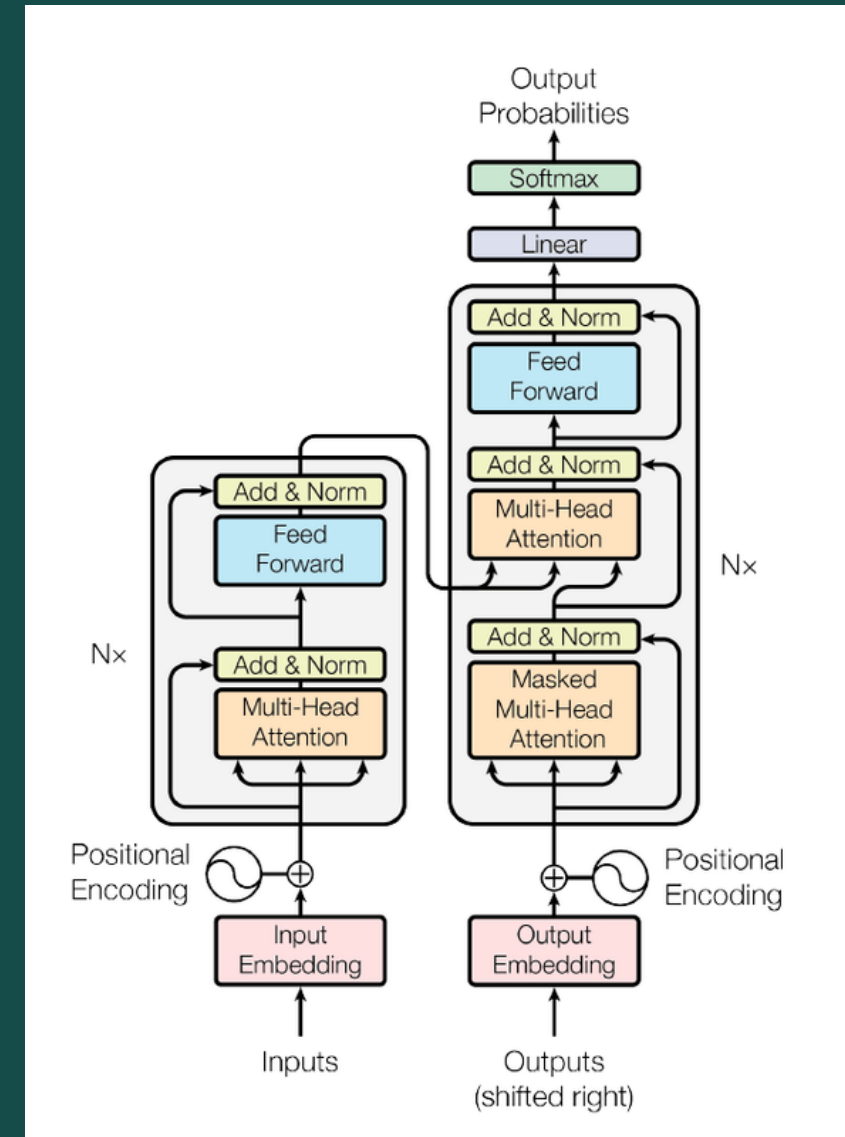
Głębokie uczenie oraz sieci neuronowe

- Pojedynczy neuron
- Architektura sieci neuronowej
- Wsteczna propagacja



Transformer

- Enkoder
 - Word2vec
 - Multi-head attention
- Dekoder
 - Masked Multi-head attention
- Wyjście transformera



Bert vs Gpt-2

BERT

- Bazowany na bloku enkodera
- Trenowany w zadaniach MLM orz NSP
- Służy do pracy nad tekstem

GPT-2

- Bazowany na bloku enkodera
- Trenowany w zadaniach CLM
- Służy do generacji wielkich tekstów

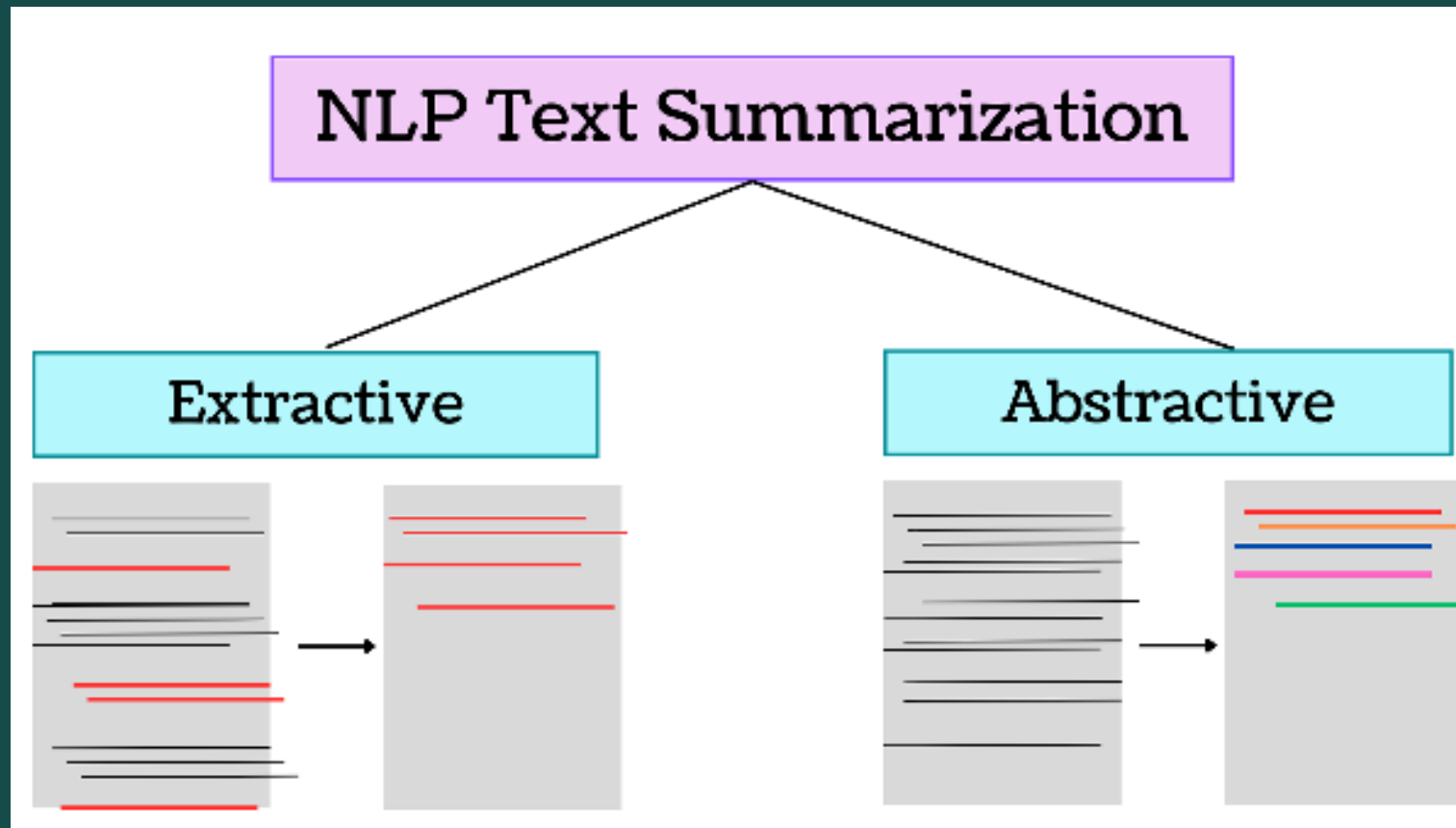
Zastosowanie sieci GPT-2 w sumaryzacji tekstu

- Sumaryzacja tekstu
- Zbiór danych i wstępne ich przetworzenie
- Wyniki modelu
- Aplikacja GUI

Wykorzystane technologie

- Python
 - TensorFlow i Pytorch
- Hugging face
- Jupyter

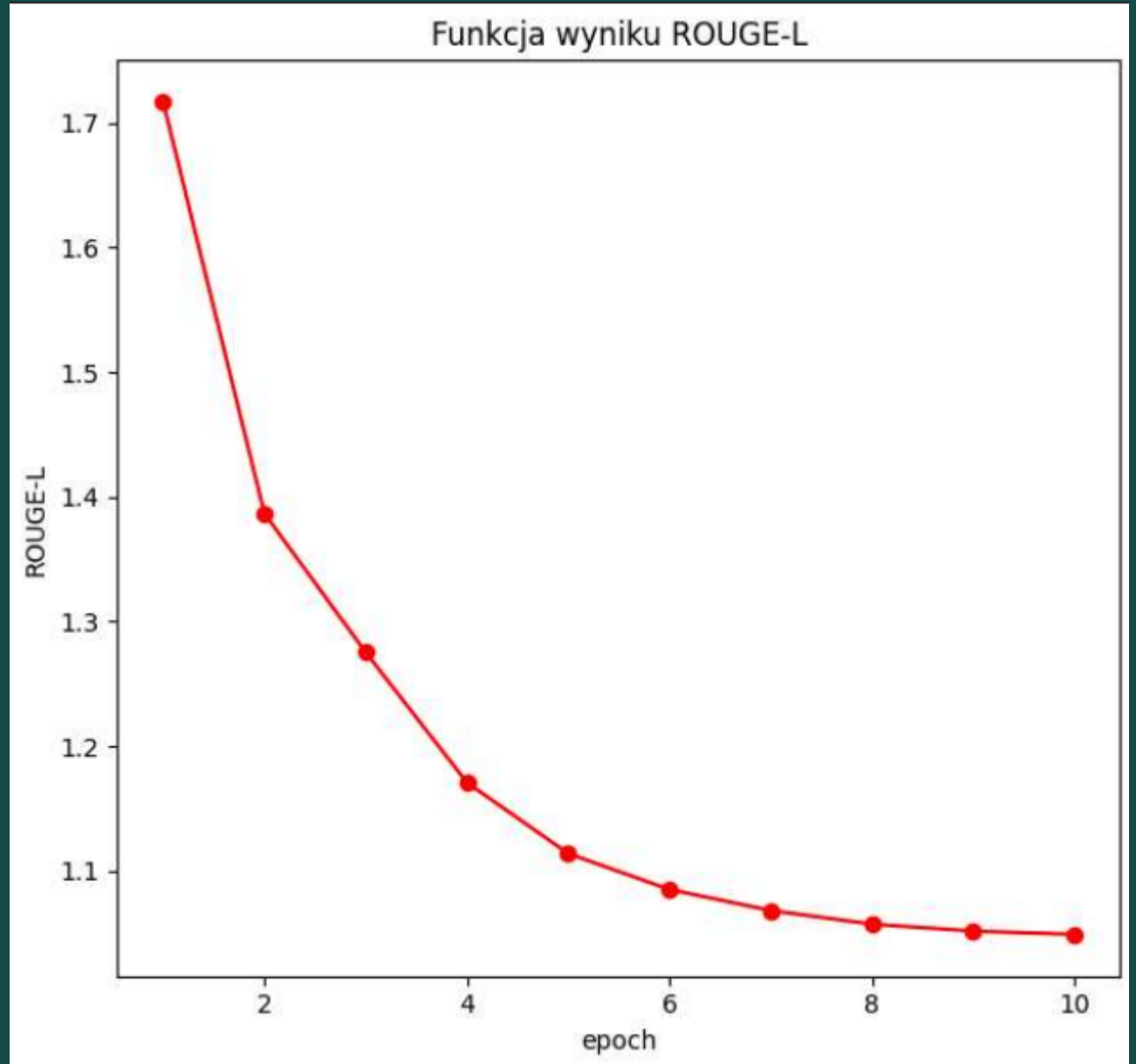
Summaryzacja tekstu



Zbiór treningowy xSum

- Zawartość
- Podział danych
- Wstępne przetworzenie danych

Wyniki treningu



Zaistniałe problemy

- Wybrany model
- Wielkość treningu
- Metoda treningu
- Metoda Ewaluacji
- Wyniki treningu i modelu

Aplikacja GUI

- Losowy tekst
- Wyniki rouge
- Konfiguracja modelu

The screenshot shows a window titled "GPT-2 Summarization". On the left is a sidebar with controls: "Prześlij tekst" and "Losowy tekst" buttons, "Max generation" (50), "Min generation" (10), "Temperature" (1.0), "top_k" (50), "top_p" (1.0), "rep_penalt" (2.5), and radio buttons for "Trained model" (selected) and "Basic model". The main area contains a text input field with a news snippet about Mandla Hlatshwayo. Below this, two summary boxes are shown: "Sumaryzacja wygenerowana:" (TLDR) and "Sumaryzacja z etykiety:" (Ground Truth). At the bottom right, a table displays ROUGE scores.

ROUGE 1	ROUGE 2	ROUGE-L	ROUGESUM
0.344	0.226	0.281	0.344

Dziękuję za uwagę