RoBERTa for question answering. ids and attention_mask are needed for forward( ).

tokenized question		question_word										
tokenized text					text_word	answer_word	answer_word	text_word				
ids	0	encoded_id (question)	2	2	encoded_id	encoded_id	encoded_id	encoded_id	2	1	...	1
attention_mask	1	1	1	1	1	1	1	1	1	0	...	0
offsets (calculated on the tokenized text; spaces all trimmed and have no effect)	[0, 0]	[0, 0]	[0, 0]	[0, 0]	[0, len(tokenized word 1)]	[len(tokenized word 1), len(tokenized word 1) + len(tokenized word 2)]	...	...	[0, 0]	[0, 0]	...	[0, 0]
start_idx						5						
end_idx							6					

BERT for question answering. ids, attention_mask and token_type_ids are needed for forward( ).

tokenized question			question_word									
tokenized text				[CLS]	text_word	answer_word	answer_word	text_word	[SEP]			
input ids original					encoded_id	encoded_id	encoded_id	encoded_id				
ids		101	encoded_id (question)	102	encoded_id	encoded_id	encoded_id	encoded_id	102	0	...	0
attention_mask		1	1	1	1	1	1	1	1	0	...	0
token_type_ids	?	0	0	0	1	1	1	1	1	?	...	?
offsets (calculated on the original text; spaces have effect)		(0, 0)	(0, 0)	(0, 0)	[0, len(tokenized word 1)]	[1+len(tokenized word 1), 1+len(tokenized word 1) + len(tokenized word 2)]	...	...	(0, 0)	(0, 0)	...	(0, 0)
start_idx						4						
end_idx							5					