**北大 x 好域**

時間：2019/12/26（四） 16:00

參與人：陳宗天教授、林奕銜、黃俊友、蕭琮峻、許程閔、溫智皓

上次[會議紀錄](https://docs.google.com/document/d/1AiGy1fuToYfoybjNY2lYI5UxxWqGOQH_7OlZz_XXzQw/edit?usp=drivesdk)，[Working Process](https://docs.google.com/document/d/1AUf4gI5kxPYSEiniz-MhMeHtMWicBmLqcMjot6EHihM/edit?usp=sharing)

**問題回應：**

1. 不確定locate的理解是否有誤，我以為是相同的space變成一個locate，請問locate除了地址之外，有其他的屬性嗎？
   1. 沒有，Locate就是依據地址區分，將同一棟建築、不同房的空間為同一個Locate。
2. 使用率公式是？
   1. 房間訂單使用小時數 / 房間總可使用小時數
   2. 房間總可使用小時數 = （空間開門時間 - 關門時間）x 日數
3. 有些空間可能因裝潢的關係所以獲利不高（國父紀念館A是類似電影院類型的裝潢，他的損益就明顯不同），未來模型完成後可以討論是否有可新增的space屬性。

**目前進度：**

1. 已計算出各個Locate、Space的總時間使用率，可參照上次會議紀錄內容。
2. 目前已計算出每個locate訓練集與測試集分別的準確率，經判斷後，認為應為underfitting，可參照[此處](https://docs.google.com/spreadsheets/d/1stnK4V_dLazeDlz_ZmiVyVp-mhNFscx-Hu_eE4BQ7YE/edit#gid=244446863)。
   1. 目前績效不好，以及訓練、測試集準確率差異較大的（如locate\_id 11, 15），大多是資料量較少的模型，待新資料結合，以及修改假訂單的產生之後，再行比較準確率。

**下一個目標 ：**

1. 將新的資料 (2019/11/18~2019/12/23) 和之前的資料做結合。（琮峻）
2. 判斷Space的enabled狀態（1為開放，0為關閉），並修改假訂單的產生時間（created\_at ~ updated\_at）。（俊友）
3. 統計每個space＂各個時段＂的一般用戶及月票用戶的使用率。（程閔、智皓）
4. 建構連結資料庫的Processing Program，作為往後快速更新資料時使用。（奕銜）

**未來可嘗試的事情：**

1. 觀察整體locate一般用戶與月票用戶的空間使用率，可以看出有些空間是一般用戶與月票用戶佔比都很高的情況，是否考慮把這個因素加入模型中。（將差別依據大小進行類別化，再加入模型看看成效）。
2. 找出潛藏、抑或新創造出可能影響準確率的屬性。