

A Food Recommendation System Based on Graph Neural Networks

張碩文
NCKU CSIE

p76134692@gs.ncku.edu.tw

蘇祐蓁
NCKU AIM

ne6131021@gs.ncku.edu.tw

陳子輝
NCKU CSIE

p76135062@gs.ncku.edu.tw

許漢權
NCKU IMI

q56135019@gs.ncku.edu.tw

Abstract

隨著近年線上外送平台如 *Foodpanda*、*UberEat* 逐漸成為主流，餐點推薦系統的重要性日益增加。有效的推薦系統不僅能夠精確地向使用者推薦合適的餐點，還可以幫助平台最大化收益並提升用戶體驗。本研究將提出一種基於圖神經網路 (Graph Neural Networks; GNN) 的推薦系統模型，通過分析平台上用戶與餐點之間的關聯性，從中萃取出有價值的資訊，進而提供更精準的餐點推薦。此模型將針對使用者的偏好與歷史行為進行學習，結合圖結構的數據特徵，實現更個性化且高效的推薦效果。最後，我們將對該模型進行實驗與評估，並與現有的推薦系統進行比較。

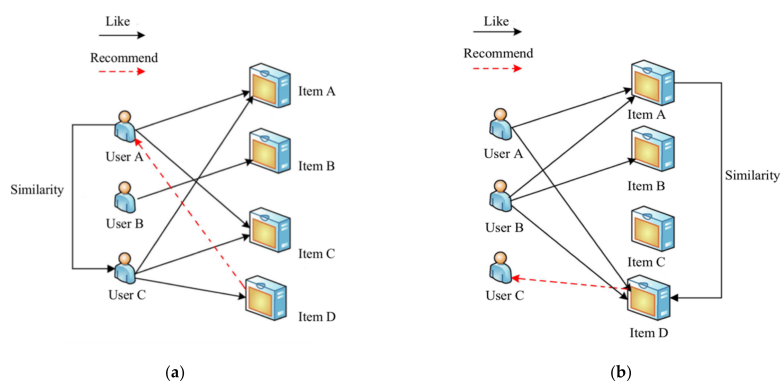


Figure 1. 二分圖於推薦系統 [1]

1. Introduction

因線上外送平台如 *Foodpanda* 和 *UberEat* 的迅速發展，餐飲產業的數位化轉型日益顯著。這些平台不僅簡化了顧客與餐廳之間的聯繫，還透過大規模數據分析來改進用戶體驗及提高平台的經濟效益。餐廳推薦系統在這一過程中扮演了至關重要的角色，能夠根據使用者的歷史消費行為、偏好和時下趨勢，及時推薦合適的餐廳，從而提升平台服務的精確性和客戶滿意度。

傳統的推薦系統方法，如基於內容的過濾 (content-based filtering) 和協同過濾 (collaborative filtering)，雖然在一定程度上能夠提供有效的推薦，但往往忽視了用戶與餐廳之間更深層次的聯繫結構。隨著圖神經網路 (Graph Neural Networks; GNN) 的發展，發現可以將平台上用戶與餐廳之間的關聯性表示為圖結構，從而在更高層次上捕捉到它們之間的複雜互動，其中這個圖結構能以二分圖 (bipartite graph) 表示。

如圖 1 所示，能將所有的節點分成，左邊的使用者節點，右邊的推薦餐廳節點，因左右兩邊的節點類型不同，該圖同時為異質圖 (attributed heterogeneous graph)。本研究的架構圖如圖 2 所示，從 *Food-*

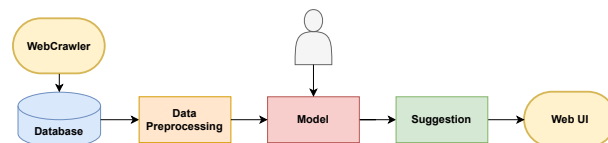


Figure 2. 推薦系統架構圖

panda、*GoogleMap* 上使用爬蟲將每個店家的資訊儲存到資料庫，並使用基本的資料前處理，對於使用者節點，其節點內容如同表格 1，會將這些資訊同時輸入模型，並得出最後的推薦結果，再以 Web UI 的方式呈現。其中本研究將使用 Zhang 等人所提出的圖卷

節點類型	節點內容
使用者	當下位置、使用時間、天氣、預算
推薦餐廳	餐廳名稱、價位、評分、評論

Table 1. 使用者與推薦餐廳節資訊

機網路 [2] 應用在店家推薦的資料集，並使用歸一化折扣累積增益 (Normalized Discounted Cumulative Gain; NDCG)、召回率 (Recall) 去衡量推薦系統之表現。

2. Proposed

2.1. Data Preprocessing

2.2. Model

本研究採用之 NIE-GCN [2] 內部可拆成以下步驟，第一步會先對每個餐廳節點 i 做向量嵌入，得到 d 維的向量 $e_i \in \mathbb{R}^d$ ，將每個餐廳節點 i 的嵌入向量蒐集後可表達成下式

$$E_I^{(0)} = \{e_{i_1}^{(0)}, e_{i_2}^{(0)}, \dots, e_{i_N}^{(0)}\} \in \mathbb{R}^{N \times d}, \quad (1)$$

其中 N 為餐廳節點的總數， d 為嵌入向量之維度，上標 0 為第 0 層傳播的狀態。得到每個餐廳節點的嵌入向量後，因使用者節點 u 之鄰居必為餐廳節點，因此透過與使用者有互動之餐廳節點來推敲使用者的習性。

第二步要建構使用者節點的初始嵌入向量 e_u^0 ，需透過使用者相鄰的餐廳節點 i 之嵌入向量 e_i^0 建構，第 0 層建構方式如下式，

$$e_u^{(0)} = \sigma\left(\sum_{i \in N_u} \frac{1}{\sqrt{|N_u||N_i|}} e_i^{(0)}\right), \quad (2)$$

其中 N_u 、 N_i 分別為使用者節點 u 和餐廳節點 i 之鄰居， $\sigma(\cdot)$ 為雙曲正切函數 (tanh)。有了 e_u 和 e_i 後，下一步要去計算與使用者節點 u 相鄰之餐廳節點 i 連線的注意力分數，來決定鄰居點的重要性，其計算方式如下式

$$\rho(u, i) = Q^T \sigma(W(e_u^{(0)} || e_i^{(0)} + b)), \quad (3)$$

其中 $W \in \mathbb{R}^{2d \times d}$ 、 $Q \in \mathbb{R}^{1 \times d}$ 、 $b \in \mathbb{R}^{1 \times d}$ ，這三個都是注意力機制中的權重矩陣， $\sigma(\cdot)$ 為雙曲正切函數 (tanh)， $||$ 為向量串接，若 $\rho(u, i)$ 的分數越高則代表 u 和 i 之間的存在更高的關聯性，最後這些注意力分數再使用 Softmax 範圍限制在 $[0, 1]$ 之間，得到 $\alpha(u, i)$ ，並以該值作為最終使用者節點 u 和餐廳節點 i 之間的注意力值。

$$\alpha(u, i) = \frac{\exp(\rho(u, i))}{(\sum_{j \in N_u} \exp(\rho(u, j)))^\beta}. \quad (4)$$

3. Experiment Results

預期 NDCG 能超過 0.9

References

- [1] Y. Xiao, C. Li, and V. Liu, "Dfm-gcn: A multi-task learning recommendation based on a deep graph neural network," *Mathematics*, vol. 10, no. 5, 2022. 1

- [2] Y. Zhang, Y. Zhang, D. Yan, Q. He, and Y. Yang, "Nie-gcn: Neighbor item embedding-aware graph convolutional network for recommendation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 54, no. 5, pp. 2810–2821, 2024. 2