



113-1 人工智慧

深度學習

預測PM2.5濃度 - 以北京為例

S1161010 林品儀

S1161113 陳毓欣

S1161129 施佩茹



目錄

• • • •

01

緒論

02

文獻探討

03

研究方法

04

實驗

05

結論

06

分工

• • • •

緒論

提供背景資訊、引出研究動機與目的

臺灣空汙可分為境內汙染及境外汙染

其中境外汙染多受中國北京、中國沿岸影響

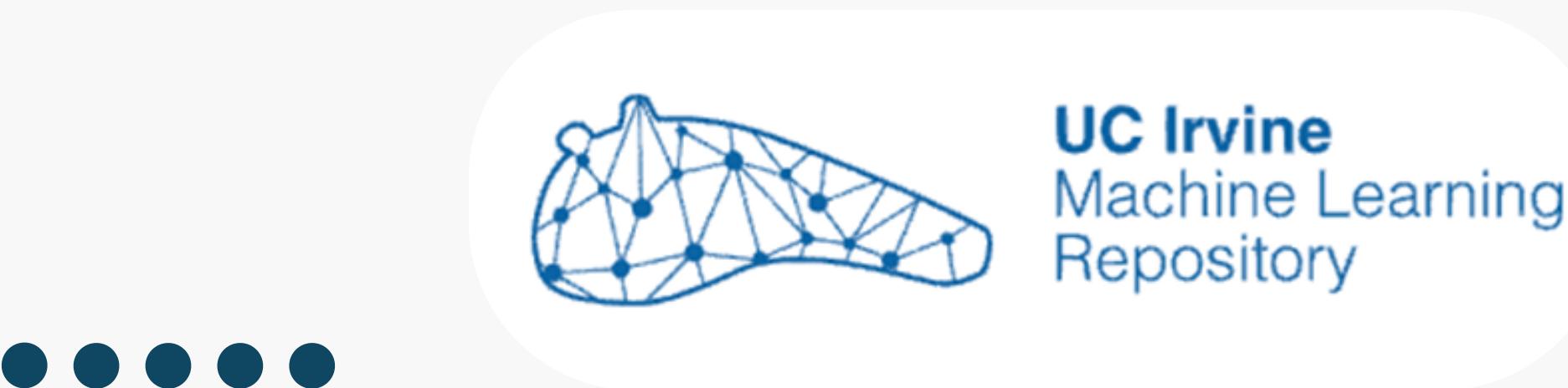
在中國的改善下，汙染物逐漸減少

臺灣的境外汙染物也相對下降



目的

本研究使用 UC Irvine Machine Learning Repository 之 Beijing PM2.5 資料集作為相關數據以預測北京未來的 PM2.5 波動。



目的

修改上學期（112 學年度第二學期）相關
機器學習報告並加入人工智慧時序性模型以
監測出更好的預測結果。



文獻探討 – CNN

.....

回顧相關研究，參考其改善模型方式

結合CNN與LSTM模型建構分類

惡意程式方法：

黃綱正(2019)

主要結合卷積神經網路與循環神
經網路模型，進行惡意程式分
類。經實驗結果 CNN 準確率達
到 86.40%，若將 CNN 結合
LSTM 模型更可提高至 87.79%
的準確率。



開發 CNN 模型預測學生是否退
學：

蔡婷安(2020)

以 CNN 模型判斷在前一學期結
束時，是否可能因學分不及格而
退學。經實驗得知 CNN 的參數
設定並沒有一定的趨勢，需透過
經驗進行嘗試得出最佳參數。除
此之外 CNN 能以更快的速度進
行更全方面的表徵學習。

.....

文獻探討 – RNN

• • • •

回顧相關研究，參考其改善模型方式

Recurrent neural
networks :

Balázs Hidasi, Alexandros
Karatzoglou, L. Baltrunas, D.
Tikk (2013)

循環神經網絡是一種具有內部迴路的人工神經網絡。這些內部迴路在網絡中引入了遞歸動態，從而在網絡的處理單元 (PEs) 之間產生延遲激活的依賴性。



Independently Recurrent
Neural Network (IndRNN):
Building A Longer and
Deeper RNN :

Shuai Li, W. Li, Chris Cook, Ce
Zhu, Yanbo Gao (2018)

IndRNN可以與非飽和激活函數（如 ReLU，修正線性單元）配合使用，仍然能夠穩健地進行訓練。多個IndRNN可以堆疊起來構建比現有RNN更深的網絡。

• • • •

文獻探討 – LSTM

回顧相關研究，參考其改善模型方式

使用多頭注意力機制的多重輸入

LSTM 股價預測模型：

葉國毅(2024)

本研究以台灣50成分股為訓練資料，利用皮爾森相關係數篩選多重輸入特徵，結合注意力機制與LSTM進行股價預測。結果顯示，該模型相較於其他基礎模型在預測效果與風險避險上均有優勢。



應用CNN-LSTM深度學習模型
於預測台灣二氧化碳排放量：

張曉妮(2024)

本研究以CNN與LSTM建立台灣二氧化碳排放量預測模型，加入經濟指標訓練並以MAPE衡量效果。結果顯示，CNN-LSTM模型表現最佳，提出的線性插值方法更準確模擬實際數據。

文獻探討 – GRU

.....

回顧相關研究，參考其改善模型方式

文獻一：

利用LSTM、RNN與GRU模型預測彼特比價格：一個深度學習方法的效能分析

李思賢(2024)

本研究利用長短期記憶網路 (Long Short Term Memory, LSTM) 、遞歸神經網路 (Recurrent Neural Networks, RNN) 及門控遞歸單元 (Gate Recurrent Unit, GRU) 等深度學習技術，結合時間序列分析和市場情緒指標，深入分析彼特比市場。

文獻二：

邁向淨零碳排放利用深度學習LSTM與GRU預測台灣電力消耗與再生能源未來用量之研究

游婉喻(2024)

本研究透過應用長短期記憶網路 (LSTM) 與門控循環單元 (GRU) 深度學習模型，預測台灣全國電力消耗和再生能源發電量，以支持國家能源轉型及淨零碳排放策略的實施。

● ● ● ●

文獻探討 – Transformer

.....

回顧相關研究，參考其改善模型方式

可微分查找矩陣乘法用於壓縮

Transformer網路：

周家興(2024)

本研究改進了LUT-NN，提出端到端訓練方法，替代傳統乘累加運算以節能。實驗結果顯示，在ImageNet數據集上，作者的方法提升了LUT-NN的準確率最多11%。



基於PM2.5長時間序列深度學習的極端事件預測：

楊信之(2022)

本研究針對PM2.5極端事件預測，提出極端事件損失函數、Transformer 模型及Composite Network架構，以提升極端情況下的預測準確度。模型基於2014-2020年台灣空氣品質數據，為非均衡和極端時序資料的預測開闢新方向。

• • • •

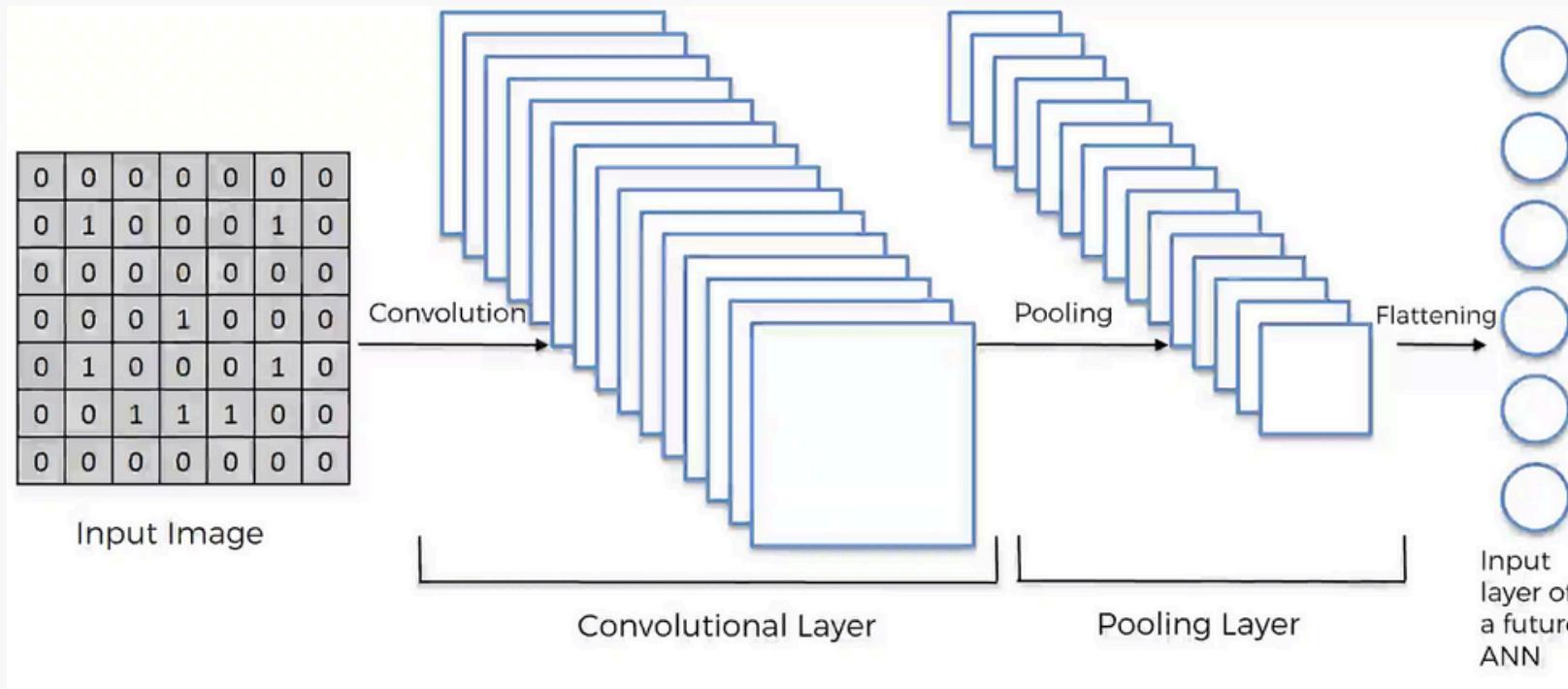


卷積神經網路-CNN

S1161010 林品儀

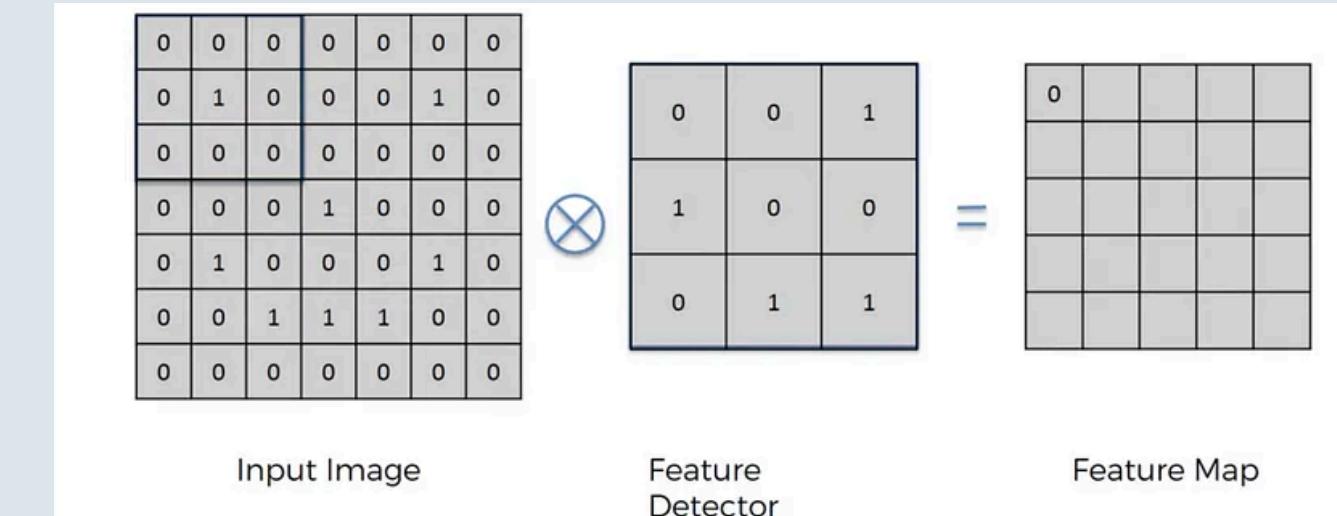


研究方法 – CNN

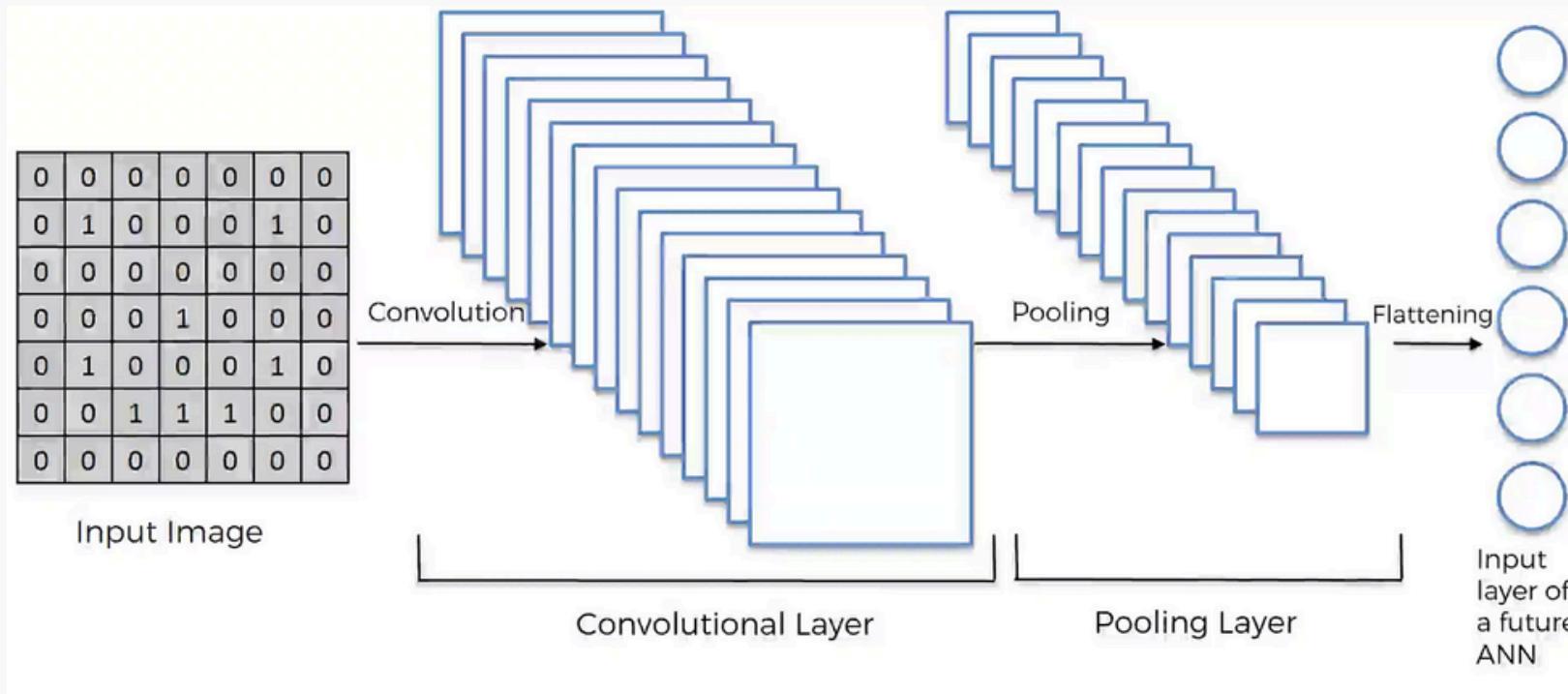


- 卷積神經網路
 - 局部特徵檢測、權重共享和池化壓縮
 - 常用於影像辨識、語音辨識和時間序列資料
- 包含卷積層（Convolution）、池化層（Pooling）及全連接層

卷積層（Convolution）
會使用過濾器（filter）在數據上滑動並進行卷積計算，逐步提取特徵並生成特徵圖。

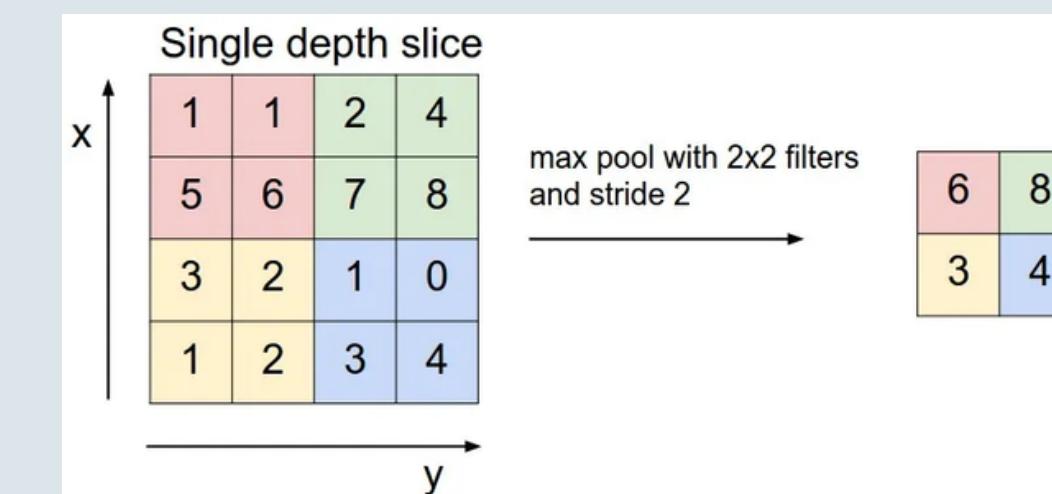


研究方法 – CNN



池化層 (Pooling)

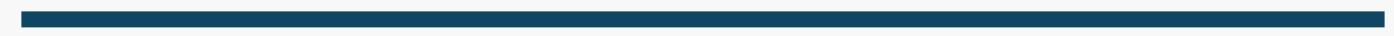
用於減少特徵圖的大小，並保留重要重要特徵，常使用最大池化 (Max Pooling) 和平均池化 (Average Pooling) 。



全連接層

整合從卷積層和池化層學到的特徵，並將這些特徵轉換成最終的分類結果。

• • • •



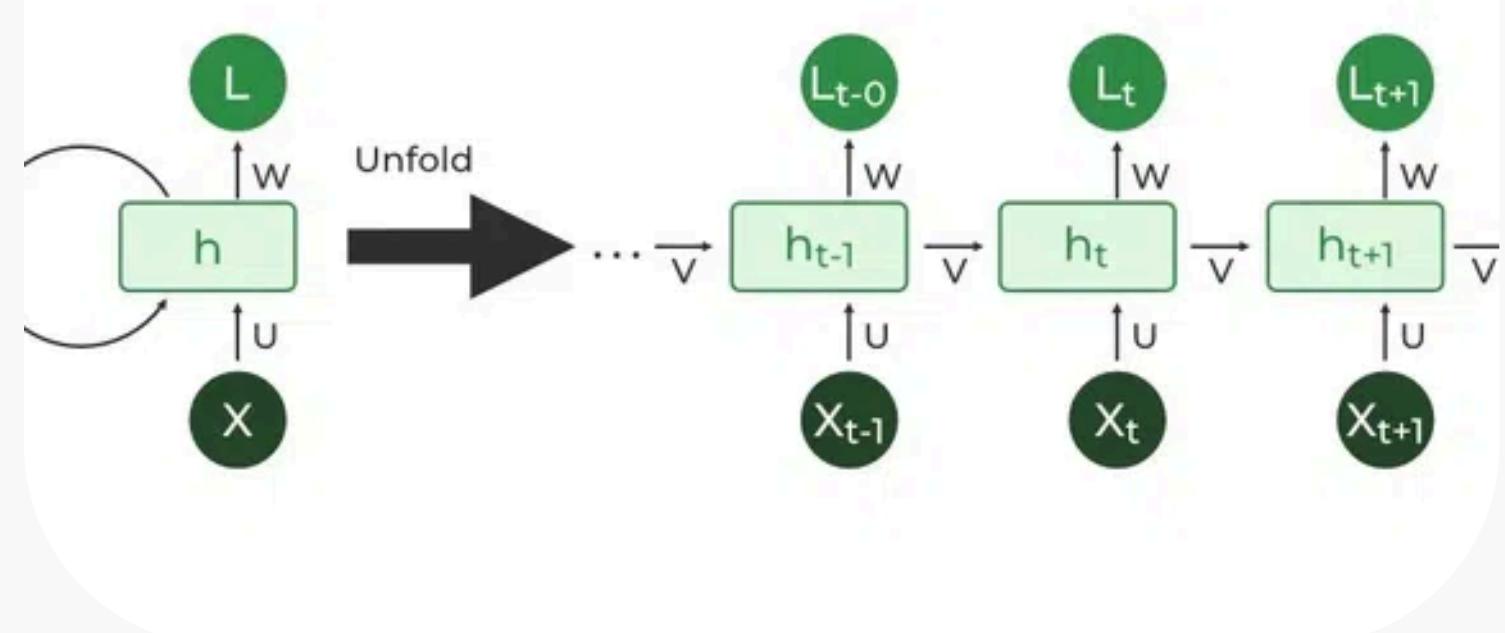
循環神經網路-*RNN*

S1161113 陳毓欣



• • • •

研究方法 – RNN

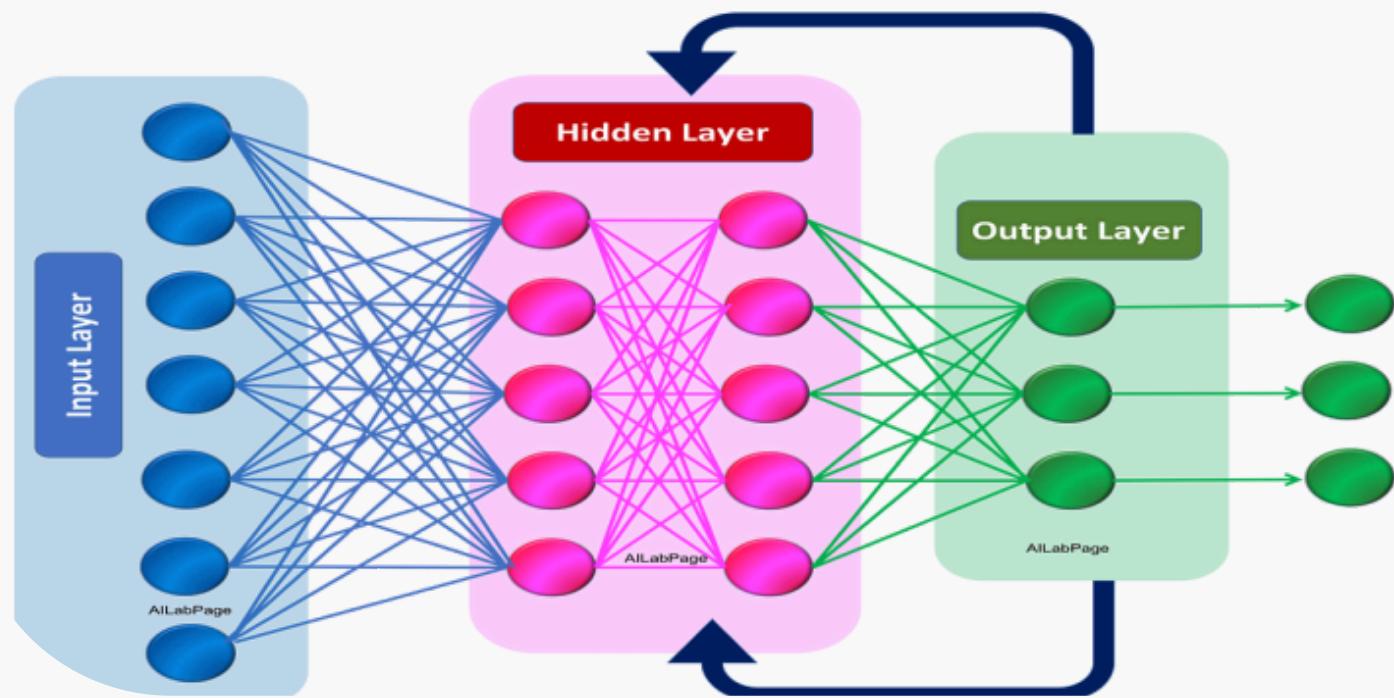


此次報告中實驗將包含測試堆疊、雙向、混用等
等以找出最佳模型及其最佳參數

- 內部循環（定向循環）
- 對於涉及語音、文字、財務資料、音訊和視
訊的任務特別有效
- 其記憶長序列的效果受到梯度消失問題的限
制
- 動態時間行為

研究方法 – RNN

Recurrent Neural Networks



架構：

- 編碼器
- 解碼器
- 雙向讀取

變種：

- 完全循環
- **Elman** 網路和 **Jordan** 網路
- 雙向循環神經網路



長短期記憶-LSTM

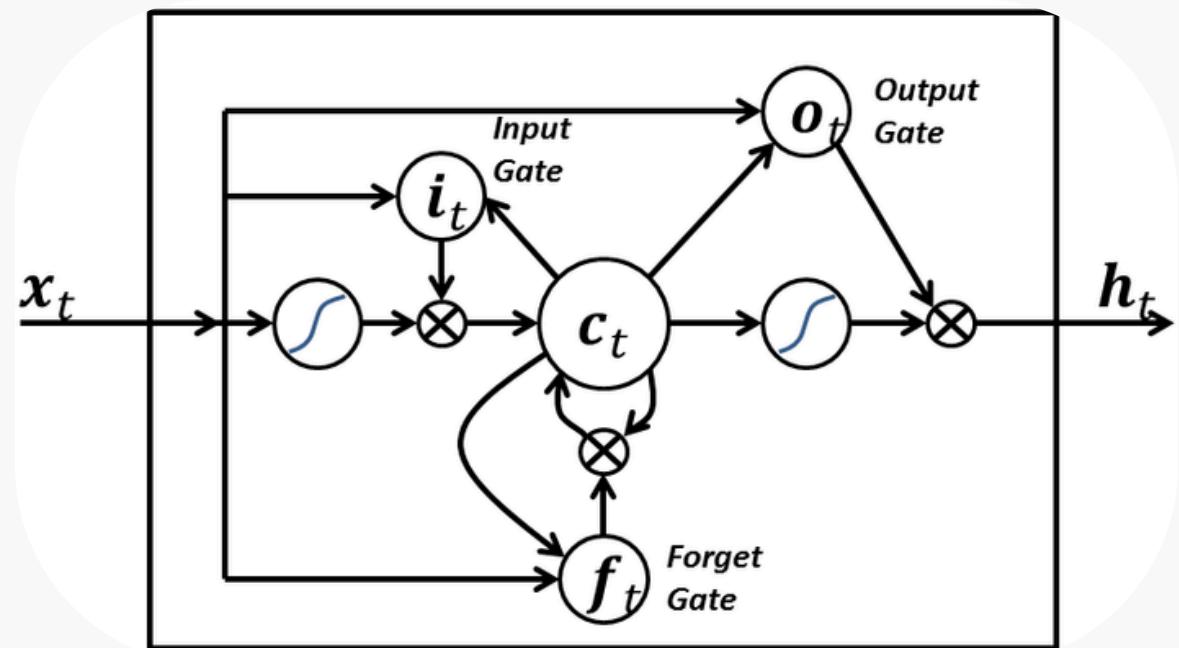
S1161113 陳毓欣



研究方法 – LSTM

———

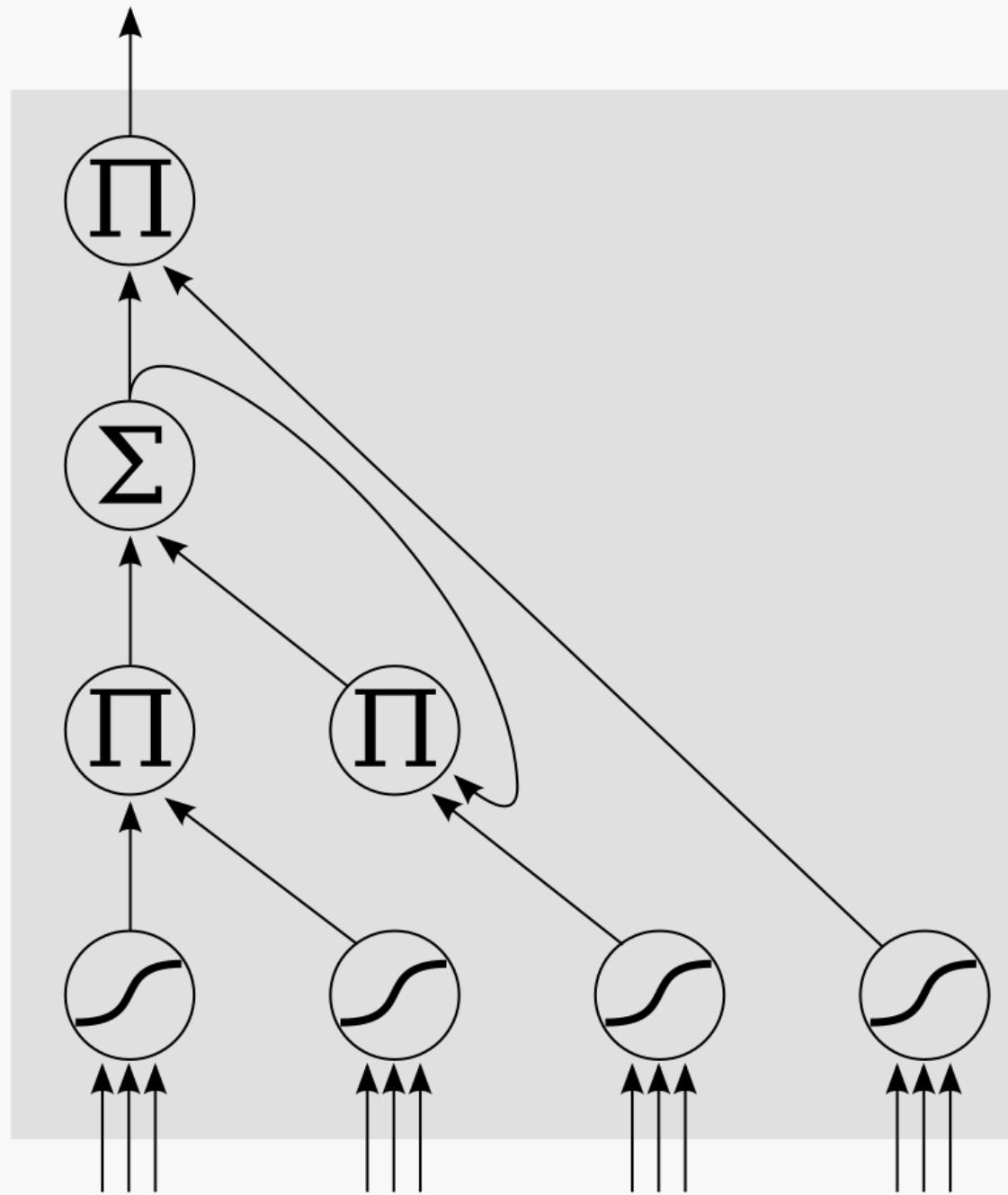
此次報告中實驗將包含測試堆疊、雙向、混用等
等以找出最佳模型及其最佳參數



- 時間循環神經網路
- 處理和預測時間序列中間隔和延遲非常長的重要事件
- 普遍用於自主語音辨識
- 非線性模型
- 可用於構造更大型的深度神經網路

由細胞單元 (cell) 、輸入門 (input gate) 、
輸出門 (output gate) 、遺忘門 (forget
gate) 組成

研究方法 – LSTM



- 輸入門產出近似於零則不進到下一層
- 遺忘門的值近似於零則將值忘掉
- 輸出門可決定區塊記憶中的input是否輸出

其中最重要的是Forget gate，其次是Input gate，最次是Output gate。

正常的倒循環類神經（例：梯度下降法）是一個有效訓練LSTM區塊記住長時間數值的方法。

它可以記憶不定時間長度的數值，區塊中有gate能決定input是否重要到能被記住及能被輸出。

.....

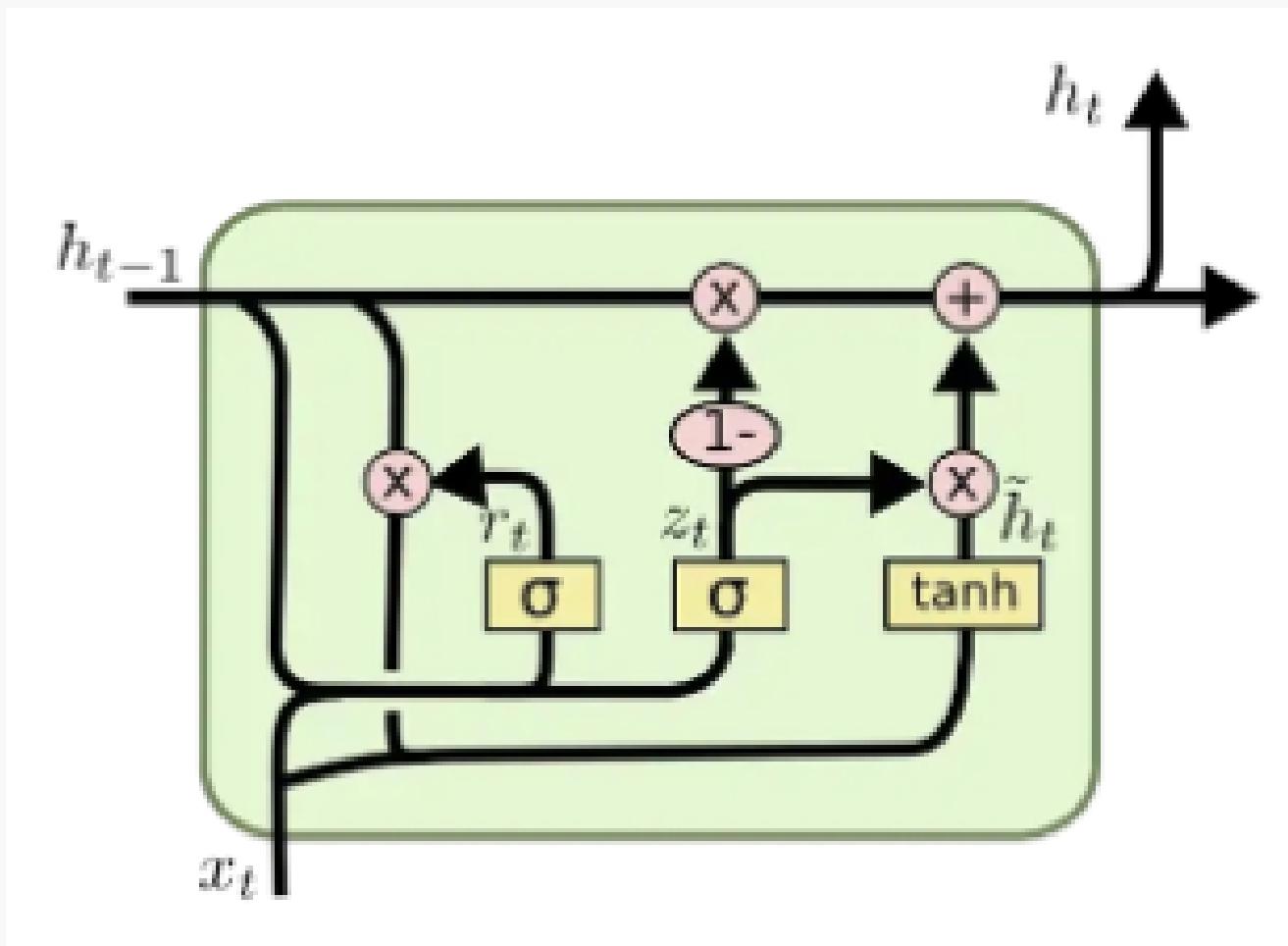


門控循環單元-GRU

S1161129 施佩茹

.....

研究方法 – Gated Recurrent Unit——



GRU也稱為門控循環單元，藉由「門控」控制訊息流，由LSTM延伸出來的模型，但相較於LSTM有以下幾點的差異：

- 將輸入門、遺忘門、輸出門，簡化為更新門跟重設門
- 更新門：控制訊息保留的程度，類似於LSTM的遺忘門和輸入門的功能。
- 重設門：決定如何從過去的隱藏狀態中重置或過濾訊息，縮減訊息處理流程。
- 直接利用隱藏狀態來同時保留長期和短期記憶
- 在LSTM中，輸出門負責調整隱藏狀態輸出的幅度，而GRU則將這個功能整合到更新門中，因此不需要一個單獨的輸出門。

研究方法 – Gated Recurrent Unit——

更新門：控制應該從前一隱藏狀態保留多少訊息

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

重置門：控制前一隱藏狀態中的哪些訊息應該被忽略

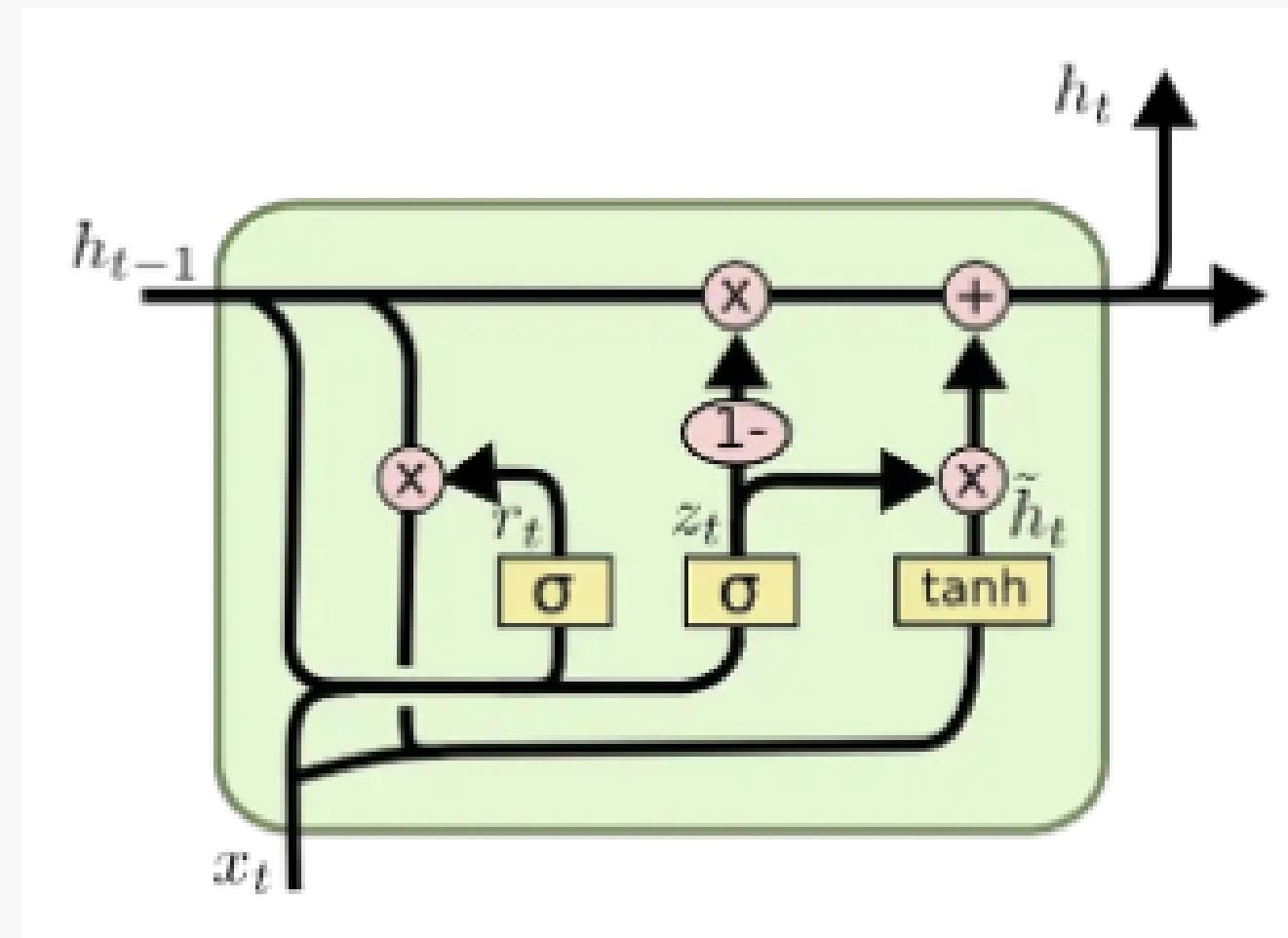
$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

候選隱藏狀態：生成了當前時間步驟的新訊息

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

隱藏狀態：是前一隱藏狀態和候選隱藏狀態的加權和，取決於更新門的狀態

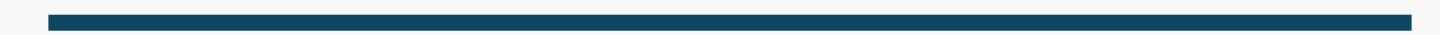
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$





Transformer

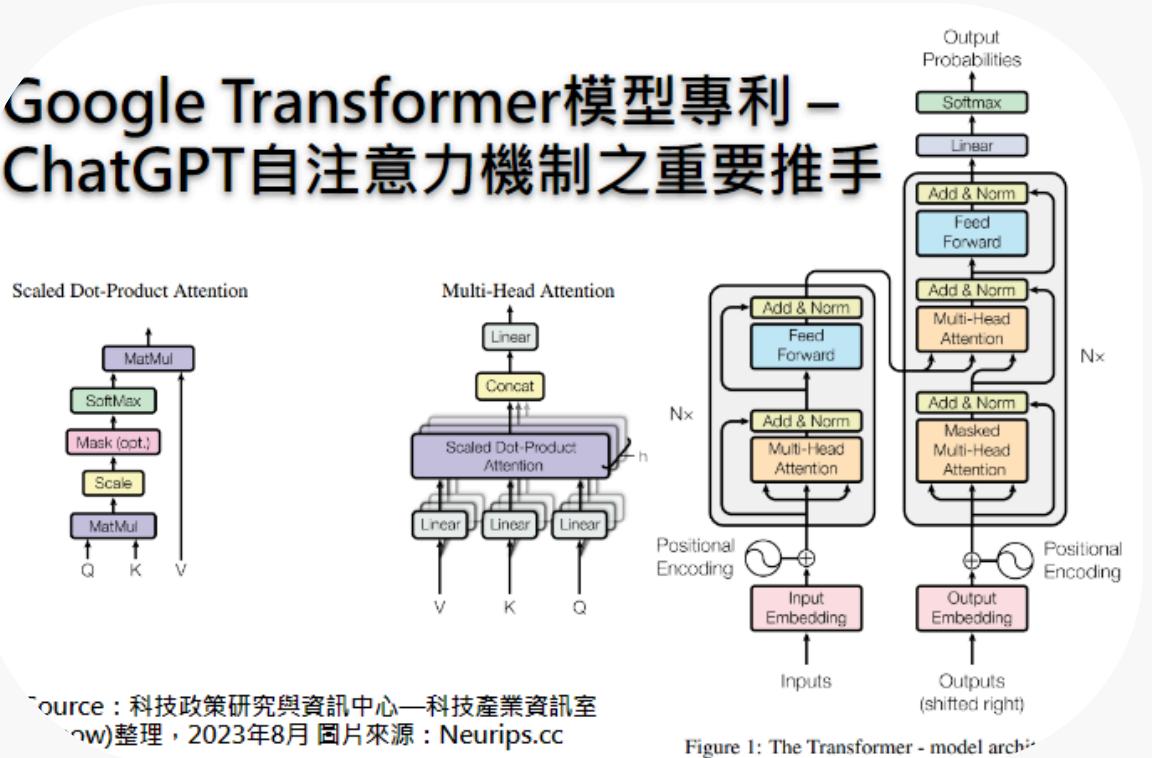
S1161113 陳毓欣



研究方法 – Transformer

此次報告中實驗將包含測試堆疊、雙向、混用等
等以找出最佳模型及其最佳參數

Google Transformer模型專利 –
ChatGPT自注意力機制之重要推手

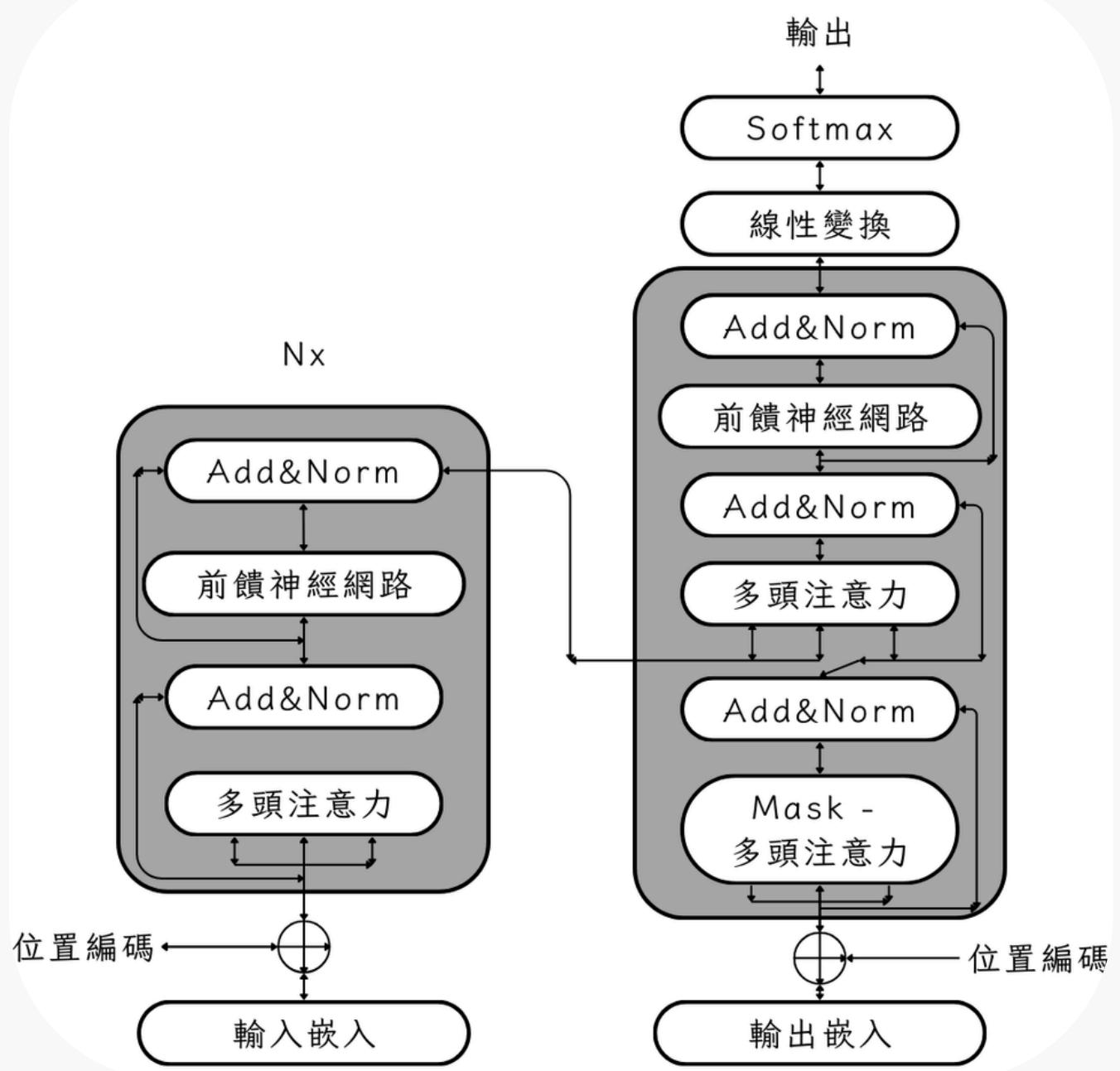


- 採用注意力機制的深度學習模型
- 可以按輸入資料各部分重要性的不同而分配不同的權重
- 主要用於自然語言處理與電腦視覺領域
- 旨在處理自然語言等順序輸入資料

與RNN不同的是，這種模型能一次性處理所有輸入資料（平行計算增加），並減少其訓練時間。

注意力機制可以為輸入序列中的任意位置提供上
下文。

研究方法 – Transformer



注意力機制：

- 提取序列中任意先前點的狀態資訊
- 能夠訪問所有先前的狀態並根據學習到的相關性度量對其進行加權

RNN vs Transformer：

- RNN添加注意力機制能提高模型的效能
- Transformer注意力機制本身就足夠強大，且不需要像RNN模型一樣再對資料進行順序迴圈處理。

實驗 - 資料集說明



特徵名稱	特徵說明	型態	特徵名稱	特徵說明	型態	特徵名稱	特徵說明	型態
pm2.5	pm2.5濃度	float64	month	月份	int64	DEWP	露點溫度	int64
No	編號	int64	day	日期	int64	TEMP	氣溫	float64
year	年分	int64	hour	小時	int64	PRES	氣壓	float64

實驗 - 資料集說明



特徵名稱	特徵說明	型態	特徵名稱	特徵說明	型態
cbwd	風向	object	Ir	累積降雪量	int64
Iws	累積風速	float64			
Is	累積降水量	int64			

實驗 - 實驗前處理

1. 將年、月、日欄位組合

```
dataset = read_csv(filename, parse_dates=[[ 'year', 'month', 'day', 'hour']],
                    index_col=0, date_parser=parse)
print('\n顯示原始資料_組合(年月日時)')
print(dataset.head(5))
```

2. 將pm2.5欄位中的0取代NA後刪除第一天24小時的資料

```
#mark all NA values with 0
dataset['pollution'].fillna(0, inplace=True)

print('\n顯示前處理資料_0取代NA&新欄位名稱')
print(dataset.head(5))

#drop the first 24 hours
dataset=dataset[24:]
print('\n顯示前處理資料_drop第一天24小時，因為pm2.5為0')
```



實驗 - 實驗前處理

3. 欄位整理並重新命名

```
#manually specify column names  
dataset.columns = ['pollution', 'dew', 'temp',  
                   'press', 'wnd_dir', 'wnd_spd',  
                   'snow', 'rain']  
dataset.index.name='date'
```

原本欄位名稱	更改後欄位名稱	欄位代表意涵
NO	已刪除	編號
year、month、day、hour	date	觀測時間
pm2.5	pollution	pm2.5濃度
DEWP	dew	露點溫度
TEMP	temp	氣溫
PRES	press	壓力
cbwd	wnd_dir	風向
Iws	wnd_spd	風速
Is	snow	降雪量
Ir	rain	降雨量

實驗 - 實驗前處理

4. 將 wnd_dir 欄位進行標籤編碼

```
# 將 wnd_dir 欄位的值進行轉換  
df['wnd_dir'] = df['wnd_dir'].replace({'CV': 0,  
                                         'NW': 1,  
                                         'SE': 2,  
                                         'NE': 3,  
                                         'SW': 4})
```

欄位名稱	代表意涵	轉換後代號
CV	無風	0
NW	西北風	1
SE	東南風	2
NE	東北風	3
SW	西南風	4

實驗 - 實驗前處理

5. 最終結果顯現

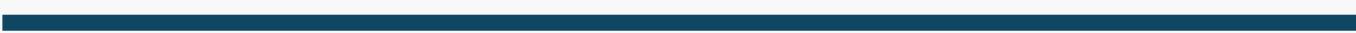
No	year	month	day	hour	pm2.5	DEWP	TEMP	PRES	cbwd	Iws	Is	Ir
1	2010	1	1	0	NA	-21	-11	1021	NW	1.79	0	0
2	2010	1	1	1	NA	-21	-12	1020	NW	4.92	0	0
3	2010	1	1	2	NA	-21	-11	1019	NW	6.71	0	0

date	pollution	dew	temo	press	wnd_dir	wnd_spd	snow	rain
2010/1/2	129	-16	-4	1020	2	1.79	0	0
2010/1/2	148	-15	-4	1020	2	2.68	0	0
2010/1/2	159	-11	-5	1021	2	3.57	0	0



卷積神經網路-實驗結果

S1161010 林品儀



實驗 - CNN

神經元個數

CNN層：
filters: 8,
kernel_size: 2

Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 4, epochs: 15
Test MAE: 1.345
Test RMSE: 1.856
Test MAPE: 8.23%
花費時間: 0:28:15.119012

實驗 - CNN(with MaxPooling)

神經元個數

CNN層：
filters: 64,
kernel_size: 3

Dense層 : 1

CNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 5.534
Test RMSE: 6.240
Test MAPE: 29.79%
花費時間: 0:08:16.291097

實驗 - 堆疊CNN

神經元個數

CNN層：
filters: 8,
kernel_size: 2
第一層隱藏層：8
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 20
Test MAE: 3.717
Test RMSE: 4.587
Test MAPE: 16.27%
花費時間: 0:34:11.236347

實驗 - 堆疊CNN(with MaxPooling)

神經元個數

CNN層：
filters: 8,
kernel_size: 3
第一層隱藏層：8
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 4, epochs: 30
Test MAE: 3.641
Test RMSE: 4.423
Test MAPE: 17.25%
花費時間: 1:54:47.362423

實驗 - 堆疊CNN

神經元個數

CNN層：
filters: 8,
kernel_size: 2
第一層隱藏層：8
第一層隱藏層：8
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 15
Test MAE: 4.372
Test RMSE: 5.290
Test MAPE: 18.01%
花費時間: 0:21:20.070984

實驗 - 堆疊CNN(with MaxPooling)

神經元個數

CNN層：
filters: 8,
kernel_size: 3
第一層隱藏層：8
第一層隱藏層：8
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 2, epochs: 30
Test MAE: 2.861
Test RMSE: 3.666
Test MAPE: 10.37%
花費時間: 4:31:56.327557

統整實驗結果 - CNN

神經元個數	CNN 使用最佳參數結果
一層 CNN層	Test MAE: 1.345, Test RMSE: 1.856, Test MAPE: 8.23% 花費時間: 0:28:15.119012
一層 CNN層(with Pooling)	Test MAE: 5.534, Test RMSE: 6.240, Test MAPE: 29.79% 花費時間: 0:08:16.291097
二層 CNN層	Test MAE: 3.717, Test RMSE: 4.587, Test MAPE: 16.27% 花費時間: 0:34:11.236347
二層 CNN層(with Pooling)	Test MAE: 3.641, Test RMSE: 4.423, Test MAPE: 17.25% 花費時間: 1:54:47.362423
三層 CNN層	Test MAE: 4.372, Test RMSE: 5.290, Test MAPE: 18.01% 花費時間: 0:21:20.070984
三層 CNN層(with Pooling)	Test MAE: 2.861, Test RMSE: 3.666, Test MAPE: 10.37% 花費時間: 4:31:56.327557

實驗 - CNN+LSTM

神經元個數

CNN層：
filters: 8,
kernel_size: 2
第一層隱藏層：64
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 0.7443
Test RMSE: 0.925
Test MAPE: 2.75%
花費時間: 0:34:20.660109

實驗 - CNN+RNN

神經元個數

CNN層：
filters: 8,
kernel_size: 2
第一層隱藏層：64
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 0.855
Test RMSE: 1.137
Test MAPE: 2.59%
花費時間: 0:19:40.546443

實驗 - CNN+GRU

神經元個數

CNN層：
filters: 8,
kernel_size: 2
第一層隱藏層：64
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 0.866
Test RMSE: 1.095
Test MAPE: 3.47%
花費時間: 0:28:59.716134

統整實驗結果 - CNN混和模型

神經元個數	LSTM 使用最佳參數結果
CNN + LSTM	Test MAE: 0.7443, Test RMSE: 0.925 Test MAPE: 2.75% 花費時間: 0:34:20.660109
CNN + RNN	Test MAE: 0.855, Test RMSE: 1.137 Test MAPE: 2.59% 花費時間: 0:19:40.546443
CNN +GRU	Test MAE: 0.866, Test RMSE: 1.095 Test MAPE: 3.47% 花費時間: 0:28:59.716134

實驗 - 並聯CNN

神經元個數

3層CNN：
filters: 8,
kernel_size: 2, 3, 5
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 20
Test MAE: 0.861
Test RMSE: 1.031
Test MAPE: 3.87%
花費時間: 0:12:05.270362

實驗 - 並聯CNN+LSTM

神經元個數

3層CNN：
filters: 8,
kernel_size: 2, 3, 5
第一層隱藏層：64
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 20
Test MAE: 1.895
Test RMSE: 2.187
Test MAPE: 7.47%
花費時間: 0:21:29.015740

實驗 - 並聯CNN+RNN

神經元個數

3層CNN：
filters: 8,
kernel_size: 2, 3, 5
第一層隱藏層：64
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 20
Test MAE: 1.949
Test RMSE: 2.335
Test MAPE: 9.43%
花費時間: 0:18:43.818856

實驗 - 並聯CNN+GRU

神經元個數

3層CNN：
filters: 8,
kernel_size: 2, 3, 5
第一層隱藏層：64
Dense層：1

CNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 20
Test MAE: 1.321
Test RMSE: 1.643
Test MAPE: 4.36%
花費時間: 0:22:05.209224

統整實驗結果 - 並聯CNN混和模型

神經元個數	LSTM 使用最佳參數結果
並聯 CNN	Test MAE: 0.861, Test RMSE: 1.031 Test MAPE: 3.87% 花費時間: 0:12:05.270362
並聯 CNN+LSTM	Test MAE: 1.321, Test RMSE: 1.643 Test MAPE: 4.36% 花費時間: 0:22:05.209224
並聯 CNN+RNN	Test MAE: 1.895, Test RMSE: 2.187 Test MAPE: 7.47% 花費時間: 0:21:29.015740
並聯 CNN+GRU	Test MAE: 1.949, Test RMSE: 2.335 Test MAPE: 9.43% 花費時間: 0:18:43.818856

實驗 - CNN小結

	CNN	CNN+LSTM	CNN+RNN	並聯CNN
隱藏層層數	CNN層： filters: 8, kernel_size: 2	CNN層： filters: 8, kernel_size: 2 第一層隱藏層：64	CNN層： filters: 8, kernel_size: 2 第一層隱藏層：64	3層CNN： filters: 8, kernel_size: 2, 3, 5 Dense層：1
MAE	1.345	0.7443	0.855	0.861
RMSE	1.856	0.925	1.137	1.031
MAPE	8.23%	2.75%	2.59%	3.87%



長短期記憶-實驗結果

S1161113 陳毓欣



實驗 - LSTM

神經元個數

LSTM層：64
Dense層：1

LSTM
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 14.809
Test RMSE: 25.118
Test MAPE: 24.56%
花費時間: 2:05:54.426071

實驗 - LSTM

神經元個數

LSTM層：64

第一層隱藏層：128

Dense層：1

LSTM
使用最佳參數結果

Best Parameters: batch_size: 64, epochs: 20

Test MAE: 14.695

Test RMSE: 24.793

Test MAPE: 24.57%

花費時間: 2:55:06.608498

實驗 - LSTM

神經元個數

LSTM層 : 64

第一層隱藏層 : 512

第二層隱藏層 : 128

Dense層 : 1

LSTM
使用最佳參數結果

Best Parameters: batch_size: 2, epochs: 20

Test MAE: 14.481

Test RMSE: 24.219

Test MAPE: 24.82%

花費時間: 2:18:22.232444

統整實驗結果 - LSTM

神經元個數	LSTM 使用最佳參數結果
LSTM層：64 Dense層：1	Best Parameters: batch_size: 32, epochs: 30 Test MAE: 14.809 Test RMSE: 25.118 Test MAPE: 24.56% 花費時間: 2:05:54.426071
LSTM層：64 第一層隱藏層：128 Dense層：1	Best Parameters: batch_size: 64, epochs: 20 Test MAE: 14.695 Test RMSE: 24.793 Test MAPE: 24.57% 花費時間: 2:55:06.608498
LSTM層：64 第一層隱藏層：512 第二層隱藏層：128 Dense層：1	Best Parameters: batch_size: 2, epochs: 20 Test MAE: 14.481 Test RMSE: 24.219 Test MAPE: 24.82% 花費時間: 2:18:22.232444

實驗 - 雙向LSTM

神經元個數

LSTM層：64
Dense層：1

雙向LSTM
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 14.809
Test RMSE: 25.118
Test MAPE: 24.56%
花費時間: 2:05:54.426071

實驗 - 雙向LSTM

神經元個數

LSTM層：64

第一層隱藏層：128

Dense層：1

雙向LSTM
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 15

Test MAE: 14.646

Test RMSE: 24.752

Test MAPE: 24.91%

花費時間: 2:26:27.707380

實驗 - 雙向LSTM

神經元個數

LSTM層：64

第一層隱藏層：512

第二層隱藏層：128

Dense層：1

雙向LSTM
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 10

Test MAE: 14.761

Test RMSE: 24.863

Test MAPE: 24.65%

花費時間: 2:09:58.299012

統整實驗結果 - 雙向 LSTM

神經元個數	雙向LSTM 使用最佳參數結果
LSTM層：64 Dense層：1	Best Parameters: batch_size: 32, epochs: 30 Test MAE: 14.809 Test RMSE: 25.118 Test MAPE: 24.56% 花費時間: 2:05:54.426071
LSTM層：64 第一層隱藏層：128 Dense層：1	Best Parameters: batch_size: 8, epochs: 15 Test MAE: 14.646 Test RMSE: 24.752 Test MAPE: 24.91% 花費時間: 2:26:27.707380
LSTM層：64 第一層隱藏層：512 第二層隱藏層：128 Dense層：1	Best Parameters: batch_size: 8, epochs: 10 Test MAE: 14.761 Test RMSE: 24.863 Test MAPE: 24.65% 花費時間: 2:09:58.299012

實驗 - 堆疊LSTM

神經元個數

LSTM層：64
Dense層：1

堆疊LSTM
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 14.615
Test RMSE: 24.822
Test MAPE: 24.22%
花費時間: 2:26:26.594409

實驗 - 堆疊LSTM

神經元個數

LSTM層：64

第一層隱藏層：128

Dense層：1

堆疊LSTM
使用最佳參數結果

Best Parameters: batch_size: 64, epochs: 30

Test MAE: 14.426

Test RMSE: 24.531

Test MAPE: 24.07%

花費時間: 2:0:43.029763

實驗 - 堆疊LSTM

神經元個數

LSTM層：64

第一層隱藏層：512

第二層隱藏層：128

Dense層：1

堆疊LSTM
使用最佳參數結果

Best Parameters: batch_size: 128, epochs: 40

Test MAE: 14.533

Test RMSE: 24.784

Test MAPE: 24.35%

花費時間: 2:41:54.132918

統整實驗結果 - 堆疊LSTM

神經元個數

LSTM層：64
Dense層：1

LSTM層：64
第一層隱藏層：128
Dense層：1

LSTM層：64
第一層隱藏層：512
第二層隱藏層：128
Dense層：1

堆疊LSTM 使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 14.615
Test RMSE: 24.822
Test MAPE: 24.22%
花費時間: 2:26:26.594409

Best Parameters: batch_size: 64, epochs: 30
Test MAE: 14.426
Test RMSE: 24.531
Test MAPE: 24.07%
花費時間: 2:0:43.029763

Best Parameters: batch_size: 128, epochs: 40
Test MAE: 14.533
Test RMSE: 24.784
Test MAPE: 24.35%
花費時間: 2:41:54.132918

實驗 - 雙向、堆疊LSTM

神經元個數

LSTM層：64
Dense層：1

堆疊LSTM
使用最佳參數結果

Best Parameters: batch_size: 64, epochs: 50
Test MAE: 14.614
Test RMSE: 24.770
Test MAPE: 24.30%
花費時間: 2:18:19.471488

實驗 - 雙向、堆疊LSTM

神經元個數

LSTM層：64

第一層隱藏層：128

Dense層：1

堆疊LSTM
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 10

Test MAE: 14.684

Test RMSE: 24.771

Test MAPE: 24.10%

花費時間: 2:45:51.873666

實驗 - 雙向、堆疊LSTM

神經元個數

LSTM層：64

第一層隱藏層：512

第二層隱藏層：128

Dense層：1

堆疊LSTM
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 15

Test MAE: 14.657

Test RMSE: 24.750

Test MAPE: 26.43%

花費時間: 3:35:37.109759

統整實驗結果 - 雙向、堆疊LSTM

神經元個數

LSTM層：64
Dense層：1

LSTM層：64
第一層隱藏層：128
Dense層：1

LSTM層：64
第一層隱藏層：512
第二層隱藏層：128
Dense層：1

雙向、堆疊LSTM
使用最佳參數結果

Best Parameters: batch_size: 64, epochs: 50
Test MAE: 14.614
Test RMSE: 24.770
Test MAPE: 24.30%
花費時間: 2:18:19.471488

Best Parameters: batch_size: 32, epochs: 10
Test MAE: 14.684
Test RMSE: 24.771
Test MAPE: 24.10%
花費時間: 2:45:51.873666

Best Parameters: batch_size: 8, epochs: 15
Test MAE: 14.657
Test RMSE: 24.750
Test MAPE: 26.43%
花費時間: 3:35:37.109759

實驗 - LSTM小結

	LSTM	LSTM	雙向LSTM	雙向LSTM	堆疊LSTM	雙向、堆疊 RNN	雙向、堆疊 RNN
隱藏層層數	64	64 (512, 128)	64	64(128)	64(128)	64(128)	64 (512, 128)
parameter	batch_size: 32, epochs: 30	batch_size: 2, epochs: 20	batch_size: 32, epochs: 30	batch_size: 8, epochs: 15	batch_size: 64, epochs: 30	batch_size: 32, epochs:30	batch_size: 64, epochs:20
MAE	14.809	14.481	14.809	14.646	14.426	17.185	17.061
RMSE	25.118	24.219	25.118	24.752	24.531	28.936	28.774
MAPE	24.56%	24.82%	24.56%	24.91%	24.07%	21.779%	28.48%

實驗 - LSTM小結

	LSTM	堆疊LSTM	雙向、堆疊LSTM	雙向、堆疊LSTM	雙向、堆疊LSTM
隱藏層層數	64 (512, 128)	64(128)	64	64(128)	64(512, 128)
parameter	batch_size: 2, epochs: 20	batch_size: 64, epochs: 30	batch_size:64, epochs:50	batch_size:32, epochs:10	batch_size:8, epochs:15
MAE	14.481	14.426	14.614	14.684	14.657
RMSE	24.219	24.531	24.770	24.771	24.750
MAPE	24.82%	24.07%	24.30%	24.10%	26.43%



RNN-實驗結果

S1161113 陳毓欣



實驗 - RNN

神經元個數

RNN層 : 64

RNN
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 50

Test MAE: 13.020

Test RMSE: 26.634

Test MAPE: 29.85%

花費時間: 0:51:05.021346

實驗 - RNN

神經元個數

RNN層 : 64

第一層隱藏層 : 128

RNN
使用最佳參數結果

Best Parameters: batch_size: 16, epochs: 50

Test MAE: 14.573

Test RMSE: 25.089

Test MAPE: 24.83%

花費時間: 0:55:05.021346

實驗 - RNN

神經元個數

RNN層 : 64

第一層隱藏層 : 512

第二層隱藏層 : 128

RNN
使用最佳參數結果

Best Parameters: batch_size: 64, epochs: 20

Test MAE: 15.072

Test RMSE: 26.239

Test MAPE: 25.20%

花費時間: 0:44:20.123702

統整實驗結果 - RNN

神經元個數

RNN層：64

RNN層：64
第一層隱藏層：128

RNN層：64
第一層隱藏層：512
第二層隱藏層：128

RNN
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 50
Test MAE: 13.020
Test RMSE: 26.634
Test MAPE: 29.85%
花費時間: 0:51:05.021346

Best Parameters: batch_size: 16, epochs: 50
Test MAE: 14.573
Test RMSE: 25.089
Test MAPE: 24.83%
花費時間: 0:55:05.021346

Best Parameters: batch_size: 64, epochs: 20
Test MAE: 15.072
Test RMSE: 26.239
Test MAPE: 25.20%
花費時間: 0:44:20.123702

實驗 - 雙向RNN

神經元個數

RNN層：64

雙向RNN
使用最佳參數結果

Best Parameters: batch_size: 4, epochs: 50

Test MAE: 14.676

Test RMSE: 24.818

Test MAPE: 27.13%

花費時間: 01:50:14.831198

實驗 - 雙向RNN

神經元個數

RNN層：64

第一層隱藏層：128

雙向RNN
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 20

Test MAE: 14.620

Test RMSE: 24.972

Test MAPE: 24.31%

花費時間: 02:14:26.333333

實驗 - 雙向RNN

神經元個數

RNN層：64

第一層隱藏層：512

第二層隱藏層：128

雙向RNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 40

Test MAE: 14.837

Test RMSE: 25.237

Test MAPE: 25.08%

花費時間: 02:01:44.864200

統整實驗結果 - 雙向RNN

神經元個數

雙向RNN
使用最佳參數結果

RNN層：64

Best Parameters: batch_size: 4, epochs: 50
Test MAE: 14.676
Test RMSE: 24.818
Test MAPE: 27.13%
花費時間: 01:50:14.831198

RNN層：64
第一層隱藏層：128

Best Parameters: batch_size: 8, epochs: 20
Test MAE: 14.620
Test RMSE: 24.972
Test MAPE: 24.31%
花費時間: 02:14:26.333333

RNN層：64
第一層隱藏層：512
第二層隱藏層：128

Best Parameters: batch_size: 32, epochs: 40
Test MAE: 14.837
Test RMSE: 25.237
Test MAPE: 25.08%
花費時間: 02:01:44.864200

實驗 - 堆疊RNN

神經元個數

RNN層：64

Dense層：1

堆疊RNN
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 5

Test MAE: 14.268

Test RMSE: 24.482

Test MAPE: 25.07%

花費時間: 01:38:26.786831

實驗 - 堆疊RNN

神經元個數

RNN層 : 64

第一層隱藏層 : 128

Dense層 : 1

堆疊RNN
使用最佳參數結果

Best Parameters: batch_size: 4, epochs: 50

Test MAE: 14.605

Test RMSE: 24.740

Test MAPE: 25.36%

花費時間: 01:23:47.104550

實驗 - 堆疊RNN

神經元個數

RNN層 : 64

第一層隱藏層 : 512

第二層隱藏層 : 128

Dense層 : 1

堆疊RNN
使用最佳參數結果

Best Parameters: batch_size: 64, epochs: 20

Test MAE: 14.830

Test RMSE: 26.418

Test MAPE: 25.32%

花費時間: 01:07:41.577229

統整實驗結果 - 堆疊RNN

神經元個數

RNN層 : 64

Dense層 : 1

RNN層 : 64

第一層隱藏層 : 128

Dense層 : 1

RNN層 : 64

第一層隱藏層 : 512

第二層隱藏層 : 128

Dense層 : 1

堆疊RNN 使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 5

Test MAE: 14.268

Test RMSE: 24.482

Test MAPE: 25.07%

花費時間: 01:38:26.786831

Best Parameters: batch_size: 4, epochs: 50

Test MAE: 14.605

Test RMSE: 24.740

Test MAPE: 25.36%

花費時間: 01:23:47.104550

Best Parameters: batch_size: 64, epochs: 20

Test MAE: 14.830

Test RMSE: 26.418

Test MAPE: 25.32%

花費時間: 01:07:41.577229

實驗 - 雙向、堆疊RNN

神經元個數

RNN層：64

Dense層：1

雙向、堆疊RNN
使用最佳參數結果

Best Parameters: batch_size: 4, epochs: 10

Test MAE: 19.066

Test RMSE: 29.324

Test MAPE: 22.456%

花費時間: 0:54:40.468135

實驗 - 雙向、堆疊RNN

神經元個數

RNN層：64

第一層隱藏層：128

Dense層：1

雙向、堆疊RNN
使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30

Test MAE: 17.185

Test RMSE: 28.936

Test MAPE: 21.779%

花費時間: 1:03:14.864198

實驗 - 雙向、堆疊RNN

神經元個數

RNN層 : 64

第一層隱藏層 : 512

第二層隱藏層 : 128

Dense層 : 1

雙向、堆疊RNN
使用最佳參數結果

Best Parameters: batch_size: 64, epochs: 20

Test MAE: 17.061

Test RMSE: 28.774

Test MAPE: 28.48%

花費時間: 1:05:04.333198

統整實驗結果 - 雙向、堆疊RNN

神經元個數

RNN層 : 64

Dense層 : 1

RNN層 : 64

第一層隱藏層 : 128

Dense層 : 1

RNN層 : 64

第一層隱藏層 : 512

第二層隱藏層 : 128

Dense層 : 1

雙向、堆疊RNN 使用最佳參數結果

Best Parameters: batch_size: 4, epochs: 10

Test MAE: 19.066

Test RMSE: 29.324

Test MAPE: 22.456%

花費時間: 0:54:40.468135

Best Parameters: batch_size: 32, epochs: 30

Test MAE: 17.185

Test RMSE: 28.936

Test MAPE: 21.779%

花費時間: 1:03:14.864198

Best Parameters: batch_size: 64, epochs: 20

Test MAE: 17.061

Test RMSE: 28.774

Test MAPE: 28.48%

花費時間: 1:05:04.333198

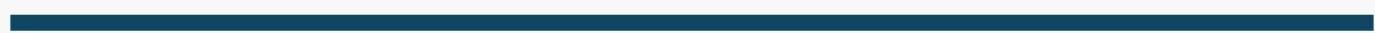
實驗 - RNN小結

	RNN	RNN	雙向RNN	雙向RNN	堆疊RNN	雙向、堆疊 RNN	雙向、堆疊 RNN
隱藏層層數	64	4 (128)	64	64(128)	64	64(128)	64(512, 128)
parameter	batch_size: 8, epochs: 50	batch_size: 16, epochs: 50	batch_size: 4, epochs: 50	batch_size: 8, epochs: 20	batch_size: 8, epochs: 5	batch_size: 32, epochs:30	batch_size: 64, epochs:20
MAE	13.020	14.573	14.676	14.620	14.268	17.185	17.061
RMSE	26.634	25.089	24.818	24.972	24.482	28.936	28.774
MAPE	29.85%	24.83%	27.13%	24.31%	25.07%	21.779%	28.48%



Trnasformer- 實驗結果

S1161113 陳毓欣



實驗 - Transformer

神經元個數

Transformer
使用最佳參數結果

Transformer
Dense層：1

Best Parameters: batch_size: 2
dropout_rate: 0.01 epochs: 1 ff_dim: 512
key_dim: 96 num_heads: 4
optimizer: 'rmsprop'
MAE: 20.406
RMSE: 33.442
MAPE: 29.77%
花費時間：01:26:15.842634

實驗 - Transformer

神經元個數

Transformer
使用最佳參數結果

Transformer
第一層隱藏層：128
Dense層：1

Best Parameters: batch_size: 2
dropout_rate: 0.01 epochs: 1 ff_dim: 128
key_dim: 96 num_heads: 4
optimizer: 'rmsprop'
MAE: 22.825
RMSE: 36.753
MAPE: 31.05%
花費時間：02:11:53.463597

實驗 - Transformer

神經元個數

Transformer

第一層隱藏層：512

第二層隱藏層：128

Dense層：1

Transformer
使用最佳參數結果

Best Parameters: batch_size: 2
dropout_rate: 0.01 epochs: 1 ff_dim: 8
key_dim: 128 num_heads: 16
optimizer: 'rmsprop'
MAE: 21.774
RMSE: 35.307
MAPE: 30.04%
花費時間 : 02:48:46.465513

統整實驗結果 - Transformer

神經元個數	Transformer 使用最佳參數結果
Transformer Dense層：1	MAE: 20.406 RMSE: 33.442 MAPE: 29.77% 花費時間：01:26:15.842634
Transformer 第一層隱藏層：128 Dense層：1	MAE: 22.825 RMSE: 36.753 MAPE: 31.05% 花費時間：02:11:53.463597
Transformer 第一層隱藏層：512 第二層隱藏層：128 Dense層：1	MAE: 21.774 RMSE: 35.307 MAPE: 30.04% 花費時間：02:48:46.465513

實驗 - Transformer+堆疊

神經元個數

Transformer+堆疊
使用最佳參數結果

Transformer+堆疊
Dense層：1

Best Parameters: batch_size: 2
dropout_rate: 0.01 epochs: 1 ff_dim: 32
key_dim: 160 num_heads: 16
optimizer: 'rmsprop'
MAE: 21.667
RMSE: 36.881
MAPE: 30.70%
花費時間：01:32:26.661354

實驗 - Transformer+堆疊

神經元個數

Transformer+堆疊
第一層隱藏層：128
Dense層：1

Transformer+堆疊
使用最佳參數結果

Best Parameters: batch_size: 2
dropout_rate: 0.01 epochs: 1 ff_dim: 8
key_dim: 64 num_heads: 2
optimizer: 'adam'
MAE: 19.936
RMSE: 35.586
MAPE: 30.39%
花費時間：02:21:26.434789

實驗 - Transformer+堆疊

神經元個數

Transformer+堆疊
第一層隱藏層：512
第二層隱藏層：128
Dense層：1

Transformer+堆疊
使用最佳參數結果

Best Parameters: batch_size: 1
dropout_rate: 0.05 epochs: 2 ff_dim: 8
key_dim: 160 num_heads: 4
optimizer: 'rmsprop'
MAE: 26.603
RMSE: 42.331
MAPE: 33.55%
花費時間：02:56:45.848053

統整實驗結果 - Transformer+堆疊

神經元個數

Transformer+堆疊
使用最佳參數結果

Transformer+堆疊
Dense層：1

MAE: 21.667
RMSE: 36.881
MAPE: 30.70%

花費時間：01:32:26.661354

Transformer+堆疊
第一層隱藏層：128
Dense層：1

MAE: 19.936
RMSE: 35.586
MAPE: 30.39%

花費時間：02:21:26.434789

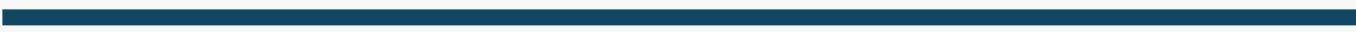
Transformer+堆疊
第一層隱藏層：512
第二層隱藏層：128
Dense層：1

MAE: 26.603
RMSE: 42.331
MAPE: 33.55%

花費時間：02:56:45.848053

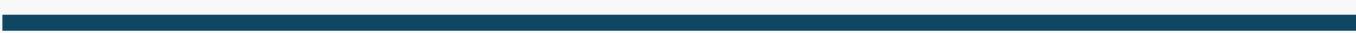
實驗 - Transformer小結

	Transformer	Transformer+堆疊
隱藏層層數	Transformer Dense層 : 1	Transformer 第一層隱藏層 : 128 Dense層 : 1
MAE	20.406	19.936
RMSE	33.442	35.586
MAPE	29.77%	30.39%



門控循環單元-實驗結果

S1161129 施佩茹



實驗 - GRU

神經元個數

GRU層：256
第一層隱藏層：16
第二層隱藏層：2
Dense層：1
StandardScaler正規化

GRU
未使用最佳參數結果

Test MAE: 14.369
Test RMSE: 23.639
Test MAPE: 27.89%

GRU層：256
第一層隱藏層：16
第二層隱藏層：2
Dense層：1

QuantileTransformer(Normal)正規化

Test MAE: 14.209
Test RMSE: 24.367
Test MAPE: 23.94%

實驗 - GRU

神經元個數

GRU層：256
第一層隱藏層：64
第二層隱藏層：2
Dense層：1
StandardScaler正規化

GRU
未使用最佳參數結果

Test MAE: 14.358
Test RMSE: 24.745
Test MAPE: 24.00%

GRU層：256
第一層隱藏層：64
第二層隱藏層：2
Dense層：1

QuantileTransformer(Normal)正規化

Test MAE: 14.413
Test RMSE: 24.424
Test MAPE: 24.15%

統整實驗結果 - GRU

神經元個數

GRU層：256
第一層隱藏層：16
第二層隱藏層：2 Dense層：1
StandardScaler正規化

GRU層：256
第一層隱藏層：16
第二層隱藏層：2 Dense層：1
QuantileTransformer(Normal)正規化

GRU層：256
第一層隱藏層：64
第二層隱藏層：2 Dense層：1
StandardScaler正規化

GRU層：256
第一層隱藏層：64
第二層隱藏層：2 Dense層：1
QuantileTransformer(Normal)正規化

GRU

使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 14.513
Test RMSE: 24.820
Test MAPE: 25.32%

Best Parameters: batch_size: 64, epochs: 30
Test MAE: 14.318
Test RMSE: 24.580
Test MAPE: 25.27%

Best Parameters: batch_size: 8, epochs: 15
Test MAE: 14.927
Test RMSE: 25.632
Test MAPE: 23.89%

Best Parameters: batch_size: 64, epochs: 30
Test MAE: 13.896
Test RMSE: 23.832
Test MAPE: 23.20%

實驗 - 雙向GRU

神經元個數

GRU層：4
第一層隱藏層：64
第二層隱藏層：2
Dense層：1
StandardScaler正規化

雙向GRU
未使用最佳參數結果

Test MAE: 14.358
Test RMSE: 23.907
Test MAPE: 25.83%

GRU層：4
第一層隱藏層：64
第二層隱藏層：2
Dense層：1

QuantileTransformer(normal)正規化

Test MAE: 14.637
Test RMSE: 24.969
Test MAPE: 23.93%

實驗 - 雙向GRU

神經元個數

GRU層：4
第一層隱藏層：256
第二層隱藏層：256
Dense層：1
StandardScaler正規化

雙向GRU
未使用最佳參數結果

Test MAE: 14.072
Test RMSE: 24.
Test MAPE: 25.00%

實驗 - 雙向GRU

神經元個數

GRU層：256
第一層隱藏層：256
第二層隱藏層：64
Dense層：1
StandardScaler正規化

雙向GRU
未使用最佳參數結果

Test MAE: 14.499
Test RMSE: 25.041
Test MAPE: 24.08%

GRU層：256
第一層隱藏層：256
第二層隱藏層：64
Dense層：1
MinMaxScaler正規化

Test MAE: 15.027
Test RMSE: 25.515
Test MAPE: 23.82%

統整實驗結果 - 雙向GRU

神經元個數	雙向GRU 使用最佳參數結果
GRU層：4 第一層隱藏層：64 第二層隱藏層：2 Dense層：1 StandardScaler正規化	Best Parameters: batch_size: 64, epochs: 5 Test MAE: 14.358 Test RMSE: 23.907 Test MAPE: 25.83%
GRU層：4 第一層隱藏層：64 第二層隱藏層：2 Dense層：1 QuantileTransformer(normal)正規化	Best Parameters: batch_size: 64, epochs: 30 Test MAE: 14.637 Test RMSE: 24.969 Test MAPE: 23.93%
GRU層：4 第一層隱藏層：256 第二層隱藏層：256 Dense層：1 StandardScaler正規化	Best Parameters: batch_size: 32, epochs: 30 Test MAE: 14.872 Test RMSE: 25.824 Test MAPE: 25.00%
GRU層：256 第一層隱藏層：256 第二層隱藏層：64 Dense層：1 StandardScaler正規化	Best Parameters: batch_size: 64, epochs: 20 Test MAE: 14.499 Test RMSE: 25.041 Test MAPE: 24.08%

統整實驗結果 - 雙向GRU

神經元個數

雙向GRU
使用最佳參數結果

GRU層：4
第一層隱藏層：64
第二層隱藏層：2
Dense層：1

QuantileTransformer(normal)正規化

Best Parameters: batch_size: 64, epochs: 30
Test MAE: 14.637
Test RMSE: 24.969
Test MAPE: 23.93%

GRU層：256
第一層隱藏層：256
第二層隱藏層：64
Dense層：1

MinMaxScaler正規化

Best Parameters: batch_size: 32, epochs: 20
Test MAE: 14.945
Test RMSE: 25.168
Test MAPE: 24.77%

實驗 - 堆疊GRU

神經元個數

GRU層：32
第一層隱藏層：2
第二層隱藏層：2
Dense層：1
StandardScaler正規化

堆疊GRU
未使用最佳參數結果

Test MAE: 14.391
Test RMSE: 23.918
Test MAPE: 25.03%

GRU層：32
第一層隱藏層：2
第二層隱藏層：2
Dense層：1

QuantileTransformer(normal)正規化

Test MAE: 14.851
Test RMSE: 25.350
Test MAPE: 24.00%

實驗 - 堆疊GRU

神經元個數

GRU層：32
第一層隱藏層：128
第二層隱藏層：8
Dense層：1
StandardScaler正規化

堆疊GRU
未使用最佳參數結果

Test MAE: 14.356
Test RMSE: 24.318
Test MAPE: 23.86%

GRU層：32
第一層隱藏層：128
第二層隱藏層：8
Dense層：1
QuantileTransformer(normal)正規化

Test MAE: 14.497
Test RMSE: 24.789
Test MAPE: 23.71%

實驗 - 堆疊GRU

神經元個數

GRU層：128
第一層隱藏層：4
第二層隱藏層：2
Dense層：1
StandardScaler正規化

GRU層：128
第一層隱藏層：4
第二層隱藏層：2
Dense層：1

QuantileTransformer(normal)正規化

堆疊GRU
未使用最佳參數結果

Test MAE: 15.293
Test RMSE: 24.109
Test MAPE: 33.34%

Test MAE: 14.546
Test RMSE: 24.962
Test MAPE: 24.43%

實驗 - 堆疊GRU

神經元個數

GRU層：128
第一層隱藏層：4
第二層隱藏層：8
Dense層：1
StandardScaler正規化

GRU層：128
第一層隱藏層：4
第二層隱藏層：8
Dense層：1

QuantileTransformer(normal)正規化

堆疊GRU
未使用最佳參數結果

Test MAE: 14.421
Test RMSE: 24.538
Test MAPE: 24.32%

Test MAE: 14.979
Test RMSE: 25.244
Test MAPE: 23.83%

統整實驗結果 - 堆疊GRU

神經元個數

GRU層：32 第一層隱藏層：2
第二層隱藏層：2 Dense層：1
StandardScaler正規化

GRU層：32 第一層隱藏層：2
第二層隱藏層：2 Dense層：1
QuantileTransformer(normal)正規化

GRU層：32 第一層隱藏層：128
第二層隱藏層：8 Dense層：1
StandardScaler正規化

GRU層：32 第一層隱藏層：128
第二層隱藏層：8 Dense層：1
QuantileTransformer(normal)正規化

雙向GRU 使用最佳參數結果

Best Parameters: batch_size: 64, epochs: 30
Test MAE: 14.300
Test RMSE: 24.508
Test MAPE: 25.25%

Best Parameters: batch_size: 16, epochs: 20
Test MAE: 14.690
Test RMSE: 25.364
Test MAPE: 23.53%

Best Parameters: batch_size: 64, epochs: 30
Test MAE: 14.557
Test RMSE: 24.663
Test MAPE: 25.06%

Best Parameters: batch_size: 16, epochs: 20
Test MAE: 14.962
Test RMSE: 25.360
Test MAPE: 23.85%

統整實驗結果 - 堆疊GRU

神經元個數

GRU層：128 第一層隱藏層：4
第二層隱藏層：2 Dense層：1
StandardScaler正規化

GRU層：128 第一層隱藏層：4
第二層隱藏層：2 Dense層：1
QuantileTransformer(normal)正規化

GRU層：128 第一層隱藏層：4
第二層隱藏層：8 Dense層：1
StandardScaler正規化

GRU層：128 第一層隱藏層：4
第二層隱藏層：8 Dense層：1
QuantileTransformer(normal)正規化

雙向GRU 使用最佳參數結果

Best Parameters: batch_size: 32, epochs: 30
Test MAE: 14.281
Test RMSE: 24.786
Test MAPE: 23.97%

Best Parameters: batch_size: 64, epochs: 15
Test MAE: 14.749
Test RMSE: 25.409
Test MAPE: 23.70%

Best Parameters: batch_size: 64, epochs: 30
Test MAE: 14.068
Test RMSE: 24.809
Test MAPE: 25.03%

Best Parameters: batch_size: 64, epochs: 20
Test MAE: 14.968
Test RMSE: 25.656
Test MAPE: 23.48%

實驗 - 雙向、堆疊GRU

神經元個數

GRU層：4
第一層隱藏層：32
第二層隱藏層：8
Dense層：1
StandardScaler正規化

堆疊GRU
未使用最佳參數結果

Test MAE: 14.393
Test RMSE: 24.099
Test MAPE: 24.09%

GRU層：4
第一層隱藏層：32
第二層隱藏層：8
Dense層：1
MinMaxScaler正規化

Test MAE: 15.279
Test RMSE: 25.339
Test MAPE: 23.88%

實驗 - 雙向、堆疊GRU

神經元個數

GRU層：4
第一層隱藏層：32
第二層隱藏層：32
Dense層：1
StandardScaler正規化

堆疊GRU
未使用最佳參數結果

Test MAE: 14.331
Test RMSE: 24.143
Test MAPE: 26.43%

實驗 - 雙向、堆疊GRU

神經元個數

GRU層：8
第一層隱藏層：8
第二層隱藏層：2
Dense層：1
StandardScaler正規化

堆疊GRU
未使用最佳參數結果

Test MAE: 14.331
Test RMSE: 24.143
Test MAPE: 26.43%

實驗 - 雙向、堆疊GRU

神經元個數

GRU層：8
第一層隱藏層：8
第二層隱藏層：64
Dense層：1
StandardScaler正規化

堆疊GRU
未使用最佳參數結果

Test MAE: 14.540
Test RMSE: 24.192
Test MAPE: 25.46%

GRU層：8
第一層隱藏層：8
第二層隱藏層：64
Dense層：1
MinMaxScaler正規化

Test MAE: 15.088
Test RMSE: 25.219
Test MAPE: 24.21%

統整實驗結果 - 雙向、堆疊GRU

神經元個數

GRU層：4 第一層隱藏層：32
第二層隱藏層：8 Dense層：1
StandardScaler正規化

GRU層：4 第一層隱藏層：32
第二層隱藏層：8 Dense層：1
MinMaxScaler正規化

GRU層：4 第一層隱藏層：32
第二層隱藏層：32 Dense層：1
StandardScaler正規化

雙向GRU
使用最佳參數結果

Best Parameters: batch_size: 64, epochs: 15
Test MAE: 14.739
Test RMSE: 24.939
Test MAPE: 25.23%

Best Parameters: batch_size: 2, epochs: 20
Test MAE: 14.636
Test RMSE: 24.963
Test MAPE: 25.66%

Best Parameters: batch_size: 16, epochs: 30
Test MAE: 14.790
Test RMSE: 25.198
Test MAPE: 24.32%

統整實驗結果 - 雙向、堆疊GRU

神經元個數

GRU層：8 第一層隱藏層：8
第二層隱藏層：2 Dense層：1
StandardScaler正規化

GRU層：8 第一層隱藏層：8
第二層隱藏層：64 Dense層：1
StandardScaler正規化

GRU層：8 第一層隱藏層：8
第二層隱藏層：64 Dense層：1
MinMaxScaler正規化

雙向GRU
使用最佳參數結果

Best Parameters: batch_size: 8, epochs: 20
Test MAE: 14.420
Test RMSE: 24.803
Test MAPE: 24.39%

Best Parameters: batch_size: 64, epochs: 30
Test MAE: 14.430
Test RMSE: 24.677
Test MAPE: 24.48%

Best Parameters: batch_size: 64, epochs: 10
Test MAE: 14.979
Test RMSE: 25.239
Test MAPE: 24.67%

實驗 - GRU小結

	GRU	雙向GRU	堆疊GRU	堆疊GRU	堆疊GRU
隱藏層層數	256 (64, 2)	4 (64, 2)	32 (2, 2)	128 (4, 8)	128 (4, 8)
正規化	QuantileTransformer(normal)	QuantileTransformer(normal)	Standard	Standard	QuantileTransformer(normal)
parameter	batch_size: 64, epochs: 30	batch_size: 64, epochs: 30	batch_size: 64, epochs: 30	batch_size: 64, epochs: 30	batch_size: 64, epochs: 20
MAE	13.896	14.637	14.300	14.068	14.968
RMSE	23.832	24.969	24.508	24.809	25.656
MAPE	23.20%	23.93%	25.25%	25.03%	23.48%
					118

實驗 - GRU小結

	GRU	雙向、堆疊 GRU	雙向、堆疊 GRU	雙向、堆疊 GRU
隱藏層層數	256 (64, 2)	4(32, 32)	8(8, 2)	8(8, 64)
正規化	QuantileTransformer (normal)	MinMaxScaler	StandardScaler	StandardScaler
parameter	batch_size: 64, epochs: 30	batch_size: 16, epochs: 30	batch_size: 8, epochs: 20	batch_size: 16, epochs: 30
MAE	13.896	14.790	14.420	14.430
RMSE	23.832	25.198	24.803	24.677
MAPE	23.20%	24.32%	24.39%	24.48%

比較所有結果 (取各模型最優比較)

模型、隱藏層、神經元個數	數據結果
一層CNN	MAE: 1.345, RMSE: 1.856, MAPE: 8.23%
CNN+LSTM	MAE: 0.7443, RMSE: 0.925, MAPE: 2.75%
CNN+RNN	MAE: 0.855, RMSE: 1.137, MAPE: 2.59%
並聯CNN	MAE: 0.861, RMSE: 1.031, MAPE: 3.87%
LSTM 64	MAE: 14.809, RMSE: 25.118, MAPE: 24.56%
LSTM 64(512, 128)	MAE: 14.481, RMSE: 24.219, MAPE: 24.81 %
雙向LSTM 64	MAE: 14.809, RMSE: 25.118, MAPE: 24.56%
雙向LSTM 64(128)	MAE: 14.646, RMSE: 24.752, MAPE: 24.91%

比較所有結果 (取各模型最優比較)

模型、隱藏層、神經元個數	數據結果
堆疊LSTM 64(128)	MAE:14.426, RMSE:24.531, MAPE: 24.07%
雙向、堆疊LSTM 64	MAE: 14.614, RMSE: 24.770, MAPE: 24.30%
雙向、堆疊LSTM 64(128)	MAE: 14.684, RMSE: 24.771, MAPE: 24.10%
雙向、堆疊LSTM 64(512, 128)	MAE: 14.657, RMSE: 24.750, MAPE: 26.43%
RNN 64	MAE: 13.020, RMSE: 26.634, MAPE: 29.85%
RNN 64(128)	MAE:14.676, RMSE: 24.818, MAPE: 27.13%
雙向RNN 64	MAE:14.620, RMSE: 24.972, MAPE: 24.31%
堆疊RNN 64	MAE: 14.268, RMSE: 24.482, MAPE: 25.07%

比較所有結果 (取各模型最優比較)

模型、隱藏層、神經元個數	數據結果
雙向、堆疊RNN 64(128)	MAE:17.185, RMSE:28.936, MAPE: 21.779%
雙向、堆疊RNN 64(512, 128)	MAE: 17.061, RMSE: 28.774, MAPE: 28.774%
Transformer	MAE: 20.406, RMSE: 33.442, MAPE: 29.77%
Transformer+堆疊 (128)	MAE: 19.936, RMSE: 35.586, MAPE:30.39%
GRU 256(64, 2) QuantileTransformer(Normal)正規化	MAE: 13.896, RMSE: 23.832, MAPE: 23.20%
雙向GRU 4(64, 2)QuantileTransformer(normal)正規化	MAE:14.637, RMSE: 24.969, MAPE: 23.93%
堆疊GRU 32(2, 2)StandardScaler正規化	MAE:14.300, RMSE: 24.508, MAPE: 25.25%
堆疊GRU 128(4, 8)StandardScaler正規化	MAE: 14.068, RMSE: 24.809, MAPE: 25.03%

比較所有結果 (取各模型最優比較)

模型、隱藏層、神經元個數

數據結果

堆疊GRU 128(4, 8)
QuantileTransformer(normal)正規化

MAE:14.968, RMSE: 25.656, MAPE: 23.48%

雙向、堆疊GRU 4(32, 32)
StandardScaler正規化

MAE:14.790, RMSE: 25.198, MAPE: 24.32%

雙向、堆疊GRU 8(8, 2)
StandardScaler正規化

MAE: 14.420, RMSE: 24.803, MAPE: 24.39%

雙向、堆疊GRU 8(8, 64)
StandardScaler正規化

MAE: 14.430, RMSE: 24.667, MAPE:24.48%

比較所有結果 (取各模型最優比較)

機器學習模型

數據結果

Prophet

MAE:58.74, RMSE:79.73

Autoregressive Integrated
Moving Average(ARIMA)

MAE:69.42, RMSE:94.21

SVR

MAE:57.04, RMSE:88.62

回歸樹

MAE:46.27, RMSE:26.05

結論

- 處理複雜時間序列與高非線性數據時，深度學習模型表現優於機器學習模型
- 深度學習模型中以 CNN 為基礎的混合模型（如 CNN+LSTM、CNN+RNN）表現最佳

分工表

S1161010

林品儀

卷積神經網絡、PPT製作

S1161113

陳毓欣

RNN、長短期記憶、
Transformer模型、
文件撰寫

S1161129

施佩茹

門控循環單元、PPT製作



Thank you

