

國立雲林科技大學工業工程與管理系碩士在職專班

碩士論文

Department of Industrial Engineering and Management

National Yunlin University of Science & Technology

Master Thesis

基於捷運及 YouBike 資料集預測 YouBike 未來借還量

YouBike Rental Predictions

Based on the MRT and YouBike Datasets

蔡聿楹

Yu-Ying Tsai

指導教授：陳奕中 博士

Advisor: Yi-Chung Chen, Ph.D.

中華民國 110 年 6 月

June 2021

# 國立雲林科技大學

## 碩士班學位論文考試委員會審定書

本論文係 蔡幸愷 君在本校 工業工程與管理系 所提論文  
「基於捷運及 YouBike 資料集預測 YouBike 未來借還量」

合於碩士資格水準，業經本委員會評審認可，特此證明。

口 試 委 員 ： 陳奕中

洪鈺欣

黃淵科

陳奕中

指 導 教 授 ： 陳奕中

陳奕中

所 長 ：

蔡幸愷

中 華 民 國 110 年 7 月 14 日

## 摘要

近年來，YouBike 微笑單車系統(以下簡稱 YouBike 系統)已成為台北地區民眾進行短距離移動時的首選，因此其使用人數逐年增加，並為 YouBike 系統的營運商帶來不小的維運壓力。在使用人數增加的情況下，站點缺車或因為滿車而無法還車的情形會更常發生，並造成民眾的不滿。有鑑於此，YouBike 營運商如何更準確的預測每日每個站點的借還車人數，並依此進行合理的腳踏車運送的排程便成為一件非常重要的事情。過往雖然已有很多學者進行 YouBike 系統的借還量預測研究，然而這些研究大多只使用歷史借還量為主進行預測，頂多輔以天氣或雨量等一些參考資料而已。這樣的方式雖然可以準確預估日常狀況的借還量，但遇到突發事件時則通常預測準確率會大幅降低，並造成 YouBike 營運商在後續腳踏車運送排程上的困擾。為了克服上述缺點，我們提出以捷運系統進出站人數預測結果為主的 YouBike 借還量預測方法。因為大部分 YouBike 的使用者皆是把 YouBike 當做來往捷運站最後一哩路的交通工具，所以捷運系統的進出站人數必然與 YouBike 借還量高度相關。更重要的是，因捷運系統的使用人數眾多，其進出站人數會比 YouBike 借還量更早反映出城市中的突發狀況。換言之，上述作法在城市有突發狀況時，理應能比傳統方法對 YouBike 借還量做更準確的預測，而我們的實驗結果也證實了這點。

**關鍵字：**YouBike、MRT、時間序列預測、類神經網路

## ABSTRACT

In recent years, the Taipei Bike Sharing System, or YouBike, has become the top choice for short-distance transit in Greater Taipei. It is seeing an annual increase in users, which has also put a sizable amount of pressure on the operator. This is because an increasing number of users will increase the chance of stations being empty or full, thereby preventing rental or return and causing user dissatisfaction. In this case, being able to more accurately predict the number of rentals per station per day and then appropriately scheduling bike dispatching is thus a crucial matter for YouBike operator. Despite many researchers conducting studies on rental and return predictions for the YouBike system in the past, these studies mostly performed predictions based on historical data, aided by reference data such as weather or rainfall at the most. Although this approach enables accurate predictions for normal daily circumstances, its accuracy declines significantly when there are sudden incidents, which is problematic for the YouBike operator's subsequent bike dispatch scheduling. To overcome this issue, we proposed a YouBike rental and return prediction method based on the numbers of people entering and leaving MRT stations. This is because most YouBike users see the bikes as "last-mile public transport vehicles" to and from MRT stations, so there must be a high correlation between YouBike rentals and the numbers of people entering and leaving MRT stations. More importantly, the MRT system has a large passenger volume, so the numbers of people entering and leaving MRT stations can reflect sudden incidents in the city earlier than YouBike rentals can. In other words, this approach should be able to predict YouBike rentals and returns more accurately than conventional methods can in the event of sudden incidents in the city, and our experiment results confirmed this.

***Keywords: YouBike, MRT, time series prediction, neural network***

## 誌謝

終於來到最後執筆撰寫致謝的這一刻。兩年時間裡，職場與學校的兼顧，台中與雲林的往返，著實並不容易。謝謝雲科大工管系提供良好的學習環境；謝謝工工班同學們營造的學習氛圍，一同相互勉勵、討論課業，有你們作為榜樣，自己才不敢懈怠，得以堅持下來完成學業；謝謝指導教授陳奕中老師在論文上的指導以及口試老師的鼓勵與建議使得論文更加完整。

重返校園的心境與過往學生時期截然不同，歷經職場更能體會學習的快樂，不僅增進學識，也磨礪了心智。這兩年的過程都將成為我人生中的養分與美好回憶。最後謝謝世恩，從最初啟發我職涯上的志向、鼓勵我下定決心進修，有你一路支持與陪伴，才得以在奔三之際完成階段性的目標。願我們始終以一起成為更好的人而努力。

蔡聿楹謹致

中華民國 110 年 6 月

## 目錄

摘要 .....	i
ABSTRACT .....	ii
誌謝 .....	iii
目錄 .....	iv
表目錄 .....	vi
圖目錄 .....	vii
第一章 緒論 .....	1
1.1 研究背景與動機 .....	1
1.2 研究目的 .....	2
1.3 研究範圍與限制 .....	3
1.4 研究流程 .....	4
第二章 文獻探討 .....	6
2.1 公共自行車轉乘分析與特性 .....	6
2.1.1 可接受之步行距離相關研究 .....	6
2.1.2 YouBike 轉乘行為特性分析 .....	7
2.2 公共自行車預測 .....	7
2.3 集群分析 .....	9
2.3.1 集群分析的方法 .....	9
2.3.2 監督式與非監督式 .....	10
2.4 時間序列預測 .....	11
2.4.1 時間序列預測的方法 .....	11
2.4.2 統計模型 .....	12
2.4.3 機器學習 .....	12
2.4.4 淺層類神經網路 .....	12
2.4.5 深度學習 .....	13
第三章 研究方法 .....	14
3.1 資料集介紹 .....	14
3.1.1 捷運資料集 .....	15

3.1.2 YouBike 資料集 .....	15
3.2 臺北捷運資料集集群分析 .....	16
3.2.1 臺北捷運資料前處理 .....	16
3.2.2 輪廓係數法 .....	20
3.2.3 集群分析法-K-means .....	21
3.2.4 依分群結果挑選捷運及 YouBike 站點 .....	22
3.3 基於捷運人流預測 Youbike 使用量 .....	24
3.3.1 MRT 進出站人次預測資料集 .....	24
3.3.2 YouBike 使用人次預測資料集 .....	24
3.3.3 倒傳遞類神經時間序列網路 .....	25
第四章 實驗結果 .....	27
4.1 捷運進出站人次預測 .....	28
4.1.1 捷運進站人次預測結果分析 .....	28
4.1.2 捷運出站人次預測結果分析 .....	31
4.2 YouBike 借還車使用量預測 .....	34
4.2.1 捷運站周邊 YouBike 還車預測結果分析 .....	35
4.2.2 捷運站周邊 YouBike 借車預測結果分析 .....	40
4.2.3 捷運站周邊不同步行距離的 YouBike 使用量預測 .....	45
第五章 結論 .....	53
5.1 研究結論 .....	53
5.2 研究建議 .....	53
參考文獻 .....	55

## 表目錄

表 3.1 臺北捷運每時進出站人次 .....	15
表 3.2 Youbike 開放資料平台原始資料格式說明 .....	16
表 3.3 行政院人事行政處公布之國定假日 .....	17
表 3.4 捷運站點人流標準化 .....	20
表 3.5 集群中心站點分析結果 .....	22
表 3.6 質心站點方圓一公里內 YouBike 站點 .....	24
表 3.7 忠孝復興站-500m 以內 YouBike 借車預測模型結果 .....	25
表 4.1 捷運忠孝復興站進站人次測試結果 .....	29
表 4.2 捷運忠孝復興站進站人次測試結果指標 .....	30
表 4.3 捷運動物園站進站人次測試結果 .....	31
表 4.4 捷運動物園站進站人次測試結果指標 .....	31
表 4.5 捷運忠孝復興站出站人次測試結果 .....	32
表 4.6 捷運忠孝復興站出站人次測試結果指標 .....	33
表 4.7 捷運動物園站出站人次測試結果 .....	34
表 4.8 捷運動物園站出站人次測試結果指標 .....	34
表 4.9 捷運忠孝復興站周邊 YouBike 站點還車量測試結果 .....	36
表 4.10 捷運忠孝復興站周邊 YouBike 站點還車量測試結果指標 .....	38
表 4.11 捷運動物園站周邊 YouBike 站點還車量測試結果 .....	39
表 4.12 捷運動物園站周邊 YouBike 站點還車量測試結果指標 .....	40
表 4.13 捷運忠孝復興站周邊 YouBike 站點借車量測試結果 .....	41
表 4.14 捷運忠孝復興站周邊 YouBike 站點借車量測試結果指標 .....	43
表 4.15 捷運動物園站周邊 YouBike 站點借車量測試結果 .....	44
表 4.16 捷運動物園站周邊 YouBike 站點借車量測試結果指標 .....	45
表 4.17 捷運忠孝復興站周邊 500 公尺外 YouBike 站點還車量測試結果 .....	47
表 4.18 捷運忠孝復興站周邊 500 公尺外 YouBike 站點還車量測試結果指標 ...	49
表 4.19 捷運忠孝復興站周邊 500 公尺外 YouBike 站點借車量測試結果 .....	50
表 4.20 運忠孝復興站周邊 500 公尺外 YouBike 站點還車量測試結果指標 .....	52



## 圖目錄

圖 1.1 研究流程圖 .....	5
圖 3.1 研究方法流程圖 .....	14
圖 3.2 各捷運站人次加總流程 .....	18
圖 3.3 忠孝復興站進站累加人次圖 .....	18
圖 3.4 忠孝復興站出站累加人次圖 .....	18
圖 3.5 臺北捷運進站人流週別累計旅次統計 .....	19
圖 3.6 臺北捷運出站人流週別累計旅次統計 .....	19
圖 3.7 輪廓係數分析圖 .....	21
圖 3.8 忠孝復興站周邊站點地圖 .....	23
圖 3.9 Google Maps 路線規劃 .....	23
圖 3.10 模型架構設定 .....	26
圖 4.1 模型建置架構 .....	27
圖 4.2 資料集分布 .....	28
圖 4.3 誤差值直方圖-捷運忠孝復興站進站 .....	29
圖 4.4 誤差值直方圖-捷運動物園站進站 .....	30
圖 4.5 誤差值直方圖-捷運忠孝復興站出站 .....	32
圖 4.6 誤差值直方圖-捷運動物園站出站 .....	33
圖 4.7 誤差值直方圖-捷運忠孝復興站周邊 YouBike 還車 .....	35
圖 4.8 誤差值直方圖-捷運動物園站周邊 YouBike 還車 .....	39
圖 4.9 誤差值直方圖-捷運忠孝復興站周邊 YouBike 借車 .....	40
圖 4.10 誤差值直方圖-捷運動物園站周邊 YouBike 借車 .....	44
圖 4.11 誤差值直方圖-捷運忠孝復興站周邊 500 公尺外 YouBike 還車 .....	46
圖 4.12 誤差值直方圖-捷運忠孝復興站周邊 500 公尺外 YouBike 借車 .....	49

# 第一章 緒論

## 1.1 研究背景與動機

為環境的永續發展，綠色交通(Green Transport)在全球蔚成風氣，並落實在西班牙巴塞隆納、法國里昂、英國倫敦等許多城市。綠色交通的概念是以對環境較低汙染、適合都市空間的交通運具為主，在其所倡導的運具之中，公共自行車租賃系統在台灣獲得成功普及。臺北市公共自行車租賃系統－「YouBike 微笑單車」(以下簡稱 YouBike)在 2009 年始示範營運，於 2012 年正式啟用，至 2018 年已突破一億使用人次。YouBike 建置的理念係鼓勵民眾使用低污染、低耗能的公共自行車作為短程運具，並且成為大眾運輸系統第一哩與最後一哩的轉乘交通方式，以雙北而言，截至 2021 年 5 月 31 日，臺北市 YouBike1.0 共設有 399 個站點，其中緊鄰捷運站並以捷運站名命名的 YouBike 站點，共計 74 個；新北市 YouBike 共設有 664 個站點，其中緊鄰捷運站並以捷運站名命名的 YouBike 站點，共計 58 個，由此可見 YouBike 站點的選址考量在作為捷運轉乘的立意上是實質貫徹的。

近年來，YouBike 確實逐漸成為台北地區民眾進行短距離移動時的首選，其使用人數逐年增加，在需求日益擴大的情況下，產生在租借車輛與歸還車輛，兩種不同情狀下的供需不平衡，YouBike 租賃站點無車可借或無位可還的情況在部分租賃站點或特定時段尤為明顯，造成民眾的不滿，為 YouBike 系統的營運商帶來不小的維運壓力。即使可透過 YouBike 公司提供的站點即時資訊，知曉各站點現況為正常租借、無車可借或車位滿載，但該即時資訊尚無法發揮 YouBike 公司提前進行車輛調度計畫的作用，也無法確保使用者在查詢當下至抵達站點的時間

差之下，皆有車可借、有位可還。面對此問題，吳冠毅(2018)建立 YouBike 預測系統，以每五分鐘的資訊來預測下一個五分鐘的借還量，此預測結果相比 YouBike 所提供的即時站點資訊，讓民眾可以更準確的推測抵達站點時的車輛數。透過該研究結果，可得知若能將預測時間拉長，有足夠的時間讓 YouBike 公司進行車輛調度計畫，將能夠幫助 YouBike 公司克服在人員及配送貨車等有限資源的情況下，仍能達到完善的車輛調度，減少尖峰時段的車輛運送人力需求，以降低人事成本，同時提升使用者滿意度進而提升租借率。

近幾年有許多對於如何更準確地預測每日每個站點的借還車人數，並依此進行合理的 YouBike 運送排程研究，但有的研究只對單一 YouBike 站點進行預測，雖然方法成果顯示有效，但沒有辦法有效幫助營運方進行整體車輛調度計畫，或是預測時間過短，無法實質在實務上有效地幫助車輛調度的安排評估，且大多只使用歷史借還量為主進行預測或輔以天氣或雨量等一些參考資料。這樣的方式雖然可以準確預估日常狀況的借還量，但在這個狀況下，建立出來的模型會與周圍的真實狀況脫節，無法對應泛化的現實問題，因此一旦遇到突發事件時則通常預測準確率會大幅降低，並造成 YouBike 營運商在後續腳踏車運送排程上的困擾。

## 1.2 研究目的

本研究提出以捷運系統進出站人數預測結果為主的 YouBike 借還量預測方法，係因 YouBike 已被相關研究驗證是臺北捷運使用者作為移動的最初及最後一哩路之運具之一。李舒媛(2018)透過悠遊卡大數據探討 YouBike 租賃及轉乘捷運之使用者行為，分析 2016 年 11 月之捷運資料與 YouBike 資料，得知租借 YouBike 並有轉乘捷運行為者，分別佔臺北市悠遊卡持有者 20.25%、新北市悠遊卡持有者 18.6%；由捷運轉乘 YouBike 者，佔臺北市悠遊卡持有者 25.55%、新北市悠遊卡持

有者24.9%，由此得知捷運系統的進出站人數必然與YouBike借還量具高度相關性。透過本研究概念，我們預期能讓YouBike的借還量預測與YouBike站點周邊的捷運旅運量進行串聯，避免突發性的大量借車或大量還車情況，造成使用者無車可借、無位可還的窘境。黃晏姍(2015)分析公共自行車系統營運特性，可知市中心及交通設施周邊的租賃站不論規模大小，都比較容易有缺車與缺位問題，因此在預測時串聯交通設施的旅運量，而非只以YouBike本身資料進行預測，其預測結果才能有效減少缺車或缺位之情況。更重要的是，因為捷運系統的使用人數眾多，其進出站人數會比YouBike借還量更早反映出城市中的突發狀況。換言之，上述作法在城市有突發狀況時，理應能比傳統方法對YouBike借還量做更準確的預測。

上述概念雖然合理，但仍會在執行上遇到一些困難，例如臺北捷運有108個營運車站，若是每個站點都建一個模型，則執行成本過高。因此我們提出在建模前加入使用k-means來分群的論點，藉由這個論點，我們預期能把捷運旅運量變化相似的捷運站點分在同一個群組，並用以同一個模型進行訓練，能有效降低需要訓練的模型數目與降低模型的計算成本。

### 1.3 研究範圍與限制

本研究取得2017年YouBike借還資料，共計21,315,093筆，並搭配取得2017年臺北捷運公開資料集中的進出站資料，共計89,439,552筆。考慮將YouBike作為捷運轉乘的第一哩與最後一哩之使用者多數為大專院校之學生，因此排除寒暑假，採用2017年3月至2017年5月之原始資料作為本研究探討使用。另捷運營運時間為早上六點至凌晨零點，並無凌晨零點至早上六點之捷運資料，因此剔除YouBike零點至早上六點之資料，不納入探討範圍。同時考量YouBike平日與假日

的使用情況並不相同，通常分別使用平日模型模式與假日模型才能增加準確度，鑒於本研究希冀深入探討的對象為使用狀況較為規律的通勤者，同時也是YouBike的忠實使用者，因此限縮研究範圍為周一至周五，並扣除國定假日，減少異變狀態。

基於模型數目與模型的計算成本，本研究將臺北捷運原始資料中的108個站點進行分群，分別在各群組裡找出一個代表站，接著找出各捷運代表站點半徑500公尺內及500至1000公尺內之YouBike租賃站點，以範圍內周邊的YouBike站建立一個預測模型。

#### 1.4 研究流程

本研究之流程說明如下：

1. 研究之研究動機及目的：藉由研究動機與目的來釐清研究問題，對於YouBike借還量的預測，非單一使用YouBike資料集，而是考量周邊環境及使用者使用情況，加入捷運人流資料，並透過類神經時間序列模型預測未來捷運人流，進而得到未來YouBike借還量的預測結果。
2. 相關文獻探討：分成四個部分進行探討，分別為「公共自行車轉乘分析與特性」、「公共自行車預測」、「集群分析」以及「時間序列預測方法」。
3. 資料蒐集與清洗：將臺北捷運進出站旅次資料及YouBike借還量資料，排除2017年3月至2017年5月以外之數據，並將該區間的原始資料分別彙整切割成每日上半天與下半年，分別統計該時間區間之捷運站進出站人次與YouBike站點之借還車次數。
4. 輪廓係數法：使用輪廓係數法決定分群數之k值。

5. K-means 集群分析法：藉由集群分析法K-means依各站點捷運人流進行分群，挑選出不同類型群體的捷運代表站點，再由捷運周邊半徑距離500公尺以內及500至1000公尺以內的YouBike站點，進行旅運量的分析後取出特徵值。
6. 實測模型：藉由機器學習模型預測Youbike未來借還量，並透過測試資料驗證模型之預估準確性。

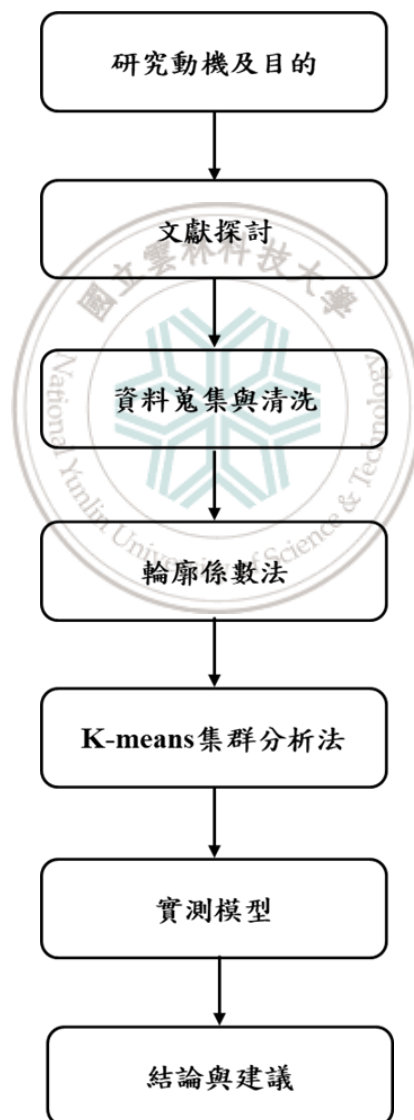


圖 1.1 研究流程圖



## 第二章 文獻探討

本章節介紹與本研究相關之文獻，包含公共自行車轉乘分析與特性、公共自行車預測、集群分析及時間序列預測方法。

### 2.1 公共自行車轉乘分析與特性

為準確預測 YouBike 的借還用量，首先我們勢必得先瞭解使用者的習性與需求，將一般民眾可以接受的步行距離或時間考量進捷運與 YouBike 之間的轉乘可能性，並且進一步探討民眾在 YouBike 轉乘時有哪些行為特性。

#### 2.1.1 可接受之步行距離相關研究

顏志偉、簡文彥、賴宗裕等人(2010)針對臺北都會區大眾捷運系統周邊土地整體開發計畫進行探討，問卷調查得知民眾可接受之步行距離約為500公尺、可接受之步行時間約為10分鐘。臺北市區內捷運站與站之間相間的距離約為1公里，因此在進行周邊土地開發計畫時，以捷運站為中心點，距離車站半徑500公尺為周邊土地開發計畫的範圍。因應臺北市及新北市現行土地使用管制規定有距離車站用地500公尺之相關規定，且車站規劃係以500公尺為步行可及服務半徑，故界定以車站半徑500公尺為基準所劃設之完整街廓為車站周邊土地整體開發之合理影響範圍。沈依潔(2014)以問卷調查臺北市民眾願意以步行方式進行活動或到達目的地之最大時間長度，發現受訪者步行活動的意願時間約為2小時，當中以郊遊踏青之時長最長，平均高達2.41個小時；從事散步運動的時間長度最短，但平均仍有1.3個小時。而將步行作為旅運方式的可接受時間，並不隨目的地不同而有顯著差異，約以10分鐘為上限，又以前往餐飲店與購物商店之步行意願時間為最長；以前往停車場之步行意願時間為最短，而民眾到達公車與大眾捷運系統

之步行時間若低於10分鐘，民眾將有較高的步行意願，由此可驗證臺北市民眾日常旅運中，步行以短程需求為主。綜觀以上我們歸納一般民眾的步行範圍接受度，步行距離約為500公尺為上限、步行時間為十分鐘以內，因此在捷運與YouBike的轉乘關係上，可以推測捷運站周邊距離站點500公尺以外的YouBike站點較無成為捷運轉乘運具的可能性。

### 2.1.2 YouBike 轉乘行為特性分析

葉耀墩、林昭賢、劉心荷(2015)針對 YouBike 騎乘用途之特性進行問卷調查，統計結果顯示民眾對於 YouBike 的使用上有幾點特性：於平日使用 YouBike 者佔 49%、假日使用 YouBike 者佔 39%；在平日的使用上，主要以 YouBike 做為通勤用途佔 37%，假日的使用則做為休閒娛樂為主佔 37%；使用 YouBike 後轉乘的運具，捷運佔 55%為最多，其次為步行佔 26%。上述調查結果顯示 大部分 YouBike 的使用者皆是把 YouBike 當做來往捷運站最後一哩路的交通工具。李舒媛(2018)串聯 2016 年 11 月雙北地區 YouBike 與臺北捷運旅次的悠遊卡紀錄，依據 YouBike 使用者的使用頻率，將 YouBike 使用者類型進行定義，分為單月使用小於 6 次、單月使用次數為 6~20 次、單月使用次數大於 20 次，依序分別為偶爾使用者、經常使用者、忠實使用者，並判斷 YouBike 旅次是否轉乘捷運。研究結果 YouBike 使用資料上顯示，雙北地區皆以偶爾使用者的人數最多，其主要使用時段為假日昏峰；忠實使用者的人數最少，使用時段以平日晨峰、昏峰，及約略的夜峰為主。從區域的差異性來看，臺北市 YouBike 使用量主要聚集於中心商務區、學校與捷運站；新北市 YouBike 使用量則聚集於政經中心、捷運站。

### 2.2 公共自行車預測

翁雅君(2015)以時間序列方法的 ARIMA 預測模組，搭配 YouBike 租賃站點



周遭環境進行分群分析，進而驗證 YouBike 使用狀況會因收費方式改變而影響。

黃政逸(2015)以台北市士林區公共自行車租借系統為例，將實際租借數據導入 Arena 模擬軟體中，結果顯示運補門檻設置高時，租借站容易出現需要等待租借、還車的情形。

李奕譚(2016)透過自我迴歸移動平均整合模式(ARIMA)、三次指數平滑法(Holt-Winters)以及倒傳遞神經網路(BP Neural Network)分別獨立建立公共自行車的預測模式，研究結果顯示，單一預測模型對使用量預測的精度有限，因此考量建立混合預測模型，再通過將模型動態迭代以得出穩定的時間序列模式及天氣擾動之神經網路模型，建構出精確度更高的混合模型。

吳冠毅(2018)利用支援向量回歸(Support Vector Regression, SVR)分別建立平日與假日之 Youbike 車位及車輛數預測模型，再採用粒子群演算法(Particle Swarm Optimization, PSO)尋找參數，提升 SVR 預測模型的準確度。林冠廷(2018)使用資料探勘技術建立預測，探討下一時段站點借車量及還車量之預測，以時間延遲類神經網路、回饋式類神經網路及自迴歸整合移動平均模型，進行三種不同模型的比較。研究結果顯示，使用分群演算法，有效找出相似的借還車模式，再使用時間延遲類神經網路進行預測，能得到較好的預測結果。

張宸睿(2019)以傳統時間序列方法與回饋式類神經網路建立預測模型，透過集群分析法 K-means 將時間序列模式進行分群，再利用小波分解模與天氣資料結合建立預測模型，其研究結果顯示類神經網路模型優於傳統時間序列模型。陳亞伶(2019)以彰化火車站前站 Youbike 為主軸，將租賃站的借車量以模糊時間序列方法預測，再將模糊預測結果以及影響 Youbike 借車量的關鍵天氣因子投入倒傳遞類神經網路做預測，得知資料分群模型大致上皆較未分群結果優異，而加入天

氣因子的倒傳遞類神經網路模型績效最佳。陳昱瑋(2019)透過臺北市內湖區的計程車搭乘量與 YouBike 借用量，觀察數據得到特徵以推斷車輛使用的原因，並使用深度學習模型來計算對車輛使用的影響原因，證實該方法的有效性和合理性。

## 2.3 集群分析

集群分析的方法係依據資料的相似度或相異度，將資料分群歸屬若干個群組。同一群體內的資料具有高度的相似性，而各群體之間則相似度低。

### 2.3.1 集群分析的方法

一般可分為階層式集群方法(Hierarchical cluster methods)和非階層式集群方法(Non-Hierarchical cluster methods)，以及結合以上兩種方法的二階段集群分析法(Two-stage cluster analysis)。

Ward, J. H. (1963)提出之華德法為階層式集群法，其方法先將每一樣本資料視為一集群，再依序將集群合併，合併順序的安排是基於方差分析概念，以離差平方和為相似性測度。同集群樣本之間的離差平方和最小，集群與集群間的離差平方和最大。J. B. MacQueen (1967)提出的 K-means 為非階層式集群方法，事先設定數目為  $k$ ，將各事物點分割成  $k$  個原始集群後，計算某一事物點到各集群重心的歐幾里得距離，接著分派此事物點至距離最近的集群，並重新計算各集群的重心，進行重複計算及分派，直到各事物點都不必重新分派到其他集群為止。Punj、Stewart(1983)歸納階層式集群法與非階層式集群法兩種方法之優劣，得出對 K-平均數法而言，若以隨機方法選取起始點，其集群效果較差，但若以集群之重心或中位數為起始點，則其集群效果要比華德法與平均連鎖法好。

二階段法是 Anderbeg(1973)所提出之概念，結合階層式及切割式演算法這兩

種方法，首先以任意階層式集群方法計算出理想之集群數目來求出  $k$  值，接著使用  $k$  值的數目來執行切割式集群分析。採用階層式集群演算時，樣本一旦合併後，就無法合併至其他群。切割式分群方法則有著  $k$  值數量需主觀決定的缺點，研究者需要經驗才能判斷合適的  $k$  值，因此使用二階段法可以同時避免單獨使用階層式或切割式演算法時之缺點。Sgarma(1996)建議若要求更精確的分類結果，可以同時使用分層法與非分層法。第一階段以華德法做分群以決定群組個數，第二階段再以 K-means 進行集群，二階段方法目的是由於第一階段華德法是屬於分層式集群分析，當二個個體一旦被分在一群，則其後永遠就在同一群，而此時在第二階段所使 K-means 就是彌補此種缺點，以達到最佳組內同質、組間異質的群數。此種分析的手法就是透過兩階段分群方法，修正 T 分層式集群法中對於集群分析一旦形成就無法對併入不適當集群觀察值進行重新分群的缺點，同時也克服了非層級式集群法需事先決定集群數目與集群中心點的問題。

### 2.3.2 監督式與非監督式

集群分析(Clustering)是機器學習當中常見的方法。機器學習演算法可略分為「監督式學習」及「非監督式學習」這兩大類，集群分析即是非監督式學習的代表。由 J. B. MacQueen 於 1967 年正式發表 K-means 演算法，是最早也是最具代表性的集群化計算方式，其相對簡單且有效率，為多數商用資料探勘軟體所採用。其公式與演算法步驟如下：(1) 預設集群的數量  $k$ 。(2) 隨機選取  $k$  個點做為集群的初始中心點。(3) 計算每個點到集群點中心位置的歐式距離(Euclidean distance)後，將每個點加入到距離最近的集群。(4) 每個集群中心內都會有被分類過來的資

料，以此計算每個集群的重心，尋找新的集群中心。(5) 重覆第三步驟和第四步驟，直到所有集群中心不再改變。

## 2.4 時間序列預測

透過簡禎富、許嘉裕 (2014).資料挖礦與大數據分析，可得知時間序列預測是推測為來過程的一種迴歸預測方法，依據規律時間間距下連續觀察的量測值，例如趨勢、季節或介入事件等，來分析這些已發生的時間序列資料之特性，歸納並估計出能反映歷史資料的時間序列模式，進行事物的趨勢預測。

### 2.4.1 時間序列預測的方法

Adams and Hand(2000)提出，資料探勘能從資料庫獲取有意義的資訊，以及對資料歸納具有結構的模式。資料探勘針對不同之目的和資料特性，有不同的方法，例如：類神經網路、集群分析、決策樹、遺傳基因演算法等。類神經網路(Artificial Neural Network, ANN)是模效生物神經網路的資訊處理系統，類神經網路的種類繁多，其中應用最普遍的是倒傳遞網路。倒傳遞網路(Back-Propagation Neural Network, BPN)是一種監督式學習網路(Supervised Learning Network)，即在網路學習過程中，將已知正確的輸入資料與輸出資料間的對應關係做為學習範例，因此在每一次的學習中，網路所得到的輸出資料都有一組正確的資料與之比對。在輸出層經過比對後得到學習誤差值，可將此誤差值從輸出層逐層回傳到內部的隱藏層中，網路則根據此回傳的誤差資訊來來監督網路加權值之調整，此學習過程可一再重複，直到誤差降為接近實際輸出值的推論輸出值。因此，倒傳遞網路適合用於預測之研究。

### 2.4.2 統計模型

Box, Jenkins 在 1961 年提出整合移動平均自迴歸模型 (Autoregressive Integrated Moving Average, ARIMA)，是常見的時間序列之一，大多應用在非平穩的時間序列上。由於 ARIMA 不僅結合自迴歸模型及移動平均模型，在進行差分將時間序列平穩化，ARIMA 包含了三個重要的參數： $p$ ,  $d$ ,  $q$ ，其中  $p$  為自迴歸項，模型會根據過去的  $p$  的歷史數值之迴歸作為當前時間點數值； $d$  為時間序列平穩化的差分次數，以時間點  $t$  的數值減去時間點  $t-1$  的數值，即稱為 1 階差分，以此類推； $q$  為移動平均項數，模型會根據過去的  $q$  項時間點數值平均誤差來預測當前時間點數值。

### 2.4.3 機器學習

機器學習方法從輸入資料中學習到的特性，並透過這些特性對預測未來資料，其原理由大量之統計理論綜合而成，能以模型本身之預測經驗改進內部參數，使預測結果能更加符合輸入資料，在許多流量預測之研究以機器學習模型預測也驗證方法之有效性。魏宏達(2017)採用 Apache Spark 搭配 Hadoop，將蒐集到的開放資料，如 YouBike 可借車數、氣象局觀測資料、人事行政局國定假日行事曆資料放入 HDFS 後執行資料前處理，並依 Spark MLlib 提供的機器學習演算法，進行不同分類技術的實驗，預測臺北市各 YouBike 站點缺車。

### 2.4.4 淺層類神經網路

近年來類神經網路 (Neural Network, NN) 已被廣泛且成功的用於時間序列的預測，類神經網路主要概念是透過神經元彼此間的連接及傳遞訊息，主要透過輸入值後，經由激活函數 (Activation Function) 的運算後將值輸出，神經網路可分為輸入層、隱藏層及輸出層。只需將合適的輸入與輸出給予模型，模型就能不斷透



過輸入資料的訓練學習，以調整輸入層、隱藏層及輸出內的各個參數，便能達成預測的目標。蘇怡瑄(2010)研究以類神經網路預測單一捷運站和單一出口人數，預期短期捷運運量的預測得到高準確率。

#### 2.4.5 深度學習

為解決 RNN 在訓練時無法學習到比較久遠之前的訊息，F.A. Gers 等人(2000)提出長短期記憶(Long Short-Term Memory, LSTM)架構，其模型架構主要根據 RNN 進行改良，改善 RNN 在長期記憶上的缺點，每一個記憶單元中都含有三個閘來控制記憶單元狀態的傳遞，分別為輸入閘(Input Gate)、遺忘閘(Forget Gate)、輸出閘(Output Gate)，來決定控制資料是否要輸入、控制之前的狀態是否要保留以及決定狀態是否要輸出。藉由這三個閘來控制狀態的傳遞，使得 LSTM 能夠有更長遠的記憶能力，因此多數用來處理複雜的時間序列資料，是近年來熱門用於人群流量預測的深度學習模型。如 Zheng 等人(2020)基於 LSTM 預測交通流量，實驗結果表明，LSTM 在預測精度上優於 ARIMA。

### 第三章 研究方法

本研究之研究方法分為兩個階段，流程如圖 3.1 所示。蒐集臺北市捷運及公共自行車租賃系統原始資料後，在第一階段先將臺北市捷運資料進行前處理，將處理後的資料進行特徵標準化，接著使用輪廓係數法，找出最佳分群數值 K，再使用集群分析法，依分群結果找出各捷運站集群的代表站點及其周邊 YouBike 站點。第二階段將前一階段所得到的質心捷運站點與周邊 YouBike 站點資料進行前處理，取得捷運歷史人流資料及其半徑距離 500 公尺內之 YouBike 站點資料，建立捷運人流預測模型，並進行 YouBike 站點歷史借用資料的處理，建立 YouBike 站點借還量預測模型。

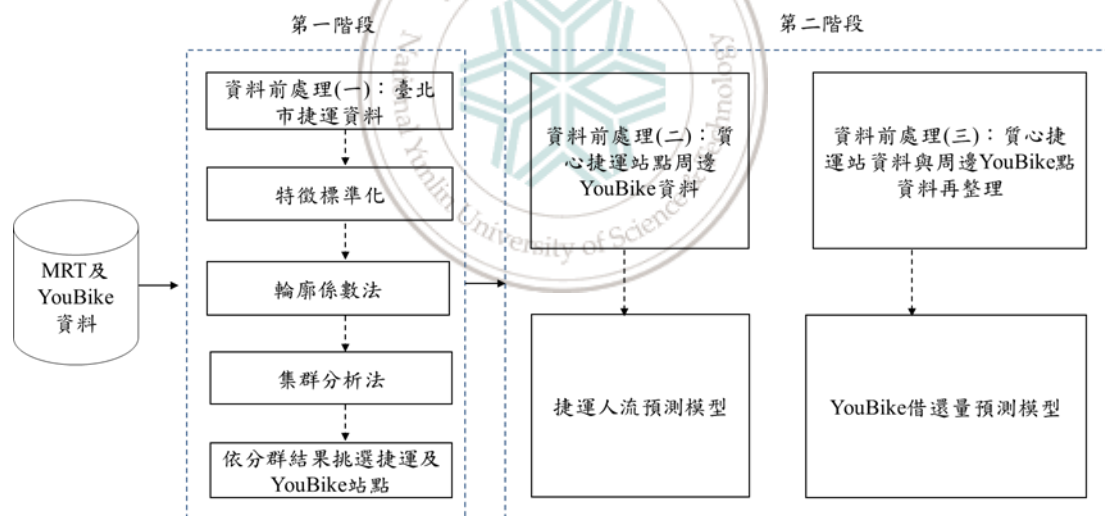


圖 3.1 研究方法流程圖

#### 3.1 資料集介紹

本研究為探究臺北捷運旅運量與 YouBike 未來借還量之關係，使用臺北市資料大平臺公開之臺北捷運每時進出站人次資料以及臺北市交通局所申請之臺北市公共自行車租賃系統借還量資料。以下對於此兩個資料集進行個別說明。

### 3.1.1 捷運資料集

該資料集取自臺北捷運公開資料集的臺北捷運每日分時各站 OD 流量統計資料，考量通常將 YouBike 作為捷運轉乘第一哩與最後一哩之使用者多數為大專院校學生，因此排除寒暑假，採用資料範圍為 2017 年 3 月 1 日至 2017 年 5 月 31 日，共計 22,534,848 筆。資料欄位如表 3.1，含有日期、時段、進站站名、出站站名與人次。

表 3.1 臺北捷運每時進出站人次

日期	時段	進站	出站	人次
2017/3/15	15	忠孝復興	公館	16
2017/3/15	15	忠孝復興	台電大樓	14
2017/3/15	15	忠孝復興	古亭	25
2017/3/15	15	忠孝復興	中正紀念堂	14
2017/3/15	15	忠孝復興	小南門	5
2017/3/15	15	忠孝復興	頂溪	49
2017/3/15	15	忠孝復興	永安市場	29
2017/3/15	15	忠孝復興	景安	31
2017/3/15	15	忠孝復興	南勢角	12
2017/3/15	15	忠孝復興	台大醫院	8
2017/3/15	15	忠孝復興	台北車站	314
2017/3/15	15	忠孝復興	中山	41
2017/3/15	15	忠孝復興	雙連	13
		.		
		.		
		.		

### 3.1.2 YouBike 資料集

該資料集取自 YouBike 借還資料，同樣採用時間區間於 2017 年 3 月 1 日至



2017 年 5 月 31 日之資料，涵蓋雙北地區，共計 5,485,630 筆，資料欄位包含交易序號、卡號、扣款時間、扣款金額、借車時間、借車站代號等共有 23 個欄位，其格式如表 3.2。

表 3.2 Youbike 開放資料平台原始資料格式說明

代號	說明
sno：	站點代號
sna：	場站名稱(中文)
tot：	場站總停車格
sbi：	場站目前車輛數量
sarea：	場站區域(中文)
mday：	資料更新時間
lat：	緯度
lng：	經度
ar：	地(中文)
sareaen：	場站區域(英文)
snaen：	場站名稱(英文)
aren：	地址(英文)
bemp：	空位數量
act	全站禁用狀態

### 3.2 臺北捷運資料集集群分析

本研究採用非監督式學習的集群分析，資料無需事先經人為定義，讓演算法依據資料的特徵值進行分群，但資料的前處理仍是進行集群分析的必要工作，因此需要對資料進行：(1)缺失、無效值的清洗 (2)資料降維 (3)變數的處理。

#### 3.2.1 臺北捷運資料前處理

本研究階段先將 2017 年 3 月 1 日至 2017 年 5 月 31 日之原始資料進行清洗，剔除表 3.3 所示之行政院人事行政處公布的國定假日及周六、周日後，依圖 3.2 方式分別計算臺北捷運 108 個站點，從 2017 年 3 月 1 日至 5 月 31 日的每周一至周五、每一小時裡的進站人次總和與出站人次總和，再進行每一站的每周一至每周

五、每時段之進站人次累加。如圖 3.3 為捷運忠孝復興站 2017 年 3 月 1 日至 5 月 31 日的每周一至周五、每一小時的進站累加人次圖；圖 3.4 則為忠孝復興站自 2017 年 3 月 1 日至 5 月 31 日的每周一至周五、每一小時的出站累加人次圖。

表 3.3 行政院人事行政處公布之國定假日

107 年國定假日		
日期	星期	假別
1 月 1 日	一	元旦
2 月 15 日	四	農曆春節
2 月 16 日	五	農曆春節
2 月 19 日	一	農曆春節
2 月 20 日	二	農曆春節
2 月 28 日	三	228
4 月 4 日	三	兒童節
4 月 5 日	四	清明節
4 月 6 日	五	彈性放假
5 月 1 日	二	勞動節
6 月 18 日	一	端午
9 月 24 日	一	中秋
10 月 10 日	三	雙十
12 月 31 日	一	元旦彈性放假

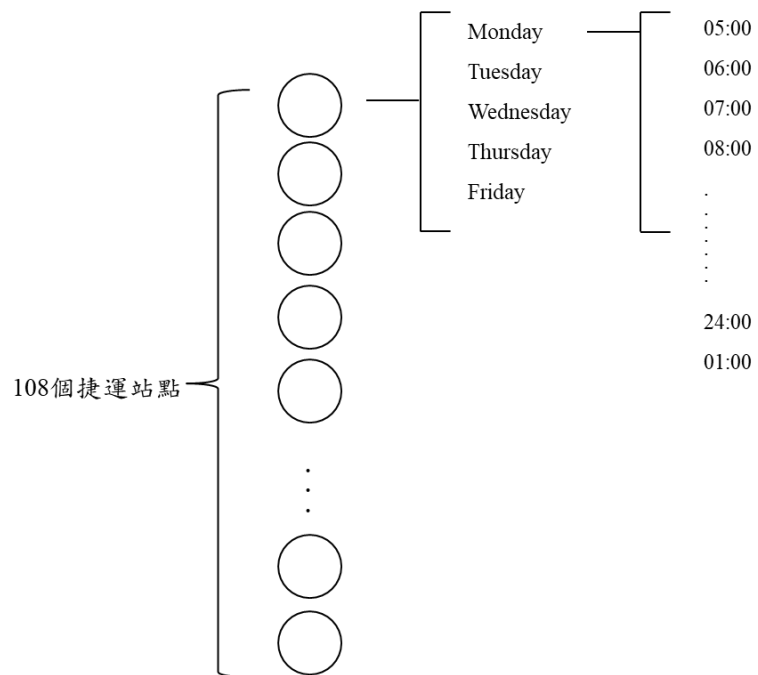


圖 3.2 各捷運站人次加總流程

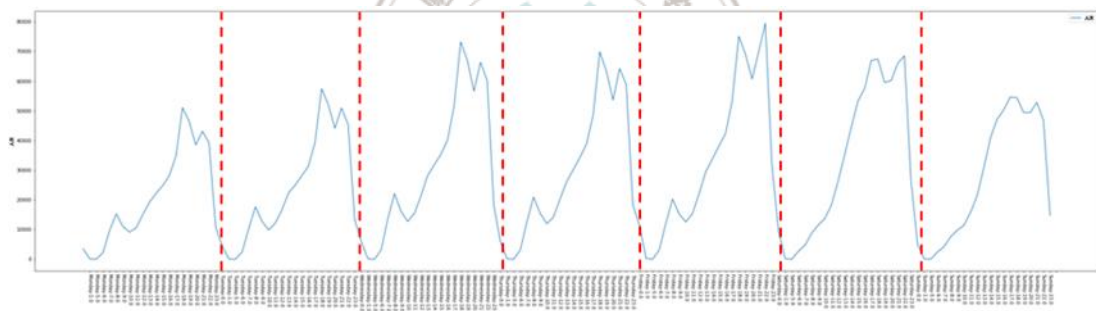


圖 3.3 忠孝復興站進站累加人次圖

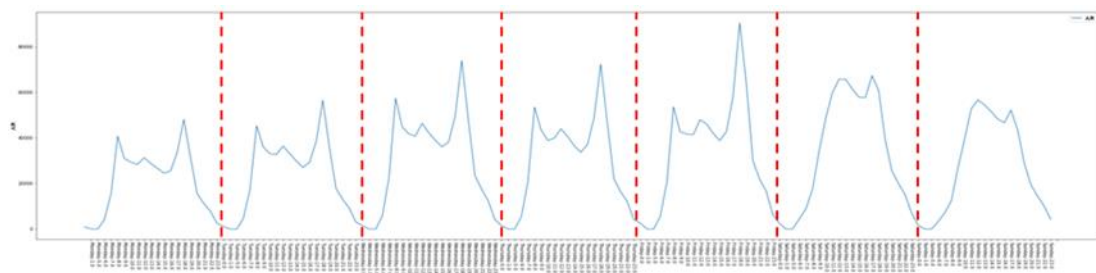


圖 3.4 忠孝復興站出站累加人次圖

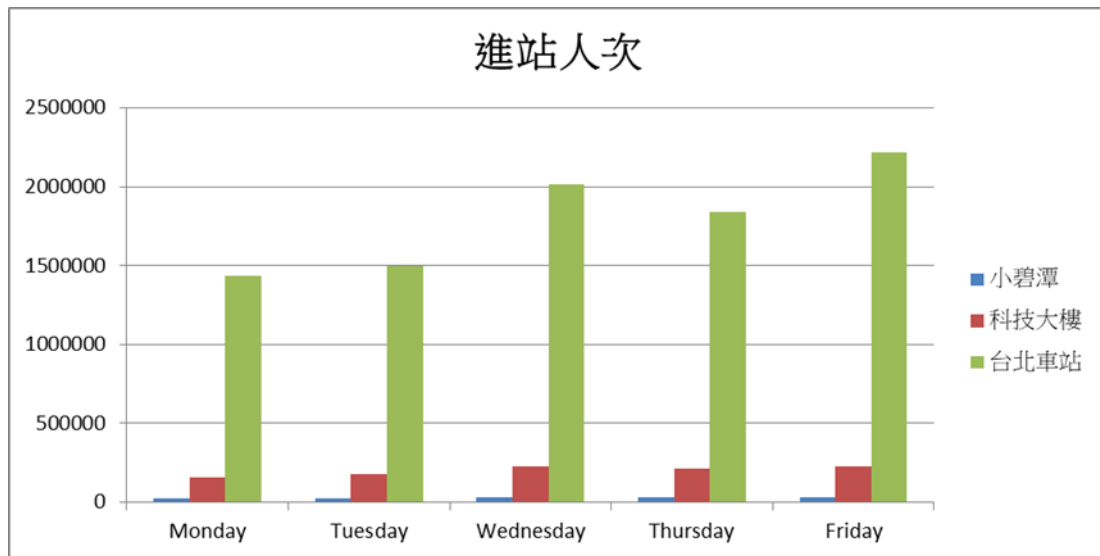


圖 3.5 臺北捷運進站人流週別累計旅次統計

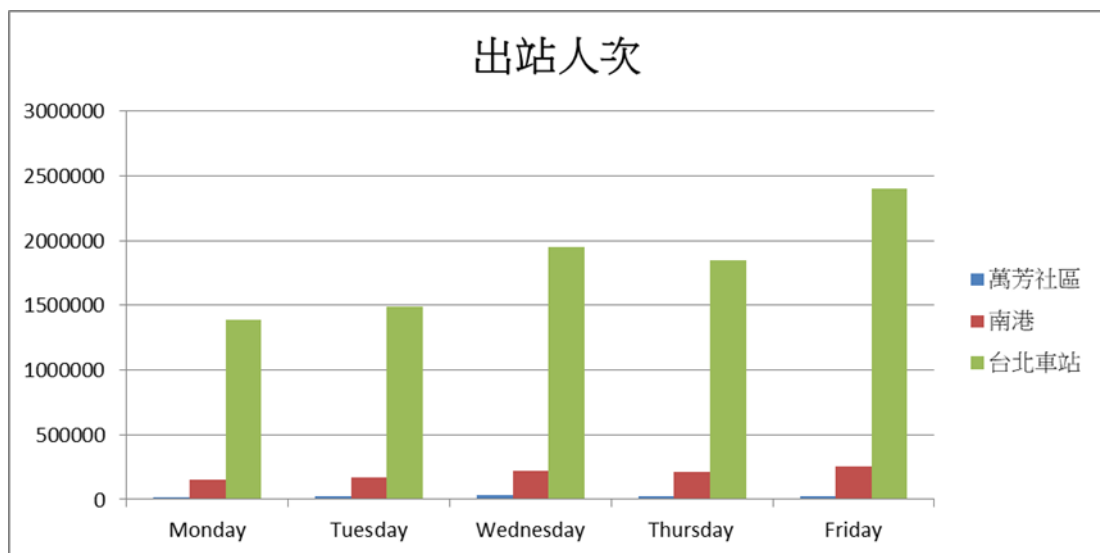


圖 3.6 臺北捷運出站人流週別累計旅次統計

從臺北捷運 2017 年 3 月 1 日至 2017 年 5 月 31 日的 108 個站點週別累計旅次統計，可以觀察到臺北捷運各站點的進出站人流實際數值差距甚大，當中累積旅次的最大值、中間值、最小值，分別為台北車站、科技大樓站和小碧潭站，如圖 3.5 與圖 3.6 所示。因此在進行 K-means 集群分析前，需先將所有站點的人流資料進行標準化 (Min-Max Normalization)，故表 3.4 將實際流量值轉換為各站點之間最

大值的比例。

表 3.4 捷運站點人流標準化

站別	Monday	Tuesday	Wednesday	Thursday	Friday
七張	0.156510113	0.175199371	0.22816477	0.209774593	0.230351153
三和國中	0.1588614	0.179225165	0.226222458	0.212640712	0.223050264
三民高中	0.159735859	0.177345656	0.228201039	0.211508432	0.223209013
三重	0.157088862	0.176028839	0.228643944	0.210546497	0.227691858
三重國小	0.159317507	0.178975881	0.227018525	0.211818094	0.222869992
中山	0.151510498	0.173471235	0.224298259	0.212383607	0.238336401
中山國中	0.157214191	0.179052181	0.227501583	0.213658756	0.222573288
中山國小	0.158395164	0.178329982	0.226484777	0.212164057	0.22462602
中正紀念堂	0.149525162	0.176807592	0.228726745	0.212917028	0.232023472
丹鳳	0.159265895	0.176242498	0.225056971	0.210766287	0.228668348
亞東醫院	0.160105328	0.179449521	0.228928281	0.21168302	0.21983385
信義安和	0.155962579	0.17733705	0.226346572	0.21353599	0.226817809
先嗇宮	0.158913805	0.179507551	0.229000735	0.211412795	0.221165114
內湖	0.157537936	0.177243982	0.226440318	0.210084164	0.2286936
公館	0.15061781	0.172125355	0.220393387	0.209769065	0.247094383

### 3.2.2 輪廓係數法

在選定方法以後，需要確認資料應分為幾群較為適當，在分群數量的確認上，本研究選擇使用輪廓係數(Silhouette Coefficient)來確立合理的群聚數量。輪廓係數最早由 Rousseeuw 於 1986 年提出，使用內聚度及分離度兩種因素，是用以評價集群效果好壞的其中一種方法，滿足集群的概念，找出「組間異質；組內均質」的值，同集群之間距離之距離以 a 表示，異群之間的點平均距離以 b 表示，得出的 S 值越大，表示集群區隔越明顯，計算完所有樣本的 Si 均值之後，可用以評估集群是否合理有效。以本研究來說，試著將人流迥異的站點區分為不同的集群，公式為： $S = \frac{b-a}{\max(a,b)}$ ，圖 3.7 所示之輪廓係數分析圖，從中計算各點間斜率，取最大值後得到當 K=3 時，分群效果最好。

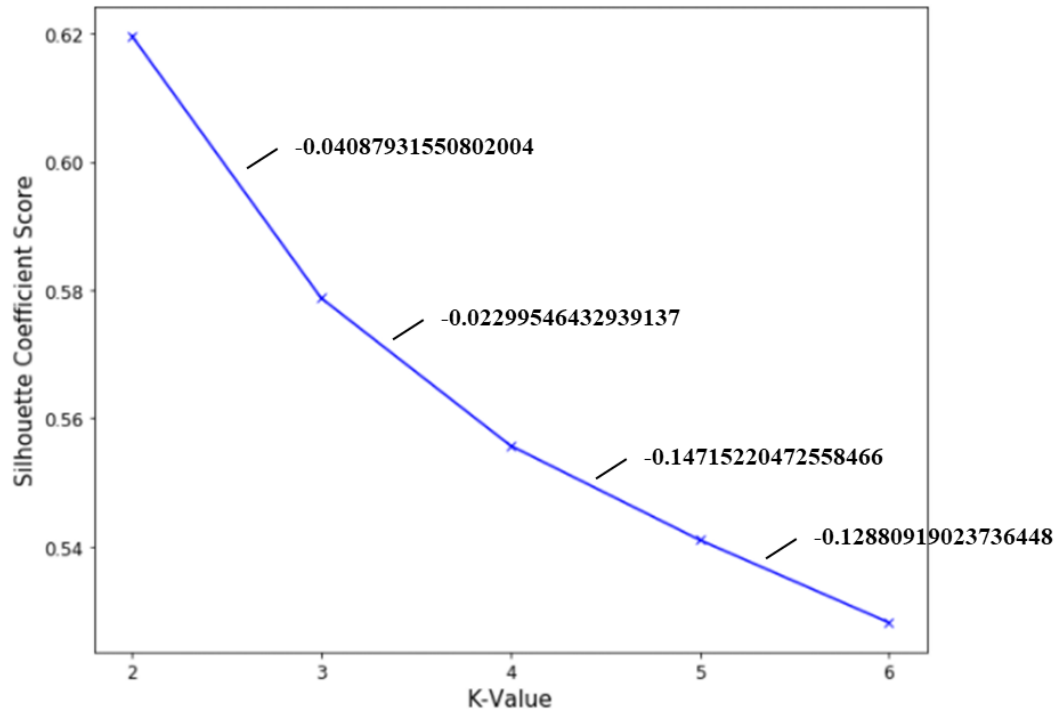


圖 3.7 輪廓係數分析圖

### 3.2.3 集群分析法-K-means

1. 預設集群的數量(K)。

$$\mu_c^{(0)} \in R^d, c = 1, 2, \dots, K$$

2. 隨機選取 K 個點做為集群的初始中心點。
3. 計算每個點到集群點中心位置的歐式距離(Euclidean Distance)後，將每個點加入到距離最近的集群。

$$S_c^{(t)} = \{X_i: \|X_i - \mu_c^{(t)}\| \leq \|X_i - \mu_{c^*}^{(t)}\|, \forall i = 1, \dots, n\}$$

4. 每個集群中心內都會有被分類過來的資料，以此計算每個集群的重心，尋找新的集群中心。因此，基於輪廓係數法得到群數為 3 後，代入 Kmeans 分群法，集群中心結果如表 3.5，分別為萬芳醫院站、動物園站以及忠孝復興站。

$$\mu_c^{(t+1)} = \frac{\text{sum}(S_c^{(t)})}{n_c} = \sum_{i=1}^{n_c} x_i \mid_{x_i \in S_c^{(t)}}$$

5. 重覆第三步驟和第四步驟，直到所有集群中心不再改變。

$$S_c^{(t+1)} = S_c^{(t)}, \forall c = 1, \dots, K$$

6. 得出分群質心站點為萬芳醫院站、動物園站、忠孝復興站。

表 3.5 集群中心站點分析結果

站別	Monday_Std	Tuesday_Std	Wednesday_Std	Thursday_Std	Friday_Std
90 萬芳醫院	0.160381	0.178243	0.231247	0.212695	0.217435
18 動物園	0.144977	0.179202	0.233885	0.210845	0.231091
54 忠孝復興	0.150177	0.171645	0.221375	0.211954	0.244849

### 3.2.4 依分群結果挑選捷運及 YouBike 站點

基於考量一般民眾可接受步行距離約為五百公尺，本研究以台北捷運的「萬芳醫院站」、「動物園站」與「忠孝復興站」三站點，透過臺北公共自行車官方網站提供的站點地圖，確認該三個捷運站周邊的 YouBike 租賃站數量及租賃站名稱，並刪去 2017 年 5 月 31 日之後所設站點，再使用 Google Maps 的路線規劃功能，找出距離為半徑 500 公尺內，有可能被民眾利用於轉乘捷運的 YouBike 站點，同時找出捷運站半徑 500 以上至 1000 公尺內的 YouBike 站點，作為模型對照組。篩選結果如下：捷運動物園站半徑 500 公尺內有捷運動物園站(2 號出口)，總計共一個 YouBike 站點；半徑 500 公尺以上至 1000 公尺內，無 YouBike 站點。捷運萬芳醫院站半徑 500 公尺內有臺北市文山運動中心，共計一個 YouBike 站點；半徑 500 公尺以上至 1000 公尺內有文山第二行政中心、捷運辛亥站，共計二個 YouBike 站點。捷運忠孝復興站半徑 500 公尺內有瑠公公園、捷



運忠孝復興站(2 號出口)、忠孝東路三段 217 巷口、龍門廣場、復興市民路口，共計五個 YouBike 站點；半徑 500 尺至 1000 尺內有內政部營建署、仁愛醫院、朱崙商場、八德市場，共計四個 YouBike 站點。以上共 13 個 Youbike 站點，將其原始欄位進行降維，保留日期、星期、時段、站點與人次紀錄進行分析，分析資料共計 2,386 筆。

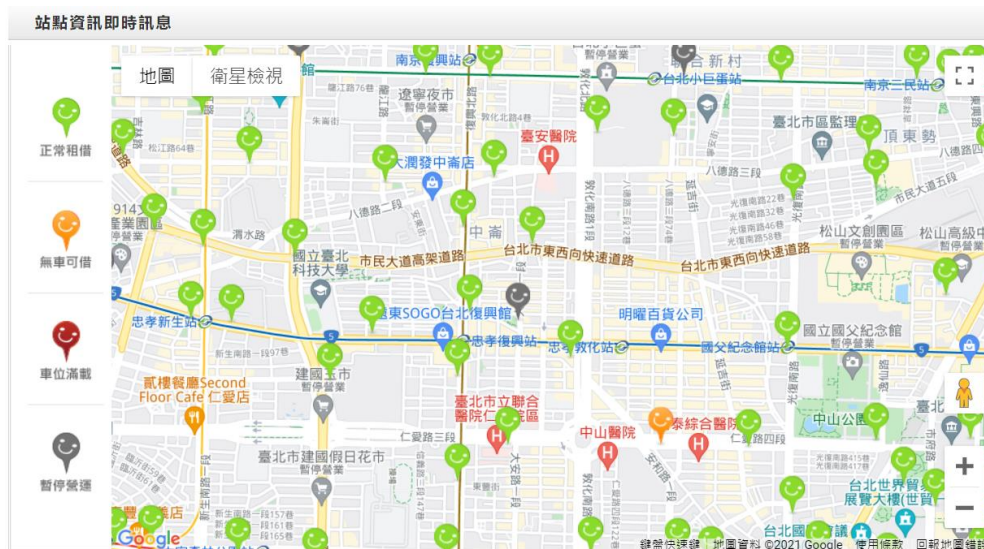


圖 3.8 忠孝復興站周邊站點地圖

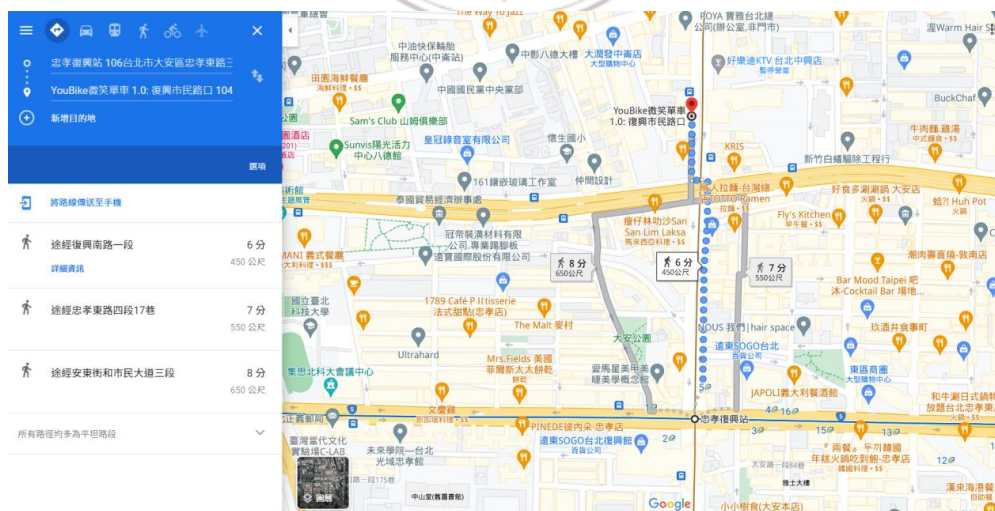


圖 3.9 Google Msp 路線規劃



表 3.6 質心站點方圓一公里內 YouBike 站點

捷運站點	YouBike 站點	距離(m)	步行時間(minute)
動物園	捷運動物園站(2 號出口)	170	2
萬芳醫院	臺北市文山運動中心	400	5
萬芳醫院	文山第二行政中心(暫停營運)	800	11
萬芳醫院	捷運辛亥站	950	12
忠孝復興	瑠公公園(暫停營運)	190	2
忠孝復興	捷運忠孝復興站(2 號出口)	240	4
忠孝復興	忠孝東路三段 217 巷口	350	6
忠孝復興	龍門廣場	400	5
忠孝復興	復興市民路口	450	6
忠孝復興	內政部營建署	700	9
忠孝復興	仁愛醫院	750	11
忠孝復興	朱崙商場	850	11
忠孝復興	八德市場	900	12

### 3.3 基於捷運人流預測 Youbike 使用量

本研究是先預估捷運的進出站人次，再結合捷運的預估結果來預測 Youbike 借還使用量。因此在本章節將說明捷運進出站人次的資料是如何定義及篩選，並且基於此相關條件選擇 Youbike 使用量的資料，最後利用類神經時間序列網路建立兩階段的預測模型。

#### 3.3.1 MRT 進出站人次預測資料集

臺北捷運 108 個站點分三群，選取出的質心站點後，將早上五點至下午兩點定義為上半天；下午三點到凌晨一點定義為下半天，以此進行半天為一單位的時間序列模型建立。

#### 3.3.2 YouBike 使用人次預測資料集

由三個質心捷運站點為中心點，找出方圓一公里以內的 YouBike 租賃車站共計 13 個，配合捷運營運時間之流量，同樣將早上五點到下午兩點定義為上半天，

下午三點到凌晨一點定義為下半天。為了將類別(Categorical)與文字(Text)資料轉換成數字，讓程式能夠更好地去理解及運算，使用 One-Hot Encoding 編碼，為每個類別新增一個欄位，用 0 或 1 表示是或否，再以此進行半天為一單位，同時加上模型一預測結果的時間序列來建立模型。

表 3.7 忠孝復興站-500m 以內 YouBike 借車預測模型結果

Date	Stage	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Sun.	First Half	Second Half	MRT_ Counts	Counts X	Counts Y
3/1	龍門廣場	0	0	1	0	0	0	0	1	0	797	25	57
3/1	捷運忠孝復興站 (2 號出口)	0	0	1	0	0	0	0	1	0	797	102	82
3/1	瑠公公園	0	0	1	0	0	0	0	1	0	797	25	38
3/1	復興市民路口	0	0	1	0	0	0	0	1	0	797	30	35
3/1	忠孝東路三段 217 巷口	0	0	1	0	0	0	0	1	0	797	27	36
3/1	龍門廣場	0	0	1	0	0	0	0	0	1	827	57	24
3/1	捷運忠孝復興站 (2 號出口)	0	0	1	0	0	0	0	0	1	827	82	97
3/1	瑠公公園	0	0	1	0	0	0	0	0	1	827	38	16
3/1	復興市民路口	0	0	1	0	0	0	0	0	1	827	35	29
3/1	忠孝東路三段 217 巷口	0	0	1	0	0	0	0	0	1	827	36	38
3/2	龍門廣場	0	0	0	1	0	0	0	1	0	782	24	53
3/2	捷運忠孝復興站 (2 號出口)	0	0	0	1	0	0	0	1	0	782	97	55
3/2	瑠公公園	0	0	0	1	0	0	0	1	0	782	16	29
3/2	復興市民路口	0	0	0	1	0	0	0	1	0	782	29	35

### 3.3.3 倒傳遞類神經時間序列網路

本研究欲建構一捷運進出站人流與 Youbike 使用量的預測模型，由於此類型資料均與其前後時間有相對應的關係，故採用類神經時間序列模型。類神經網路架構一般有三種層，包含輸入層、隱藏層、輸出層，每一層網路中都是由眾多節

點所相互連結組合而成，其中輸入層主要接收外部訊號並將訊號往下傳入神經網路，隱藏層接收輸入層所傳遞的訊號，並對此訊號進行運算，而這些訊號的運算處理過程是無法看見的，最後來到輸出層。輸出層接收前層網路處理後的訊號，將結果傳送。至於時間序列網路的架構，會將每次模型輸出的回饋結果加入下一段時間的輸入值，依此類推產生具有前段時間特徵的預測結果。

在研究分析模型的使用上，主要使用 Matlab 的 Neural Net Time Series 工具建立，如圖 3.11 所示。本研究模型架構隱藏層神經元數設定為 10，Number of Delay 根據資料是用前一個時段預測現在時段，所以設定為 1 進行建構。

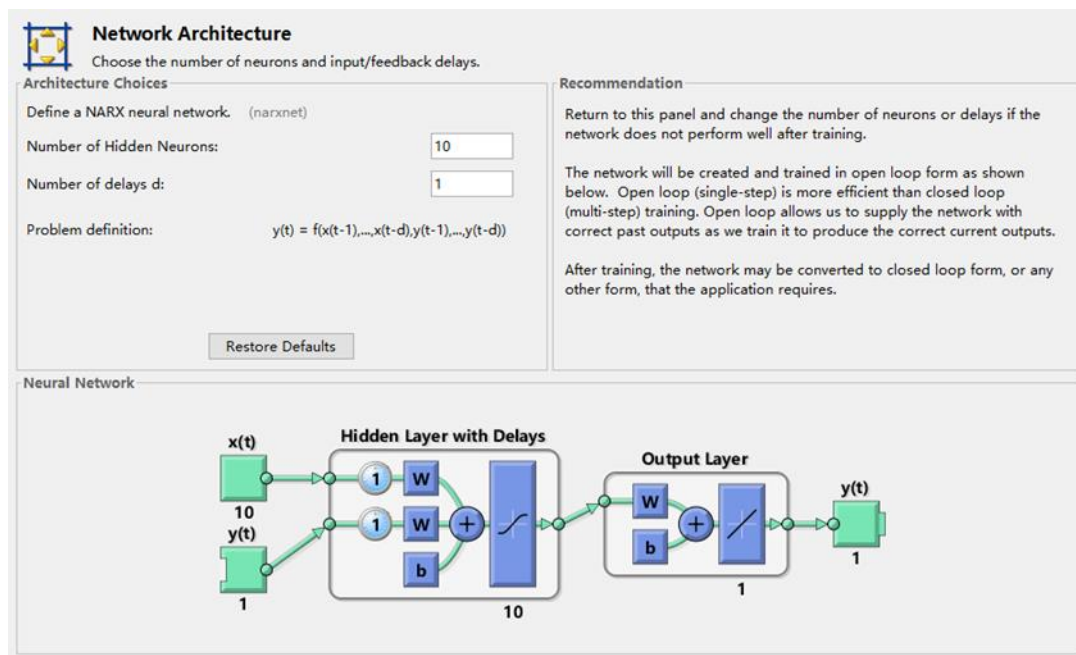


圖 3.10 模型架構設定

## 第四章 實驗結果

本研究提出以捷運系統進出站人數的預測結果來進行 YouBike 借還量預測，這樣的觀念係基於大部分 YouBike 的使用者把 YouBike 作為來往捷運站最初或最後一哩路的交通工具，而其行為模式通常分成兩種：(1)騎乘 YouBike 至捷運站周邊歸還單車後，進入捷運站(2)出了捷運站後，在周邊租借 YouBike 騎乘返家。因此我們將捷運及 YouBike 的使用時間，區隔定義為 05:00AM~15:00PM、15:00PM~01:00 AM 兩個半天的時段，並以預測出的上半天捷運進站人流，來預測下一個半天的 YouBike 歸還量，如圖 4.1 之模型建置流程。

各項類神經網路的模型建置過程，依圖 4.2 統一將預處理好的資料集按照固定比例隨機區分，分別為 70%的訓練資料，是模型建立最主要的資料來源，提供模型以監督式訓練的學習特徵；10%的驗證資料是用來檢視每個 Epoch 的模型學習過程狀況，並根據驗證指標結果回饋，進行倒傳遞的權重參數更新；最後 10%為測試資料，為每次權重更新產生後作為測試依據，但不為學習訓練的參數更新回饋，較能夠表現該模型實際應用的可能狀況。

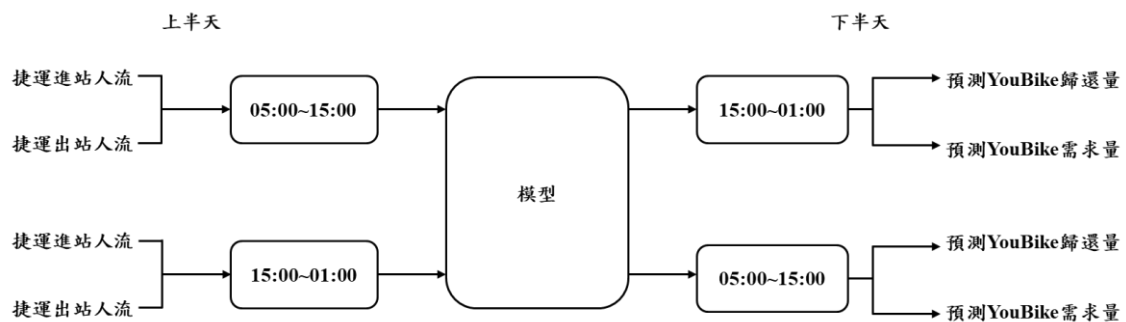


圖 4.1 模型建置架構

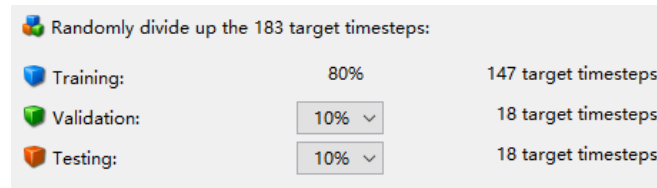


圖 4.2 資料集分布

## 4.1 捷運進出站人次預測

依據台北捷運分群結果的質心站點為忠孝復興站、萬芳醫院站、動物園站，將該三個捷運站點自 2017 年 3 月 1 日至 5 月 31 日之每日進出站人次數，進行每一小時的人次加總，並將早上五點到下午兩點歸納為上半天，下午三點到凌晨一點歸納為下半天，依此進行半天為一單位的时间序列來建立模型預測下個半天的進出站結果。以下針對人次量相對最高的忠孝復興站及人次量相對最少的動物園站作為分析對象，驗證模型在不同類型資料集的分析結果。

本研究針對捷運站的進站人次預測狀況，是利用 Error Histogram 訓練的結果是否為常態分布，檢視類神經的學習過程是否有效學習；其中 X 軸是誤差值，Y 軸是資料筆數，橘色線是誤差為 0 的位置，代表絕對猜對，藍色是訓練資料的誤差筆數，綠色是驗證資料的誤差筆數，紅色是測試資料的誤差筆數。

### 4.1.1 捷運進站人次預測結果分析

忠孝復興捷運站的進站人次預測為常態分布(圖 4.2)，證明模型是有效的學習，並從測試集(表 4.1)觀察，主要幾個指標分別 MAE 為 52.857 及 RMSE 為 62.627(表 4.2)，由於忠孝復興捷運站每天平均約有 968.940 人進站，因此誤差值是可以被接受的範圍內。

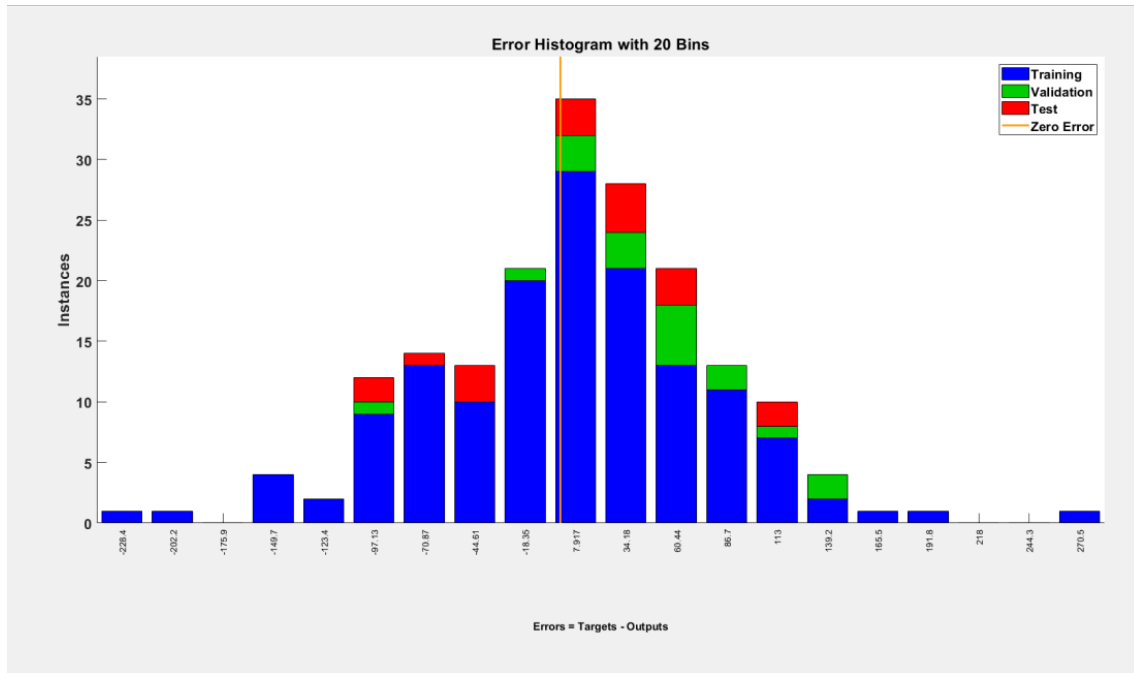


圖 4.3 誤差值直方圖-捷運忠孝復興站進站

表 4.1 捷運忠孝復興站進站人次測試結果

下半年預測人次	預測結果	誤差	平方差
1366	1348.31093	17.6890728	312.903296
1280	1377.08315	97.083147	9425.13743
1125	1179.42985	54.4298506	2962.60863
1192	1204.47526	12.475265	155.632237
639	596.650642	42.3493576	1793.46809
1606	1559.69881	46.301191	2143.80029
609	627.905358	18.9053579	357.412556
1055	1146.75556	91.7555581	8419.08244
616	614.428507	1.57149303	2.46959035
523	612.386561	89.3865605	7989.9572
1620	1581.8632	38.1368008	1454.41557
1211	1168.4271	42.5729024	1812.45202
1243	1146.31196	96.6880395	9348.57699
1329	1361.24996	32.249957	1040.05973
590	550.912038	39.0879618	1527.86876
1474	1371.43714	102.562857	10519.1397
547	571.605476	24.6054762	605.429461
502	605.580599	103.580599	10728.9405

表 4.2 捷運忠孝復興站進站人次測試結果指標

MAE	MSE	RMSE
52.857	3922.186	62.627

捷運動物園站的進站人次預測結果雖沒有達到常態分布(圖 4.4)，但誤差分布集中於橘線兩側，因此證明模型也是有效的學習，從測試集(表 4.3)觀察主要幾個指標，分別 MAE 為 50.947 及 RMSE 為 74.094(表 4.4)，由於動物園站每天平均約有 441.826 人次進站，因此誤差值仍是可以被接受的範圍。

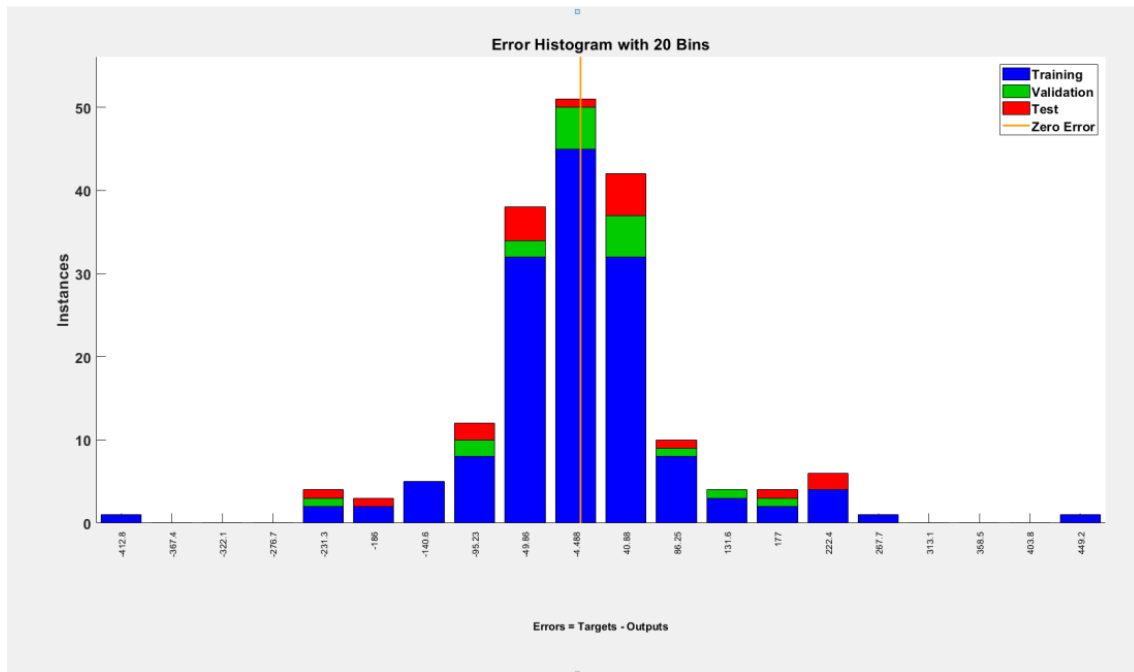


圖 4.4 誤差值直方圖-捷運動物園站進站

表 4.3 捷運動物園站進站人次測試結果

下半天預測人次	預測結果	誤差	平方差
382	429.78763	47.788	2283.658
408	384.94275	23.057	531.637
295	337.784012	42.784	1830.472
238	266.392304	28.392	806.123
329	319.073576	9.926	98.534
351	333.845311	17.155	294.283
436	456.94262	20.943	438.593
361	386.375202	25.375	643.901
449	509.595164	60.595	3671.774
620	491.739903	128.260	16450.652
305	277.936488	27.064	732.434
615	624.131456	9.131	83.383
243	304.461698	61.462	3777.540
411	537.977796	126.978	16123.361
344	561.971565	217.972	47511.603
301	291.281662	9.718	94.446
501	559.683464	58.683	3443.749
310	308.241917	1.758	3.091

表 4.4 捷運動物園站進站人次測試結果指標

MAE	MSE	RMSE
50.947	5489.957	74.094

#### 4.1.2 捷運出站人次預測結果分析

忠孝復興捷運站的出站人次預測為常態分布(圖 4.5)，證明模型是有效的學習，並從測試集(表 4.5)觀察，主要幾個指標分別 MAE 為 44.278 及 RMSE 為 56.864(表 4.6)，由於忠孝復興捷運站每天平均約有 960.462 人次出站，因此誤差值是可以被接受的範圍內。



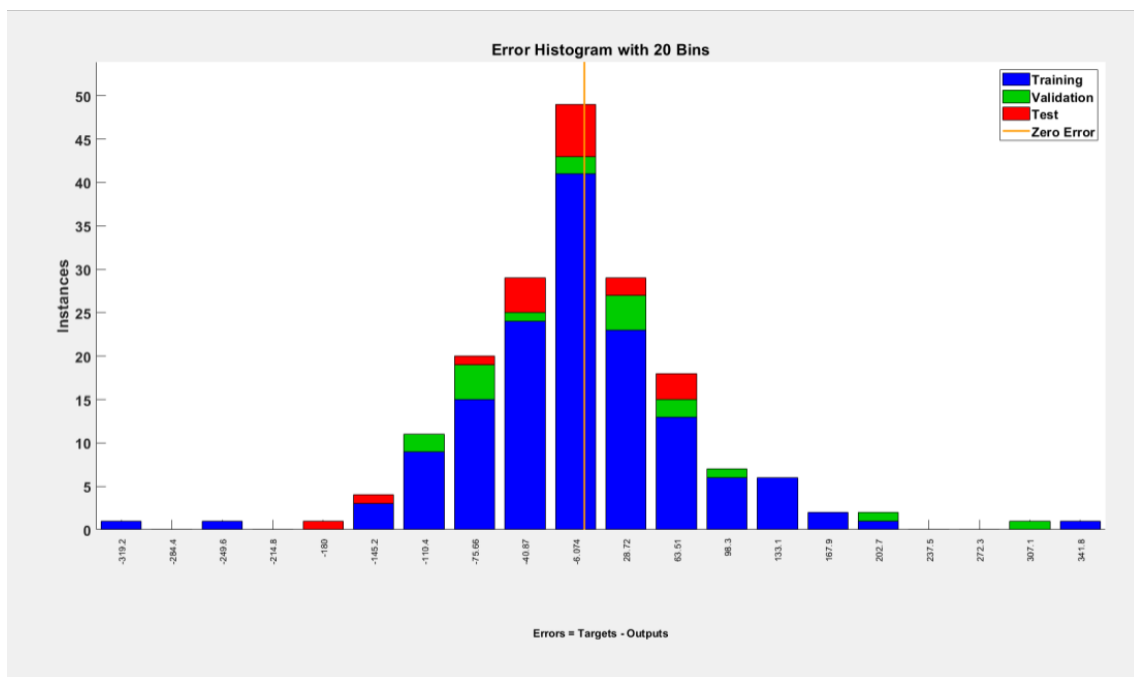


圖 4.5 誤差值直方圖-捷運忠孝復興站出站

表 4.5 捷運忠孝復興站出站人次測試結果

下半天預測人次	預測結果	誤差	平方差
954	947.532	6.468	41.829
1194	1266.741	72.741	5291.187
1242	1157.846	84.154	7081.871
794	821.482	27.482	755.238
825	851.091	26.091	680.757
866	867.487	1.487	2.212
1249	1227.716	21.284	453.020
900	880.003	19.997	399.861
1063	944.519	118.481	14037.835
1295	1243.821	51.179	2619.242
772	866.650	94.650	8958.593
846	853.010	7.010	49.142
921	922.472	1.472	2.168
1275	1190.829	84.171	7084.701
954	967.172	13.172	173.491
1031	1000.517	30.483	929.197
787	867.645	80.645	6503.682
791	847.037	56.037	3140.119

表 4.6 捷運忠孝復興站出站人次測試結果指標

MAE	MSE	RMSE
44.278	3233.563	56.864

捷運動物園站的出站人次預測結果雖沒有達到常態分布(圖 4.6)，但誤差分布集中於橘線兩側，因此證明模型也是有效的學習，從測試集(表 4.7)觀察主要幾個指標，分別 MAE 為 21.544 及 RMSE 為 27.289(表 4.8)，由於動物園站每天平均約有 365.777 人次出站，因此誤差值仍是可以被接受的範圍。

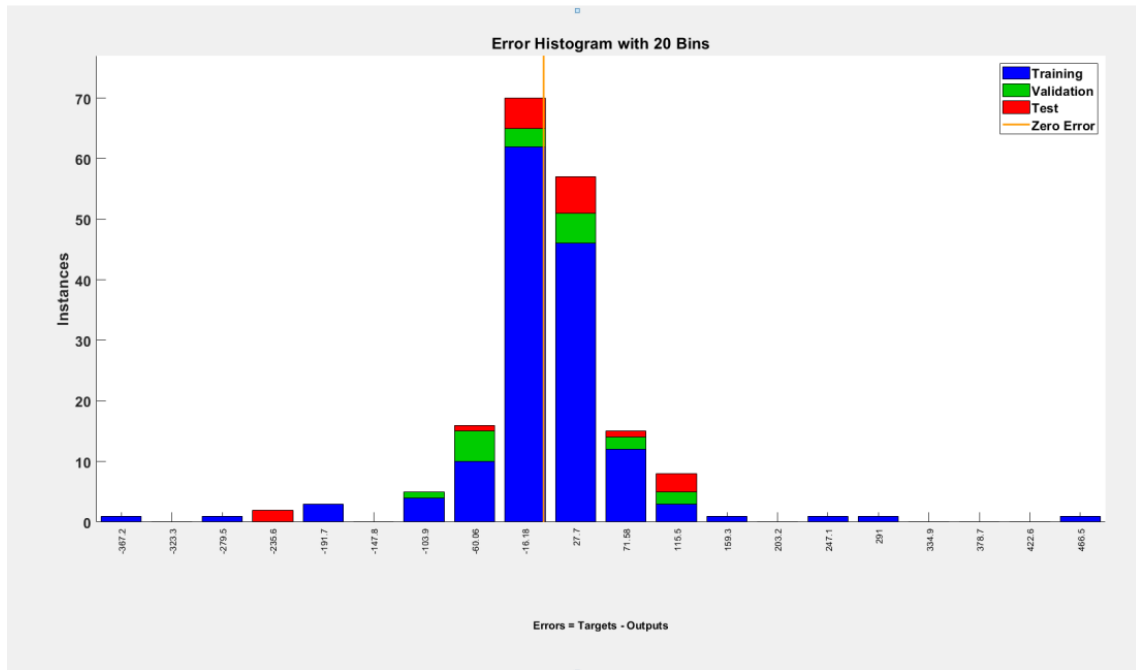


圖 4.6 誤差值直方圖-捷運動物園站出站

表 4.7 捷運動物園站出站人次測試結果

下半天預測人次	預測結果	誤差	平方差
297	289.430076	7.570	57.304
276	276.407491	0.407	0.166
339	294.360618	44.639	1992.674
295	292.721572	2.278	5.191
232	270.942996	38.943	1516.557
288	306.686063	18.686	349.169
298	261.108201	36.892	1361.005
281	285.145709	4.146	17.187
322	298.363888	23.636	558.666
259	283.216088	24.216	586.419
299	291.297344	7.703	59.331
311	327.342342	16.342	267.072
273	287.007507	14.008	196.210
339	276.45489	62.545	3911.891
292	286.954399	5.046	25.458
382	348.722138	33.278	1107.416
300	335.280512	35.281	1244.715
298	285.817807	12.182	148.406

表 4.8 捷運動物園站出站人次測試結果指標

MAE	MSE	RMSE
21.544	744.713	27.289

經過以上兩個不同群體代表的捷運進出站測試結果，均能夠在合理範圍內預測進出捷運站的人次，因此驗證此類神經時間序列模型在捷運資料集的預測是可行的方法。

#### 4.2 YouBike 借還車使用量預測

本階段分別透過 K-means 分群後產生的三個質心捷運站點為中心，分析各站點半徑 500 公尺內的 YouBike 租賃車站作為借還車的預測對象，每筆

YouBike 資料中均結合對應的捷運站之區間進出站人次預測結果數值，分別有動物園站：捷運動物園站(2 號出口)；萬芳醫院站：臺北市文山運動中心；忠孝復興站：瑠公公園、捷運忠孝復興站(2 號出口)、忠孝東路三段 217 巷口、龍門廣場、復興市民路口，依此共分為三組 YouBike 租賃站點建立借還車預測模型。以下會針對 YouBike 租賃車站數相對最高的忠孝復興站周邊及車站數相對最少的動物園站作為分析對象，驗證模型在不同類型資料集的分析結果。

#### 4.2.1 捷運站周邊 YouBike 還車預測結果分析

忠孝復興站周邊的 YouBike 還車量預測為常態分布(圖 4.7)，證明模型是有效的學習，並從測試集(表 4.9)觀察，主要幾個指標分別 MAE 為 12.033 及 RMSE 為 16.301(表 4.10)，不過由於忠孝復興站周邊 YouBike 車站每天平均約有 41.145 台的還車量，因此誤差值是可以被接受的範圍內。

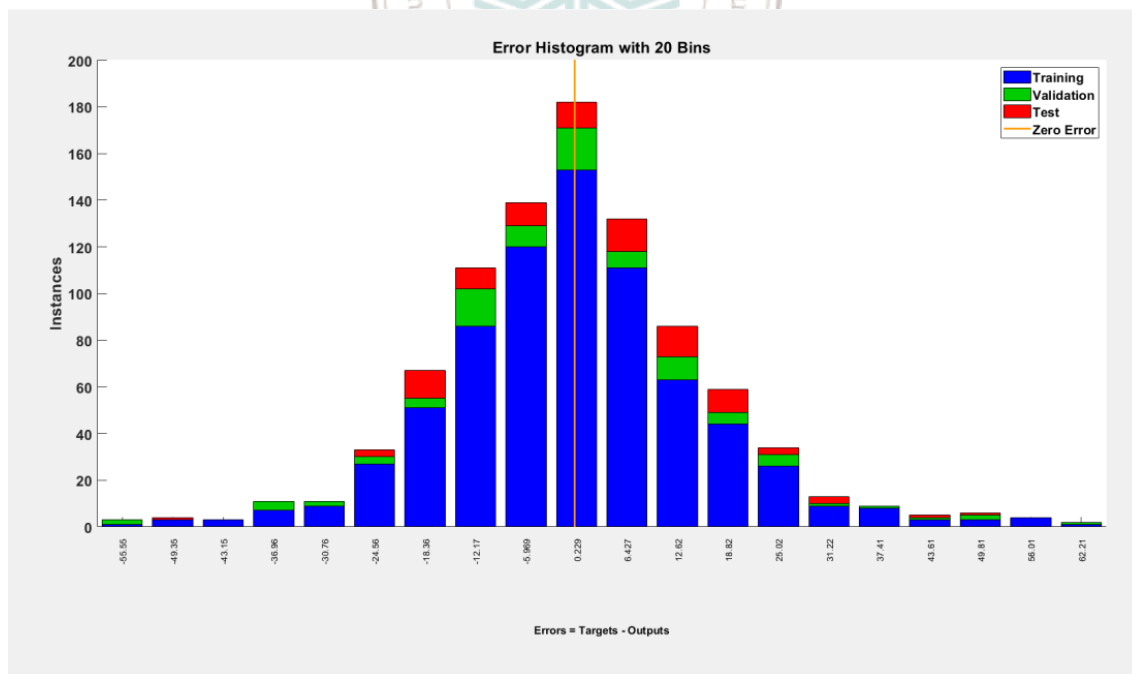


圖 4.7 誤差值直方圖-捷運忠孝復興站周邊 YouBike 還車

表 4.9 捷運忠孝復興站周邊 YouBike 站點還車量測試結果

下半天預測人次	預測結果	誤差	平方差
11	39.830	28.830	831.161
52	42.025	9.975	99.504
37	64.883	27.883	777.446
52	49.326	2.674	7.148
54	71.297	17.297	299.198
18	37.066	19.066	363.511
53	34.313	18.687	349.213
38	43.785	5.785	33.461
4	28.091	24.091	580.383
12	4.374	7.626	58.149
21	26.747	5.747	33.028
9	21.205	12.205	148.963
49	32.712	16.288	265.301
15	37.809	22.809	520.260
25	36.020	11.020	121.433
24	32.950	8.950	80.095
13	40.493	27.493	755.855
60	44.238	15.762	248.453
80	90.544	10.544	111.166
25	38.620	13.620	185.507
59	39.153	19.847	393.910
25	34.335	9.335	87.137
45	55.519	10.519	110.651
45	44.122	0.878	0.772
27	29.228	2.228	4.965
14	13.104	0.896	0.802
11	48.792	37.792	1428.263
16	13.663	2.337	5.461
35	31.518	3.482	12.123
18	38.557	20.557	422.601
103	44.839	58.161	3382.719
67	44.188	22.812	520.378
28	19.464	8.536	72.860
31	32.286	1.286	1.654
26	33.836	7.836	61.404

62	42.986	19.014	361.548
39	37.363	1.637	2.679
70	79.949	9.949	98.991
73	68.348	4.652	21.643
46	43.749	2.251	5.067
13	12.506	0.494	0.244
29	33.259	4.259	18.139
22	35.822	13.822	191.057
37	34.499	2.501	6.254
48	46.354	1.646	2.708
14	35.476	21.476	461.229
66	74.780	8.780	77.092
23	27.432	4.432	19.642
24	37.944	13.944	194.441
40	37.116	2.884	8.317
61	44.362	16.638	276.823
24	27.962	3.962	15.699
53	38.176	14.824	219.759
53	49.815	3.185	10.143
67	31.667	35.333	1248.448
15	43.861	28.861	832.972
11	29.849	18.849	355.291
62	60.827	1.173	1.377
13	23.777	10.777	116.147
26	20.288	5.712	32.627
39	36.020	2.980	8.879
58	56.987	1.013	1.027
82	60.000	22.000	484.001
52	45.411	6.589	43.413
48	44.048	3.952	15.615
61	67.142	6.142	37.728
48	41.335	6.665	44.416
18	29.536	11.536	133.075
46	56.500	10.500	110.252
38	40.028	2.028	4.112
60	34.803	25.197	634.894
26	33.038	7.038	49.528
39	41.968	2.968	8.811



21	72.060	51.060	2607.082
48	41.391	6.609	43.679
1	7.524	6.524	42.566
14	3.550	10.450	109.195
43	30.085	12.915	166.785
57	54.020	2.980	8.881
21	21.136	0.136	0.019
82	64.041	17.959	322.540
49	55.039	6.039	36.470
17	34.116	17.116	292.957
51	50.337	0.663	0.440
44	31.573	12.427	154.437
47	45.213	1.787	3.195
37	75.252	38.252	1463.220
45	54.970	9.970	99.407
65	53.137	11.863	140.724
36	32.373	3.627	13.158
24	36.081	12.081	145.941

表 4.10 捷運忠孝復興站周邊 YouBike 站點還車量測試結果指標

MAE	MSE	RMSE
12.033	265.711	16.301

捷運動物園站周邊的 YouBike 還車量預測分布為右偏態(圖 4.9)，但模型仍有學習相對應特徵，從測試集(表 4.9)觀察，主要幾個指標分別 MAE 為 3.923 及 RMSE 為 4.717(表 4.10)，捷運動物園站周邊 YouBike 車站每天平均約有 7.065 台的還車量，不過誤差值仍是在可以被接受的範圍內。

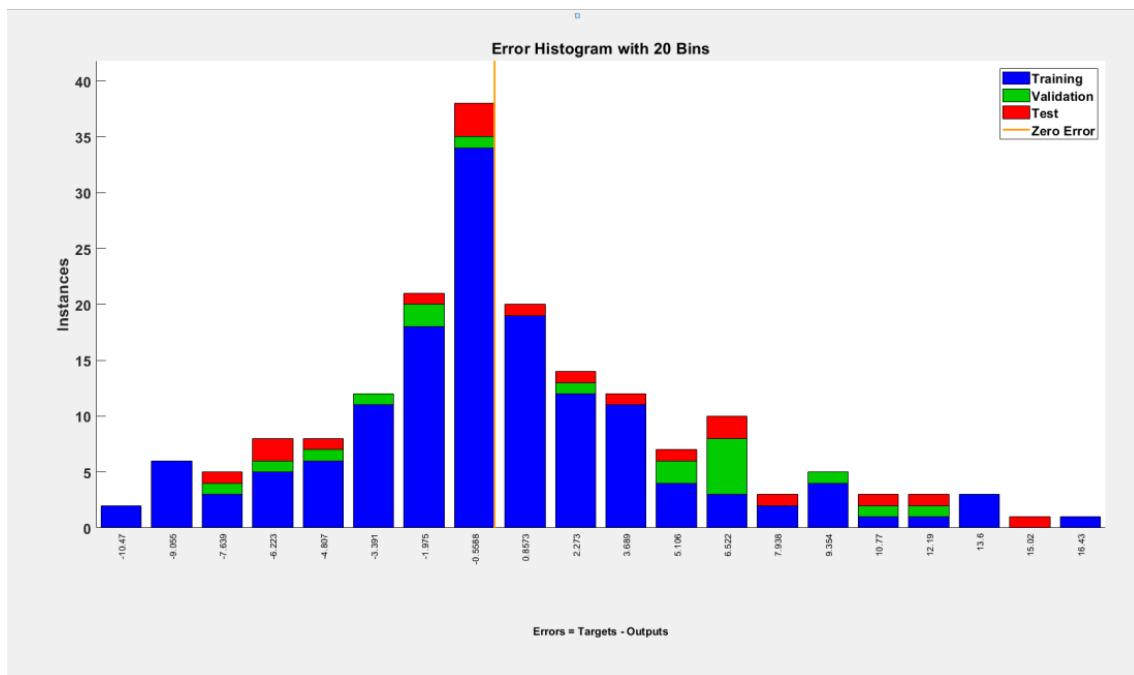


圖 4.8 誤差值直方圖-捷運動物園站周邊 YouBike 還車

表 4.11 捷運動物園站周邊 YouBike 站點還車量測試結果

下半天預測人次	預測結果	誤差	平方差
1	3.03071682	2.030717	4.123811
2	3.57916062	1.579161	2.493748
18	17.9752758	0.024724	0.000611
5	11.3745561	6.374556	40.63497
5	3.30908581	1.690914	2.859191
4	3.51299654	0.487003	0.237172
3	5.30940735	2.309407	5.333362
0	3.21692585	3.216926	10.34861
1	9.91683219	8.916832	79.5099
6	3.03206986	2.96793	8.808609
11	5.75209691	5.247903	27.54049
11	3.02972872	7.970271	63.52522
8	2.21765553	5.782344	33.43551
3	8.16145521	5.161455	26.64062
0	2.46189183	2.461892	6.060911
4	2.33168703	1.668313	2.783268
7	2.2589831	4.741017	22.47724
1	8.98212961	7.98213	63.71439

表 4.12 捷運動物園站周邊 YouBike 站點還車量測試結果指標

MAE	MSE	RMSE
3.923	22.252	4.717

#### 4.2.2 捷運站周邊 YouBike 借車預測結果分析

忠孝復興站周邊的 YouBike 借車量預測狀況為常態分布(圖 4.10)，證明模型是有效的學習，並從測試集(表 4.13)觀察，主要幾個指標分別 MAE 為 14.427 及 RMSE 為 20.853(表 4.14)，由於忠孝復興站周邊 YouBike 車站每天平均約有 41.537 台的借車量，因此誤差值是可以被接受的範圍內。

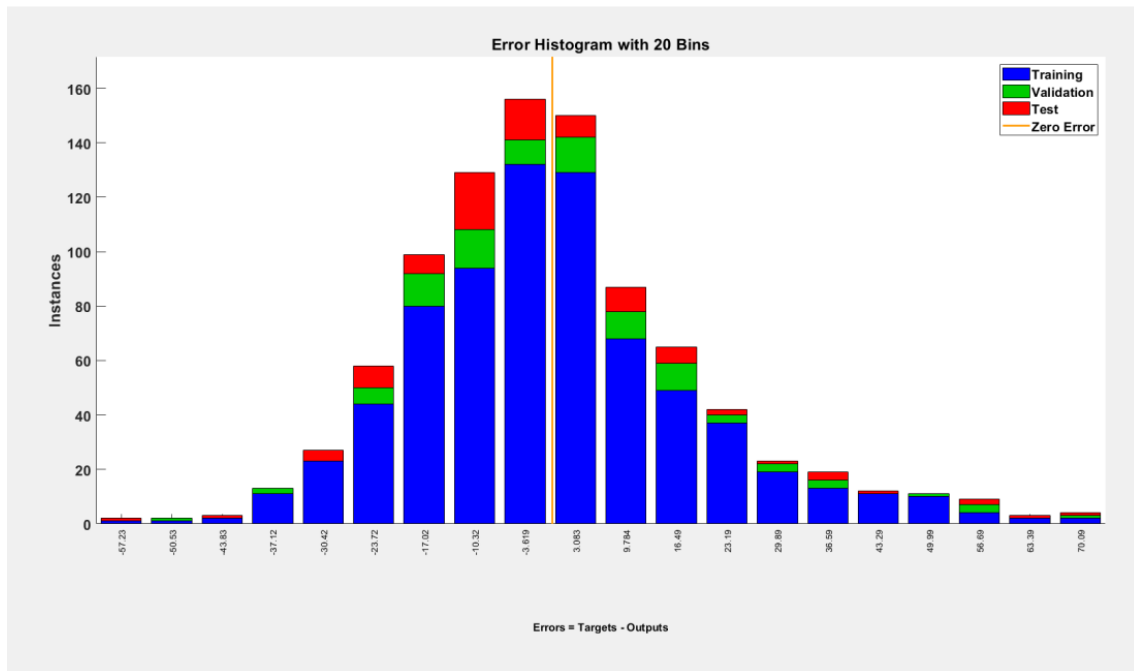


圖 4.9 誤差值直方圖-捷運忠孝復興站周邊 YouBike 借車

表 4.13 捷運忠孝復興站周邊 YouBike 站點借車量測試結果

下半天預測人次	預測結果	誤差	平方差
97	32.677	64.323	4137.505
65	48.536	16.464	271.060
53	51.161	1.839	3.380
70	48.070	21.930	480.914
7	17.238	10.238	104.808
36	37.320	1.320	1.743
21	25.995	4.995	24.951
20	21.217	1.217	1.481
7	12.053	5.053	25.529
31	12.086	18.914	357.730
36	39.423	3.423	11.714
25	49.593	24.593	604.800
32	41.136	9.136	83.459
31	36.165	5.165	26.681
83	60.304	22.696	515.092
34	32.605	1.395	1.946
41	46.979	5.979	35.748
37	37.244	0.244	0.059
33	54.701	21.701	470.938
115	49.002	65.998	4355.714
36	30.154	5.846	34.180
12	38.115	26.115	681.984
5	33.605	28.605	818.262
20	18.877	1.123	1.261
36	41.685	5.685	32.320
36	52.899	16.899	285.576
23	53.435	30.435	926.298
32	33.456	1.456	2.119
44	49.818	5.818	33.846
41	52.960	11.960	143.036
36	42.784	6.784	46.020
46	49.121	3.121	9.739
29	29.359	0.359	0.129
27	53.212	26.212	687.077
45	44.877	0.123	0.015

23	47.235	24.235	587.330
37	48.717	11.717	137.290
77	64.911	12.089	146.149
69	50.552	18.448	340.315
72	38.862	33.138	1098.107
38	39.384	1.384	1.914
12	51.769	39.769	1581.559
63	34.118	28.882	834.164
42	32.075	9.925	98.515
34	43.844	9.844	96.896
82	65.745	16.255	264.232
34	49.513	15.513	240.646
29	17.041	11.959	143.021
7	47.283	40.283	1622.718
13	8.480	4.520	20.434
27	43.744	16.744	280.373
103	43.051	59.949	3593.854
39	40.664	1.664	2.768
32	41.786	9.786	95.756
16	18.139	2.139	4.574
79	35.268	43.732	1912.518
45	44.015	0.985	0.971
66	50.660	15.340	235.306
40	37.398	2.602	6.773
39	32.206	6.794	46.152
44	33.911	10.089	101.787
35	35.972	0.972	0.945
59	52.272	6.728	45.267
29	30.592	1.592	2.534
47	37.120	9.880	97.622
27	34.535	7.535	56.771
92	52.415	39.585	1566.962
73	61.761	11.239	126.321
28	52.808	24.808	615.423
18	27.648	9.648	93.077
25	12.638	12.362	152.817
10	26.444	16.444	270.400
17	11.166	5.834	34.030

110	34.868	75.132	5644.854
43	34.552	8.448	71.370
21	38.865	17.865	319.168
17	25.543	8.543	72.979
36	16.022	19.978	399.115
40	35.241	4.759	22.651
35	32.898	2.102	4.417
67	52.230	14.770	218.161
30	39.523	9.523	90.697
40	49.997	9.997	99.949
36	26.088	9.912	98.240
27	40.130	13.130	172.397
35	37.311	2.311	5.339
60	43.922	16.078	258.506
23	24.825	1.825	3.331
21	34.342	13.342	178.008
25	27.057	2.057	4.232
40	51.532	11.532	132.986

表 4.14 捷運忠孝復興站周邊 YouBike 站點借車量測試結果指標

MAE	MSE	RMSE
14.427	434.833	20.853

捷運動物園站周邊的 YouBike 借車量預測分布為右偏態(圖 4.11)，但模型仍有學習相對應特徵，從測試集(表 4.15)觀察，主要幾個指標分別 MAE 為 3.981 及 RMSE 為 4.707(表 4.16)，捷運動物園站周邊 YouBike 車站每天平均約有 7.321 台的借車量，不過誤差值仍是在可以被接受的範圍內。



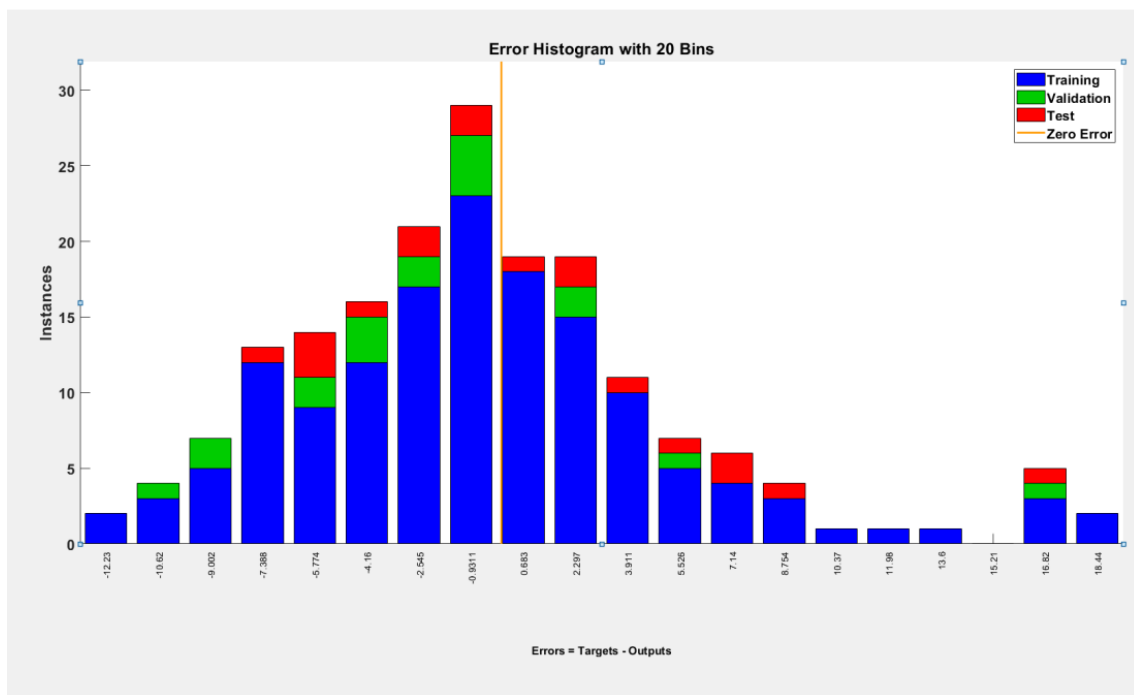


圖 4.10 誤差值直方圖-捷運動物園站周邊 YouBike 借車

表 4.15 捷運動物園站周邊 YouBike 站點借車量測試結果

下半天預測人次	預測結果	誤差	平方差
8	4.010120326	3.990	15.919
0	6.910004408	6.910	47.748
1	11.48804524	10.488	109.999
19	14.35854927	4.641	21.543
12	18.16167321	6.162	37.966
7	9.080739254	2.081	4.329
10	11.34882156	1.349	1.819
8	5.255645201	2.744	7.531
8	5.933255092	2.067	4.271
6	7.988924249	1.989	3.956
16	11.2992655	4.701	22.097
7	7.905511982	0.906	0.820
11	7.938907249	3.061	9.370
20	12.19641886	7.804	60.896
3	8.801951143	5.802	33.663
1	3.22865408	2.229	4.967
6	4.244345276	1.756	3.082
8	5.014939414	2.985	8.911

表 4.16 捷運動物園站周邊 YouBike 站點借車量測試結果指標

MAE	MSE	RMSE
3.981	22.160	4.707

經過以上兩個不同群體代表的 YouBike 借還車測試結果，均能夠在合理範圍內預測借車量與還車量，因此驗證此類神經時間序列模型在 YouBike 資料集的預測可行性。但可能因為忠孝復興站周邊的 YouBike 站點較多，資料複雜度較完整，所以忠孝復興站周邊的 YouBike 車站預測結果相較於動物園站預測結果較佳。此外，也驗證能夠將第一階段捷運進出站模型的人次預測結果作為特徵值，建立第二階段 YouBike 借還車模型，並且此兩階段模型的預測是可行的方法。

#### 4.2.3 捷運站周邊不同步行距離的 YouBike 使用量預測

作為實驗模型對照組，本研究同時找出捷運站半徑 500 以上至 1000 公尺內的 YouBike 站點，並亦建立借還車預測模型。

捷運忠孝復興站 500 公尺外的 YouBike 還車量預測分布為右偏態(圖 4.12)，但模型仍有學習相對應特徵，從測試集(表 4.17)觀察，主要幾個指標分別 MAE 為 16.530 及 RMSE 為 21.551(表 4.18)，忠孝復興站 500 公尺外的 YouBike 車站每天平均約有 31.666 台的還車量，誤差值仍是在可以被接受的範圍內。

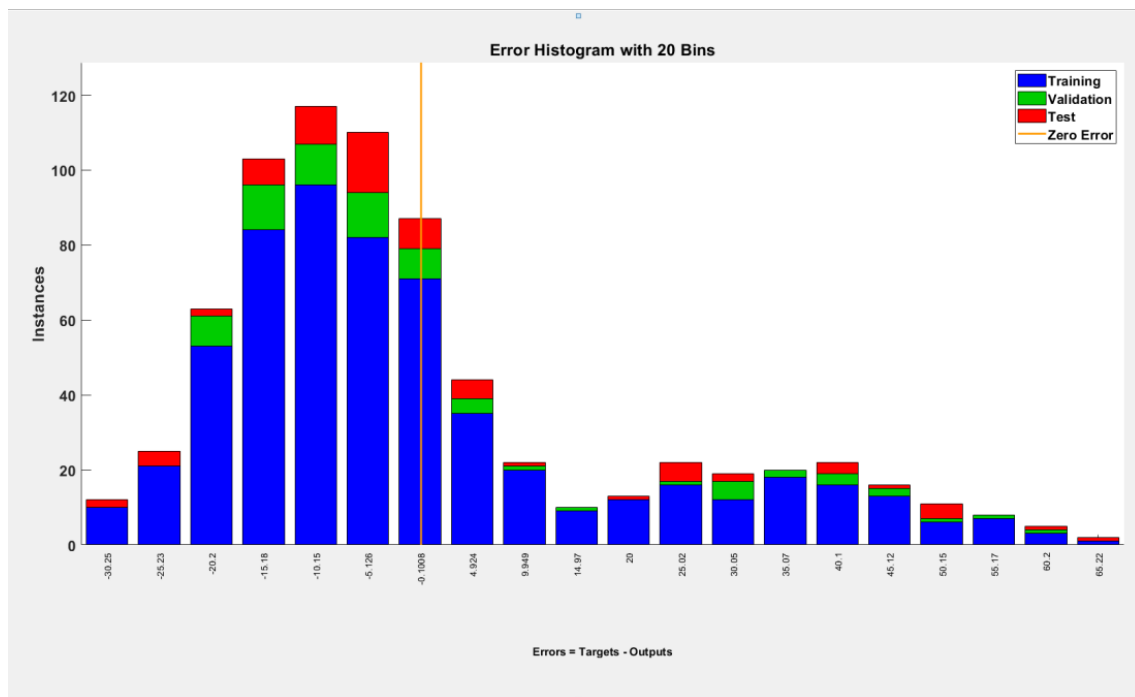


圖 4.11 誤差值直方圖-捷運忠孝復興站周邊 500 公尺外 YouBike 還車



表 4.17 捷運忠孝復興站周邊 500 公尺外 YouBike 站點還車量測試結果

下半天預測人次	預測結果	誤差	平方差
35	30.200	4.800	23.040
55	25.383	29.617	877.178
18	27.656	9.656	93.234
23	40.203	17.203	295.939
58	40.338	17.662	311.945
14	25.539	11.539	133.148
7	28.637	21.637	468.173
22	30.312	8.312	69.083
23	22.927	0.073	0.005
40	31.434	8.566	73.384
7	30.034	23.034	530.561
52	38.740	13.260	175.820
23	34.293	11.293	127.536
11	25.677	14.677	215.401
21	32.237	11.237	126.274
84	29.780	54.220	2939.834
29	32.927	3.927	15.419
19	36.272	17.272	298.314
93	37.649	55.351	3063.681
19	29.611	10.611	112.602
68	28.097	39.903	1592.243
23	30.093	7.093	50.307
69	30.142	38.858	1509.952
72	25.827	46.173	2131.938
61	28.642	32.358	1047.035
8	25.383	17.383	302.163
14	28.300	14.300	204.489
22	34.293	12.293	151.123
33	38.484	5.484	30.078
85	38.548	46.452	2157.765
14	28.244	14.244	202.902
21	22.901	1.901	3.615
15	25.359	10.359	107.301
22	31.441	9.441	89.130
64	30.590	33.410	1116.213
11	30.108	19.108	365.106

17	34.293	17.293	299.055
16	29.502	13.502	182.311
16	25.523	9.523	90.696
16	24.642	8.642	74.677
23	40.128	17.128	293.384
36	39.828	3.828	14.651
29	34.293	5.293	28.018
26	29.939	3.939	15.513
27	37.807	10.807	116.797
7	29.332	22.332	498.715
23	33.760	10.760	115.779
5	35.487	30.487	929.461
6	27.647	21.647	468.591
36	30.351	5.649	31.906
25	30.363	5.363	28.762
24	26.827	2.827	7.990
39	35.988	3.012	9.074
25	35.952	10.952	119.937
88	33.765	54.235	2941.478
33	30.020	2.980	8.878
24	25.534	1.534	2.353
74	25.513	48.487	2350.956
11	25.383	14.383	206.861
19	29.060	10.060	101.205
30	30.954	0.954	0.909
10	30.906	20.906	437.040
12	34.293	22.293	496.986
14	30.149	16.149	260.779
18	29.955	11.955	142.912
69	30.085	38.915	1514.410
35	37.304	2.304	5.307
38	34.293	3.707	13.740
11	32.847	21.847	477.306
30	31.183	1.183	1.399
42	29.396	12.604	158.867
19	25.383	6.383	40.740
10	30.130	20.130	405.228

表 4.18 捷運忠孝復興站周邊 500 公尺外 YouBike 站點還車量測試結果指標

MAE	MSE	RMSE
16.530	464.447	21.551

捷運忠孝復興站 500 公尺外的 YouBike 借車量預測分布為右偏態(圖 4.13)，但模型仍有學習相對應特徵，從測試集(表 4.19)觀察，主要幾個指標分別 MAE 為 11.916 及 RMSE 為 15.041(表 4.20)，忠孝復興站 500 公尺外的 YouBike 車站每天平均約有 30.296 台的借車量，誤差值仍是在可以被接受的範圍內。

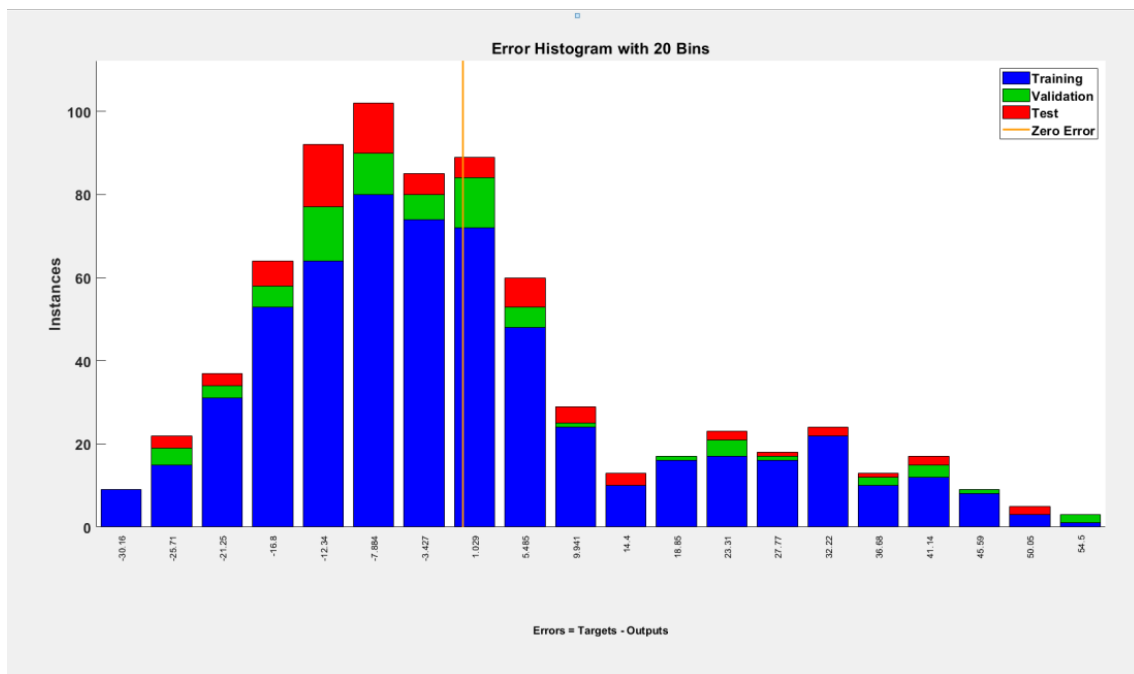


圖 4.12 誤差值直方圖-捷運忠孝復興站周邊 500 公尺外 YouBike 借車



表 4.19 捷運忠孝復興站周邊 500 公尺外 YouBike 站點借車量測試結果

下半天預測人次	預測結果	誤差	平方差
31	26.961	4.039	16.313
26	30.045	4.045	16.365
33	23.899	9.101	82.825
11	29.011	18.011	324.401
43	27.769	15.231	231.995
12	31.996	19.996	399.822
28	29.864	1.864	3.475
20	28.308	8.308	69.022
15	30.901	15.901	252.857
15	22.223	7.223	52.168
19	18.381	0.619	0.383
33	29.994	3.006	9.037
17	24.249	7.249	52.555
55	27.107	27.893	778.046
16	30.161	14.161	200.540
40	34.389	5.611	31.480
34	29.416	4.584	21.015
31	24.596	6.404	41.015
42	37.403	4.597	21.129
41	36.588	4.412	19.462
78	36.156	41.844	1750.918
38	26.612	11.388	129.693
3	29.920	26.920	724.713
65	32.677	32.323	1044.794
28	38.151	10.151	103.047
38	28.120	9.880	97.609
4	28.426	24.426	596.634
8	26.944	18.944	358.864
38	29.604	8.396	70.499
19	29.895	10.895	118.698
11	28.401	17.401	302.797
25	34.506	9.506	90.371
17	31.915	14.915	222.455
35	27.515	7.485	56.033
16	33.593	17.593	309.508
30	28.469	1.531	2.343

25	29.274	4.274	18.267
22	33.763	11.763	138.370
34	27.947	6.053	36.638
19	28.343	9.343	87.300
16	22.000	6.000	35.996
32	29.426	2.574	6.628
21	21.324	0.324	0.105
22	30.004	8.004	64.065
23	23.708	0.708	0.501
23	36.669	13.669	186.839
20	31.385	11.385	129.627
16	27.743	11.743	137.907
28	23.084	4.916	24.164
18	30.778	12.778	163.285
71	37.787	33.213	1103.136
66	26.830	39.170	1534.273
34	29.569	4.431	19.635
23	32.877	9.877	97.556
21	32.386	11.386	129.634
54	31.938	22.062	486.739
30	20.251	9.749	95.034
23	27.180	4.180	17.474
61	29.132	31.868	1015.557
22	32.606	10.606	112.494
31	26.964	4.036	16.292
46	29.670	16.330	266.661
19	30.341	11.341	128.611
55	30.457	24.543	602.381
33	34.166	1.166	1.359
23	30.019	7.019	49.270
15	21.243	6.243	38.974
18	23.396	5.396	29.115
58	37.011	20.989	440.528
48	36.011	11.989	143.732
22	23.910	1.910	3.649
46	27.735	18.265	333.618
18	32.711	14.711	216.400

表 4.20 捷運忠孝復興站周邊 500 公尺外 YouBike 站點還車量測試結果指標

MAE	MSE	RMSE
11.916	226.229	15.041

經忠孝復興站 500 公尺距離以上的 YouBike 借還車測試結果，從 Error Histogram 的表現驗證在同為忠孝復興站周邊 YouBike 車站，500 公尺距離以內的結果較佳，超過 500 公尺的距離可能因步行距離的考量、距離內有更多的替代方式等，因此可以得到預測範圍定義在 500 公尺內會有較好的效果。



## 第五章 結論

### 5.1 研究結論

本研究以捷運資料集與 YouBike 資料集作為分析對象，透過多步驟的資料前處理分別建立不同模型，第一階段以常日週一到週五的捷運進出人流分布進行資料分群，利用輪廓係數、K-means 找出代表群體的站點，再透過這些站點查找到週邊 500 公尺內的 YouBike 站點，以此作為第二階段的分析對象。第二階段先透過資料處理將 YouBike 大量數據分整理為每日上下半天統計，也以此作為 YouBike 借還車的預測時段目標，並且將第一階段各中心捷運站的模型預測進出站人次結果加入，建立一個結合捷運進出站人流與 YouBike 借還車關係的時間序列模型。

就模型預測結果，在第一階段能有效地表現進出站的人次狀況，即使以三個集群中複雜度最高的忠孝復興站與條件最簡單的動物園站來看，相對於平均進出站人次，結果表現亦具參考性。第二階段的部分，雖是基於第一階段的結果作為重要特徵提供模型學習，但看表現也是在可以接受的範圍，不過 500 公尺距離以內的分析結果相對於 500 公尺距離以外的分析結果更具參考性。

更重要的是，因為捷運系統的使用人數眾多，其進出站人數會比 YouBike 借還量更早反映出城市中的突發狀況。換言之，上述作法在城市有突發狀況時，理應能比傳統方法對 YouBike 借還量做更準確的預測，而我們的實驗結果也證實了這點。

### 5.2 研究建議

本研究的實驗分析若有更多時間資源，可以在第一階段的模型加入更多同

一集群的捷運站特徵資訊，預測同一集群代表捷運站的進出站人；在第二階段的實驗亦可以將時間區段增加不同組數，目前以半天作為一個單位，可以嘗試更密集的 6 小時、4 小時或是最小 2 小時為單位的時間區間，作為預測評估的結果比較。



## 參考文獻

1. 吳冠毅(2018)，分類式支援向量迴歸結合粒子群演算法預測 Youbike 車位及車輛數，臺北市立大學資訊科學系研究所碩士論文。
2. 李舒媛(2018)，以悠遊卡大數據探討 YouBike 租賃及轉乘捷運之使用者行為，淡江大學運輸管理學系運輸科學碩士班碩士論文。
3. 黃晏姍(2015)，臺北市公共自行車系統營運特性分析，淡江大學運輸管理學系運輸科學碩士班碩士論文。
4. 顏志偉、簡文彥、賴宗裕、陳芊灼、蘇偉強、黃千倚(2012)，臺北都會區大眾捷運系統周邊土地整體開發計畫之探討，捷運技術半年刊第 46 期。
5. 沈依潔(2014)，民眾步行行為意向之研究—以台北市民為例，國立交通大學運輸科技與管理學系碩士論文。
6. 葉耀墩、林昭賢、劉心荷(2015)，公共自行車 YouBike 營運資料分析及精進作為，新北市政府年度自行研究報告。
7. 翁雅君(2015)，透過時間序列方法分析 YouBike 使用狀況：以信義區為例，國立雲林科技大學資訊管理系碩士論文。
8. 黃政逸(2015)，以系統模擬探討公共自行車之運補策略-以台北市公共自行車為例，逢甲大學工業工程與系統管理學系碩士論文。
9. 李奕譚(2016)，基於公共自行車開放數據之使用量預測研究，國立臺灣大學土木工程學研究所碩士論文。
10. 林冠廷(2018)，運用兩階段資料探勘建構 YouBike 站點需求量與還車量預測之研究，國立雲林科技大學工業工程與管理系碩士論文。



11. 張宸睿(2019)，運用回饋式類神經網路與 K-means 演算法建構 YouBike 站點需求量預測之研究，國立雲林科技大學工業工程與管理系碩士論文。
12. 陳亞伶(2019)，運用模糊時間序列方法與類神經網路於 Youbike 租賃站需求量預測之研究，國立雲林科技大學工業工程與管理系碩士論文。
13. 陳昱瑋(2019)，利用深度學習模型探究交通工具的使用量及其影響因素-以計程車與 Youbike 為例，國立雲林科技大學工業工程與管理系碩士論文。
14. Ward, J.H.(1963), "Hierarchical grouping to optimize an objective function," Journal of the American Statistical Association, Vol. 58, No. 301, pp. 236-244.
15. J. B. MacQueen (1967), "Some methods for classification and analysis of multivariate observations", in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol.1, pp. 281-297.
16. Punj, G. and Stewart, D. W. (1983) Cluster analysis in marketing research: review and suggestions for application, Journal of Marketing Research, pp. 134-148.
17. Anderbeg, M. R. (1973), Cluster Analysis for Applications. New York: Academic Press.
18. Sgarma, S. (1996), "Applied Multivariate Techniques" , John Wiley & Sons, Inc, pp. 212-216.
19. 簡禎富、許嘉裕(2014)，資料挖礦與大數據分析，前程文化。
20. Adams N. M., and Hand D. J. (2000), Improving the practice of classifier performance assessment. Neural Computation, 12, pp. 305-311.
21. Box, G.E.P., G.M. Jenkins (1976), Time Series Analysis Forecasting and Control, 2nd edn., Holden-Day, San Francisco.
22. 陳威昇(2003)，類神經網路訓練程序之些許建議，國立成功大學航空太空工程

研究所碩士論文。

23. 魏宏達(2017)，使用資料探勘分類技術優化 YouBike 運補作業，國立中央大學資訊管理學系在職專班碩士論文。
24. 蘇怡瑄(2020)，運用類神經網路預測捷運車站之運量，國立交通大學交通運輸研究所碩士論文。
25. F. A. Gers and J. Schmidhuber (2020), “Recurrent nets that time and count” in Proceed-ings of the 2000 IEEE-INNS-ENNS Inte rnational Joint Conference on Neural Networks(IJCNN), vol. 3, pp. 189-194.
26. J. Guo, Z. Xie, Y. Qin, L. Jia, and Y. Wang (2020), “Traffic Flow Forecast Through Time SeriesAnalysis Based on Deep Learning” IEEE Access, vol. 8, pp. 82562-82570.

