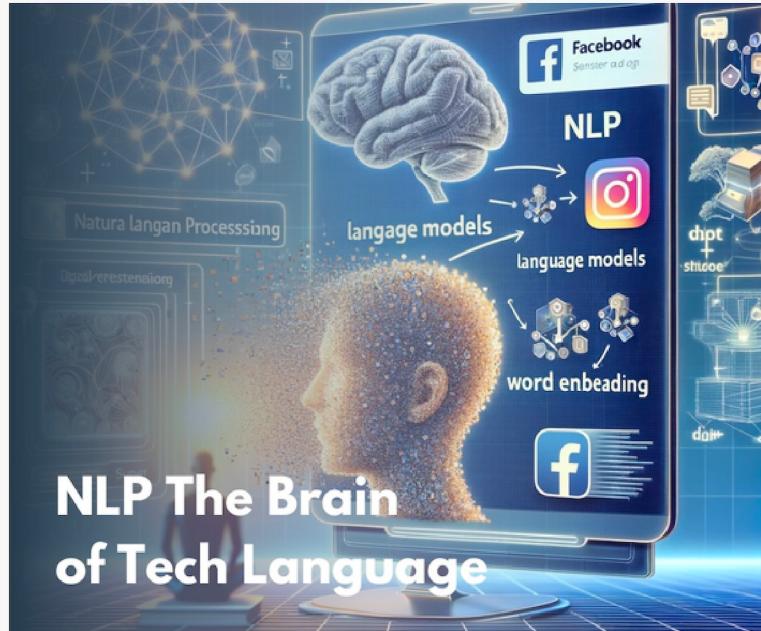


解碼NLP：智能科技背後的語言學大腦

你有沒有想過，當你告訴你的手機「嘿，Siri，幫我找附近的咖啡廳」時，它是如何理解你的話的？或者當你在Google上查詢問題時，它是如何提供那麼精準的答案的？這就是NLP的魔力——賦予電腦處理人類語言的能力。

NLP - 語言的大腦

NLP是機器學習與深度學習的結合，不僅處理單詞，還涵蓋從語義理解到情感分析的一切。通過NLP，電腦不僅可以回答你的Google查詢，甚至能跨越語言障礙，自動進行翻譯！



語言模型 - 預測未來之鑰

為了處理自然語言，我們就需要Language Model的幫助！語言模型就如同先知，以數以萬計的文本為食，透過分析，來預測下一個詞彙可能是什麼。當你詢問Siri或Google Assistant「明天需要帶傘嗎？」，它們分析你的語言並提供相應的天氣預報；當你寫郵件時，許多文字處理軟件會根據你當前的輸入來預測下一詞彙或短語。

語言模型的秘密武器-WORD EMBEDDINGS

語言模型會將文字轉化為一系列的數字串（Word embeddings），想像每個詞匯都有一串獨特的數字代碼，這些代碼集結成了每個詞彙的獨特DNA。這些代碼的相似性反映了詞彙之間的關聯，從而顯示出其獨特的意義、用法，甚至是情感色彩，這是語言模型之所以能夠精準預測和理解語言的核心原因！

Sentence	0.2	0.4	0.32	0.89	0.77	0.12	0.11	0.99
Coffee	0.54	0.8	0.35	0.34	0.56	0.22	0.56	0.43
or								
Tea	0.23	0.39	0.55	0.91	0.6	0.3	0.61	0.8
?	0.7	0.45	0.56	0.43	0.22	0.16	0.33	0.5
Average	0.42	0.51	0.45	0.64	0.54	0.17	0.4	0.68
Embeddings concatenated								
Sentence	0.2	0.4	0.32	0.89	0.77	0.12	0.11	0.99
Coffee	0.54	0.8	0.35	0.34	0.56	0.22	0.56	0.43
or								
Tea	0.23	0.39	0.55	0.91	0.6	0.3	0.61	0.8
?	0.7	0.45	0.56	0.43	0.22	0.16	0.33	0.5
Dimension of word embeddings								

近年來，許多尖端的語言模型陸續被開發出來：

- BERT：由Google開發，能更好地理解語言的上下文含義，改進了搜尋結果的相關性。
- GPT系列：由OpenAI開發，以其強大的生成能力和廣泛的應用而聞名。

所以下次當你用語音助手或搜索引擎時，記得，有一個小小的NLP和語言模型在幕後助你一臂之力。