

整合大型語言模型與影像辨識技術： 提升個人化專輯網站的智慧互動與產品品質

Group2:

112034562 許嫻軒、112034536 石文德

112034505 高子倫、112034608 張芷瑄



目錄

- 背景介紹與問題描述
- 影像辨識模型
- 大型語言模型
- 網頁設計
- 結論

A 3D-rendered illustration of a music player and headphones. The music player is a dark grey rectangular device with a screen and buttons. It is surrounded by various musical elements: a pair of white headphones, a white speaker, a white teapot, and several musical notes and sound waves. The background is a warm, brownish-orange color with a large, glowing circular light source. The overall style is clean and modern.

背景介紹與問題描述

背景介紹

- 專輯網站已成為許多音樂愛好者分享和購買音樂作品的平台之一
- 用戶在使用網站時可能**需要更多的指導**，而現有的系統可能無法提供即時的支援
- 由於光碟品質的不確定性，用戶在收到專輯時可能會發現**光碟毀損**的情況

問題描述

Who:
專輯網站的用戶、管理者和開發團隊

When:
隨時為顧客提供服務

What:
大型語言模型和影像辨識技術



Where:
應用於購買專輯網站，但也可擴展至其他網路購物平台

Why:
提高個人化專輯網站的用戶體驗和產品品質

How:
增加智慧聊天機器人和光碟毀損檢測功能

目標介紹

更人性化的聊天機器人

- GPT-3.5
- Breeze-7b

識別光碟是否毀損功能

- YOLOv9
- CNN

影像辨識模型

資料前處理

資料來源

我們收集了 253 張實體的光碟，當中包含了有刮痕的光碟，以及正常的光碟。



資料處理與擴增

我們使用了**資料擴增 (Data Augmentation)** 的技術，對原始的圖片進行了複製並**鏡像翻轉**、**轉向**、**增加噪音**和**增加曝光度**，並以這些被處理過的圖片當作一張新圖片，藉此達到**增加模型泛化性**和**增加資料集大小**的目的。





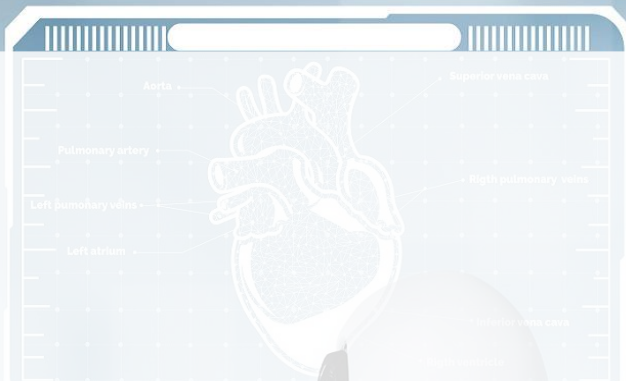
影像辨識模型

YOLOv9

RESEARCH
INNOVATION
TECHNOLOGY
IDEA
SOLUTION
COLLABORATION

HEALTH CARE
TREATMENT
MEDICAL CARE
INVENTION
DIAGNOSE
DISCOVERY

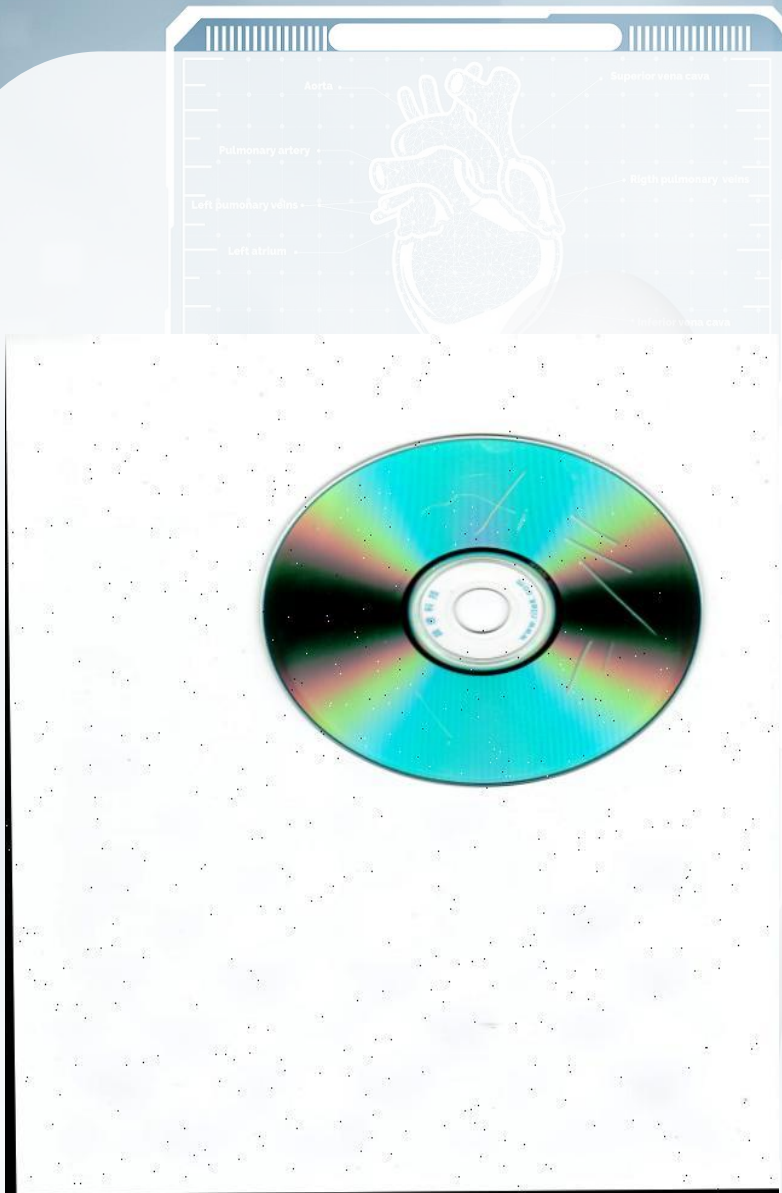
HEALTH
INSURANCE
FINANCIAL
BENEFITS
DISEASE
COVERAGE
ACCIDENT



Yolo v9



- YOLO (You Only Look Once) 是一種基於圖像全局訊息進行預測的目標檢測系統
- 引入了可編程梯度訊息 (PGI) 和廣義高效層聚合網路 (GELAN) 等開創性技術



Yolo v9

```
↳ 0.7674618620305441 0.22055026923052715 0.08346702800763595 0.05673178903121974  
0 0.631978348026726 0.19321158529091403 0.10497894867918127 0.09279200887879457  
0 0.790518648710998 0.32198550420341243 0.09819314668204485 0.11384485226293624  
0 0.7608135147199067 0.387787766205465 0.08295849801081764 0.08016474385167989
```

- 透過 label 軟體取得刮痕邊界框的座標，並輸出成 txt 檔
- 藉由 Pytorch 框架設定模型的相關超參數，並開始訓練

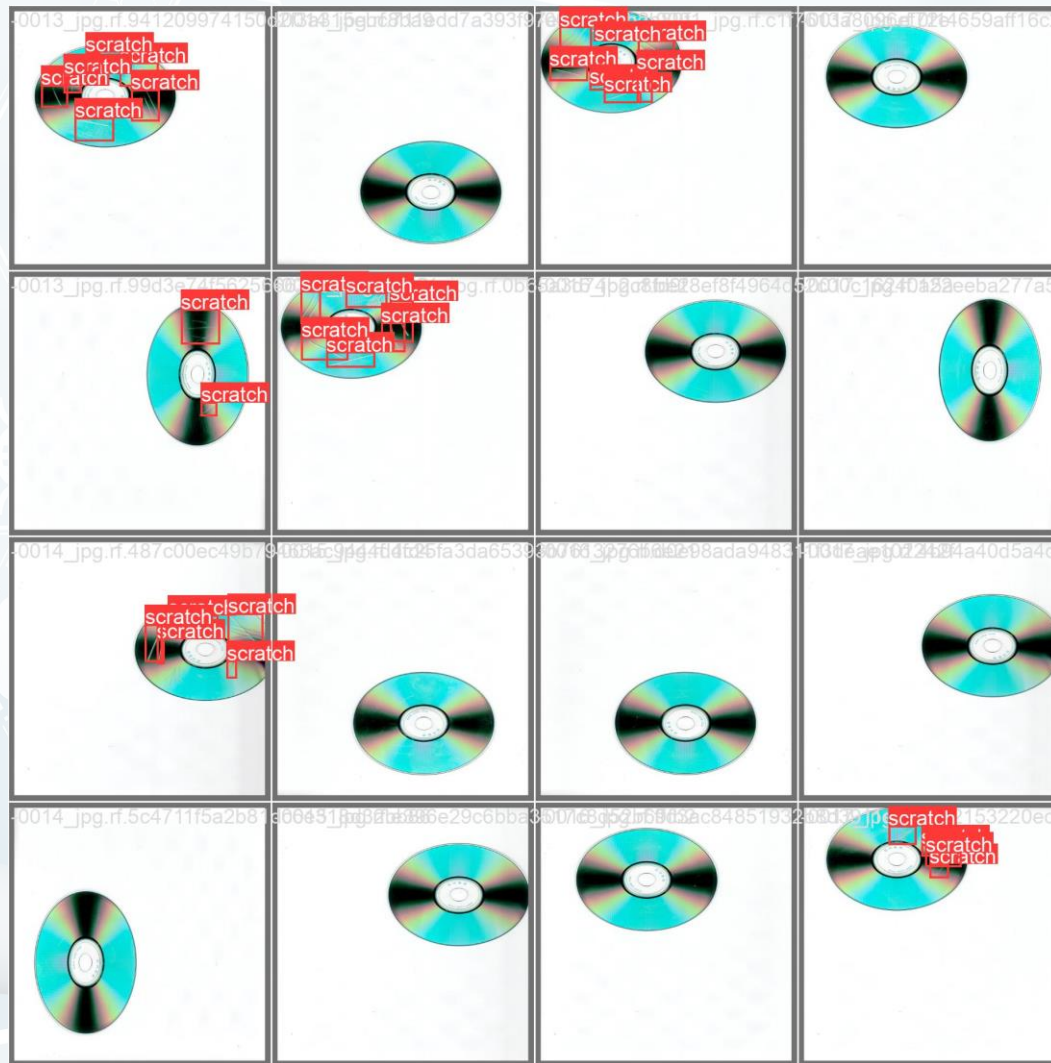
Yolo v9 結果

Accuracy 0.871

Recall 0.823

Precision 0.855

F1 Score 0.838



Yolo v9 結果

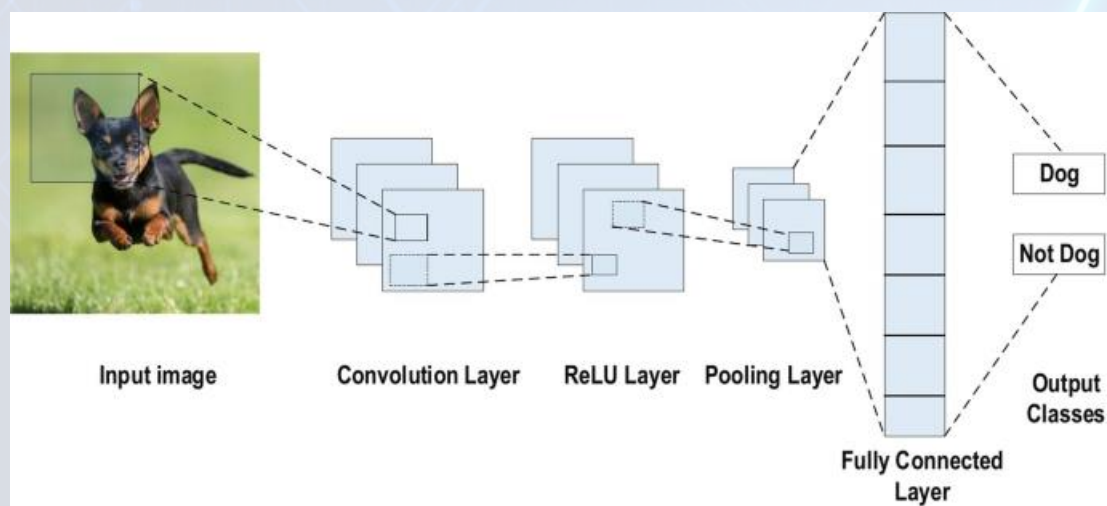
訓練後的模型能夠識別光碟上的刮痕，並在每一個獨立偵測到的物件提供其信心水準



影像辨識模型

CNN

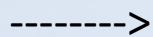
CNN 模型介紹



- CNN 是一種深度學習演算法，能夠從圖像中自動提取特徵並進行分類任務
- 透過卷積層提取特徵、池化層縮減特徵圖的尺寸，降低模型的複雜度，最後由全連接層將提取的特徵轉換為最終的預測結果

CNN

> class2
> class1



-0002_page-0001_jpg.r...fe88180de6734d2cd.jpg
-0005_page-0001_jpg.r...460c4235f97af584e.jpg
-0006_jpg.rf.a45e6fdc4...9cd4a3a79723f5dc2.jpg
-0006_page-0001_jpg.r...b90cd82761449c09.jpg
-0007_jpg.rf.a9da75c8b...565a4ee6cc0e5837f.jpg
-0007_page-0001_jpg.rf...b992a899138cc9ed.jpg
-0008_jpg.rf.b5663c23...1a44f9b54d4150499.jpg
-0008_page-0001_jpg.r...8a3204c61312a5e71.jpg
-0009_jpg.rf.1722a5b0b...46209b5d2d347603.jpg
-0009_page-0001_jpg.rf.8baea9bb4cd83cd75e3aa
-0010_jpg.rf.5d7750845...2d3dad76e8a3aa1bb.jpg
-0010_page-0001_jpg.rf...78e663ad5b223f1f3.jpg

將資料分為有刮痕和無刮痕兩類進行訓練，最後輸出預測的標籤

田口方法超參數調整

	Learning rate	Batch size	Optimizer	Accuracy
1(實驗 1)	0.1	4	Adagrad	0.502
2(實驗 2)	0.1	8	Adam	0.547
3(實驗 3)	0.1	16	SGD	0.611
4(實驗 4)	0.1	32	RMSProp	0.662
5(實驗 5)	0.01	4	Adam	0.597
6(實驗 6)	0.01	8	Adagrad	0.542
7(實驗 7)	0.01	16	RMSProp	0.722
8(實驗 8)	0.01	32	SGD	0.754
9(實驗 9)	0.001	4	SGD	0.651
10(實驗 10)	0.001	8	RMSProp	0.702
11(實驗 11)	0.001	16	Adagrad	0.607
12(實驗 12)	0.001	32	Adam	0.765
13(實驗 13)	0.0001	4	RMSProp	0.602
14(實驗 14)	0.0001	8	SGD	0.559
15(實驗 15)	0.0001	16	Adam	0.501
16(實驗 16)	0.0001	32	Adagrad	0.454


CNN結果

Accuracy	0.765
Recall	0.755
Precision	0.811
F1 Score	0.747

分析出圖片的真實類別與預測類別

```
Filename: class1/-0013_jpg.rf.4b3fc317c0bd602af3815cc8f2071c6d.jpg, True label: class1, Predicted label: class1
Filename: class1/-_jpg.rf.442a3c0fe55a718d02664859db4bc2ac.jpg, True label: class1, Predicted label: class1
Filename: class1/-_jpg.rf.84e1ddf26715975785a80e4277725a32.jpg, True label: class1, Predicted label: class1
Filename: class2/-0001_jpg.rf.1aa76940e91a1758271ee3569a71fa8c.jpg, True label: class2, Predicted label: class1
Filename: class2/-0001_page-0001_jpg.rf.e3a5ac7c248c2c94e25f22c9be8d445b.jpg, True label: class2, Predicted label: class1
```

結果比較

	CNN	 YOLOv9
Accuracy	0.765	0.871
Recall	0.755	0.823
Precision	0.811	0.855
F1 Score	0.747	0.838

結果分析

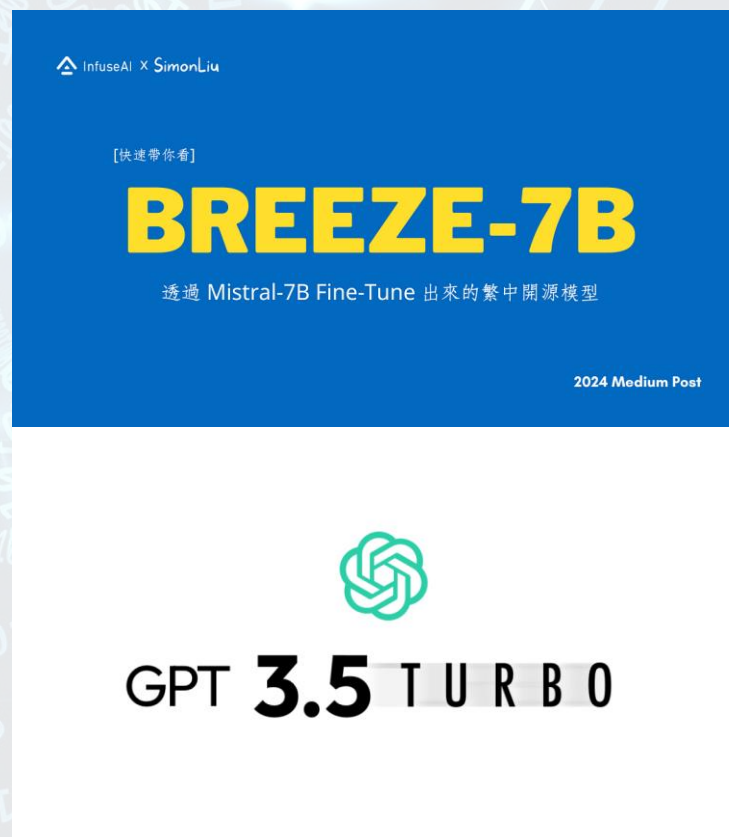
- CNN和YOLO之間準確性差異的原因在於它們的設計目標、應用場景和模型優化策略的不同。影響了它們在特定任務上的表現優劣。
- 模型的性能也會受到數據集特徵的影響，不同的數據集可能有不同的物體分佈或大小範圍，進而影響準確性。
- 當數據集中物體大小和位置變化大時，YOLO可能會表現更好。

大型語言模型

與 Chatbot 結合

與 Chatbot 結合

- Breeze-7B 模型
 - 資料處理
 - 程式碼運行
- GPT 3.5-turbo 模型
 - 資料處理
 - API 使用
 - Fine-tuning



大型語言模型

Breeze-7B 聯發科技

MEDIATEK

篤行一路1號

Breeze-7B

• 資料處理

- 讀取資料夾中所有 PDF 檔案作為訓練集
- 包含商品編號、歌手、專輯描述等

```
from langchain.document_loaders import PyPDFDirectoryLoader
from langchain.indexes.vectorstore import VectorstoreIndexCreator

file_path = "./folder"

loader = PyPDFDirectoryLoader(file_path)

# 創建 VectorstoreIndexCreator 實例時不提供 embedding 參數，使用默認值
index = VectorstoreIndexCreator().from_loaders([loader])
```

商品編號

商品名稱

歌手名稱

描述

單價 專輯類型 庫存

1 走心的歌

陳勢安

他的歌聲從呼吸到表情，每字每句都能徹徹底底的唱進你的心裡，唱到你走心。大勢情歌王陳勢安，出道 18 年，睽違 3 年的全新專輯《走心的歌》終於在 2024 年誕生了！不久前，隨著熱播影集《華麗計程車行》搶先推出的首波主打歌〈破風〉，以充滿慣性前進感的節奏和搖滾的旋律，讓歌迷更加期待新專輯的發行。這張也是陳勢安出道後的第八張專輯，邀請頂尖製作團隊，與 JerryC、張簡君偉、森白 B.K、GJ 蔣卓嘉、許媛婷、佳旺、龐家成等人共同完成專輯。專輯收錄先前推出的多首冠軍單曲以及全新歌曲，曲風也相當多

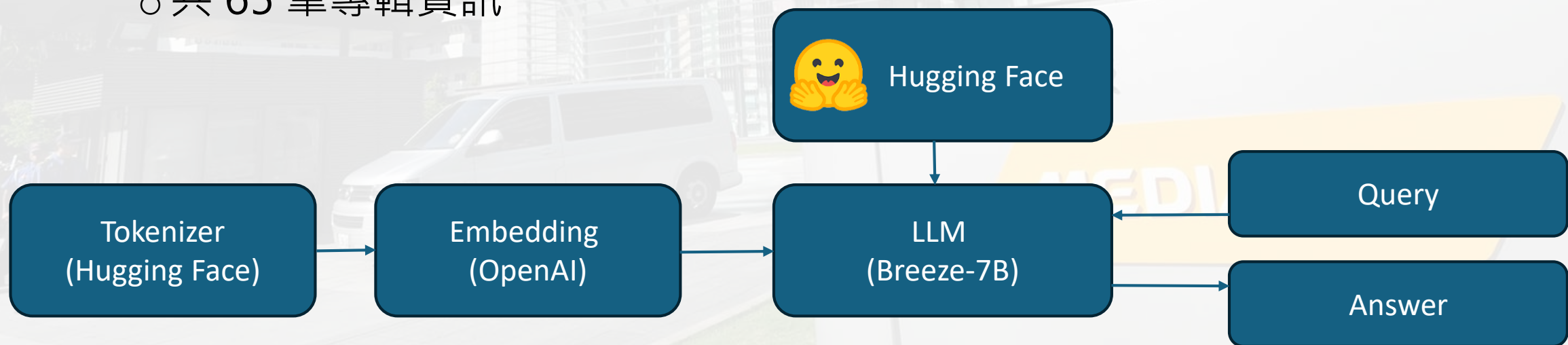
華語流行男歌手

43

Breeze-7B

- 模型架構

- 使用 LangChain 框架串聯
- 共 65 筆專輯資訊



Breeze-7B 結果

```
index.query_with_sources("從cd_goods_inform.pdf中，給我一些歌手名稱為陳勢安的專輯", chain_type="map_reduce")
```

✓ 6.7s

```
{'question': '從cd_goods_inform.pdf中，給我一些歌手名稱為陳勢安的專輯',  
'answer': '陳勢安的專輯【 Almost Human】\n',  
'sources': 'folder\\cd_goods_inform.pdf'}
```

錯誤，此為歌手陳昊森的專輯

```
index.query_with_sources("給我一些歌手名稱為陳勢安的專輯", chain_type="map_reduce")
```

✓ 8.6s

```
{'question': '給我一些歌手名稱為陳勢安的專輯',  
'answer': '陳勢安的專輯有《走心的歌》、《刻在你心底的名字》和《揭不開的痛》。 \n',  
'sources': 'folder\\cd_goods_inform.pdf'}
```

僅對一個，其餘為其他陳姓歌手專輯

Breeze-7B 結果

```
index.query_with_sources("我想聽一些西洋音樂，有西洋流行的專輯嗎？", chain_type="map_reduce")
```

✓ 6.3s

```
{'question': '我想聽一些西洋音樂，有西洋流行的專輯嗎？',  
'answer': ' 我不知道有沒有西洋流行的專輯，請參考SOURCES。 \n',  
'sources': 'folder\\cd_goods_inform.pdf'}
```

```
index.query_with_sources("我想聽一些西洋音樂，有類型為西洋流行的專輯嗎？", chain_type="map_reduce")
```

✓ 5.1s

```
{'question': '我想聽一些西洋音樂，有類型為西洋流行的專輯嗎？',  
'answer': ' 有，像是Ed Sheeran的EP專輯、The Rolling Stones的專輯、Avril Lavigne的專輯都屬於西洋流行類型。 \n',  
'sources': 'folder\\cd_goods_inform.pdf'}
```

需下非常明確的指令



大型語言模型

ChatGPT

GPT-3.5-turbo

- 資料處理

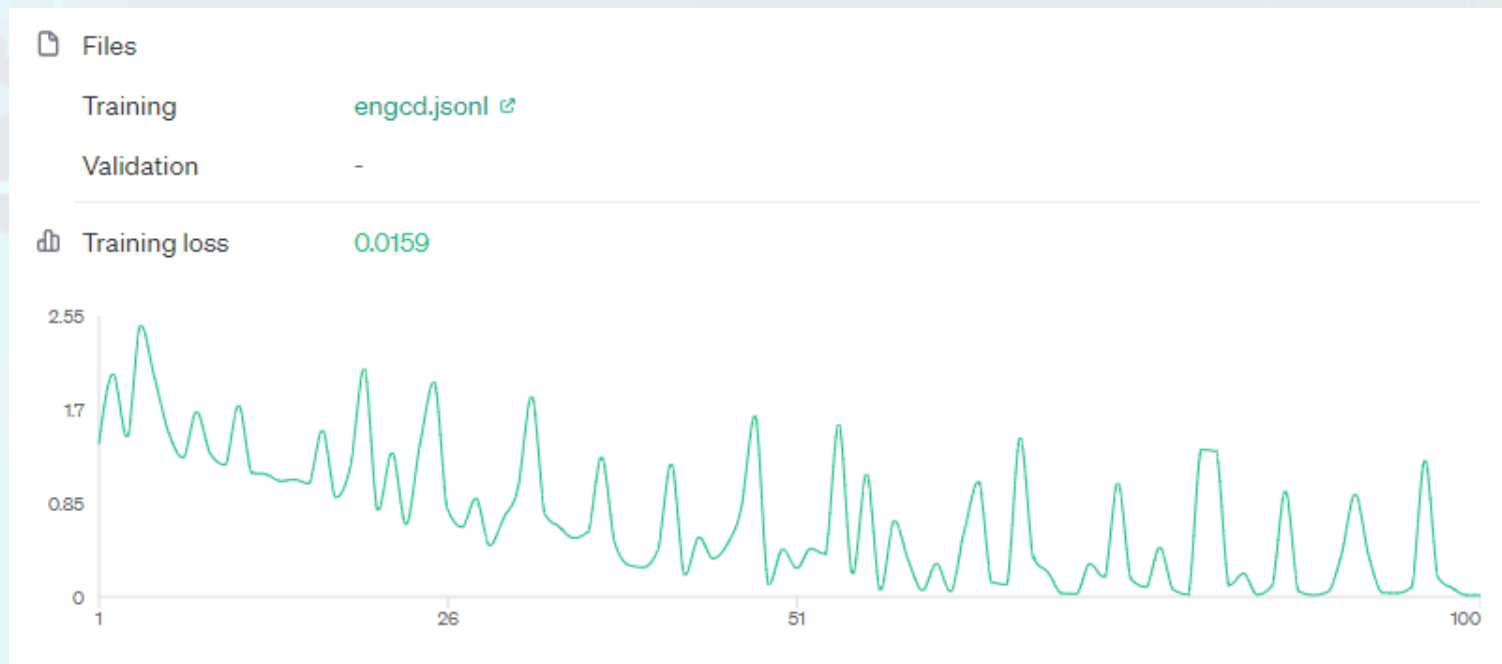
- 規定為 .jsonl 檔
- 需有 system / user / assistant 對話範例

```
{
  "messages": [
    {
      "role": "system",
      "content": "You are Marv, a humorous but matter-of-fact chatbot for an online music store"
    },
    { "role": "user",
      "content": "你們架上的西洋歌手專輯有哪些?" },
    {
      "role": "assistant",
      "content": "哦，對西洋音樂感興趣？那我來推薦一些吧。商品編號:3, 商品名稱:Beautiful Girl, 歌手名稱"
    }
  ]
}
```

GPT-3.5-turbo

- 模型架構

- 使用 OpenAI API fine-tuning(ft:gpt-3.5-turbo-0125:personal::9Wm2uSvO)
- 參數：共訓練 31,090 tokens, 10 epochs, Batch size=1, LR multiplier = 2
- 共 65 筆對話範例



GPT-3.5-turbo 結果



聊天機器人
with ChatGPT

對話紀錄


(Chatbot): 有什麼問題或是想要推薦專輯都可以問我啲~

(You): 請推薦我架上可以購買的西洋歌曲

(Chatbot): 推薦您一款西洋流行音樂專輯，《Purpose》由加拿大歌手Justin Bieber演唱。這是他的第四張錄音室專輯，於2015年11月13日由Def Jam唱片公司和School Boy Records發行。這張專輯主要聚焦在Justin的個人成長與改變，包含了超級熱門的單曲像是《Sorry》和《What Do You Mean?》。這是一張融合R&B和流行音樂元素的專輯，展現了Justin音樂風格的轉變與成熟，非常適合喜歡西洋流行音樂的您。
更多好聽的專輯請前往四大唱片選購

Data sourced from ChatGPT

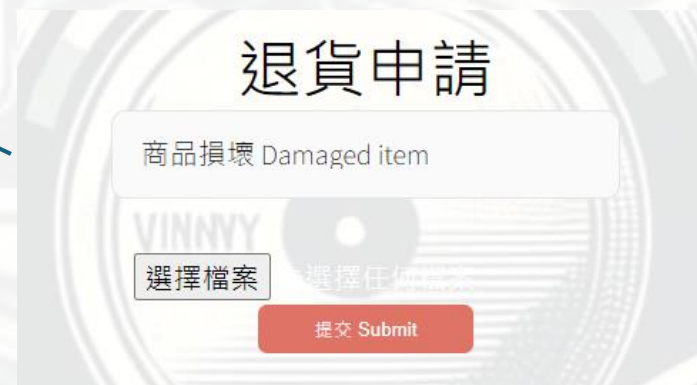
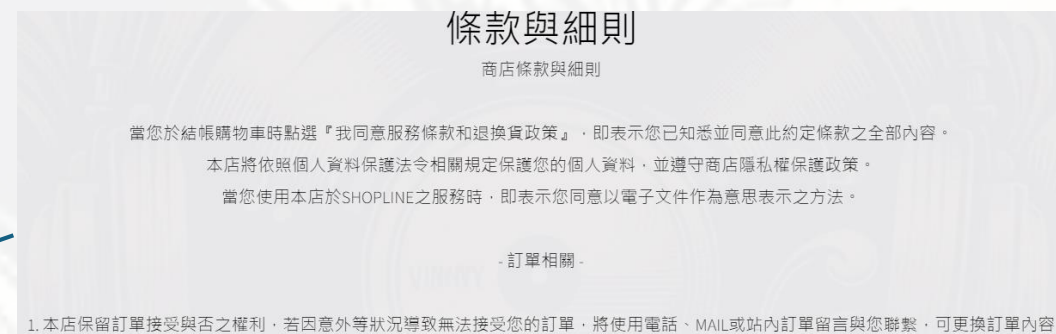
大型語言模型結果比較

	Breeze-7B	 GPT-3.5-turbo
聽得懂問句	V	V
回答正確性	較低	較高
答案完整程度	較低	較高
資料外洩風險	較低	較高

網頁設計-常見問題



- Q：購物條款與細則
A：請參考本網站購物條款與細則
- Q：不知道要怎麼挑選我喜歡的專輯？
A：請參考聊天機器人
- Q：收到的CD有問題怎麼辦？
A：收到商品7天內可退換貨，可先至退貨申請確認商品可申請退貨
- Q：如何追蹤我的訂單？
A：登入帳號後，可在「訂單狀態」頁面查看您的訂單進度。
- Q：商品大約需要多久送達？
A：一般來說，國內訂單約需3-5個工作天，國際訂單約需7-14個工作天。
- Q：是否提供禮品包裝服務？
A：是的，我們提供精美的禮品包裝服務，您可在結帳時選擇此選項。





網頁設計-聊天機器人


聊天機器人

with ChatGPT

對話紀錄

 (Chatbot) : 有什麼問題或是想要推薦專輯都可以問我呦~

 (You) : 你好，我想購買華語女歌手的專輯

 (Chatbot) : 根據您的喜好，我推薦您這張專輯：蔡依林的《OH LA LA LA》。此專輯是蔡依林參與詞曲創作，為速食業廣告量身打造的單曲，其中美式Motown曲風充滿活力，帶來了旋律像許多小精靈在開心的蹦蹦跳跳的感覺，非常適合提升你的心情並讓你跟著音樂舞動。加上MV的馬卡龍色系的服裝場景設計，繽紛可愛，以及蔡依林自帶的魅力，這張專輯定能帶給你許多的樂趣和正能量。

更多好聽的專輯請前往四大唱片選購

Data sourced from ChatGPT

快速選擇

推薦我一些好聽的歌

我想聽華語歌曲

請推薦我韓國的歌

你們有沒有賣女歌手的專輯？

網頁設計-退貨申請



未來規劃

- 後台批量檢查
- 更多退貨申請原因
- 訓練模型時，可以改成使用拍照的方式來做為資料集



退貨申請

選擇退貨理由 Select a reason

選擇退貨理由 Select a reason

商品損壞 Damaged item

商品錯誤 Incorrect item

其他 Other

結論

- 我們透過**整合大型語言模型**和**影像辨識技術**，提升了個人化專輯網站的用戶體驗和產品品質
- 使用 **GPT-3.5** 結合 Chatbot，提供更人性化的即時回應
- 使用 **Yolo v9**，將網站新增**光碟毀損偵測**的功能
- 大型語言模型比較中，GPT-3.5 的表現優於 Breeze-7B
- 影像辨識模型比較中，Yolo v9 的表現優於 CNN



Thank You!