



# TAIWAN COVID-19 DIAGNOSED PREDICT

大數據基礎與實務應用--期末專題

10727211 林彥輝



# TABLE OF CONTENTS

01

• **MOTIVATION**

Abstract &  
Introduction

02

• **DATA ANALYSIS**

Data Introduction &  
Data Analysis

03

• **PROJECT GOALS**

Product Introduction &  
Results &  
Conclusion

# MOTIVATION

隨著疫情升溫，從2019年的冬天到現在，台灣的疫情人數來到了第二波的高峰，政策也從清零轉為共存。隨著每天衛生福利部疾病管制署（CDC）的報導每日確診人數，我們會開始期待每一天的確診人數變化、死亡人數變化。

有鑒於此，想要透過這些數據，搭配新潮流的AI，讓這些AI學習每日確診人數的上升或下降趨勢，進一步預測未來的確診人數以及相關資訊。



# DATA ANALYSIS

COVID-19全球疫情地圖



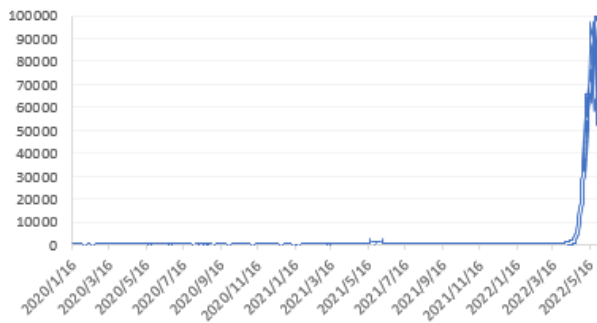
[https://covid-19.nchc.org.tw/dt\\_owl.php?dt\\_name=4&countrycode=TWN](https://covid-19.nchc.org.tw/dt_owl.php?dt_name=4&countrycode=TWN)

# 資料欄位

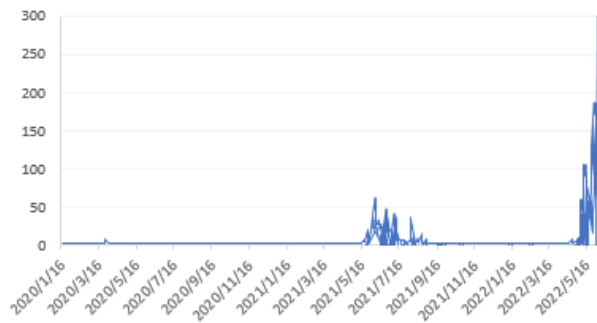
[illegible]

# 連續型資料分析

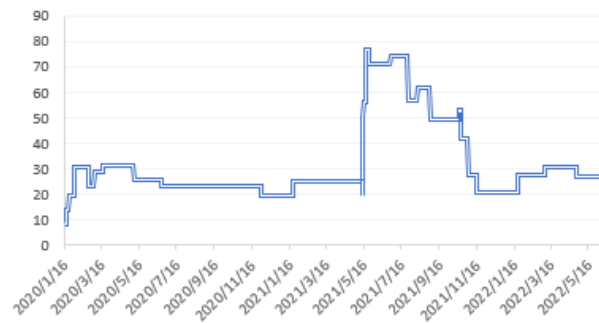
## NEWLY\_DIAGNOSED



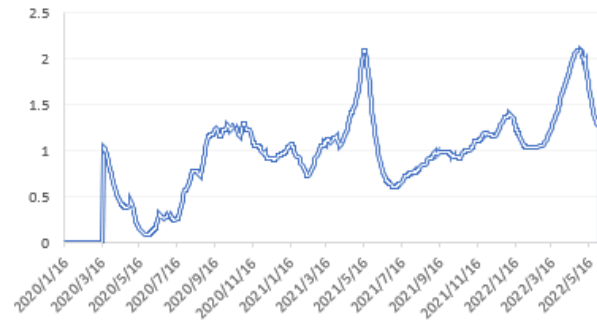
## NEW\_DEATHS



## EPIDEMIC\_CONTROL\_INDEX



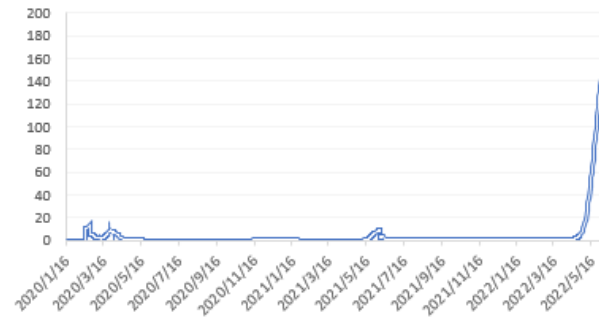
## INFECTION\_RATE



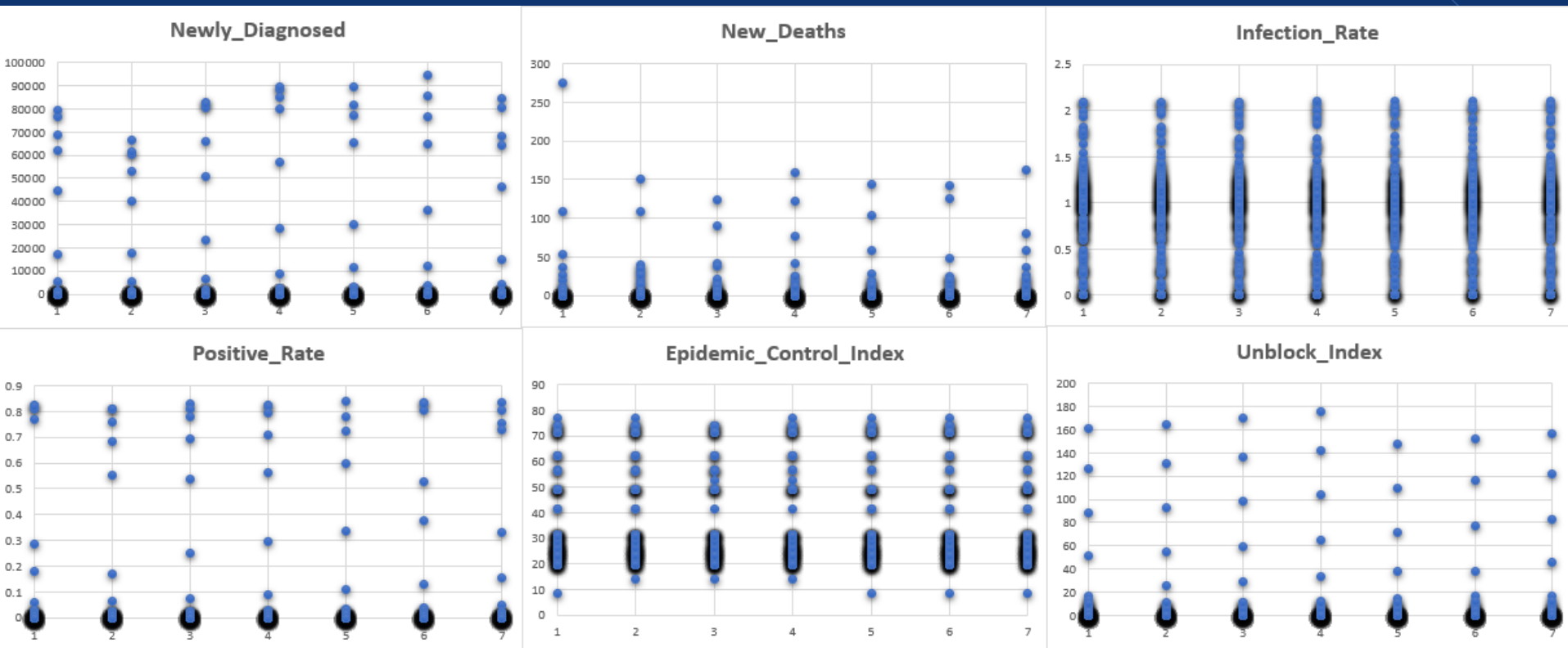
## POSITIVE\_RATE



## UNBLOCK\_INDEX



# 離散型資料分析



# TAIWAN COVID-19 DIAGNOSED PREDICT





# PRE-PROCESSING

## ● 清洗資料 (缺值、值異常)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	
1	id	iso_code	洲名	國家	日期	總確診數	新增確診	七天移動	總死亡數	新增死亡	七天移動	每百萬人	每百萬人	傳染率	新增檢驗	總檢驗	每千人	七天移動	陽性率	每確診案	疫苗總接	疫苗總接	疫苗總接	疫苗總接	七天移動	每百人接	每百人接	疫情控管	總人口數	中位數年
2	159357	TWN	亞洲	台灣	2022/6/9	2693787	72846	70803.7	3584	211	151.857	112923	150.241	0	0	0	0	0	0	0	5.6E+07	2.1E+07	113528	85613	234.98	88.35	0	2.4E+07	42.2	
3	159356	TWN	亞洲	台灣	2022/6/8	2620941	80070	71387.1	3373	159	142.286	109870	141.396	0	1.2E+07	87922	502.577	86535	0.825	1.2	5.6E+07	2.1E+07	110749	100953	234.5	88.22	0	2.4E+07	42.2	
4	159355	TWN	亞洲	台灣	2022/6/7	2540871	82990	72555.4	3214	124	137	106513	134.731	0	1.2E+07	99524	498.892	87667	0.828	1.2	5.6E+07	2.1E+07	104057	108069	234.03	88.03	26.85	2.4E+07	42.2	
5	159354	TWN	亞洲	台灣	2022/6/6	2457881	53010	72218	3090	151	132.143	103034	129.533	0	1.2E+07	105479	494.72	88934	0.812	1.2	5.6E+07	2.1E+07	116694	112921	233.6	87.86	26.85	2.4E+07	42.2	
6	159353	TWN	亞洲	台灣	2022/6/5	2404871	62077	73228.4	2939	276	126.143	100812	123.203	0	1.2E+07	65045	490.298	88623	0.826	1.2	5.6E+07	2.1E+07	0	112926	233.11	87.65	26.85	2.4E+07	42.2	
7	159352	TWN	亞洲	台灣	2022/6/4	2342794	68128	75298.7	2663	0	102.286	98209.7	111.633	0	1.2E+07	77953	487.571	89988	0.837	1.2	0	0	0	117691	0	0	26.85	2.4E+07	42.2	
8	159351	TWN	亞洲	台灣	2022/6/3	2274666	76505	77085.6	2663	142	125.571	95353.8	111.633	0	1.2E+07	79520	484.303	92194	0.836	1.2	5.6E+07	2.1E+07	70556	122456	232.76	87.43	26.85	2.4E+07	42.2	

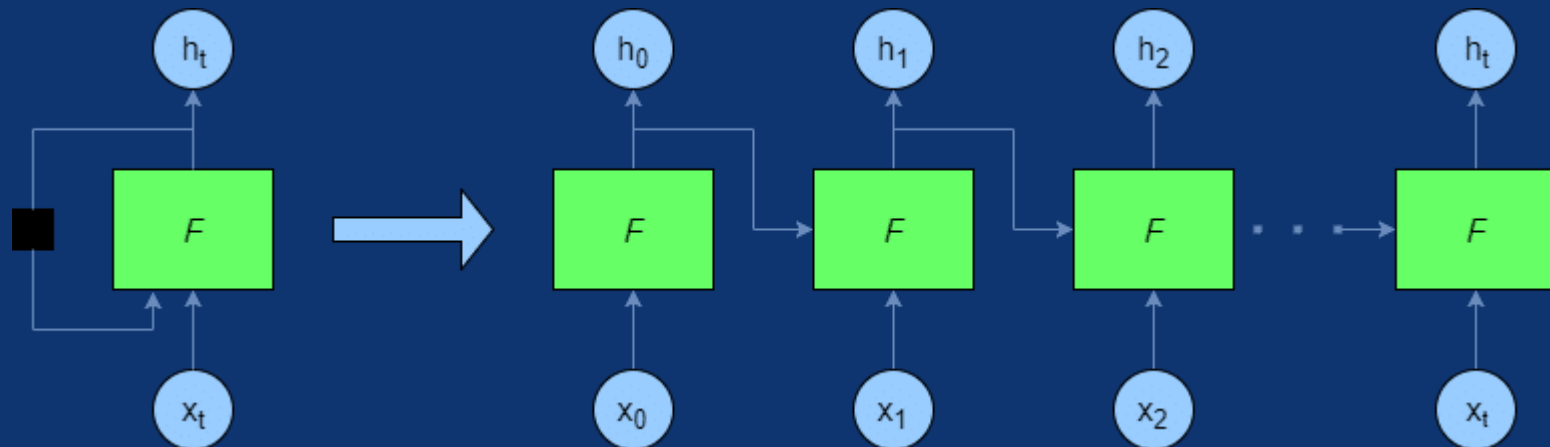
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	
1	id	iso_code	洲名	國家	日期	總確診數	新增確診	七天移動	總死亡數	新增死亡	七天移動	每百萬人	每百萬人	傳染率	新增檢驗	總檢驗	每千人	七天移動	陽性率	每確診案	疫苗總接	疫苗總接	疫苗總接	疫苗總接	七天移動	每百人接	每百人接	疫情控管	總人口數	中位數年
23	159336	TWN	亞洲	台灣	2022/5/19	1070561	89420	71384.4	1235	59	38.143	44877.8	51.771	1.56	1E+07	107367	422.506	98329	0.726	1.4	5.4E+07	2E+07	81127	75623	226.26	84.29	26.85	2.4E+07	42.2	
24	159335	TWN	亞洲	台灣	2022/5/18	981141	85082	67955.1	1176	41	32.143	41129.4	49.298	1.59	9971521	117361	418.005	95907	0.709	1.4	5.4E+07	2E+07	78230	75435	225.92	84.16	26.85	2.4E+07	42.2	
25	159334	TWN	亞洲	台灣	2022/5/17	896059	65802	63962.3	1135	38	27.429	37562.7	47.579	1.63	9854160	109558	413.086	92385	0.692	1.4	5.4E+07	2E+07	78452	75693	225.59	84.05	26.85	2.4E+07	42.2	
26	159333	TWN	亞洲	台灣	2022/5/16	830257	61714	61821.9	1097	29	23.714	34804.3	45.986	1.67	9744602	95813	408.493	90621	0.682	1.5	5.4E+07	2E+07	75661	75130	225.26	83.93	26.85	2.4E+07	42.2	
27	159332	TWN	亞洲	台灣	2022/5/15	768543	68719	71424.9	1068	19	24.571	32217.3	44.77	1.72	9648789	79382	404.476	87351	0.818	1.2	5.4E+07	2E+07	0	75472	224.94	83.84	26.85	2.4E+07	42.2	
28	159331	TWN	亞洲	台灣	2022/5/14	699824	128954	61607.9	1049	81	21.857	29336.6	43.974	1.77	9569407	89041	401.149	84632	0.728	1.4	0	0	0	76312	0	0	26.85	2.4E+07	42.2	
29	159330	TWN	亞洲	台灣	2022/5/13	570870	0	43185.9	968	0	10.286	23930.8	40.578	1.8	9480366	89780	397.416	82103	0.526	1.9	5.4E+07	2E+07	126793	77153	224.57	83.75	26.85	2.4E+07	42.2	
30	159329	TWN	亞洲	台灣	2022/5/12	570870	65415	48352.6	968	17	11.714	23930.8	40.578	1.85	9390586	90416	393.653	80667	0.599	1.7	5.3E+07	2E+07	79806	77151	224.04	83.55	26.85	2.4E+07	42.2	
31	159328	TWN	亞洲	台灣	2022/5/11	505455	57132	43291	951	8	10	21188.6	39.866	1.9	9300170	92702	389.862	77040	0.562	1.8	5.3E+07	2E+07	80039	80068	223.7	83.44	26.85	2.4E+07	42.2	
32	159327	TWN	亞洲	台灣	2022/5/10	448323	50819	39197.3	943	12	9.571	18793.7	39.53	1.95	9207468	97211	385.976	72867	0.538	1.9	5.3E+07	2E+07	74509	86971	223.37	83.36	26.85	2.4E+07	42.2	
33	159326	TWN	亞洲	台灣	2022/5/9	397504	128935	37792.7	931	35	9	16663.3	39.027	2.01	9110257	72923	381.901	68286	0.553	1.8	5.3E+07	2E+07	78060	87600	223.06	83.27	26.85	2.4E+07	42.2	
34	159325	TWN	亞洲	台灣	2022/5/8	268569	0	19373.4	896	0	4	11258.4	37.56	1.99	9037334	60354	378.844	67288	0.288	3.5	5.3E+07	2E+07	0	86998	222.73	83.2	26.85	2.4E+07	42.2	
35	159324	TWN	亞洲	台灣	2022/5/7	268569	0	21812.3	896	0	4.429	11258.4	37.56	2	8976980	71336	376.314	65852	0.331	3	0	0	0	86414	0	0	26.85	2.4E+07	42.2	
36	159323	TWN	亞洲	台灣	2022/5/6	268569	36167	23973.7	896	10	4.857	11258.4	37.56	2.02	8905644	79724	373.324	63691	0.376	2.7	5.3E+07	2E+07	126783	85830	222.31	83.14	26.85	2.4E+07	42.2	
37	159322	TWN	亞洲	台灣	2022/5/5	232402	29984	20565.1	886	5	3.714	9742.27	37.141	2.04	8825920	65032	369.982	61159	0.336	3	5.3E+07	2E+07	100221	85004	221.78	83.04	26.85	2.4E+07	42.2	
38	159321	TWN	亞洲	台灣	2022/5/4	202418	28476	17925.7	881	5	3.286	8485.35	36.931	2.07	8760888	63487	367.256	60098	0.298	3.4	5.3E+07	2E+07	128365	85302	221.36	82.97	26.85	2.4E+07	42.2	
39	159320	TWN	亞洲	台灣	2022/5/3	173942	40987	15131.4	876	8	2.857	7291.64	36.722	2.09	8697401	65146	364.594	60496	0.25	4	5.3E+07	2E+07	78911	81601	220.82	82.9	26.85	2.4E+07	42.2	
40	159319	TWN	亞洲	台灣	2022/5/2	132955	0	10181.3	868	0	1.714	5573.46	36.386	2.08	8632255	65946	361.863	60377	0.169	5.9	5.3E+07	2E+07	73844	84149	220.49	82.84	26.85	2.4E+07	42.2	

## ● 選取欄位 (日期、每日新增人數)

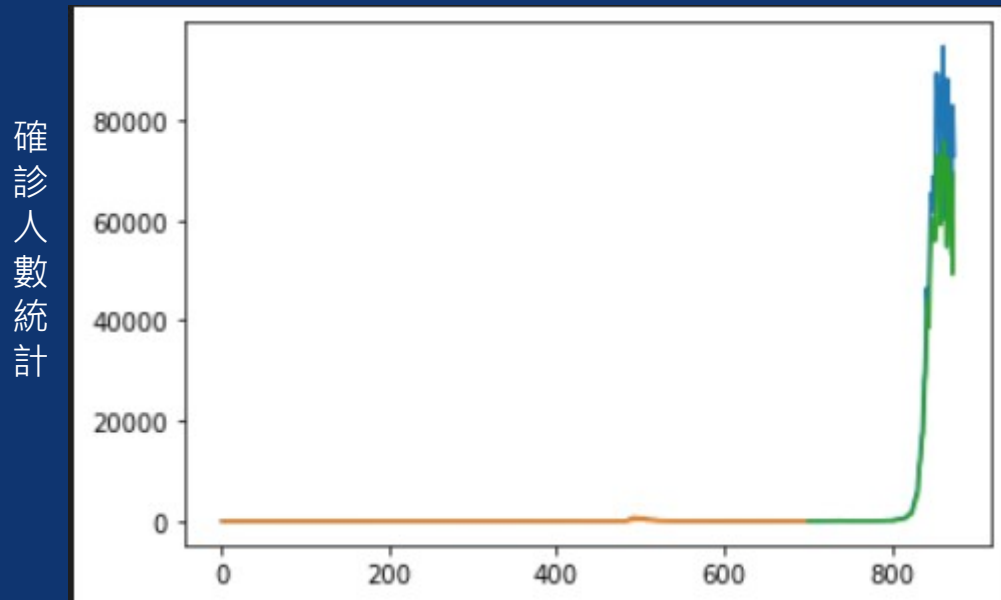


# LSTM

循環神經網路（**Recurrent Neural Network**，**RNN**）是一種用於處理序列數據的神經網路。相比一般的神經網路來說，他能夠處理序列變化的數據。比如某個單詞的意思會因為上文提到的內容不同而有不同的含義，**RNN**就能夠很好地解決這類問題。



# RESULTS



日期(since 2020/01/06)

圖例：  
Training Data  
Testing Data  
Predict Data



# CONCLUSION

- 透過這次的專題，學習到如何針對大筆的數據做分析，搭配政府的公開數據，習得如何針對這些資料做整理並視覺化這些數據。
- 搭配COVID-19的專題，讓我第一次使用了LSTM的模型來分析這些數據，並且從中調整這些機器學習的參數，最終準確率達到相當好的程度。



# REFERENCE

- [https://covid-19.nchc.org.tw/dt\\_owl.php?dt\\_name=4&countrycode=TWN](https://covid-19.nchc.org.tw/dt_owl.php?dt_name=4&countrycode=TWN)
- <https://www.kaggle.com/code/psheon/08-lstm-for-prediction/notebook>
- <https://github.com/mc6666/MyNeuralNetwork>





**THANKS!**

