

# 語言處理的革命者：BERT vs GPT

BERT（雙向編碼器表示轉換器）和GPT（生成預訓練轉換器）都是基於Transformer架構並革新了語言處理，但它們在訓練方法、目標和適用應用方面有顯著差異，今天小編就來帶你們深入探討吧。



## 起源與基本原理

- BERT，由Google於2018年推出，旨在理解句子中單詞的上下文和含義，同時考慮到它前後出現的單詞。BERT結合無監督和監督學習在大量文本語料庫上進行訓練。
- GPT，由OpenAI於2018年推出，是一種能夠自行生成文本的生成模型。其目標是創建一種語言模型，該模型可以生成連貫且適當的上下文文本。

## 訓練和目標

- BERT 使用雙向訓練策略。它通過掩碼語言模型（MLM）和下一句預測（NSP）任務來學習上下文相關的詞表示。這種雙向方法使BERT能夠從句子的兩個方向理解單詞的語義和上下文。
- GPT 使用自回歸語言模型進行訓練。它通過預測前面單詞的基礎上生成下一個單詞來生成文本。這種單向方法使GPT在生成性任務（如文本和對話生成）中表現出色，因為它通過之前生成的內容來構建連貫、合乎邏輯的文本。

## 上下文理解和適用性

- BERT 在理解不同文本部分之間的關係至關重要的任務中表現出色，如命名實體識別和問答。其雙向特性提供了豐富的上下文理解。

| BERT vs GPT |                        |                        |
|-------------|------------------------|------------------------|
| 特徵          | BERT                   | GPT                    |
| 語言模型        | 預訓練語言模型                | 對話生成語言模型               |
| 原理          | TRANSFORMER編碼器         | TRANSFORMER解碼器         |
| 輸入          | 單一或雙向文本輸入              | 對話式文本輸入                |
| 訓練方式        | 雙向上下文遮蔽語言模型 (MLM)      | 單向上下文標準的語言模型           |
| 訓練數據        | 大型文本語料庫 (EX WIKIPEDIA) | 對話式文本語料庫 (對話紀錄/社群媒體對話) |

# 語言處理的革命者：BERT vs GPT

## BERT vs GPT

| 特徵    | BERT                    | GPT                  |
|-------|-------------------------|----------------------|
| 主要用途  | 文本理解和分析<br>(情感分析、實體識別等) | 文本生成<br>(對話系統、內容創作等) |
| 交互方式  | 不專注於交互式對話，更<br>多用於理解和分析 | 設計用於交互式對話<br>和回應生成   |
| 上下文理解 | 較有限制<br>針對單一句子理解        | 能夠理解<br>更長對話的上下文形式   |
| 輸出生成  | 能夠生成句子<br>向量表示          | 能夠生成<br>連續對話文本       |
| 應用領域  | 文本分類、問答系統、<br>文本摘要      | 聊天機器人、內容創作           |

- GPT 在生成性任務中表現良好，利用其產生流暢和連貫文本的能力。然而，由於其單向性質，它可能在需要輸入輸出對齊的任務中效果較弱。GPT模型可以針對特定的下游任務進行微調，適應特定應用的要求。

BERT像是你的知識百科全書，而ChatGPT則更像是你的聊天伙伴。雖然他們各有所長，但都以不同的方式豐富著我們的數位生活。

## 情境挑戰 / 語言處理的革命者：BERT的深度理解 VS GPT的創作魅力

假設你是一位數據科學家，手頭有一項任務：

你需要從滿是專業術語的醫學研究報告中，識別出特定的醫學名詞，例如疾病和藥物名稱。你會選擇哪一位AI助手來完成這項工作呢？

A：是對上下文理解如同讀心術的BERT

B：擅長連貫創作，如同故事家的GPT？

在這種情況下，BERT將會是更合適的選擇。

因為BERT的雙向訓練策略使它能够更好地理解和解析上下文中的詞語，從而更準確地識別和提取相關的醫學名詞。BERT的模型可以被訓練來關注句子中特定詞語的上下文信息，提高識別精準度。

相比之下，如果任務是生成一段醫學領域的自然語言文本，如自動撰寫醫學報告的初稿，GPT會是更好的選擇。GPT的生成性質使它能够根據給定的開頭或提示，創造出流暢、連貫且有邏輯的文本。這在生成醫學報告的初稿時尤其有用，因為它可以根據已提供的信息生成接續的文本內容。