

2018

國泰大數據競賽

初賽相關說明

2018.08.28-2018.09.18

競賽主題

商品推薦系統

Recommender System



初賽說明

- 壽險商品結構複雜，影響客戶購買商品的因素千百種，因此如何推薦適當商品給客戶往往是行銷的關鍵點。
- 本競賽階段提供某年客戶特徵與購買商品之擬真資料，參賽者依據主辦單位提供之訓練組資料集，預測測試組資料集的客戶會購買哪一類商品。
- 希望考驗參賽團隊透過團隊合作進行資料分析、創意發想，建立精準預測客戶購買行為模型。

➔ 競賽主題

- 近年隨著大數據與人工智慧議題崛起，精準行銷已成為各企業重要發展技術，本屆競賽主題將聚焦於各大產業廣泛運用的「**商品推薦系統**」。

➔ 初賽重要時間

Step 1

公布
初賽資料



08/28
(中午12:00)

Step 2

初賽期間
參賽隊伍進行資料分析



08/28 - 09/18
(中午12:00) (中午12:00止)

Step 3

初賽繳件
截止日



09/18
(中午12:00止)

Step 4

公布
晉級複賽名單



09/25
(下午6:00)

【小提醒】「專案計劃書」可依設計方案去發想大綱內容。

➔ 評分標準

評分依據 ^{註1}	評分項目	比重(%)
分析結果	預測結果與實際結果比對準確率	100
分析說明書 ^{註2}	資料處理與特徵選取	-
	模型選擇與驗證成效說明	-

【註1】：提交檔案內容格式皆須符合繳件格式規定，否則不予評分。

【註2】：如參賽者評分(準確率)相同時，將參考「分析說明書」決定晉級複賽資格。

➔ 評分方式

依據參賽團隊預測客戶
最有可能購買的商品類別，
比對實際購買商品類別，
計算**正確預測比率(Accuracy)**





資料表說明

- 訓練組資料筆數：229,504筆
- 測試組資料筆數：10,000筆
- 目的：預測Submmit_Sample_testing_Set表中的客戶ID(CUST_ID)最有可能購買的商品類別

資料表名稱	資料表說明
train_buy_info	訓練集，客戶購買商品時的資料
train_cust_info	訓練集，客戶的行為與狀態屬性資料
train_tpy_info	訓練集，客戶過去各類別購買狀態
test_buy_x_info	測試集，客戶購買商品時的資料(未含預測變數Y)
test_cust_x_info	測試集，客戶的行為與狀態屬性資料(未含預測變數Y)
test_tpy_x_info	測試集，客戶過去各類別購買狀態(未含預測變數Y)
Submmit_Sample_testing_Set	測試組資料，提交範例格式



繳件說明

- 截止時間：09/18 (中午12:00)前
- 繳件信箱^{註1}：cathaybigdata@cathlife.com.tw
- 繳件主旨：【初賽】**隊伍名稱**
- 繳件文件^{註2}：
「**分析結果**」與「**分析說明書**」皆須符合繳件格式，並夾帶於E-mail信件附件。

【註1】：請確認內容無誤後，再繳交報告至信箱，勿多次傳送檔案。主辦方於收到文件並確認無誤後，E-mail通知隊長。

【註2】：若資料過大，請壓縮處理後，再夾帶檔案於信件中。



分析結果格式

□ 提交檔案(csv檔)：

- 提交範例檔案第一欄為CUST_ID，第二欄為BUY_TYPE，格式為csv檔案。範例如右：

CUST_ID	BUY_TYPE
1558728	a
1969074	a
2091970	b
2034048	c
2939970	a
...	...

- 提交範例請符合csv格式，其他檔案格式一律不接受。

□ 提交程式碼(txt檔)

- 包含資料處理、建模、註解，請統一另貼於記事本(txt檔)繳交，檔案名稱為「隊伍名稱_CODE」。



分析說明書格式

繳交文件	格式說明	說明
封面檔 (1頁)	文件命名規則	隊伍名稱_封面
	封面內容	◆ 須包含隊伍名稱、參賽者姓名
	字體	◆ 中文字體：標楷體 ◆ 英文及數字：Arial
	儲存格式	Word檔與Pdf檔皆須提供。
內文檔 ^{註1} (3頁內容+3頁附件) ^{註2}	文件命名規則	隊伍名稱_分析說明書
	內文架構	◆ 須包含以下架構 1. 資料處理與特徵選取 2. 模型選擇與驗證成效說明
	字體	◆ 中文字體：標楷體 ◆ 英文及數字：Arial ◆ 字型大小：14、固定行高，行高20
	儲存格式	Word檔與Pdf檔皆須提供。

【註1】：「分析說明書」之內文檔不可透露參賽者身份，違反者主辦單位有權扣分或取消參賽資格。

【註2】：「分析說明書」之內文檔主要內容至多3頁，附件至多3頁，總計最多6頁。



THE END

我們複賽見!