Q1:

1. BERT use fine-tuning to train a model to fit different task. This means that in the process of fine-tune, we use language data with answer and question to train the gradient descend. We expect the model to output two integers for answer. One means the start of the answer in the context. The other means the end. We need to random initialize two vector and use them to do inner product with output of BERT pretrained model. After softmax the position with highest inner product become answer.
2. GPT uses the provided prompt as a guide to generate a relevant response. The prompt provides direction for the model to follow in order to generate an answer that is consistent with the question. Based on the prompt and the context, GPT generates a candidate answer by selecting the most relevant passage or sentence(s) from the context that contains the information needed to answer the question. At last, GPT then uses additional language processing techniques to refine the candidate answer, such as identifying any synonyms or related terms that may be relevant and ensuring that the answer is grammatically correct and makes sense in the context of the passage. Due to the way of generation, GPT’s answer may not be the same sentence in the provided context, but the combination of different part of context

第一組使用「請從最後一篇的文章中找出最後一個問題的答案」(左)

「從最後一篇文章擷取問題的正確答案」（右）

一張含有 文字 的圖片

自動產生的描述一張含有 文字 的圖片

自動產生的描述

可以見得右邊的正確率較高，有些錯誤的答案也離正確解答更近，我認為是因為「擷取」比「找出」更精確一些，畢竟模型是從文本中截出一段是答案的文字

第二組使用「請從最後一篇的文章中找出最後一個問題的答案」(左)

「在P中，Q的答案是A」（右）

一張含有 文字 的圖片

自動產生的描述一張含有 文字 的圖片

自動產生的描述

在有些答案中右邊是比左邊更加接近正解的，但同時也產出一些匪夷所思的錯誤。錯誤可能來自於prompt的表達比較不直觀，也不完全是中文，導致模型預測錯誤。不過這樣的隨機性也導致有些結果是更接近正解的。詳細的比例可能需要更多測資才能顯現。

第三組使用「請從最後一篇的文章中找出最後一個問題的答案」(左)

「According to the passage, print the answer of the question 」（右）

一張含有 文字 的圖片

自動產生的描述一張含有 文字 的圖片

自動產生的描述

可以看到右邊的表現不太好，可能因爲prompt為英文，讓模型無法確切接收任務的資訊