

Autor:

Krzysztof Małczak (nr albumu 53378)

Promotor:

Krzysztof Rychlicki-Kicior

Tytuł pracy:

“Wpływ zastosowania agentów na efektywność modeli językowych.”

Alternatywny: “Porównanie efektywności modeli językowych z oraz bez wykorzystania agentów”

Abstrakt:

Celem pracy jest analiza i porównanie efektywności modeli językowych (LLM) w kontekście generowania treści z oraz bez wykorzystania agentów (LLM based agents). Modele językowe są coraz częściej wykorzystywane w różnych gałęziach przemysłu, jednak ich wydajność może być ograniczona ze względu na stale rosnący poziom skomplikowania zadań, do których są wykorzystywane. Oddelegowanie poszczególnych elementów zadania do dedykowanych agentów zarządzanych przez moduł centralny (zazwyczaj LLM) może znacząco podnieść jakość odpowiedzi. W pracy przeanalizowano wyniki generowania treści w obu podejściach, porównano dokładność, spójność, wydajność oraz użyteczność generowanych treści. Na podstawie przeprowadzonych eksperymentów sformułowano wnioski dotyczące skuteczności oraz potencjalnych zastosowań każdej z metod.

Koncepcja części praktycznej (eksperymenty):

1. Metodyka badań:

- Określenie jakie modele będą przedmiotem analizy;

- Określenie jakie scenariusze oraz prompty będą rozpatrywane (np. analiza tekstów prawniczych, wyszukiwanie konkretnych informacji w tekście, generowanie treści kreatywnych);
- Określenie sposobu mierzenia efektywności np. trafność odpowiedzi, czas generowania, satysfakcja użytkownika, złożoność implementacyjna;
- Przygotowanie pół-automatycznych testów porównawczych z i bez wykorzystania agentów w wyżej wymienionych scenariuszach;

2. Analiza wyników

- Analiza wyników uzyskanych w testach (np. w formie tabel, wykresów itp.);
- Analiza, czy i w jakich sytuacjach wykorzystanie agentów poprawia efektywność LLM'ów, a kiedy mogą być zbędni lub nawet pogarszać jakość generowanych treści;

3. Dyskusja i wnioski:

- Omówienie potencjalnych przyczyn różnic w efektywności;
- Wskazanie obszarów, w których warto prowadzić dalsze badania;
- Podsumowanie, jakie wnioski można wyciągnąć z porównania, oraz jakie są implikacje dla zastosowań komercyjnych/badań

Stos technologiczny (DRAFT):

1. Języki programowania:

- JavaScript/TypeScript
- Python

2. Narzędzia:

- Jupyter Notebook (analiza danych, wykresy, tabele, itp.)

3. Frameworki:

- LangChain
- LangGraph
- TBD