

Question

Context: Do you have a car?

Query

Ground-Truth: Yes, I have a car.



**Retrieve
utterances**

No, I have a car.
I have a cold.
I don't know.

**Give scores by
human evaluation**

False candidates

No, I have a car.	2
I have a cold.	1
I don't know.	4

**Remove
high-score
utterances**

No, I have a car.	2
I have a cold.	1
I don't know.	4