

murmuras 

Gustaw, Szymon, Kamil, Natalia

Cele projektu

1. Problem: Dostarczanie danych marketingowych i badawczych pozyskiwanych z ekranów użytkowników smartfonów
2. Przykładowo: informacje statystyczne o spersonalizowanych reklamach i kuponach promocyjnych lub dane o poglądach politycznych szacowane na podstawie konwersacji.
3. Wymagania: Program ma być generyczny i radzić sobie w analizowaniu treści z dowolnych aplikacji.
4. Zachowanie prywatności użytkowników. Wszystkie dane prywatne muszą być przetwarzane tylko lokalnie i nie mogą być wysyłane na serwer.



Rozwiązanie - Jak to zrobimy?

- Obecnie firma posiada rozwiązanie proof of concept, jednak jest ono niewydajne i niezdeployowane.
- Zrobimy prototyp skupiający się na problemie z reklamami
- Istniejące wewnątrz firmy rozwiązanie dostarczające tekst widziany przez użytkownika wraz z bounding boxami
- Preprocessing - klasyfikacja i grupowanie pól tekstowych
- Użycie lightweightowego modelu LLM do zebrania interesujących nas danych (np. wysokość zniżki w kuponie)
- Deployment w postaci aplikacji mobilnej działającej w tle

Technologie i koncepty

- LLM, NLP
- Hugging Face
- Android Studio, Android Services
- TensorFlow Lite
- klasteryzacja, klasyfikacja



Wyzwania

1. Ograniczone zasoby sprzętowe - modele LLM wymagają sporo pamięci i mocy obliczeniowej
2. Opracowanie wiarygodnych benchmarków - ten dostarczony nam przez firmę opiera się na porównaniu z modelem wzorcowym (ChatGPT-4o) i jest przez to ograniczony

Co dalej?

1. Generalizacja na inne zagadnienia, takie jak wspomniane danych o poglądach
2. Zaawansowany preprocessing - istniejący pipeline robi screenshoty i rozpoznaje na nich boxy z tekstem - tracimy tutaj wiele kontekstowych informacji o layoucie - możliwość zaimplementowania własnego parsowania uwzględniającego więcej informacji

Dziękujemy za uwagę