**Moduł I. Wypracowanie modelu przetwarzania głosu do postaci tekstowej**

**(Szacowany termin rozpoczęcia prac: 01.01.2024, okres realizacji ok. 10 miesięcy)**

W ramach niniejszego modułu wypracowane zostaną algorytmy pozwalające na konwersję głosu do postaci tekstu. Zbadane zostaną zarówno tradycyjne podejścia (bazujące na analizie i przetwarzaniu jak i metodyki oparte na sztucznej inteligencji (uczeniu maszynowym oraz uczeniu głębokim). Aby osiągnąć zakładany cel zbudowane zostaną bazy danych (niezbędne w procesie trenowania modeli oraz ich ewaluacji) – każda z nich będzie składać się z co najmniej 100 różnych artykułów – w przypadku plików głosowych, będą one miały nie mniej niż 3 sekundy (zapewnione zostaną zarówno głosy męskie jak i kobiece, jak również uwzględniony zostanie różny wiek potencjalnych użytkowników).

W ramach Etapu mają wykonane zostać:

1. Analiza Frameworków Google Cloud Vision i Google Cloud Speech-to-Text dla rozwiązań webowych
2. Implementacja Apple Core ML
3. Implementacja Google Cloud Speech-to-Text

**Moduł II. Wypracowanie modelu przetwarzania obrazów do postaci tekstowej**

**(Szacowany termin rozpoczęcia prac: 01.01.2024, okres realizacji ok. 10 miesięcy)**

W ramach niniejszego modułu wypracowane zostaną algorytmy pozwalające na konwersję obrazów do postaci tekstu. Zbadane zostaną zarówno tradycyjne podejścia (bazujące np. na detekcji krawędzi oraz szkieletyzacji) jak i metodyki oparte na sztucznej inteligencji (uczeniu maszynowym oraz uczeniu głębokim).

Po przeprowadzeniu budowy bazy danych, wnioskodawca planuje wypracowanie własnego modelu uczenia maszynowego, który będzie odpowiedzialny za precyzyjne rozpoznawanie obiektów znajdujących się na zdjęciu.

Wynikiem działania algorytmu będzie zapytanie w formie tekstowej, które w dalszej kolejności będzie mogło być przekazane do systemu bazodanowego

W ramach Etapu wykonane mają zostać:

1. Analiza Frameworku ML Kit dla Android
2. Analiza Frameworków Core ML i Vision dla iOS
3. Analiza Frameworków Google Cloud Vision dla rozwiązań webowych
4. Implementacja Apple Vision
5. Implementacja Android ML Kit
6. Implementacja Google Cloud Vision

**Moduł III. Produktyzacja modeli konwersji obrazu i głosu do postaci tekstu**

**(Szacowany termin rozpoczęcia prac: 01.10.2024, okres realizacji ok. 7 miesięcy)**

W ramach niniejszego etapu opracowane zostanie rozwiązanie wdrażające wypracowane modele konwersji głosu i obrazu do postaci zapytania do systemu bazodanowego. Mianowicie, w aplikacji zostaną wdrożone algorytmy, które pozwolą na wykorzystanie modeli przygotowanych w ramach pierwszego etapu – czyli, użytkownik systemu docelowego będzie wprowadzał obraz lub wypowie odpowiednie słowa, natomiast w aplikacji w miejscu przeznaczonym na zapytanie pojawią się słowa określające obiekt znajdujący się na danej próbce graficznej lub transkrypcja wypowiedzi użytkownika (w obu przypadkach przekonwertowane do postaci odpowiedniego zapytania)

W ramach Etapu wykonane mają zostać:

1. Implementacja procesu wysłania zapytania i otrzymania odpowiedzi z modelu wykonanego w Etapie nr 1

**Moduł IV. Produktyzacja modeli badania elementów bliskoznacznych w postaci wyszukiwarki produktowej**

**(Szacowany termin rozpoczęcia prac: 01.12.2024, okres realizacji ok. 10 miesięcy)**

Zakres planowanych prac B+R

Zadaniem, które będzie realizowane w ramach niniejszego etapu jest przeprowadzenie wdrożenia modelu badania słów bliskoznacznych (generalizacji) oraz modelu detekcji korelacji pomiędzy produktami (na podstawie ich nazwy, właściwości lub grupy produktowej) wypracowanych w ramach etapu numer 3 (zadania badawczego) – w celu sugerowania produktów podobnych lub dokładnie spełniających potrzeby zawarte w wiadomości wysłanej przez użytkownika poprzez komunikator stworzony tymże zadaniu. Wnioskodawca wykorzysta przygotowane zawczasu modele w ramach interfejsu konwersacyjnego do przeprowadzenia procesu realizacji zamówienia, który będzie stanowił główną oś opracowywanej solucji. Oznacza to, że na podstawie wysłanych przez użytkowników wiadomości, system będzie w stanie określić, czy dana wiadomość zawiera zapytanie o produkty oraz zwracać zawarte w nich produkty w formie rekomendacji, nawet jeżeli ich nazwa będzie zniekształcona lub użyte zostanie inne słowo kluczowe. Ponadto, system będzie w stanie sugerować odpowiednie produkty (zamienniki lub pochodzące z tej samej rodziny) w sytuacji niedostępności pożądanego produktu albo gdy system nie będzie w stanie jednoznacznie określić produktu zawartego w wysłanej przez użytkownika wiadomości.

W ramach Modułu wykonane mają zostać:

1. Przygotowanie UX i makiet Lo-Fi dla wersji Web i Mobile
2. Projektowanie UI dla wersji Web i Mobile
3. Implementacja Back-End
4. Implementacja Front-End
5. Implementacja iOS

**Moduł V.** **Produktyzacja modeli badania elementów bliskoznacznych w formie komunikatora do składania zamówień**

**(Szacowany termin rozpoczęcia prac: 01.01.2024, okres realizacji ok. 10 miesięcy)**

Zadaniem, które będzie realizowane w ramach niniejszego etapu jest przeprowadzenie wdrożenia modelu badania słów bliskoznacznych (generalizacji) oraz modelu detekcji korelacji pomiędzy produktami (na podstawie ich nazwy, właściwości lub grupy produktowej) wypracowanych w ramach etapu numer 3 (zadania badawczego) – w celu sugerowania produktów podobnych lub dokładnie spełniających potrzeby zawarte w wiadomości wysłanej przez użytkownika poprzez komunikator stworzony w zadaniu bieżącym Wnioskodawca wykorzysta przygotowane zawczasu modele w ramach interfejsu konwersacyjnego do przeprowadzenia procesu realizacji zamówienia, który będzie stanowił główną oś opracowywanej solucji. Oznacza to, że na podstawie wysłanych przez użytkowników wiadomości, system będzie w stanie określić, czy dana wiadomość zawiera zapytanie o produkty oraz zwracać zawarte w nich produkty w formie rekomendacji, nawet jeżeli ich nazwa będzie zniekształcona lub użyte zostanie inne słowo kluczowe. Ponadto, system będzie w stanie sugerować odpowiednie produkty (zamienniki lub pochodzące z tej samej rodziny) w sytuacji niedostępności pożądanego produktu albo gdy system nie będzie w stanie jednoznacznie określić produktu zawartego w wysłanej przez użytkownika wiadomości.

W ramach Etapu wykonane mają zostać:

1. Wyświetlanie podpowiedzi produktów podczas wprowadzania wiadomości tekstowej do komunikatora
2. Przygotowanie UX i makiet Lo-Fi dla wersji Web i Mobile
3. Projektowanie UI dla wersji Web i Mobile
4. Implementacja Back-End
5. Implementacja Front-End
6. Implementacja iOS
7. Implementacja Android
8. Wyświetlanie podpowiedzi produktów podczas dodawania obrazu do konwersacji
9. Przygotowanie UX i makiet Lo-Fi dla wersji Web i Mobile
10. Projektowanie UI dla wersji Web i Mobile
11. Implementacja Back-End
12. Implementacja Front-End
13. Implementacja iOS
14. Implementacja Android
15. Wyświetlanie podpowiedzi produktów podczas nagrywania wiadomości głosowej
16. Przygotowanie UX i makiet Lo-Fi dla wersji Web i Mobile
17. Projektowanie UI dla wersji Web i Mobile
18. Implementacja Back-End
19. Implementacja Front-End
20. Implementacja iOS
21. Implementacja Android
22. Wyświetlanie podpowiedzi produktów w następstwie wysłania wiadomości w komunikatorze
23. Przygotowanie UX i makiet Lo-Fi dla wersji Web i Mobile
24. Projektowanie UI dla wersji Web i Mobile
25. Implementacja Back-End
26. Implementacja Front-End
27. Implementacja iOS
28. Implementacja Android

V. Wybranie przez kupującego wygenerowanych w systemie podpowiedzi produktów, w celu dodania do zapytania

1. Przygotowanie UX i makiet Lo-Fi dla wersji Web i Mobile
2. Projektowanie UI dla wersji Web i Mobile
3. Implementacja Back-End
4. Implementacja Front-End
5. Implementacja iOS
6. Implementacja Android