

# 社群媒體貼文分析意見系統

專題編號：110-CSIE-S003

執行期限：110 年第 1 學期至 111 年第 1 學期

指導教授：王正豪

專題參與人員：108590033 陳柏錫  
108590041 楊勝朋  
108590018 顏維鴻

## 一、摘要

本專題的目的是透過分析 ppt 文章的留言讓使用者能夠知曉民眾對特定主題文章的想法是偏向支持或是反對，並將分析結果透過網頁顯示。分析的套件我們使用的是 **bert-fast**，**Fast-Bert** 能夠應用於多類別分析作業。

## 二、緣由與目的

### (一)緣由:

近年來網路發展迅速，自今，社群媒體已成為資訊傳播的霸主。資訊成長速度從爆發式演變成倍增式的增加，各種多樣的資訊充斥於網路中，如總統大選、烏俄戰爭、比特幣、網紅藝人八卦、音樂發行、政府政策、新專利發布等。假消息眾多，因此，搜尋並掌握適切的資訊，已成為現今科學研究的基石。

目前社群平台雖然有意見調查系統，但效果不彰，本系統是直接針對熱門議題先做分析後，顯示在搜尋列當中，讓大家直接點擊目前熱門議題，並了解對議題的正反面意見。

### (二)目的:

本專題題目為「社群媒體貼文分析意見系統」，主要辨識留言者的意見進行詞性分類後再對其結果辨識後，進行正反面意見的統計

本系統針對判別詞性及語意的部份，是透過 **Hugging Face** 上公開提供的

中文預訓練模組作為 **Fast-Bert** 基底再透過我們自己的資料集做訓練。

## 三、使用的工具及技術

### (一)Fast-Bert:

**Fast-Bert** 是 **utterworks** 在 **github** 發布的套件，**Fast-Bert** 開發的目的是應用於原生 **bert** 無法處理的多分類作業，使用 **HuggingFace** 上提供的中文預訓練模組作為基底，再透過我們的資料集進行訓練。

### (二)React:

**React** 是一種實現使用者介面的 **JavaScript** 函式庫，宣告式的 **view**，可以讓程式容易除錯、元件式開發可以組合成複雜的使用者介面，讓我們可以更容易建構網頁。

### (三)MySQL:

**MySQL** 是一個開放源碼的小型關聯式資料庫管理系統，開發者為瑞典 **MySQL AB** 公司。**MySQL** 因為效能高、成本低、可靠性好，尤其是開放原始碼這強大的功能，所以被廣泛地應用在 Internet 上的中小型網站中，隨著 **MySQL** 不斷成熟，如今也逐漸應用在例如維基百科、Google、Facebook 等大型網站。

### (三)Node.js:

**Node.js** 是能夠在伺服器端運行 **JavaScript** 的開放原始碼、跨平台執行環境。**Node.js** 由 OpenJS Foundation(原為 Node.js Foundation，已與 JS Foundation 合併)持有和維護，亦為 **Linux** 基金會的專案。**Node.js** 採用 **Google** 開發的 **V8** 執行程式碼，使用事件驅動、非阻塞和非同步輸入輸出模型等技術來提高效能，可優化應用程式的傳輸量和規模。

#### 四、系統的架構

(一)前端:傳送爬蟲請求給後端，並從資料庫讀取爬蟲後的資料顯示在網頁上。

(二)後端:爬蟲完的資料，傳給分析程式，並取的分析完的資料傳回資料庫。

(三)資料庫:存取爬蟲後的資料。

(四)分析程式:拿爬蟲後的資料進行標籤分析，把標籤完的資料傳到後端。

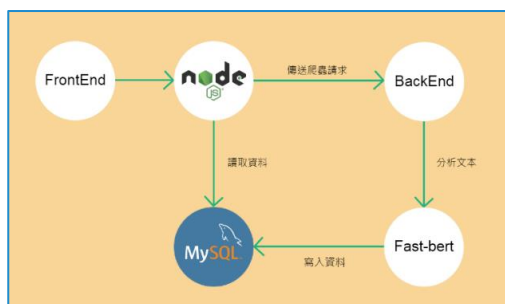


圖 4.1 系統架構圖

#### 八、參考文獻

[1] Bert

<https://zhuanlan.zhihu.com/p/365232061>

[2] Fast-Bert

<https://github.com/utterworks/fast-bert>

[3] HuggingFace Library

<https://github.com/huggingface/transformers>

[4] MySQL

<https://zh.wikipedia.org/wiki/MySQL>

[5]Node.js

<https://zh.wikipedia.org/wiki/Node.js>