

Znakowanie dowodów w zadaniu odpowiadania na pytania

Instrukcja znakowania

Jako ludzie posługujący się językiem naturalnym, często zadajemy pytania, na które chcemy znać odpowiedzi. Niestety, doświadczenie uczy nas, że nie każda otrzymana odpowiedź jest prawdziwa, np. ze względu na nieprecyzyjne pytanie, ograniczoną wiedzę pytanego lub inne nieporozumienie. Aby upewnić się o wartości uzyskanej informacji, możemy poprosić rozmówcę o rozszerzenie odpowiedzi, podanie uzasadnienia lub źródła wiedzy. Celem przeprowadzonego znakowania jest budowa zbioru danych, z którego automatyczne systemy mogą nauczyć się odpowiadać na pytanie „A skąd wiesz?”.

Czym jest “dowód” w odpowiadaniu na pytania?

Dla danego pytania, odpowiedzi i tekstowej bazy wiedzy, dowodem (ang. *evidence*) nazywamy taki fragment tekstu wyodrębniony z polskiej Wikipedii, z którego hipotetyczny czytelnik mógłby wywnioskować odpowiedź na pytanie.

Na przykład:

Pytanie: *Ile tzw. wojen punickich (kartagińskich) stoczył Rzym z Kartaginą w latach 264–146 p.n.e.?*

Odpowiedź: *Trzy.*

Dowód: *Ostatnia, trzecia wojna punicka w latach 149–146 p.n.e. miała miejsce w afrykańskich terytoriach Kartaginy i zakończyła się jej zdobyciem oraz zburzeniem.*
[\[https://pl.wikipedia.org/wiki/Wojny_punickie\]](https://pl.wikipedia.org/wiki/Wojny_punickie)

Możliwych dowodów w bazie wiedzy może być wiele i mogą się one różnić drobnymi lub istotnymi szczegółami od pytania i od siebie nawzajem. W powyższym przykładzie na podstawie pytania moglibyśmy spodziewać się odpowiedzi „*Rzym stoczył z Kartaginą trzy wojny punickie w latach 264-146 p.n.e.*”. Choć podany dowód brzmi nieco inaczej, wystarcza on do wywnioskowania odpowiedzi przez czytelnika.

W szczególności dopuszczalne są następujące różnice:

- Dowód, w odróżnieniu od odpowiedzi, składa się z co najmniej jednego pełnego zdania¹, ale może także składać się z kilku (następujących po sobie).
- Dowód może nie określać odpowiedzi wprost, a wymagać prostego rozumowania w oparciu o wiedzę ogólną, np. skoro trzecia wojna punicka była ostatnią, to wszystkich musiało być trzy.
- Dowód może nie zawierać wszystkich informacji podanych w pytaniu, np. o rozpoczęciu wojen punickich w 264 r. p.n.e.
- Dowód może za to zawierać informacje nie podane w pytaniu, np. o zburzeniu Kartaginy.

¹ Z wyjątkiem wyliczeń, tabel, itp. Patrz ‘Nietypowe przypadki’ dalej.

- W przypadku bazy wiedzy składającej się z wielu dokumentów wymagane jest określenie identyfikatora (adresu) dokumentu zawierającego dowód (np. strony w Wikipedii).

Możliwe są także sytuacje, gdy dowód w sensie podanym powyżej nie istnieje. Mamy z tym do czynienia, gdy baza tekstowa w ogóle nie zawiera podanej odpowiedzi na pytanie, lub też gdy do określenia odpowiedzi wymagane jest uwzględnienie informacji z wielu dokumentów.

Dwa zadania znakowania dowodów

Przewidziano dwa sposoby wskazywania dowodów do każdego pytania.

W **Zadaniu A** anotator sam znajduje w bazie wiedzy dowody dla każdego pytania. Dla każdego pytania należy wskazać do 5 różnych (najlepiej pochodzących z różnych dokumentów) dowodów – jeśli istnieją w bazie. Możliwe są różne podejścia, ale najłatwiej będzie postępować następująco:

1. Wybrać z pytania i/lub odpowiedzi słowa kluczowe, np. “wojny punickie”,
2. Wpisać słowa te do wyszukiwarki tekstowej, np.
<https://pl.wikipedia.org/w/index.php?search=&title=Specjalna%3ASzukaj&go=Przejd%C5%BA>
3. Przejrzeć treść zwróconych artykułów w poszukiwaniu odpowiedzi,
4. Zapisać znalezione dowody,
5. Jeśli to konieczne, ponowić wyszukiwanie z innymi słowami kluczowymi (synonimy itp...)

Dla potrzeb zadania A wprowadzamy ograniczenie długości wybieranego fragmentu. Jako dowód należy wskazać najkrótszy fragment dokumentu, który spełnia wymagania (co najmniej jedno pełne zdanie). Może on stanowić ciąg zdań, ale nie może być ich więcej niż 5. Jeśli konieczne byłoby wybranie dłuższego fragmentu, należy poszukać bardziej zwięzłego dowodu.

Gdy dostępnych jest wiele dowodów i pojawia się możliwość wyboru, należy wskazywać możliwe zróżnicowane sformułowania. Duplikaty tego samego zdania w różnych artykułach nie powinny być znakowane.

Zadanie B jest znacznie prostsze. Celem znakującego jest określenie, czy wskazany fragment spełnia kryteria dowodu określone powyżej, czy też nie. Odpowiedź jest zatem binarna: prawda (1) lub fałsz (0).

Nietypowe przypadki

Tutaj opisujemy różnego rodzaju nietypowe przypadki, na które natknęliśmy się podczas znakowania. Zapraszamy do dołączenia do naszego discorda jeśli znalazłeś/aś inny nietypowy przypadek albo po prostu masz wątpliwość jak coś oznaczyć: (...)

- Co robić, jeśli według wiedzy znakującego podana odpowiedź na pytanie jest nieprawidłowa?

Pytania zostały ręcznie przejrane pod tym względem i odpowiedzi powinny być prawdziwe. Jeśli pytanie lub odpowiedź są niejednoznaczne, należy przyjąć taką ich interpretację, która wydaje się najprawdźsiwsza. W ogólności jednak celem pracy nie jest ocena prawdziwości odpowiedzi i dowodów należy szukać także dla tych nieprawidłowych.

- Co robić, jeśli nie udaje się znaleźć żadnych dowodów?

Taka sytuacja jest możliwa. Jeżeli po dłuższych poszukiwaniach nie znajduje się żaden dowód, należy pozostawić to pytanie niewypełnione.

- Co robić, jeśli fragment stanowi dowód jedynie, gdy znany jest szerszy kontekst, a traci sens po wyrwaniu z kontekstu?

Można założyć, że czytelnikowi znany jest tytuł artykułu. Natomiast gdy potrzeba innych informacji, należy tak rozszerzyć zakres fragmentu, żeby znalazły się one we wskazanym dowodzie. W przypadku zadania B fragmenty niezrozumiałe bez kontekstu należy odrzucić.

- Czy można zaakceptować dowody, które wymagają prostych przekształceń liczbowych, np. odpowiedzi na pytanie o wiek na podstawie podanego roku?

Tak.

- Czy można wybrać tekst z tabeli, wyliczenia lub podpisu do ilustracji w Wikipedii?

Tak, należy wtedy tylko zwrócić uwagę na to, by informacja w formie tekstowej zachowała czytelność (na ile to możliwe). Jeśli treść w takim fragmencie nie jest wyrażona pełnymi zdaniami, wymóg co najmniej jednego zdania w dowodzie nie obowiązuje.

- Czy można używać treści stron ujednoznaczniających?

Tak, strony ujednoznaczniające często zawierają zwięzłe informacje, które dobrze nadają się na dowody.

- Czy można zmieniać treść skopiowanego tekstu, żeby spełniał narzucone wymagania?

Nie, pomijając kwestie formatowania (usuwanie nadmiarowych spacji, nowych linii etc.).

- Jeśli konieczna jest wiedza ogólna dla uznania dowodu za właściwy, jaki jej poziom można założyć?

Można uznać, że czytelnik posiada wiedzę typową dla osób, które ukończyły kształcenie ogólne w zakresie obowiązkowym w Polsce, czyli ukończona szkoła podstawowa/gimnazjum. Nie oznacza to jednak, że zna wszystkie fakty obecne w programie nauczania. Raczej minimalną wiedzę/umiejętności/termin, np. jeśli pytanie

dotyczy gatunku “roślin”, a w tekście organizm określony jest jako należący do “flory” to zakładamy, że czytelnik rozumie związek między tymi pojęciami. Oczywiście uznajemy, że czytelnik nie zna odpowiedzi na samo pytanie – w przeciwnym wypadku podawanie dowodów byłoby zbędne.

- Czy można w dowodzie uwzględnić dwa oddzielne fragmenty tekstu?

Nie, fragment wybrany jako dowód musi być ciągły. W przypadku zwykłego tekstu oznacza to zdania następujące bezpośrednio po sobie, ale zasada ta obowiązuje również do wybieranych wyliczeń, tabel itp. Jeśli np. do odpowiedzi na pytanie jest konieczny nagłówek listy i jeden z jej elementów, to uwzględniamy również wszystkie elementy pomiędzy nimi.

Zadanie A: wybieranie dowodów z tekstu

Anotacja zadania A odbywa się w arkuszu Google dostępnym tutaj: (...)

W arkuszu znajduje się siedem zakładek, w każdej z nich po 1000 pytań. Każde pytanie jest powtórzone pięciokrotnie, aby umożliwić dodanie do pięciu różnych dowodów.

Opis kolumn:

1. Nr – numer pytania, nie należy go zmieniać,
2. Pytanie – treść pytania,
3. Odpowiedź – lista możliwych odpowiedzi na pytanie,
4. Nr fragmentu – indeks fragmentu dla danego pytania, pełni funkcję wyłącznie pomocniczą,
5. Link – miejsce, w które należy wkleić adres URL dokumentu w którym został znaleziony fragment tekstu z dowodem,
6. Fragment – miejsce, w które należy wkleić treść znalezionego fragmentu z dowodem.

Zadanie B: ocena słuszności dowodów

Anotacja zadania B odbywa się w arkuszu Google dostępnym tutaj: (...)

W arkuszu znajduje się dziesięć zakładek, w każdej z nich po 700 pytań. Każde pytanie jest powtórzone kilkunastokrotnie i jest to nich przypisane wiele różnych dowodów.

Opis kolumn:

7. Nr – numer pytania, nie należy go zmieniać,
8. Pytanie – treść pytania,
9. Odpowiedź – lista możliwych odpowiedzi na pytanie,
10. Tytuł – tytuł artykułu z Wikipedii, w którym znaleziono dowód,

11. Fragment – treść artykułu, która razem z Tytułem służy za dowód,
12. Poprawny - miejsce, w którym należy ocenić, czy dowód jest poprawny (1) czy nie (0).

Ważne! W tym zadaniu nie zwracamy uwagi na format dowodu, tzn. nie musi być minimalnym fragmentem zawierającym odpowiedź, nie musi się on składać z pełnych zdań oraz może być ich więcej niż pięć. Istotne jest jedynie to, czy na podstawie dowodu można wywnioskować odpowiedź.