

1. 以下有關 BERT 的敘述何者正確？
- A. BERT 採用的是 Transformer 架構裡 Encoder 部分
 - B. BERT 採用的是 Transformer 架構裡 Decoder 部分
 - C. BERT 採用的是 Transformer 架構裡 Encoder+Decoder
 - D. BERT 不是採用 Transformer 架構

ANS：BERT 輸出是每一個 token 的 embeddings，採用的是 Encoder 的架構。

2. BERT 裡對於英文文本輸入的每一個 token 是
- A. Word
 - B. Character
 - C. Word piece
 - D. Bigram

ANS：C，一個英文字有可能被分成兩段，每一段是一個 token。

3. BERT 裡在決定每一個 token 的 embedding 時，
- A. 只參考在它前面的 token
 - B. 只參考在它後面的 token
 - C. 只參考自己
 - D. 參考前面和後面的 token

ANS：D，BERT 是採用 Transformer 的 Encoder 架構，其中每一個 token 會參考前後 tokens 的 attention 來決定自己的 embeddings。

4. 以下何種是 BERT 的應用？
- A. NER
 - B. 判斷兩個句子是否同義？
 - C. 判斷一篇文章的類別
 - D. 從文本裡找出問題的答案
 - E. 以上皆是

ANS：E，這些都是 BERT 可能的下游任務。

5. 在使用 BERT 來做文章分類時，一般根據的是
- A. [CLS]的 embedding
 - B. 第一個 token 的 embedding
 - C. [SEP]的 embedding
 - D. [PAD] 的 embedding

ANS：A，[CLS]的 embeddings 是參考到所有的 tokens，故可以根據此 embeddings 加上 linear 和 Softmax 層來做分類。

6. 以下有關 BERT 的敘述何者正確？

- A. BERT 英文的輸入都是 word
- B. BERT 中文的輸入都是詞
- C. BERT 會考慮字詞的次序
- D. BERT 的輸出是一個向量

ANS：C，其中 BERT 英文的輸入是 word pieces，中文輸入一般是中文字，其輸出是每一個 token 一個向量。

7. 從文本建立知識圖譜時會用到那些技術？（多選）

- A. NER
- B. RE
- C. Q&A
- D. Machine translation

ANS：A 和 B，NER 找出知識圖譜的節點，RE 找出關係。

8. 下列何者有關 RE 的敘述是正確的？

- A. 輸入需包括文本
- B. 輸入需包括頭（head）尾（tail）的實體
- C. 輸出為數種關係的一種
- D. 以上皆是

ANS：D，輸入包括文本、頭（head）尾（tail）的實體，輸出則是數種關係的一種。

9. 有關 BERTopic 的敘述何者正確？

- A. 比 LDA 慢
- B. 適合較長的文章
- C. 適合一篇文章有多個主題
- D. 使用 c-TF-IDF 來決定每一主題的 word distribution

ANS：D，BERTopic 可以平行進行，一般而言比 LDA 快，且其作法是假設每一篇文本是一個主題，故適合短文本。

10. 以下何者不是 BERTopic 的必要元素？

- A. BERT 的 embeddings
- B. 需要分群演算法（Clustering）
- C. 需要 c-TF-IDF 來決定每一個字在每一主題的重要性
- D. 全數需要

ANS : A , 任何 embeddings 都可以跟 BERTopic 搭配。