# Wstęp

W obecnych czasach, gdzie każdy ma dostęp do ogromnej ilości artykułów, dokumentów itd. problemem staje się czytanie ich ze zrozumieniem oraz wyszczególnianie najistotniejszych informacji. W związku z tym powstało narzędzie do tworzenia podsumowań tekstu, które rozwiązuje ów problem.

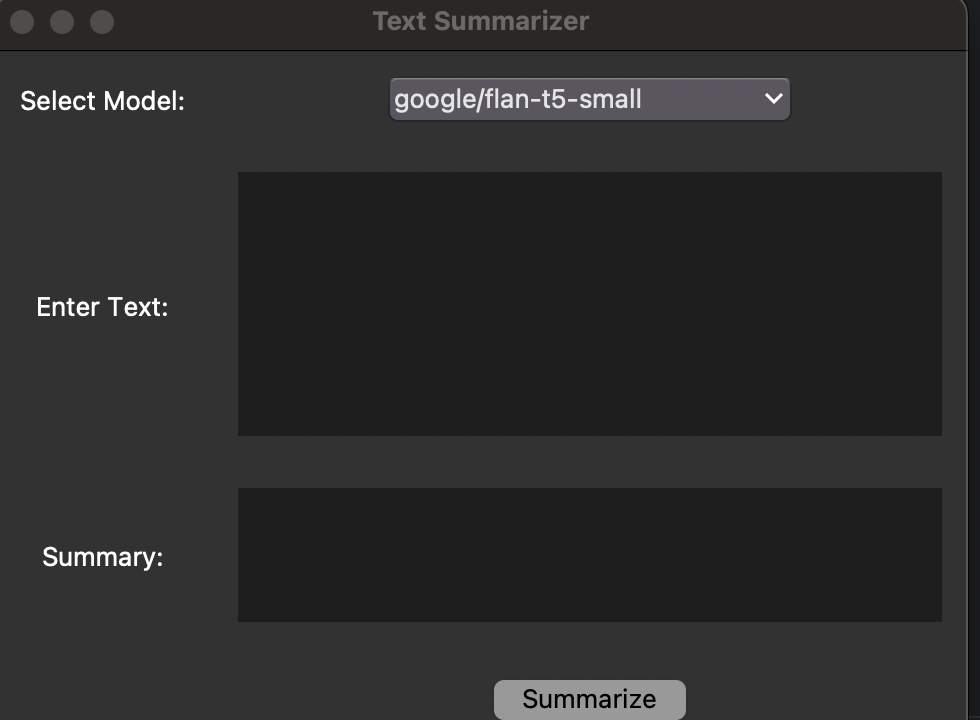
# Opis problemu

Zadaniem narzędzia jest rozwiązanie problemu przeładowania użytkownika ilością informacji zawartych w artykule. Bardzo często brakuje czasu, aby przeczytać dokładne artykuł i wyciągnąć z niego najistotniejsze informacje. Narzędzie do tworzenia podsumowań ma na celu skrócenie czasu potrzebnego do przeczytania oraz zrozumienia tekstu jednocześnie nie tracąc wątku. Dodatkowo została sprawdzona teza: Rodzaj modelu wpływa na dokładność podsumowywanych tekstów.

# Budowa aplikacji

Aplikacja sama w sobie jest bardzo prosta. Składa się ona z jednego okienka, na którym mamy możliwość dostarczenia tekstu do streszczenia. Ciekawą funkcjonalnością aplikacji, jest możliwość wyboru modelu, który posłuży do streszczania tekstu.

W okienku aplikacji mamy dwa pola tekstowe, jeden do wklejenia tekstu do streszczenia, drugi do wyświetlenia rezultatu. Ponad polami tekstowymi znajduje się combo box, do wybrania odpowiedniego modelu. Na ekranie umieszczono również przycisk, który wywołuje akcję generowania tekstu.



# Użyte rozwiązania

Jak wspomniane wyżej, aplikacja korzysta z gotowego modelu dostarczonego przez Google, o nazwie **FLAN-T5.** Jest to model wielkojęzyczny typu text-to-text, który został udostępniony przez Google na zasadach open source pod koniec 2022 r. Został dopracowany pod kątem wielu zadań, ale można go jeszcze bardziej udoskonalić.

W swojej aplikacji zdecydowaliśmy się na użycie trzech wielkości modelu:

* Flan-t5-small
* Flan-t5-base
* Flan-t5-large

Modele te różnią się jedynie ilością parametrów. Rezultaty dla większego modelu będą bardziej dokładne niż rezultaty dla małego, ale nie oznacza to, że nie będą one w żaden sposób trafne.

Jako drugi typ modelu wybraliśmy model typowo do generowania streszczeń. Wykorzystaliśmy również 3 modele tego typu:

* Facebook/bart
* Sshleifer/distilbart
* google/pegasus

# Analiza rozwiązania

## Przygotowanie przypadku testowego

W testach zostało przeprowadzone porównanie jak poszczególne modele radzą sobie z podsumowywaniem tekstu. Dodatkowo została sprawdzona teza: Wielkość modelu wpływa na dokładność podsumowywanych tekstów.

Do przeprowadzenia testów programu został wybrany złożony tekst artykułu z portalu internetowego CNN: "Google launches Gemini, its most-advanced AI model yet, as it races to compete with ChatGPT" [1].

## Przeprowadzenie badania wśród ludzi

Aby móc porównać wyniki z programu została wybrana grupa trzech osób która miała za zadanie przeczytać tekst a następnie opowiedzieć na parę pytań. Pytania były następujące:

* Czego dotyczył ten tekst?
* Jakie są twoim zdaniem najważniejsze ogólne informacje tekstu?
* Jakie są twoim zdaniem najważniejsze szczególne informacje z tekstu?

Ankietowani zostali poproszeni, aby udzielali jak najkrótszych odpowiedzi na pytania. Poniżej znajdują się zestawienia odpowiedzi udzielonych przez ankietowanych.

### Wyniki - ankietowany 1

#### Pytanie 1 - Czego dotyczył ten tekst?

This text is about Google's introduction of a new artificial intelligence model called Gemini. It is an ambitious attempt to compete with OpenAI models, aiming to revolutionize the use of Google's consumer applications and Android smartphones.

#### Pytanie 2 - Jakie są twoim zdaniem najważniejsze ogólne informacje tekstu?

* Google introduces Gemini as its largest and most advanced AI model.
* Gemini is "multimodal," accepting various types of media simultaneously, including text, images, audio, video, and programming code.
* The initiative includes upgrading the AI chatbot Bard and integrating Gemini into widely used Google products like the search engine and Chrome web browser.

#### Pytanie 3 - Jakie są twoim zdaniem najważniejsze szczególne informacje z tekstu.

* Gemini comes in three sizes: Nano (optimized for mobile devices), Pro (default model for various tasks and customers), and Ultra (the most sophisticated model still undergoing safety testing).
* The launch showcases Google's advances in cloud computing, with Gemini trained using powerful cloud-based processors.

### Wyniki - ankietowany 2

#### Pytanie 1 - Czego dotyczył ten tekst?

The article discusses Google's new endeavor, Gemini, which is an artificial intelligence model. It is an attempt to compete with OpenAI and introduce new features to popular Google products, such as the search engine and Chrome browser.

#### Pytanie 2 - Jakie są twoim zdaniem najważniejsze ogólne informacje tekstu?

* Google introduces Gemini as its largest AI model capable of handling various media simultaneously.
* The initiative aims to rebuild Google's position after the recent success of OpenAI's ChatGPT. Gemini models come in different sizes, but their safety is still under testing.

#### Pytanie 3 - Jakie są twoim zdaniem najważniejsze szczególne informacje z tekstu.

* Gemini is multimodal, accepting text, images, audio, video, and programming code.
* The models are available in three sizes, with the most advanced Ultra model currently undergoing safety tests.

### Wyniki - ankietowany 3

#### Pytanie 1 - Czego dotyczył ten tekst?

The article discusses Google's new artificial intelligence model, Gemini. This extensive undertaking aims to compete with OpenAI in the field of generative artificial intelligence and usher in a new era of using consumer applications and Android smartphones.

#### Pytanie 2 - Jakie są twoim zdaniem najważniejsze ogólne informacje tekstu?

* Google presents Gemini as its largest and most capable AI model.
* The model is multimodal, capable of handling various types of media simultaneously. Google plans to apply Gemini in different products, from the search engine to the Chrome browser.

#### Pytanie 3 - Jakie są twoim zdaniem najważniejsze szczególne informacje z tekstu.

* Gemini is being tested on different levels, from mobile devices to advanced Ultra models.
* Google emphasizes progress in cloud computing, using a new generation of processors to speed up AI model training.

### Zbiorcze przedstawienie wyników

|  |  |  |
| --- | --- | --- |
| Czego dotyczy ten tekst? | | |
| Ankietowany 1 | Ankietowany 2 | Ankietowany 3 |
| Google | Google | Google |
| Gemini | AI Competition | Gemini |
| New AI model | Google Product Integration | Integration with google product |
|  | Gemini |  |

|  |  |  |
| --- | --- | --- |
| Jakie są twoim zdaniem najważniejsze ogólne informacje tekstu? | | |
| Ankietowany 1 | Ankietowany 2 | Ankietowany 3 |
| Largest AI model | Response to OpenAI's Success | largest AI model |
| Multimodal | Largest AI Model | Multimodal |
| Integration with google product | different size |  |

|  |  |  |
| --- | --- | --- |
| Jakie są twoim zdaniem najważniejsze szczególne informacje z tekstu? | | |
| Ankietowany 1 | Ankietowany 2 | Ankietowany 3 |
| Three sizes | multimodal | cloud computing |
| Nano | three size | AI model training |
| Pro |  |  |
| Ultra |  |  |
| Cloud computing |  |  |

Następnie zostały wybrane słowa klucze które będą sprawdzane czy się pojawiają w podsumowaniach wygenerowanych przez nasz program.

### Wybór modeli

### Zestawienie wyników podsumowań dla poszczególnych modeli

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Słowa klucze | google/flan-t5-small | google/flan-t5-base | google/flan-t5-large | google/flan-t5-xxl | facebook/bart-large-cnn | sshleifer/distilbart-cnn-12-6 | google/pegasus-xsum |
| Google |  |  |  |  |  |  |  |
| Gemini |  |  |  |  |  |  |  |
| New AI model |  |  |  |  |  |  |  |
| Integration with google product |  |  |  |  |  |  |  |
| Largest AI model |  |  |  |  |  |  |  |
| Multimoda |  |  |  |  |  |  |  |
| AI Competition |  |  |  |  |  |  |  |
|  | 29% | 43% | 43% | 57% | 57% | 86% | 57% |

### Podsumowanie wyników

Wyniki wskazują, że najlepiej tekst został podsumowany przez model sshleifer/distilbart-cnn-12-6, który zwarł w podsumowaniu 6 na 7 słów kluczowych. Wyniki potwierdziły też postawioną przez nas tezę: Rodzaj modelu wpływa na dokładność podsumowywanych tekstów. Modele text2text takie jak google-flan dokonywały gorszych podsumowań tekstów niż modele summarize, których służą konkretnie do podsumowywania tekstu. Warto dodać, że nawet model google/flan-t5-xxl, który jest znacznie większy niż modele summarize nie dokonał tak dokładnego podsumowania. Dlatego wybranie większego modelu, który będzie realizował odmienne zadania nie spowoduje, że będzie je wykonywał w sposób lepszy. W rozmiarze modelu google/flan-t5-xxl można wybrać dwa modele do różnych zadań i da to o wiele lepsze wyniki.

# Podsumowanie

Jak widać, narzędzie radzi sobie w bardzo dobry sposób. Jak zakładano na początku, widać znaczną różnicę w wygenerowanych streszczeniach. Narzędzia typu text-to-text, radzą sobie znacznie gorzej w porównaniu do narzędzi których celem jest pisanie podsumowań. Przy wyborze narzędzia, bardzo ważne jest uwzględnienie do czego będzie ono dokładnie używane.

# Bibliografia

[1] https://edition.cnn.com/2023/12/06/tech/google-launches-gemini-compete-with-chatgpt/index.html

[2] https://github.com/edubey/text-summarizer