

一、資料集

1. id - 每條推文的唯一識別符
2. text - 推文的文本
3. location - 發送推文的地理位置（可能為空）
4. keyword - 推文中的特定關鍵字（可能為空）
5. target - 僅在 train.csv 中，表示推文是否為真實災難(1) 或不為真實災難(0)
6. 任務：判斷推文是否為真實災難

二、模型架構與超參數設定

1. 模型架構

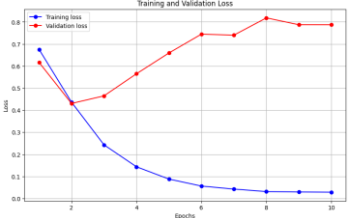
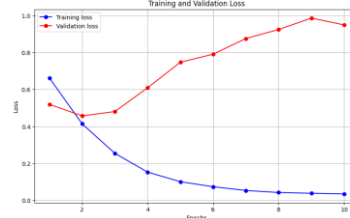
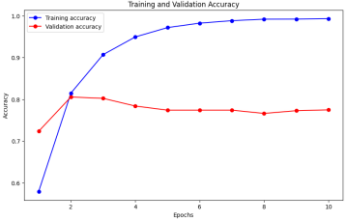
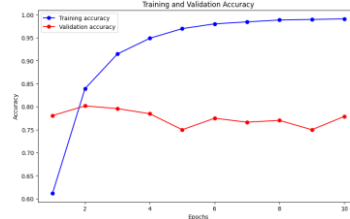
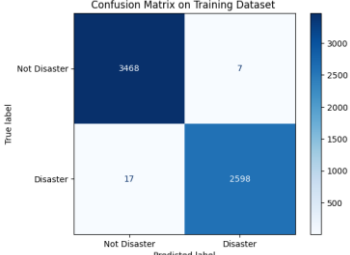
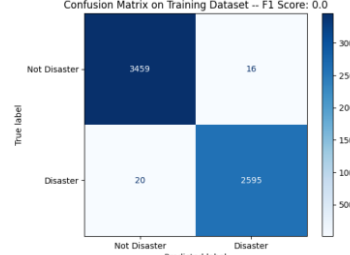
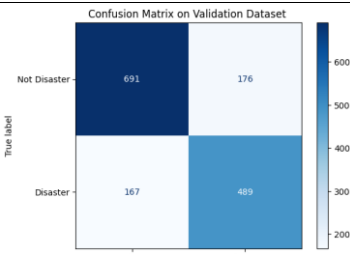
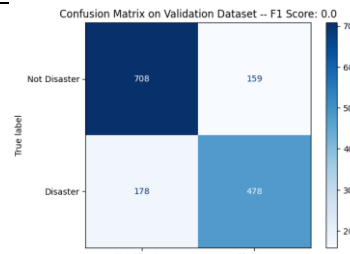
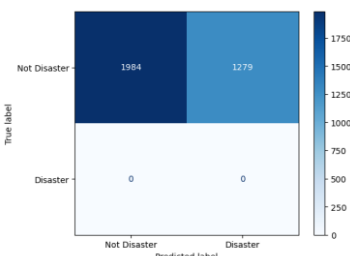
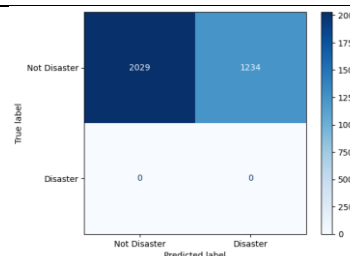
結構元件	GRU	LSTM
RNN 層	Bidirectional(GRU(128, dropout=0.1, recurrent_dropout=0.4))	Bidirectional(LSTM(128, dropout=0.1, recurrent_dropout=0.4))
中間全連接層	Dense(64, ReLU) + Dropout(0.1)	Dense(64, ReLU) + Dropout(0.1)
輸出層	Dense(1, Sigmoid)	Dense(1, Sigmoid)

- 嵌入層（Embedding）使用 mask_zero=True 處理 padding，並接續一層 Dropout 以減少 overfitting。
- Optimizer: Adam
- 損失函數：binary_crossentropy

2. 參數設定

參數	值
max_features（字典大小）	21577
pad_len（每筆輸入序列長度）	25
embed_dim（嵌入維度）	256
dropout	0.1
recurrent_dropout	0.4
learning_rate	0.0001
batch_size	16
epochs	10

三、Experimental Results

項目	GRU	LSTM
訓練時間	107.17 s	132.47 s
Train Accuracy	99.61 %	99.41 %
Validation Accuracy	77.48 %	77.87 %
Validation Recall	0.7454268292682927	0.7286585365853658
Precision	0.7353383458646616	0.750392464678179
參數量	5,836,673	5,934,465
Loss		
Accuracy		
Training Confusion Matrix		
Validation Confusion Matrix		
Testing Confusion Matrix		
Test Accuracy	60.8 %	62.18 %

四、結果分析

1. 過擬合現象：兩個模型在訓練集的準確率均高達 99.5%，但在驗證與測試集上則下降明顯，顯示有可能發生過擬合。有嘗試更改過參數和模型，但此現象依舊嚴重。
2. GRU 表現整體較佳，在參數使用量、訓練速度優於 LSTM。在模型效能上 LSTM 略優於 GRU。
 - i. 參數使用量
 - GRU 較 LSTM 少了 97792(5934465 vs. 5836673)個參數
 - LSTM 有 3 個門 (input、forget、output)
 - GRU 只有 2 個門 (reset、update)，GRU 將 forget gate、output gate 合併為一個 update gate。透過這樣的方式減少了模型參數數量，從而降低了訓練時間
 - 在這個資料集中，GRU 在減小模型複雜度和加快模型訓練速度的同時，模型效能與 LSTM 差不多。
 - ii. 訓練速度：GRU 的訓練時間為 107.17 秒，顯著低於 LSTM 的 132.47 秒，顯示 GRU 在訓練效率上較佳，這也符合 GRU 結構較簡單、參數較少的特性。
 - iii. 模型效能比較
 - 在訓練集上，兩者皆達到非常高的準確率（約 99.5%），但這可能存在 overfitting 現象
 - 驗證集上，LSTM 的準確率略高（77.87% vs. 77.48%）
 - 在 Recall 上 GRU 較優（表示找出災難推文的能力較好），而 Precision 則 LSTM 較好（表示判斷為災難的推文正確率較高）
 - 測試集的準確率亦顯示 LSTM 稍加（62.18% vs. 60.80%）
 - iv. 結論
 - 若優先考慮效能與資源消耗，GRU 是較佳選擇。
 - 若需微幅提升驗證與測試準確率，LSTM 較佳。
 - v. 註：由於 test_ans.csv 中 target 全部為 0，因此無法計算 Test 的 Recall 和 Precision（這兩個指標需要正類樣本才能評估模型在判斷「災難推文」時的能力）。