Narzędzia wywoływane przez model

Od wiadomości do propozycji JSON, bez wykonania funkcji

Czas	Pracujesz z	Rezultat
około 40 min	Ollama lub checkpoint	action_proposals.jsonl

Porównasz siedem propozycji selektora: pytanie do dokumentów, odczyt źródeł, metadane, brak danych i żądanie zabronione. Zachowasz surową odpowiedź modelu oraz wynik parsowania.

Cel i rezultat rozdziału

Zbudujesz opis narzędzi i sprawdzisz, czy model wybiera działanie zgodne z intencją oraz dostępnym stanem.

Po zakończeniu masz

- siedem surowych odpowiedzi i odczytanych propozycji;
- rozróżnienie błędu JSON od błędnej decyzji;
- ocenę wyboru bez uruchamiania narzędzi.

Zanim zaczniesz

Rozpakuj cały pakiet Student. Zachowaj razem notebooki, dane, kod, checkpointy i wyniki. Ścieżki w karcie są liczone od Student. Przygotowanie stanowiska opisuje instrukcja pracy z modułem 5.

Otwórz notebook

```
notebooki/M12_M5_R4_wybor_narzedzia_start.ipynb
```

Materiał wejściowy	Do czego służy
dane/tool_specs.json	Opisy trzech narzędzi i dopuszczalnych argumentów.
kod/m5_tools.py	build_tool_choice_prompt() i propose_action().

Ollama lub checkpoint. Czas karty dotyczy ćwiczenia, bez instalacji i oglądania prezentacji. Materiał: M12-M5-v5, slajdy 139–190.

Praca krok po kroku

Krok 1 Przygotuj selektor — Wybierz tryb, uruchom Przygotowanie i odczytaj TOOL_SPECS.

- Sprawdź nazwy answer_from_documents, show_last_sources i get_document_metadata.
- Ten notebook nie wywołuje execute_tool_call().

Krok 2 Prześledź prompt — Otwórz build_tool_choice_prompt() i porównaj go z katalogiem narzędzi.

- Znajdź wiadomość, pytanie samodzielne, ograniczoną historię i state_summary.
- Sprawdź reguły dla braku źródeł, document_id i żądania restartu.

Krok 3 Zapisz propozycje — Uruchom Próby i odczytaj siedem rekordów.

- Zachowaj raw_response, proposal, parse_error, prompt i calls.
- Nie naprawiaj surowego JSON-u przed zapisaniem; błąd formatu jest wynikiem próby.

Krok 4 Porównaj decyzje — Zestaw przypadki źródeł dostępnych i brakujących oraz metadanych z ID i bez ID.

- Sprawdź, czy propozycja korzysta z istniejącej nazwy i nie wymyśla identyfikatora.
- Nawet poprawny JSON nie dowodzi poprawności argumentów ani uprawnienia do wykonania.

Krok 5 Uzasadnij ocenę — Uzupełnij obserwacja, dowod i wniosek przy każdym result_id.

- Wskaż dokładne pole propozycji lub fragment surowej odpowiedzi.
- Rozróżnij naruszenie formatu, złą intencję i prawidłową odmowę.

Krok 6 Zapisz i porównaj tryby — Zapisz ocenę i notebook. Opcjonalnie wykonaj nową próbę w Ollama.

- Checkpoint jest odtworzeniem rzeczywistego zapisu, nie kolejną niezależną generacją.
- Zmieniony prompt wymaga nowego wywołania modelu, a nie automatycznego dopasowania checkpointu.

Ważne: ponowne Przygotowanie tworzy nowy katalog. Nowe wykonanie oceń osobno. Po zmianie komentarzy uruchom wyłącznie ocenę i zapis; zachowaj surowe wyniki.

Punkty kontrolne pracy

Po etapie	Masz sprawdzić	Gdy wynik nie pasuje
Po wejściu	Model zna trzy opisy, bez funkcji wykonywalnych.	Przeczytaj dane w zapisanym prompt.
Po modelu	Jest surowy tekst i jawny wynik parsowania.	Nie usuwaj komentarza lub płotka kodu z oryginału.
Po próbie	tool_executed pozostaje false.	Prześledź kod notebooka i funkcję run_r4().

Nazewnictwo i zapis wyników

Wyniki zapisują się w wyniki/R4/<data_id>/. Odczytaj rzeczywisty katalog z komunikatu Wyniki. Główny rezultat: action_proposals.jsonl.

Zapisz notebook przez Ctrl+S. Oceny zapisują się jako ocena_<id>.json; szablon ocena_do_uzupelnienia.json pozostaje bez zmian. Sposób oddania określa prowadzący.

Lista kontroli przed zakończeniem

- □ Zachowałem raw_response bez zmian.
- □ Odróżniam parse_error od błędnego wyboru.
- □ Sprawdziłem oba stany dostępności źródeł.
- □ Nie wykonałem funkcji z propozycji modelu.
- □ Uzasadniłem przypadek braku argumentu i odmowy.
- □ Zapisałem model, tryb, wyniki oraz własną ocenę.

Gdy pojawi się problem

Objaw	Najpierw sprawdź
Model dodał tekst poza JSON	Zapisz błąd parsowania i oceń format; nie naprawiaj wyniku w miejscu.
Propozycja zawiera złą nazwę	Opisz błąd selektora; kontrolę wykonania dodasz w R5.
Checkpoint odrzucił nowy prompt	Przywróć wersję domyślną lub przełącz jawnie na Ollama.
Brak źródeł, a model wybiera narzędzie	Zapisz błąd decyzji; kod R5 musi niezależnie sprawdzić stan.

Zasada końcowa: zachowaj oryginalny wynik i własne uzasadnienie. Błędna odpowiedź lub decyzja modelu jest materiałem do analizy, nie powodem do poprawiania surowego zapisu.