Program służący do analizy i generowania podsumowań tekstu i ulepszania go za pomocą sztucznej inteligencji

**Wykonali:**

Bartłomiej Garus, Adrian Guściora, Dawid Jopek

**Wstęp**

W ramach tego projektu napisaliśmy skrypt w języku Python, który potrafi dokonać podsumowania tekstu oraz wyświetlić najczęściej występujące słowa.

**Abstrakt:**

Użyte metody podsumowań tekstu:

* "extractive summarization" - polega to na wybieraniu najważniejszych zdań z tekstu i łączeniu ich w krótkie podsumowanie.
* "abstractive summarization", która polega na tworzeniu zupełnie nowego zdania jako podsumowania, ale jest to znacznie bardziej skomplikowane i wymaga bardziej zaawansowanych technologii przetwarzania języka naturalnego.

W tym projekcie wykorzystaliśmy biblioteki NLTK, spaCy i OpenAI do automatycznego tworzenia tekstów.

Projekt składa się z dwóch etapów: podsumowania tekstu i ulepszania tekstu za pomocą sztucznej inteligencji. W pierwszym etapie wykorzystano biblioteki NLTK i spaCy do wydobycia najważniejszych słów z tekstu i podsumowania go w kilku zdaniach.

W drugim etapie wykorzystano OpenAI do ulepszania podsumowania tekstu, tak aby był on pozytywny emocjonalnie.

**Rozwinięcie**

Nasz program składa się z kilku części. Pierwszym etapem jest przetwarzanie tekstu. Wykorzystujemy biblioteki NLTK i spaCy do tokenizacji tekstu na pojedyncze słowa.

Wybieramy tylko te słowa, które są istotne w kontekście podsumowania. Wykorzystujemy również lematyzację, aby uprościć analizę tekstu.

Następnie nasz program wykorzystuje algorytm oceny ważności słów, aby określić, które słowa są najważniejsze dla podsumowania.

Algorytm ten bierze pod uwagę częstość występowania słów w tekście oraz ich znaczenie w kontekście. Dzięki temu możemy określić, które zdania w tekście są najważniejsze.

Kolejnym etapem jest generowanie podsumowania. Nasz program wykorzystuje zbiór ważnych zdań, aby utworzyć podsumowanie tekstu. Wykorzystujemy API OpenAI, aby ulepszyć podsumowanie, korzystając z sztucznej inteligencji.

**Podsumowanie**

Nasz program okazał się skuteczny w generowaniu podsumowań tekstu. Testowaliśmy go na różnych tekstach i wyniki były zadowalające.

Ulepszanie tekstu za pomocą sztucznej inteligencji również działało według naszych oczekiwań i poprawiło jakość podsumowań.

Było to bardzo ciekawe zadanie choć momentami bywało męczące jak coś nie działało.

Dalszą metodą rozwoju byłoby dodanie innych języków. Można by było dodać testy aby zoptymalizować działanie programu aby robił to szybciej i dawał lepsze rezultaty.

Skrypt umożliwia łatwe i szybkie podsumowanie tekstu oraz wyznaczenie najczęściej występujących słów. Dzięki wykorzystaniu API OpenAI, użytkownik ma możliwość uzyskania dodatkowych informacji na temat wprowadzonego tekstu.

**Literatura:**

<https://www.nltk.org/> dostęp 05.2023r.

<https://docs.python.org/3/library/heapq.html> dostęp 05.2023r.

<https://platform.openai.com/docs/api-reference?lang=python> dostęp 05.2023r.

<https://pypi.org/project/spacy/> dostęp 05.2023r.

**Zrzut ekranu z użycia programu**



