**GAI FINAL PROJECT**

**Chest X-ray Image**

**Report Generation**

組員:

E34106010 黃偉峰

E64102096 張庭瑄

AN4106068 田容甄

AN4101026 呂宏家

# Data Preprocessing Analysis

由於此次任務的資料集較小，於是我們採用一些技術來增加資料的多樣性和數量。在這個流程中，我們對原始X光影像進行了多種預處理操作：

1. 高斯噪聲和高斯模糊：我們在影像中加入高斯噪聲再應用高斯模糊，這樣可以模擬出不同程度的影像品質變化，有助於模型在面對不同影像質量時仍能保持良好的性能。
2. 原始影像：保留原始影像作為參考，確保在數據擴增過程中不丟失原始數據的信息。
3. 翻轉影像：將影像進行水平翻轉，這有助於模型學習影像在不同方向上的變化特徵，增加模型的泛化能力。

做完上述操作，再對這些影像進行銳化處理，使得影像中的細節更加明顯，這對於模型識別微小病變特徵相當重要。

通過這些方法，我們能夠生成多種不同版本的影像，從而有效擴大資料集，提高模型的訓練效果和泛化能力。

A diagram of a diagram

Description automatically generated

上圖為預處理流程圖

# Model & Training Method

此次任務為給定胸部X光影像產生相對應的醫學報告，故需要透過Vision model作為Encoder而Language model作為decoder，實作上我們使用Hugging face所提供的VisionEncoderDecoderModel API來搭建此次任務所需要的model。

Vision Model: [nickmuchi/vit-finetuned-chest-xray-pneumonia](https://huggingface.co/nickmuchi/vit-finetuned-chest-xray-pneumonia)

此model基底使用的是Google Vision Transformer，其主要應用於圖像分類任務，而我們這次使用的Model他在chest-xray-pneumonia之上進行finetune而成，為我們找到能夠串接在VisionEncoderDecoderModel當中Loss和Accuracy相對較高的模型，適合用來作為VisionEncoderDecoderModel中的Encoder。

A screenshot of a computer code

Description automatically generated

Language Model: [emilyalsentzer/Bio\_ClinicalBERT](https://huggingface.co/emilyalsentzer/Bio_ClinicalBERT)

Language Model部分我們使用了對於生物醫學和臨床文本進行微調的 BERT 模型。Bio\_ClinicalBERT是專門針對醫學領域的文本數據進行預訓練和微調的模型，因此非常適合用來生成與胸部X光影像相對應的醫學報告。

Evaluation Score during Training:

A graph of different colored lines

Description automatically generated最高ROUGE分數:

* L-P: 0.5269976253694
* L-R: 0.1796542877144
* L-F: 0.1796542877144
* 2-P: 0.3481254640674
* 2-R: 0.1141826570828
* 2-F: 0.1141826570828

# Analysis

Encountered difficulties:

1. 算力不足

Sol:跟別人借更好的顯示卡。

1. 資料集過小

Sol:透過高斯模糊、水平翻轉來增加資料的多樣性。

1. Model不支援VisionEncoderDecoder

Sol:尋找其他模型串接。

1. 構想藍圖做不出來

Sol:原先預想透過Pretrain過後的yolov8做ROI提取，但是yolov8的使用上我們較不熟悉，並且在Vision model Encoder需要做較多的更改，可能無法使用hugging提供的API，故放棄改用其他方法。

1. 報告中的醫學名詞艱澀難懂，無從判斷模型生成的好壞

Sol:單純使用ROUGE分數作為評判。

# PROPOSAL

智能胸部X光報告轉換系統

A diagram of a person with text

Description automatically generated

我們希望建立一個可以減輕醫生負擔並且也能讓一般病患理解的X光片解讀系統。

**使用流程:**

1. 拍攝X光片放入系統中
2. 胸腔疾病辨識系統分析
3. 產生專業醫學報告
4. 由醫生對於內容進行審核
5. 交由LLM將艱澀難懂的醫學報告轉換為一般大眾可讀懂的資訊

**系統優勢：**

1. 減輕醫生負擔：系統自動生成專業醫學報告，醫生只需審核，大大減少了報告撰寫時間。
2. 提高診斷效率：自動化流程提高了胸部X光片的分析速度，讓病患能更快得到診斷結果。 增強病患理解：透過LLM將專業報告轉化為淺顯易懂的個人化衛教報告，幫助病患更好地理解自己的健康狀況。
3. 提升醫療品質：標準化的報告生成流程確保了報告的質量和一致性，提高了醫療服務的整體品質。

**未來發展：**

1. 擴展到其他醫學影像領域：除了胸部X光片，系統可進一步應用於CT、MRI等其他醫學影像分析。
2. 持續優化模型：不斷收集和學習新的數據，提升疾病辨識系統和LLM的準確性和效能。
3. 多語言支持：開發多語言版本，適應不同地區和語言需求，推廣至全球市場。
4. 整合電子病歷系統：與現有的電子病歷系統（EMR）整合，實現病患數據的全面管理和共享。