

ZZSN

Dokumentacja wstępna do projektu 9 Chatbot

Autorzy: Maciej Kaczkowski, Piotr Kitłowski

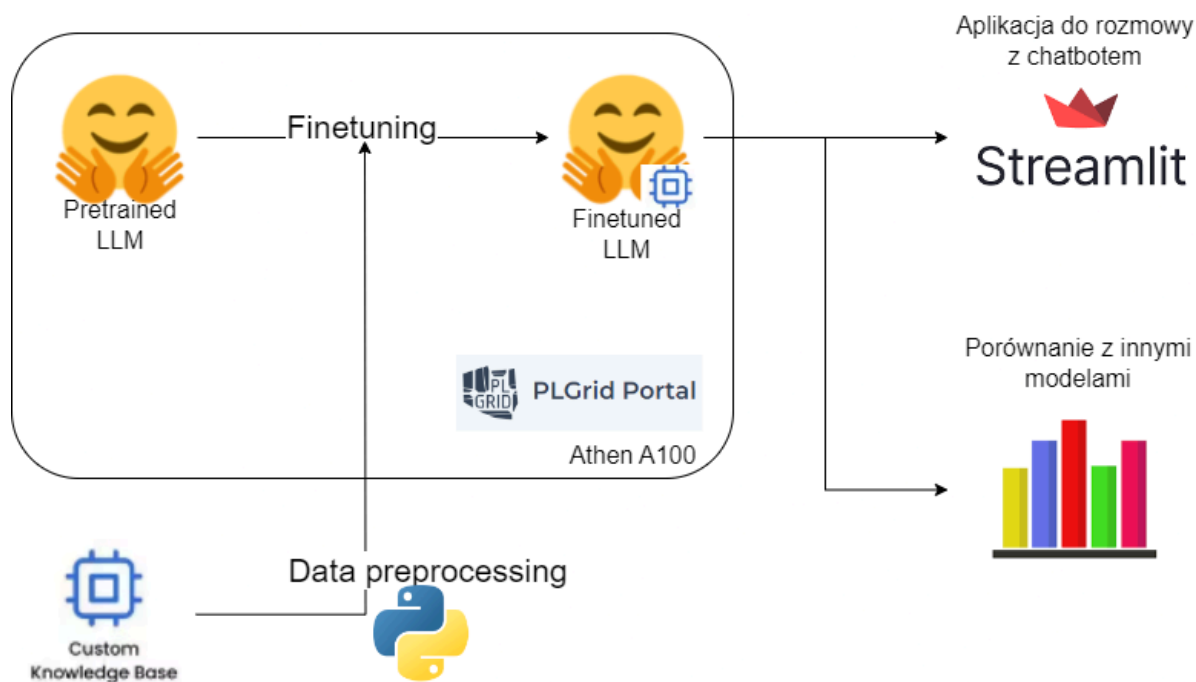
Spis treści

1. Opis zadania
2. Przewidywany sposób rozwiązania problemu
3. Wyjaśnienie poczynionych założeń (typ modelu, wybrane zbiory danych, narzędzia, zakładana funkcjonalność itp.).

Opis zadania:

Przygotuj własnego chatbota. Wytrenuj go do rozmowy na wybrany temat główny (tj. dotrenuj wybrany gotowy model lub wytrenuj od zera własny z wykorzystaniem zbioru danych o wybranej tematyce). Przygotuj aplikację pozwalającą na rozmowę z chatbotem. Porównaj go z dostępnymi modelami językowymi. Omów otrzymane wyniki.

Przewidywany sposób rozwiązania zadania:



Obrazek 1: Wstępny zarys architektury projektu

Korzystając z pretrenowanych modeli LLM / (np. Mistral, llama2) dotrenujemy model na zbiorze danych - taki zbiór zostanie wcześniej odpowiednio spreparowany i przygotowany. Następnie, tak przygotowany model zostanie zaprezentowany użytkownikowi za pomocą narzędzi typu gradio, streamlit lub langfuse. Dodatkowo zostanie przygotowana analiza porównawcza z innymi modelami w kontekście jakości chatbota. Ocena jakości zostanie wykonana przy użyciu typowych dla zagadnienia benchmarków. Proces treningowy zostanie zrealizowany z wykorzystaniem infrastruktury Athena.

Proces treningowy oraz inne potrzebne elementy zostaną napisane w Pythonie z wykorzystaniem w biblioteki Pytorch.

Założenia:

- Korzystamy z gotowego modelu open source (poniżej przykładowy model)

https://huggingface.co/docs/transformers/model_doc/llama2

- Dotrenowywujemy model na publicznie dostępnym zbiorze danych
- Posiadamy odpowiednio mocną infrastrukturę, pozwoli dotrenować model w odpowiednim czasie
- Dane zostaną pobrane z portalu kaggle (poniżej przykładowy dataset spełniający oczekiwania)

<https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce>