

Text summarizer w Pythonie

Dominika Borowska, Aleksandra Brymerska

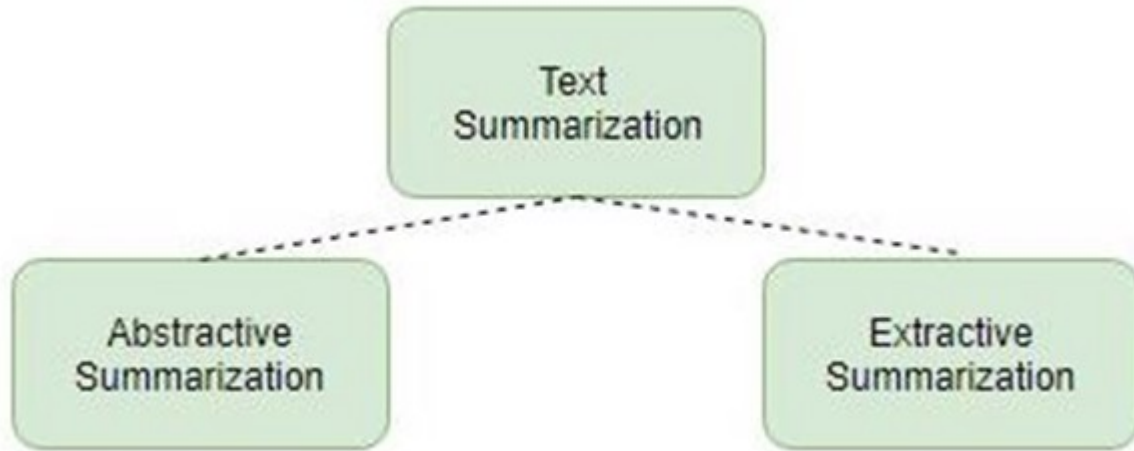
Cel projektu

Celem projektu było stworzenie narzędzia do streszczania tekstu z artykułów dostępnych w Internecie.

Text summarizer

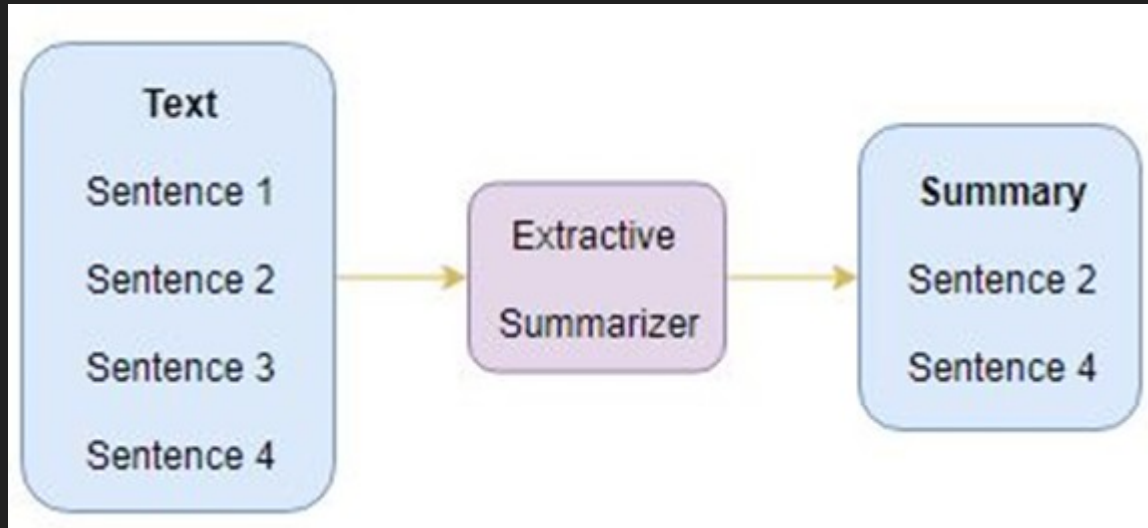
Text summarization (streszczenie tekstu) w NLP to proces tworzenia streszczeń z dużych ilości danych przy zachowaniu istotnych elementów informacyjnych i wartości treści. Język streszczenia powinien być zwięzły i prosty, tak aby przekazywał czytelnikowi znaczenie.

Podjęcia do streszczania tekstu



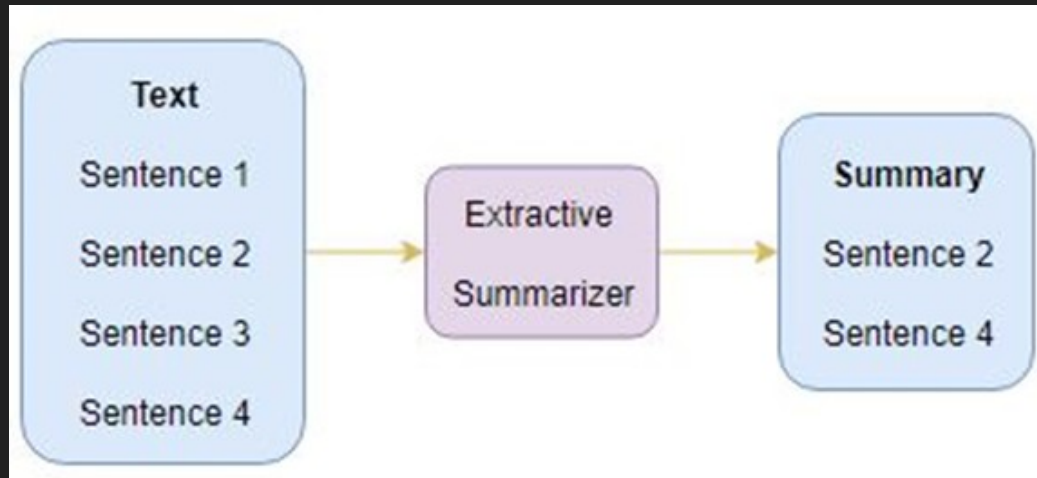
Podsumowanie ekstrakcyjne

Podsumowywanie ekstrakcyjne polega na wyodrębnianiu podstawowych słów z oryginalnego dokumentu i łączeniu ich w celu stworzenia podsumowania



Podsumowanie abstrakcyjne

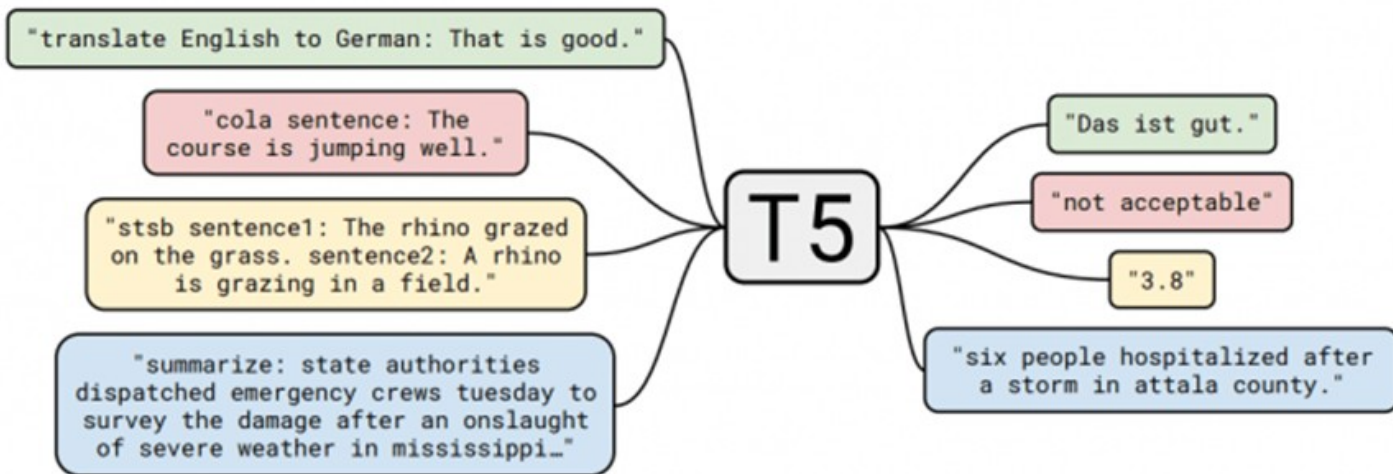
Podsumowanie abstrakcyjne koncentruje się na najważniejszych informacjach w tekście oryginalnym i tworzy nowy zestaw zdań do podsumowania.



Transformer Simple T5

T5 (Text-To-Text Transfer Transformer) to model transformatora (transformera), który jest szkolony w sposób kompleksowy z tekstem jako wejściem i zmodyfikowanym tekstem jako wyjściem, w przeciwieństwie do modeli w stylu BERT, które mogą wyprowadzać tylko etykietę klasy lub rozpiętość wejścia.

EXPLORING THE LIMITS OF TRANSFER LEARNING



Użyte narzędzia i biblioteki

W projekcie wykorzystano język Python oraz Jupyter Notebook.

W projekcie skorzystano z bibliotek:

- *BeautifulSoup*,
- *Requests*,
- *pandas* i *numpy*,
- *SpaCy*
- *SimpleT5*

Web scraping

W projekcie skorzystano z artykułów umieszczonych na stronie <https://www.dailymail.co.uk>

The logo for MailOnline, featuring the word "Mail" in a bold, black, serif font and "Online" in a blue, sans-serif font, all contained within a white rectangular box.

MailOnline

Streszczenie tekstu z wykorzystaniem transformatora i uczenia n

Ukraine will battle to a bloody victory but the Russian invasion will end with diplomacy, Ukrainian's Volodymyr Zelensky said in an hour-long TV interview to mark his third anniversary as President. "We did not start this war. But we have to finish it," Zelensky said in an interview that was recorded Friday and broadcast Saturday. "Victory will be bloody in battle. But the end will be in diplomacy. We want everything back. Russia does not want to give anything away," Zelensky said. The broadcast also included an interview with Zelensky's wife Olena Zelenska, who said the war had not changed her husband. "I can't say that he has changed. As he was a reliable husband and man, so he is," she said. But she lamented not being able to see him for months. "Our family, like all Ukrainian families, is now torn. We didn't see each other for two and a half months, we only talked on the phone. Thank you for this opportunity [this interview], because we are spending time together now -- dating on TV," she said.

Nagłówek ze strony CNN: After bloody battles, Ukraine war will end with diplomacy, Zelensky vows in anniversary interview

Podsumowanie wykonane przez model: Ukraine will battle to a bloody victory but the Russian invasion will end with diplomacy, Ukrainian President Volodymyr Zelensky says in an hour-long TV interview

Porównując to z podsumowaniem stworzonym przez poprzedni model: Ukraine will battle to a bloody victory but the Russian invasion will end with diplomacy, Ukrainian's Volodymyr Zelensky said in an hour-long TV interview to mark his third anniversary as President.

Dziękujemy za uwagę.