

澳門旅遊需求研究——基於機器學習及互聯網搜索熱度

王俊皓 苑皓 阮智豪

[摘要] 本研究提出了四種模型，分別是傳統的時間序列模型 ARIMA 以及機器學習類模型隨機森林、LSTM 以及 CNN-LSTM，分別對澳門 2008 年 1 月至 2019 年 12 月的入境總人數序列進行分析，並基於互聯網搜索熱度數據（百度指數、谷歌趨勢），比較不同的搜索引擎熱度代表的海內外遊客搜索偏好的數據。研究結果表明了結合谷歌和百度搜索熱度下的 CNN-LSTM 模型的性能最優，且 CNN-LSTM 模型並不對數據的預處理過程有嚴格的要求，更加適合本研究的場景，即澳門旅遊需求預測。

[關鍵詞] 互聯網搜索熱度 長短期記憶網絡 卷積長短期記憶神經網絡 旅遊需求 機器學習

一、研究背景及意義

（一）研究背景

澳門是一個以旅遊產業為主要經濟支柱的地區，相關產業主要包括博彩、酒店住宿、餐飲、零售以及旅遊服務等。在中國出境旅行不斷深化發展的大環境下，內地遊客前往澳門旅遊的人數逐年上升。自澳門回歸之後，內地遊客逐漸成為了澳門旅遊產業收入的主力軍。來自司爾亞司數據信息有限公司（CEIC）的數據顯示，澳門的旅遊業的收入在 2018 年為 403,58 百萬美元，比上年同期增長了約 10%，並且呈現一個逐年上升的態勢。^①

在受到 COVID-19 疫情的衝擊，2020 年 3 月中旬各國旅遊業遇到了滑鐵盧危機。全球出國旅行人次在 2020 年初的三個月內幾乎腰斬。到了 5 月份隨着疫情逐漸發酵，疫情的嚴重性逐步放大後，這一數據不可避免地下降了 98%。在全球旅遊業幾乎凍結的情況下，專家們預測在 2020 年上半年這段時間，全球的國際貿易活動將會蒙受近 3,200 億美元的損失，這個預計數字是次貸危機給全球帶來的經濟損失的三倍之多。全年的海外出行遊客人次將有可能降至 58% 至 78% 的區間，因為疫情的控制情況及各國嚴格的封閉防疫政策不容樂觀。^②

作者簡介：王俊皓（通訊作者），澳門城市大學金融學院博士研究生；苑皓，澳門科技大學商學院碩士研究生；阮智豪，澳門城市大學商學院博士研究生。

^①澳門旅遊業收入數據來自 CEIC 數據庫，<https://www.ceicdata.com/en>。

^②張靜：〈世界旅遊組織：新冠疫情或將導致今年國際旅客銳減 80%〉，聯合國新聞網頁，2020 年 5 月 7 日，<https://news.un.org/zh/story/2020/05/1056832>。

隨着 COVID-19 疫情的到來，來訪澳門的入境旅遊人數在 2020 年急劇下降。因為防疫政策的關係，《人民日報（海外版）》在 2020 年 9 月 12 日的報道指出，口岸的封閉導致了澳門第二季度的 GDP 下跌了 67.8%，^①澳門整個地區在簽注暫停的一段時間中，幾乎進入了停滯狀態，過往熱鬧非常的大三巴牌坊上空無一人，各大娛樂場的遊客也是寥寥無幾。由此可見，對一個旅遊產業豐富的地區來說，旅遊需求的預判是一個亟需解決的問題。

（二）研究意義

旅遊需求預測是政府制定旅遊業未來宏微觀經濟計劃的關鍵指標之一。作為預測需求的依據，對旅遊目的地之數據特徵、影響因素進行分析，並在預測優度高的基礎上有針對性地探究未來的宏觀政策、創新與發展旅遊項目，可不斷深化和拓寬旅遊業的整體有效市場，使其走向可持續發展的道路。

在科學技術飛快發展的今天，旅遊從業者可透過一些較為精確的模型框架對未來旅遊人數進行需求預測，以便用來進行產品定價、經營計劃等等。本研究採用了近年來廣泛運用於預測的神經網絡模型——長短期記憶網絡（long short-term memory networks）和互聯網搜索熱度指數（internet search popularity index）來預測遊客數量。該模型的預測能力主要歸功於兩點：一是互聯網搜索熱度作為互聯網時代的產物，有效地體現了旅客在出行之前使用搜尋引擎進行計劃的行為，極大地提高了模型的擬合優度；其次，隨機森林模型、LSTM 模型與 CNN-LSTM 模型在複雜的非線性問題上，亦能夠發揮出其固有的長期記憶能力，也因此成為近些年來十分出色的時間序列預測方法之一。目前為止，只有少數研究採用隨機森林模型、LSTM 模型與 CNN-LSTM 模型預測旅遊需求，本研究利用該模型對澳門遊客入境人數進行預測，用於迴歸的相關互聯網搜索關鍵詞涵蓋了旅遊的各個方面，包括餐飲、住宿、娛樂、購物、旅遊和交通。與以往研究不同的是，本研究採用的搜索熱度指標既有反映我國旅客現狀的百度指數，又有反映國外旅客現狀的谷歌趨勢指數，特別符合澳門的遊客結構組成。

二、文獻綜述

本研究搜索了大量旅遊需求相關的文獻，發現主要被用來研究旅遊影響因素以及程度的研究方法主要有時間序列模型（time series model），或者是使用多元線性迴歸（multiple linear regression）對各旅遊需求影響因素進行迴歸分析。除了傳統的經濟學模型以及統計學方法外，也有不少學者利用灰色預測法（grey prediction model）來對旅遊目的地之旅遊人數進行預測，而近年興起的人工神經網絡模型（artificial neural network model），因為它具有良好的自學習、聯繫存儲和高速尋找優化解的功能，所以能夠在多領域解決現代電腦難以解決的實際問題，表現了良好的智慧性，在近幾年的研究方法中是一個首選的方法，如，Xu 等採用人工神經網絡模型研究中國的二氧化碳何時達到峰值，得到的預測結果

^①汪靈犀：〈全力以赴 澳門旅遊業再出發〉，《人民日報（海外版）》（北京），2020 年 9 月 12 日，頁 4。

是在 2029 年、2031 年或 2035 年，中國二氧化碳排放峰值分別會達到 10.08、10.78 和 11.63 億噸，並建議中國需要盡可能使用碳密集度低的燃料減少二氧化碳的排放量，有利於國家的可持續發展。^①

李世霞等以互聯網搜索熱度為研究對象，以青島市為研究背景，探尋影響其線上熱度指標的因素；^②秦夢等通過人工神經網絡模型作為研究方法，對海南三亞的線上互聯網搜索熱度進行收集聚合，對當地的外來旅遊人數進行預測，結果發現在加入百度指數後，可以整體上提高模型的預測準確程度。^③He 等學者基於六個國家和地區到訪澳門旅客人數的日數據，比較 SARIMA 模型、CNN 模型、LSTM 模型及 SARIMA-CNN-LSTM 模型預測六個國家和地區每日到訪澳門旅遊需求的性能，發現 SARIMA-CNN-LSTM 模型比其餘三種單個模型有更好的預測精確度和性能表現。^④綜上可以發現，國內在旅遊需求預測方面對互聯網熱度指標的研究正處於萌發的狀態。

Camacho 和 Pacce 運用動態因子模型（dynamic factor model）和實時數據庫，實時預測和分析西班牙旅客的登記入住逗留情況，發現包含谷歌搜索指數的動態因子模型的評價指標最為優秀，說明有關谷歌的旅遊搜索查詢有助改善西班牙的旅遊預測。^⑤Brynjolfsson 和 Reichman 採用“人群平方”（crowd-squared）的方法，基於人群的旅遊需求意圖、偏好和意見選擇搜索趨勢數據，比較後發現預測效果優於替代數據選擇方法以及使用各種基準數據選擇方法。^⑥Rivera 使用在 SQV 數據（11 個不同場合）通過動態線性模型預測波多黎各的酒店非居民登記數量，與僅使用一次檢索到的 Google 趨勢數據相比，動態線性模型可以更好地推斷酒店非居民登記數量與 Google 趨勢 SQV 之間的關聯。^⑦綜上所述，國外在旅遊需求預測方面對互聯網熱度指標的研究方法比國內更加豐富多樣，研究成果更加成熟。

神經網絡在某些場景的客流量預測能力廣泛受到學者的好評，如在公共交通人群密集度的預測上，Bai 等人以廣州的公共交通為研究對象，採用深度信念網絡（deep belief nets）預測公交某時刻的人群密度，證明了此方法在客流量預測上的高效性。^⑧李梅同樣

^① Xu, Guangyue, Peter Schwarz, and Hualiu Yang. “Determining China’s CO2 Emissions Peak with a Dynamic Nonlinear Artificial Neural Network Approach and Scenario Analysis.” *Energy Policy*, vol. 128, no. 5, 2019, pp. 752-762.

^② 李世霞、田至美：〈基於百度指數的旅遊目的地網絡關注度影響因素分析——以青島為例〉，《首都師範大學學報（自然科學版）》（北京），第 1 期（2014），頁 56—59。

^③ 秦夢、劉漢：〈百度指數、混頻模型與三亞旅遊需求〉，《旅遊學刊》（北京），第 10 期（2019），頁 116—126。

^④ He, Kaijian, and Lei Ji, Chi Wai Don Wu, Kwok Fai Geoffrey Tso. “Using SARIMA-CNN-LSTM Approach to Forecast Daily Tourism Demand.” *Journal of Hospitality and Tourism Management*, vol. 49, no. 3, 2021, pp. 25-33.

^⑤ Camacho, Maximo, and Matías José Pacce. “Forecasting Travellers in Spain with Google’s Search Volume Indices.” *Tourism Economics*, vol. 24, no. 4, 2018, pp. 434-448.

^⑥ Brynjolfsson, E., and Tomer Geva, Shachar Reichman, “Crowd-Squared: Amplifying the Predictive Power of Search Trend Data.” *MIS Quarterly*, vol. 40, no. 4, 2016, pp. 941-962.

^⑦ Rivera, Roberto. “A Dynamic Linear Model to Forecast Hotel Registrations in Puerto Rico using Google Trends Data.” *Tourism Management*, vol. 57, no. 6, 2016, pp. 12-20.

^⑧ Bai, Yun, et al. “A Multi-pattern Deep Fusion Model for Short-term Bus Passenger Flow Forecasting.” *Applied Soft Computing*, vol. 58, no.5, 2017, pp. 669-680.

在此研究領域，採用了不同的神經網絡如 LSTM 等模型，分別對北京西單區域等站點的公交車進行客流量預測。^①把視角從公交車轉移到地鐵，^②張偉林採用了循環神經網絡（recurrent neural network）以及 LSTM 對地鐵的人群密集度進行預測，並提出了數據融合框架。上述學者並沒有在神經網絡模型的結構、演算法上進行深入研究或改良，而是更多地注重對模型在不同研究領域上的應用。這些研究方法確實在某些領域上對需求預測提供了極大的幫助，即成功擬合了高效的預測模型。神經網絡模型通常稱為“黑箱演算法”（black box algorithm），在擬合數據足夠以及模型參數調配合適的情況下，可得到比較優質的預測效果。^③本研究亦是把 LSTM 模型和優化的 CNN-LSTM 模型研究的重心放在對澳門旅遊需求預測準確度上，對其內部原理並無進行深入的改良研究。

三、研究方法及評價指標

首先，本研究對不同關鍵詞與澳門旅遊人數的相關係數進行比較，篩選出相關性足夠高的優良關鍵詞，再對它們進行預處理等；其次，將傳統的時間序列模型 ARIMA 與機器學習類模型隨機森林、LSTM 以及 CNN-LSTM 之間的、在單一時間序列作為特徵時的性能作出比較；最後，設立多組隨機森林模型、LSTM 模型以及 CNN-LSTM 模型，各採用不同的搜索熱度組合進行對照試驗，求出所有模型的評價指標均方根誤差（RMSE）及百分比絕對誤差（MAPE），以確定最優情況下的模型。

為了比較 LSTM 模型、CNN-LSTM 模型、隨機森林模型以及 ARIMA 模型在本研究對澳門旅遊需求預測的性能，本研究選用了兩種評價指標來衡量，分別是 RMSE（root mean square error）和 MAPE（mean absolute percentage error），都是常用的評測模型指標。RMSE 代表預測值以及真實的數據之間的差距，也就是標準差。標準差反映了數據偏離實際水平的程度——假如 RMSE=20，則可以認為模型的迴歸效果與實際數據之間相差 20，因此誤差越大，RMSE 的值越大，可以很好地反映不同模型之間的性能差距；而 MAPE 的核心在於預測值與真實值的絕對差再除以真實值，以此來體現預測值的偏差程度，^④ RMSE 及 MAPE 的計算公式如下：

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (x_t - \hat{x}_t)^2} \quad (1)$$

$$MAPE = \sum_{t=1}^N \left| \frac{x_t - \hat{x}_t}{x_t} \right| \times \frac{100}{N} \quad (2)$$

^①李梅：〈基於深度學習的短時公交客流預測研究〉，博士論文，北京交通大學，2019年，頁56—64。

^②張偉林：〈基於深度學習的地鐵短時客流預測方法研究〉，碩士論文，中國科學院大學，2019年，頁46—51。

^③喬若羽：〈基於神經網絡的股票預測模型〉，《運籌與管理》（合肥），第10期（2019），頁132—140。

^④賈俊平：《統計學》，北京：中國人民大學出版社，2009年，頁127—132。

四、研究設計

(一) 數據收集

本研究以澳門特別行政區為研究地點，澳門的月度入境旅客人次數據均取自澳門特別行政區政府統計暨普查局，包括內地以及國外的所有旅客，時間跨度為 2008 年 1 月至 2019 年 12 月，共 12 年 144 個月。本研究剔除了 2020 年一整年的數據，因為在新冠疫情的影響下，這一年的入境人數並不正常，且極大地受關口封閉等防疫政策所影響。鑑於本研究並沒有討論重大事項對旅遊人數的影響，而是更多地注意互聯網搜索熱度的影響因素，因此本研究也決定不把 2020 年的數據導入模型進行研究。這部分數據將被劃分為訓練集和測試集，訓練集包含了 2008 年 1 月至 2018 年 12 月一共 122 個觀測樣本，測試集包括了 2019 年的 12 個月，以用來驗證模型的準確性。

為了比較不同搜索引擎帶來的結果，本研究選取了全球兩大搜索引擎的搜索熱度數據進行研究，分別是百度指數以及谷歌（Google）趨勢。百度是中國市場佔有率最大的搜索引擎，其搜索熱度很大程度上代表了中國旅客的搜索行為以及意向，而谷歌則是國外最熱門的搜索引擎，它極大地代表了海外遊客的關鍵詞檢索行為。這兩個搜索引擎的搜索熱度都通過各自的專門網址對公眾開放。雖然谷歌趨勢和百度指數各自通過不同的算法去計算他們客戶的搜索熱度，但它們都反映了特定關鍵詞的受歡迎程度和用戶在特定時間的興趣指數。因此，為了比較這兩種搜索引擎對澳門旅遊需求預測的貢獻程度，本研究分別下載了月度的各個關鍵詞的搜索熱度數據，時間跨度與上述澳門入境總人數對應，都為 2008 年到 2019 年共 144 個月的觀測值。

(二) 互聯網搜索熱度指標的處理

本研究遵循五個階段，對百度和谷歌提供的搜索關鍵詞熱度數據、以澳門為研究背景進行關鍵詞的選取、篩選以及最後熱度數據的整合：

(1) 本研究根據 24 個基礎的搜索關鍵詞的搜索熱度數據進行收集，關鍵詞內容涵蓋了旅客出行的基本需求，例如旅遊、交通、購物等關鍵詞（表 1）。^①

(2) 根據 24 個基礎關鍵詞，在百度指數以及谷歌趨勢中進行檢索，得到本研究需要的搜索熱度數據。其次，在上述操作過程中，搜索引擎會在每一個關鍵詞的頁面，提供相關的一些其他關鍵詞：以關鍵詞“澳門旅遊”為例，網頁上將會推薦不少類似於“澳門一日遊”、“澳門旅遊攻略”、“澳門自由行”等相關關鍵詞，並按熱度排行榜進行排序。本研究以 24 個基礎關鍵詞為起點，探尋更顯著更有利於後續模型的關鍵詞，並把它記錄下來。整個過程就是不斷地搜索和選擇搜索引擎推薦的新關鍵詞，其中，一些熱度較低的關鍵詞以及缺失熱度數據的關鍵詞將被去除。最後總的關鍵詞熱度獲取數量為 102 個，其中百度指數佔 57 個，谷歌趨勢佔 45 個。

^① Li, Xiu, and Bing Pan, Rob Law, Xiankai Huang. “Forecasting Tourism Demand with Composite Search Index.” *Tourism Management*, vol. 59, no. 7, 2017, pp. 57-66.

(3) 接下來需要篩選出這 102 個關鍵詞中相對更合適的關鍵詞，以提升是次澳門旅遊需求預測模型的優度。首先將所有下載的數據在時間上對應，因為百度提供的數據是日度的，而谷歌趨勢以及澳門統計暨普查局提供的數據是月度的，為此本研究採取了加權平均的方法，把百度指數提供的日度數據按每個月加總再求取平均值，得到百度指數的月度數據。之後，再求出每個關鍵詞搜索熱度與澳門入境總人數的皮爾森相關係數，以及每個關鍵詞熱度的不同滯後期與入境總人數的皮爾森相關係數。對於每個關鍵詞，將會有 4 個相關係數，對應 4 個分別是 0—3 個月的滯後期。最後，根據相關係數的大小設定了一個閾值，百度指數的閾值為 0.7 以上，谷歌趨勢的閾值為 0.4 以上，因為谷歌關鍵詞搜索熱度與澳門旅遊人數的相關性往往不高，但百度的關鍵詞普遍較高——可能是因為內地遊客前往澳門旅遊相比海外遊客的距離更近，簽證手續更加便捷，因此在感興趣搜索過後，實現旅行的可能性也更大。最後，本研究篩選出了 33 個百度指數的關鍵詞（表 2），以及 16 個谷歌趨勢的關鍵詞（表 3）。

(4) 為了更準確地預測出澳門旅遊需求人數，本研究只選擇滯後期大於 1 的搜索熱度數據，因為搜索引擎公開的數據通常是在月底統計的，因此滯後期為 0 的數據並不能使模型更有說服力。

(5) 在最後的一個數據處理階段，本研究將兩個搜索引擎的關鍵詞熱度進行整合加總，得到兩個時間序列，再對這兩個時間序列以及澳門入境總人數一共三個時間序列取它們的對數形式，透過其關係圖可以發現因為其相關性較高，三條曲線都呈現出一定的相同趨勢（圖 1）

表 1 24 個基礎關鍵詞

No.	關鍵詞	No.	關鍵詞	No.	關鍵詞
1	澳門旅遊	9	澳門航線	17	澳門購物
2	澳門天氣	10	澳門公交車	18	威尼斯人
3	澳門地圖	11	澳門通關時間	19	拱北口岸
4	澳門交通規則	12	澳門酒吧	20	豬肉脯
5	澳門美食	13	澳門水舞間	21	新濠天地
6	澳門西多士	14	澳門旅行團	22	葡國雞
7	澳門景點	15	大三巴	23	澳門賽車
8	澳門簽證	16	澳門表演	24	巴黎人

表 2 百度指數篩選後關鍵詞

No.	關鍵詞	Lag	No.	關鍵詞	Lag
1	澳門	2	18	澳門葡京	1
2	澳門 DFS	3	19	澳門簽注	3
3	澳門幣	3	20	澳門手信	3
4	澳門兌人民幣匯率	1	21	澳門塔	3
5	澳門城市大學	1	22	澳門特產	3
6	澳門大學	1	23	澳門天氣	3
7	澳門代購	3	24	澳門天氣預報	1
8	澳門電影院	3	25	澳門星際	2
9	澳門購物攻略	3	26	澳門一日遊	3
10	澳門酒店	3	27	澳門住宿	3
11	澳門科大醫院	2	28	大三巴	3
12	澳門科技大學	1	29	港澳通行證	3
13	澳門旅遊景點	3	30	拱北	2
14	澳門美食	3	31	拱北口岸	3
15	拱北口岸通關時間	1	32	鉅記手信	3
16	橫琴口岸	1	33	美高梅	1
17	西多士	1	/	/	/

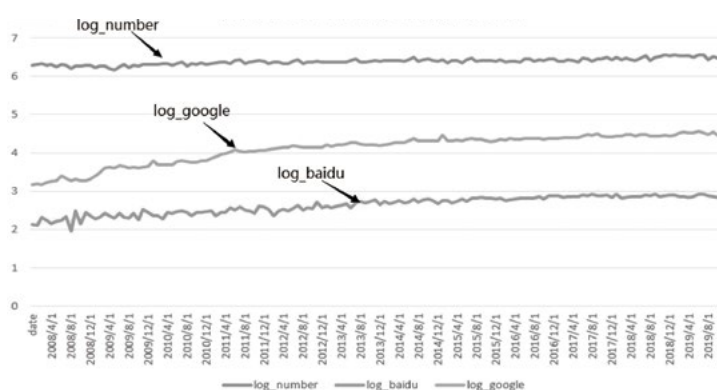
資料來源：百度指數。

表 3 谷歌趨勢篩選後關鍵詞

No.	關鍵詞	Lag	No.	關鍵詞	Lag
1	macau passport	3	9	despegar	3
2	macau_togel	3	10	yelp	3
3	things to do in macau	3	11	airbnbs	1
4	how to get to macau	2	12	tripadvisor	3
5	what to do in macau	3	13	agoda	1
6	tile macao	2	14	booking	1
7	tripadvisor hk	3	15	macau food	1
8	airbnb_macau	2	16	airbnb_hk	3

資料來源：谷歌趨勢。

圖 1 澳門旅遊人數與百度、谷歌搜索熱度之序列圖



在上述處理數據的第三個階段，本研究對 102 個關鍵詞分別按它們的不同滯後期求出其與澳門入境旅遊人數的相關係數（表 4、5），結果顯示百度的相關係數比谷歌顯著。在所有收集的關鍵詞搜索熱度中，相關性超過 0.7 的關鍵詞，百度有 33 個，而谷歌超過 0.4 的也只有 16 個。

表 4 谷歌趨勢部分關鍵詞與旅遊人數相關係數

關鍵詞	澳門旅遊人數
Number of tourists	1.0000
Z_Macau_travel	-0.3970
Z_Macau_travel (t-1)	-0.3875
Z_Macau_travel (t-2)	-0.4008
Z_Macau_travel (t-3)	-0.2472
Macau	0.2300
Macau (t-1)	0.2763
Macau (t-2)	0.1674
Macau (t-3)	0.1319
Macau_weather	-0.0687
Macau_weather (t-1)	-0.0962
Macau_weather (t-2)	-0.1189
Macau_weather (t-3)	-0.0845
Macau_ferry	-0.1592
Macau_ferry (t-1)	-0.1301
Macau_ferry (t-2)	-0.1287
Macau_ferry (t-3)	-0.1371

表 5 百度指數部分關鍵詞與旅遊人數相關係數

關鍵詞	澳門旅遊人數	澳門	澳門 (t-1)
澳門旅遊人數	1.0000	0.6837	0.7112
澳門	0.6837	1.0000	0.7302
澳門 (t-1)	0.7112	0.7302	1.0000
澳門 (t-2)	0.7355	0.7153	0.8284
澳門 (t-3)	0.6959	0.7098	0.8160
澳門 DFS	0.8305	0.7308	0.7380
澳門 DFS (t-1)	0.7990	0.7127	0.7581
澳門 DFS (t-2)	0.7889	0.7152	0.7533
澳門 DFS (t-3)	0.7990	0.7016	0.7442
澳門巴黎人	0.5542	0.5007	0.6001

以上是本研究在實證分析之前對互聯網搜索熱度的一些預處理過程，後續將對這些數據通過傳統時間序列模型 ARIMA 及機器學習類模型隨機森林、LSTM 以及 CNN-LSTM 分別進行擬合，再利用擬合好的模型進行預測，比較傳統模型與新興機器學習類模型之間的差別，以及不同的互聯網搜索引擎提供的搜索熱度對模型的預測能力是否有差異。

五、實證分析

(一) ARIMA 模型對澳門旅遊需求預測

(1) ADF 單位根檢驗

若需使用時間序列類模型進行預測，那麼導入模型的時間序列必須是平穩的，否則將導致偽迴歸的後果。因此需要對上述已預處理的澳門入境總人數進行 ADF 檢驗，查看它們是否已經平穩，結果顯示，從圖形向上的均值趨勢來看，澳門入境總人數時間序列都是不平穩的，其統計量為 -0.1097，已經超過了任何置信區間，無法拒絕原假設，代表旅遊入境人數是不平穩的（表 6），需要對澳門入境總人數進行處理。鑑於之前已經將其轉換成了對數形式，接下來需要對它進行一階差分的處理。在進行了一階差分後，再對其進行一次 ADF 單位根檢驗（表 7），發現澳門入境人數已經趨於平穩。

表 6 澳門入境總人數單位根檢驗結果

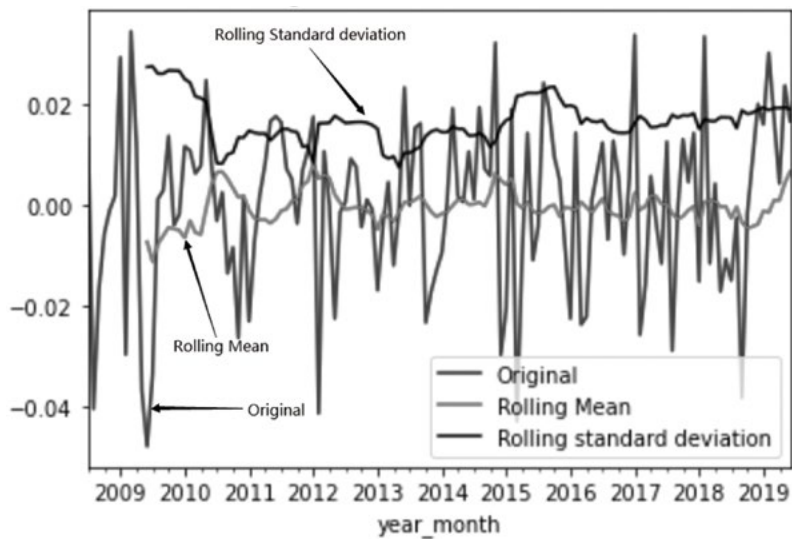
ADF 統計量	1% 臨界值	5% 臨界值	10% 臨界值	P 值
-1.097	-3.482	-2.884	-2.578	0.716

表 7 澳門入境總人數的對數單位根檢驗結果

ADF 統計量	1% 臨界值	5% 臨界值	10% 臨界值	P 值
-3.740	-3.482	-2.884	-2.578	0.004

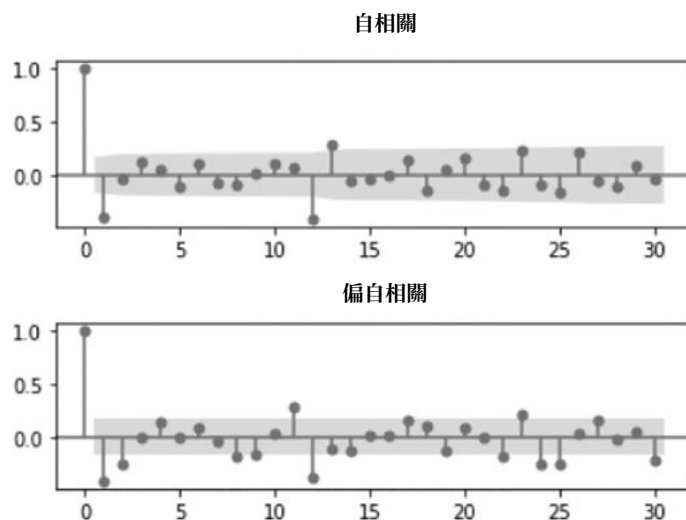
澳門入境人數在差分後，其均值以及方差已經趨於一個常數，且值極小因此可以忽略不計（圖 2）。

圖 2 澳門旅遊人數差分後的各類指標



此外，數據的自相關（autocorrelation）以及偏自相關（partial autocorrelation）係數都有截尾的特點，且明顯在一階具有高度的相關性，因此本研究選擇 $q=1$ ， $p=1$ 的 ARIMA(1,1,1) 模型，且因為數據進行過一階差分，所以 $d=1$ （圖 3）。

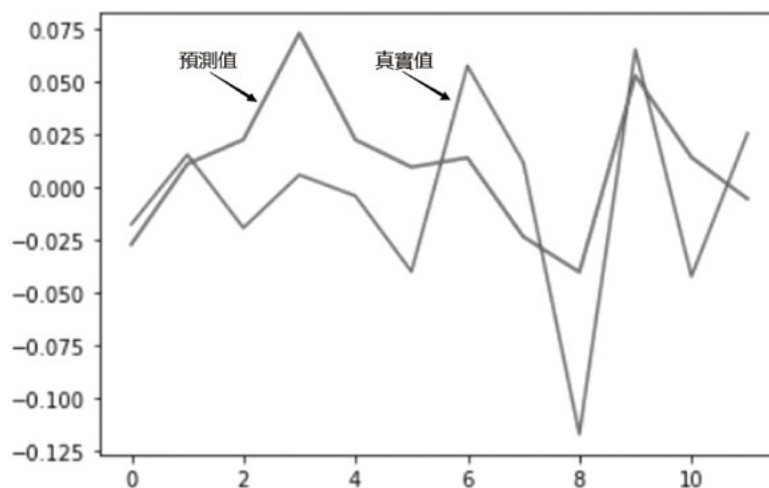
圖 3 澳門旅遊人數的自相關和偏自相關



(2) 預測分析

在確保了數據的平穩後，將開始通過 ARIMA 模型對澳門旅遊總人數進行模型擬合預測。首先對 2008 年到 2019 年的月度數據進行一個劃分，隨後用前面 11 年的數據作為訓練集，導入模型進行擬合，再利用擬合好的模型對 2019 年的 12 個月份進行預測（圖 4）。

圖 4 ARIMA 模型對澳門旅遊總人數的預測結果



此外，通過 Python 計算出了此模型的百分比絕對誤差為 69.982%，可以說是一個誤差極大的預測模型。產生這個結果的可能原因有幾點：第一，數據樣本過小。澳門統計暨普查局提供的入境數據只有 2008 年至今的月度數據，在數據量不夠龐大的情況下，不利於傳統模型進行建模預測。第二，此前許多研究往往在擬合模型時把測試集數據也納入進訓練集，此做法的結果雖然會使模型的預測能力大大提高，但是此模型預測能力非常局限，

即只能在樣本中的測試集發揮高性能，若讓其預測一段其未接觸過的數據，準確度一樣不會十分優秀。最後，傳統的時間序列預測模型 ARIMA 在預測能力上有所欠缺，但在經濟學等領域上使用作出一些迴歸分析還是比較出色的。綜上所述，ARIMA 模型雖然能夠基本完成澳門旅遊需求的預測，但是其精確度不高，在此領域上並不是一個最好的選擇。

(二) 隨機森林模型對澳門旅遊需求預測

本節將搭建四個隨機森林模型進行對比分析（表 8）：第一個模型僅包含了澳門入境人數的時間序列作為特徵變量；第二個模型包含了入境人數以及谷歌搜索熱度兩個時間序列作為特徵變量；第三個模型包含了入境人數以及百度搜索熱度兩個時間序列作為特徵變量；最後一個模型包含入境人數以及兩個搜索引擎的搜索熱度作為特徵變量，一共三個變量。

表 8 四個隨機森林模型訓練變量分配

模型名稱	訓練變量
RF Model 1	澳門入境總人數
RF Model 2	澳門入境總人數 + 谷歌搜索熱度
RF Model 3	澳門入境總人數 + 百度搜索熱度
RF Model 4	澳門入境總人數 + 百度搜索熱度 + 谷歌搜索熱度

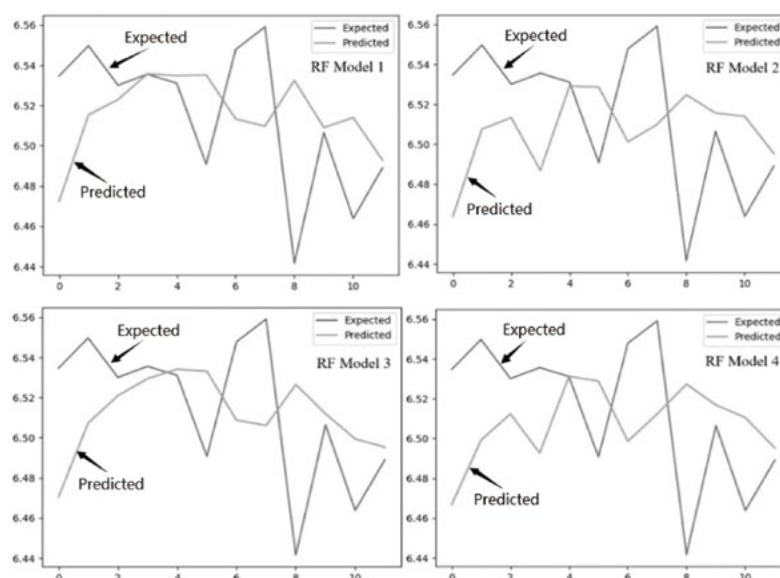
本次研究使用 Python 作為研究工具，對於四個模型統一選用的數據集為 2008 年 1 月至 2019 年 12 月澳門入境旅遊總人數，以及百度、谷歌所提供的搜索熱度，共計 144 行。鑑於要保證實驗研究的高有效性，將對數據集如上一節同樣方法進行劃分，以免隨機森林模型最後出現過擬合的情況，即避免無效的高準確率。相當於用前 11 年的數據導入模型訓練，根據模型通過 11 年數據調整好的權重，預測最後一年的數據，模型訓練及預測如下：

本次實驗劃分了四組模型進行分別預測，目的是為了更好和更直觀地發現不同變量之間的差異，能夠給預測的準確度帶來甚麼影響。在本研究的三個變量中，百度的搜索熱度指標在一定程度上代表了內地對澳門的感興趣程度，且前往澳門出境旅遊的人群中，內地遊客是一個非常大的佔比。除此之外，境外慕名而來的外國遊客也是不可忽視的一個群體，外國遊客往往使用聞名海外的谷歌瀏覽器，因此谷歌的搜索熱度代表了外國人群對澳門在一定時間內的感興趣程度。

第一組模型 RF Model 1 只包含了旅遊入境人數這一特徵變量，目的是為了與傳統計量學模型進行比較，探究在旅遊需求預測的問題上，哪種研究方法更適合處理此類問題。第二組模型 RF Model 2 使用了谷歌搜索熱度及入境人數作為特徵進行訓練，目的是與第三組模型 RF Model 3 使用百度及入境人數作為特徵的模型進行對比，探究作為兩大搜索引擎，

哪個更適合應用在澳門旅遊需求預測問題上，以及哪類搜索引擎提供的數據對模型預測精度的貢獻更大。最後，再使用所有三個時間序列，探究在本次需求預測問題上，兩大搜索熱度之間的關係是互相促進、互相拖累，還是存在多重共線性，以下是 Python 得到的四個隨機森林的預測結果（圖 5）。

圖 5 四個隨機森林模型預測結果



（三）LSTM 模型對澳門旅遊需求預測

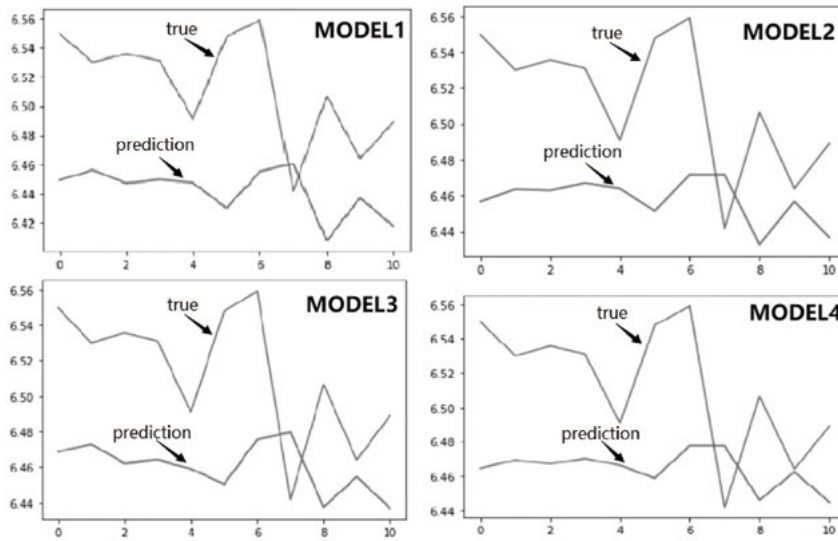
本節將搭建四個 LSTM 模型進行對比分析（表 9），四個 LSTM 模型中的訓練變量設計與上一節四個隨機森林模型相似。

表 9 四個 LSTM 模型訓練變量分配

模型名稱	訓練變量
LSTM Model 1	澳門入境總人數
LSTM Model 2	澳門入境總人數 + 谷歌搜索熱度
LSTM Model 3	澳門入境總人數 + 百度搜索熱度
LSTM Model 4	澳門入境總人數 + 百度搜索熱度 + 谷歌搜索熱度

四個 LSTM 模型所使用的數據、模型訓練及預測方法與上一節隨機森林模型方法相似，以下是 Python 得到四個 LSTM 模型的預測結果（圖 6）。

圖 6 四個 LSTM 模型預測結果



(四) CNN-LSTM 模型對澳門旅遊需求預測

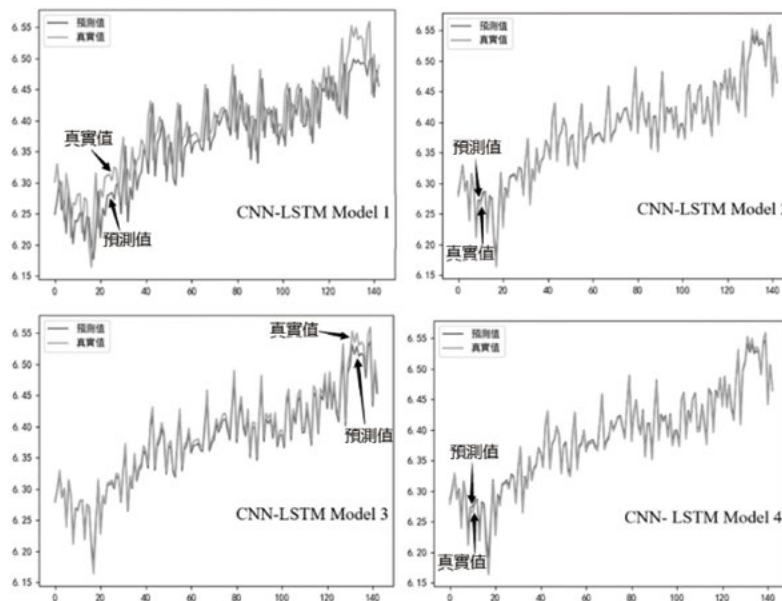
本節將搭建四個 CNN-LSTM 模型進行對比分析 (表 10)，CNN-LSTM 模型是在上一節 LSTM 模型的基礎上與 CNN 模型相結合，實現 LSTM 模型優化的機器學習方法，使得預測結果更精確。四個 CNN-LSTM 模型中的訓練變量設計與前兩節四個隨機森林模型與四個 LSTM 模型相似。

表 10 四個 CNN-LSTM 模型訓練變量分配

模型名稱	訓練變量
CNN-LSTM Model 1	澳門入境總人數
CNN-LSTM Model 2	澳門入境總人數 + 谷歌搜索熱度
CNN-LSTM Model 3	澳門入境總人數 + 百度搜索熱度
CNN-LSTM Model 4	澳門入境總人數 + 百度搜索熱度 + 谷歌搜索熱度

四個 CNN-LSTM 模型所使用的數據、模型訓練及預測方法與上一節隨機森林模型方法相似，以下為使用 Python 得到四個 CNN-LSTM 模型的預測結果 (圖 7)。

圖 7 四個 CNN-LSTM 模型預測結果



(五) 比較不同模型研究對澳門旅遊需求預測的性能

模型的好壞通常需要評價指標來衡量。以下列出了四個 LSTM 模型以及第一節的 ARIMA 模型分別的 RMSE 及 MAPE 兩個指標（表 11），具體公式如下：

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (x_t - \hat{x}_t)^2} \quad (3)$$

$$MAPE = \sum_{t=1}^N \frac{|x_t - \hat{x}_t|}{x_t} \times \frac{100}{N} \quad (4)$$

其中 N 代表了本研究的觀測值得數量， x_t 為澳門旅遊人數的真實數量， \hat{x}_t 代表了旅遊需求的預測值。

表 11 十三個模型的預測結果評價指標

模型	MAPE(%)	RMSE
ARIMA	69.982	0.044
RF Model 1	49.498	0.042
RF Model 2	61.535	0.047
RF Model 3	50.207	0.042
RF Model 4	58.428	0.045
LSTM_model 1	1.249	0.088
LSTM_model 2	0.931	0.067

LSTM_model 3	0.887	0.063
LSTM_model 4	0.850	0.061
CNN-LSTM Model 1	0.669	0.05
CNN-LSTM Model 2	0.51	0.004*
CNN-LSTM Model 3	0.125	0.009
CNN-LSTM Model 4	0.074*	0.006

註：* 標出了最低的 MAPE 指標及 RMSE 指標。

CNN-LSTM 模型四的絕對誤差百分比 (MAPE) 最低，而 ARIMA 模型和隨機森林模型在此指標的表現上較差。與 ARIMA 模型同樣只用了旅遊人數進行擬合的 LSTM 模型 1 的 MAPE 只有 1.249%，CNN-LSTM 模型 1 的 MAPE 只有 0.669%。CNN-LSTM 模型 2 的 RMSE 指標在十三個模型中表現最佳，可能說明了 RMSE 更加適合衡量 CNN-LSTM 模型的預測能力，但所有模型 RMSE 指標相差不大。然而，因為其數據的平穩性要求，進行擬合數據經過了差分處理，值相對較小，因此本研究認為此處的 RMSE 指標在不同的模型之間不具有可比較性。隨機森林模型 1、LSTM 模型 1、CNN-LSTM 模型 1 與 ARIMA 的比較證明了相較於傳統的時間序列模型，機器學習類模型在本問題的預測能力上表現更出色。

隨機森林模型、LSTM 以及 CNN-LSTM 模型 2 和 3 相比於模型 1 分別加入了谷歌和百度的搜索熱度，得到了更低的 MAPE 及 RMSE，說明了搜索熱度指標在機器學習類模型處理澳門旅遊需求預測的問題上是相對有效的指標之一，且百度指數比谷歌趨勢作為特徵表現更好。

CNN-LSTM 模型 4 是兩個指標都最優的模型，MAPE 為 0.074%，說明模型 4 預測的旅遊需求人數只與真實值偏離了 0.074%，說明百度和谷歌搜索熱度之間是互補的關係，共同作用下提高了模型的健壯性及準確度。

(六) 穩定性分析

本研究對以上提出的四組 LSTM 模型和四組 CNN-LSTM 模型進行穩定性分析，通過對數據集中不同的訓練集和測試集的劃分進行交叉檢驗，測試集選取為前 12 個月、第 24 到 36 個月、第 60 到 72 個月以及第 96 到第 108 個月，其餘月份作為訓練集分別執行四次模型預測，計算其 MAPE 及 RMSE 的標準差 (表 12)，結果顯示，使用了全部序列的模型 4 最穩定，而只採用了入境人數作為訓練特徵的模型 1 最不穩定，標準差最大。但總體而言，四組 LSTM 模型和四組 CNN-LSTM 模型的 MAPE 及 RMSE 的標準差都較小，代表了模型的穩定性較強。

表 12 四組 LSTM 模型和四組 CNN-LSTM 模型在不同測試數據集下的指標標準差

模型	Std. MAPE(%)	Std. RMSE
LSTM_model 1	0.00185	0.01202
LSTM_model 2	0.00076	0.00518
LSTM_model 3	0.00098	0.00664
LSTM_model 4	0.00070	0.00483
CNN-LSTM Model 1	0.00168	0.01173
CNN-LSTM Model 2	0.00072	0.00503
CNN-LSTM Model 3	0.00085	0.00676
CNN-LSTM Model 4	0.00064*	0.00416*

註：* 標出了最低的 MAPE 指標及 RMSE 標準差。

六、結論

本研究提出了四種模型，分別是傳統的時間序列模型 ARIMA 以及機器學習類模型隨機森林、LSTM 以及 CNN-LSTM，分別對澳門 2008 年 1 月至 2019 年 12 月的入境總人數序列進行擬合。首先為了使數量較大的旅遊人數數據更加直觀，本研究先對其取了對數。其次對於時間序列模型 ARIMA，首先要確保其數據的平穩性，避免陷入偽迴歸的結果，因此在對旅遊人數序列進行差分後進行 ADF 檢驗，得到一個平穩的旅遊人數時間序列，再通過處理後的數據對劃分好的訓練集，即 2008 年 1 月至 2019 年 1 月的數據進行擬合，再對剩下 11 個月的數據進行預測，對比預測值和真實值，求出評價指標 RMSE 及 MAPE。而對於隨機森林模型、LSTM 模型以及 CNN-LSTM 模型則省去了數據平穩的處理過程，因為即使在非平穩的數據下，隨機森林模型、LSTM 模型以及 CNN-LSTM 模型也有很好的擬合性。在計算出相互之間的評價指標後發現，對比這四組對照模型，機器學習類模型的預測力在本研究中很大程度地優於傳統的時間序列模型，其中，優化的 CNN-LSTM 模型預測性能上表現最為出色。此外，本研究還提出了三種分別基於隨機森林模型、LSTM 模型以及 CNN-LSTM 模型的預測框架，基於上述的互聯網搜索熱度數據——百度指數以及谷歌趨勢——前兩個模型只採用了單獨的搜索熱度，目的是為了比較不同的搜索引擎熱度代表的海內外遊客搜索偏好的數據，瞭解機器學習類模型在進行澳門旅遊需求預測時的準確度優劣。最後，再提出一個互聯網搜索熱度整合模型證明兩種互聯網搜索熱度之間的互補關係。研究結果證明，百度搜索熱度在本研究的預測精度上優於谷歌，其原因也側面突顯了內地旅客為澳門目前的主要遊客收入來源，百度收集到的數據相較於谷歌來說也更加完善豐富。在將兩類搜索熱度同時導入機器學習類模型進行擬合預測後，得到了本研究

預測準確度最高的 CNN-LSTM 模型，證明了互聯網搜索熱度在澳門旅遊需求預測上是一個新穎的影響因素時間序列數據，其有效性也證明了網絡關注度是未來互聯網大數據時代旅遊業需要關注的新指標之一。

綜上所述，本研究結果表明了結合谷歌和百度搜索熱度下的 CNN-LSTM 模型的性能最優，且 CNN-LSTM 模型並不對數據的預處理過程有嚴格的要求，更加適合本研究的場景，即澳門旅遊需求預測。

然而，雖然本研究對澳門需求的預測結果符合預期，但是還存在着以下幾點不足：首先，本研究的數據雖然時間跨度長達 12 年，但是月度數據對於在龐大體量的大數據前仍遊刃有餘的機器學習來說數據量還是太少；其次本研究的特徵方程並沒有繼續加強，如標記假期、季節、特殊天氣、特殊活動、經濟指標等虛擬特徵，它們的加入或可以進一步提高本研究的預測效果；最後，雖然疫情導致 2020 年上半年澳門出入境受限，當時的旅遊人數幾乎跌近冰點，測試集剔除了這部分區間的數據，但本研究對澳門旅遊市場仍然具有一定的借鑑意義。

[責任編輯 陳超敏]

[校對 王旭佳 黃耀岷 陳嘉欣]