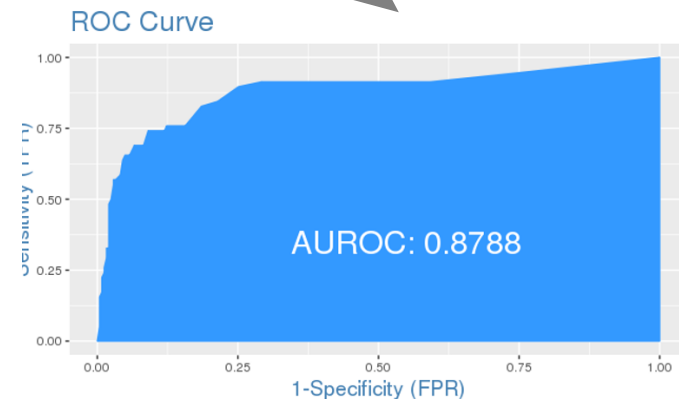


使用14日廣告成效，預測廣告第60天回本率

- 1、模型建立介紹
- 2、資料流程與設計
- 3、功能介面呈現與設計



報告人：張文賢
報告日期：2019.09.12

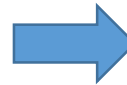
分析樣本：

1. 遊戲(將膽、一劍傾城、蜀門)
2. 時段(2018/04-2019/03、2019/10-2019/03、2018/01-2019/03)



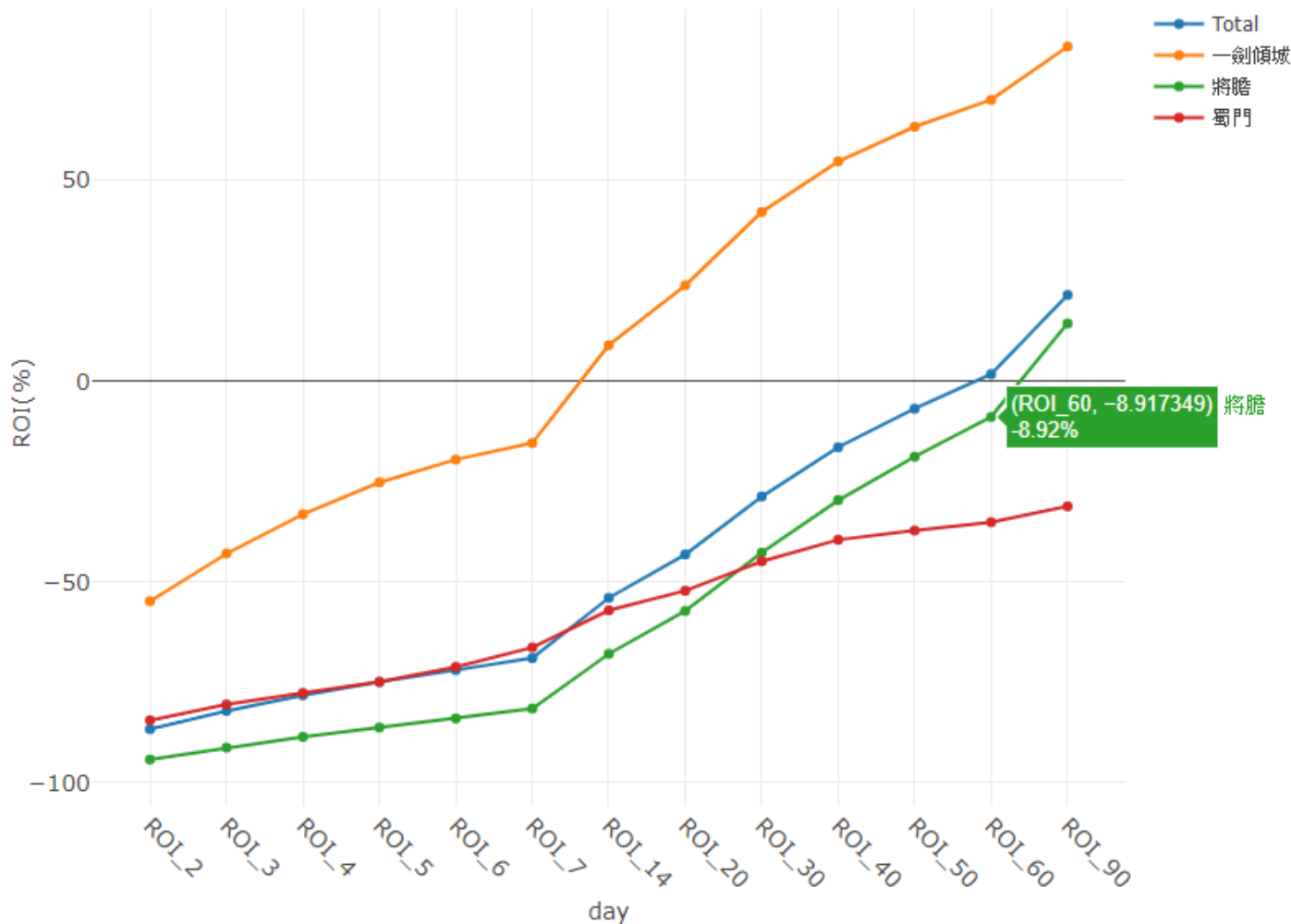
選取有效樣本：

- 排除安裝數等於0
- Impression小於1,000的廣告
- 次日ROI 不等於 -100%
- 排除14日ROI > 0 %的廣告



建模預測
資料分析

不同遊戲，ROI累積狀況：



資料維度及概況：

* 廣告fb_id (素材+TA)

投放方式、系統、性別、年齡、地區

曝光、安裝、CTR、CR、CPI、留存率(1-60日)、LTV(1-60日)、ROI(1-60日)

$$Diff_{7_1} = 7\text{日}LTV - 1\text{日}LTV$$

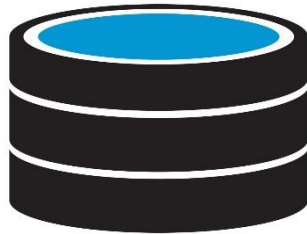
$$Diff_{14_7} = 14\text{日}LTV - 7\text{日}LTV$$

$$Ratio_{7_1} = 7\text{日}LTV / 1\text{日}LTV$$

$$Ratio_{14_7} = 14\text{日}LTV / 7\text{日}LTV$$

DateDay = 廣告投放天數

廣告60日是否能回本 (是：1，否：0)

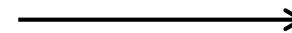


將膽

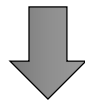
2018/04-2019/03

回本比：

回本：未回本 = 19.2%(192)：80.8%(810)



1,002筆資料



Training data：

回本隨機取70%的資料，共 $192 * 0.7 = 134$ 筆未回本隨機取70%的資料，共 $810 * 0.7 = 567$ 筆

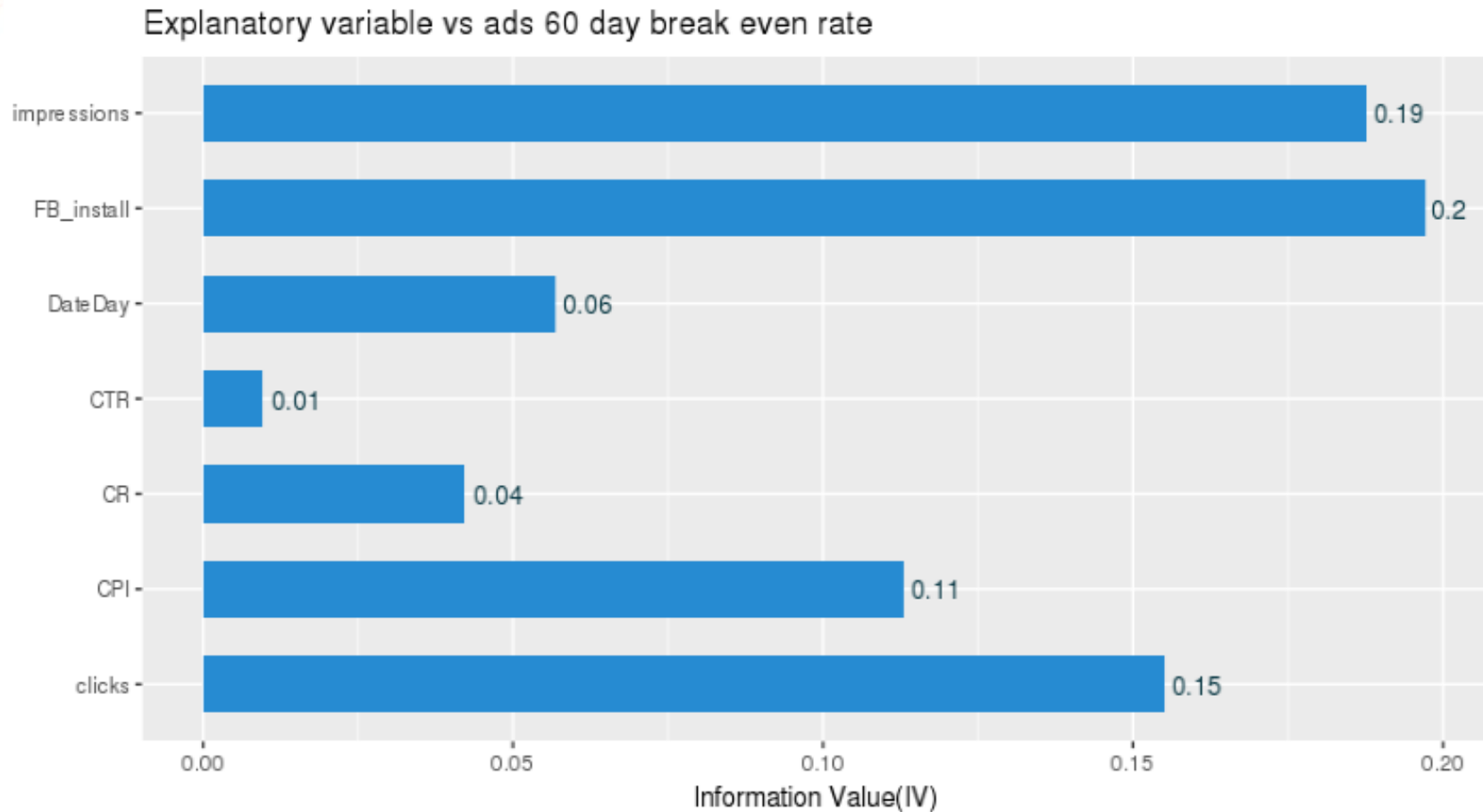
701筆資料

Test data：

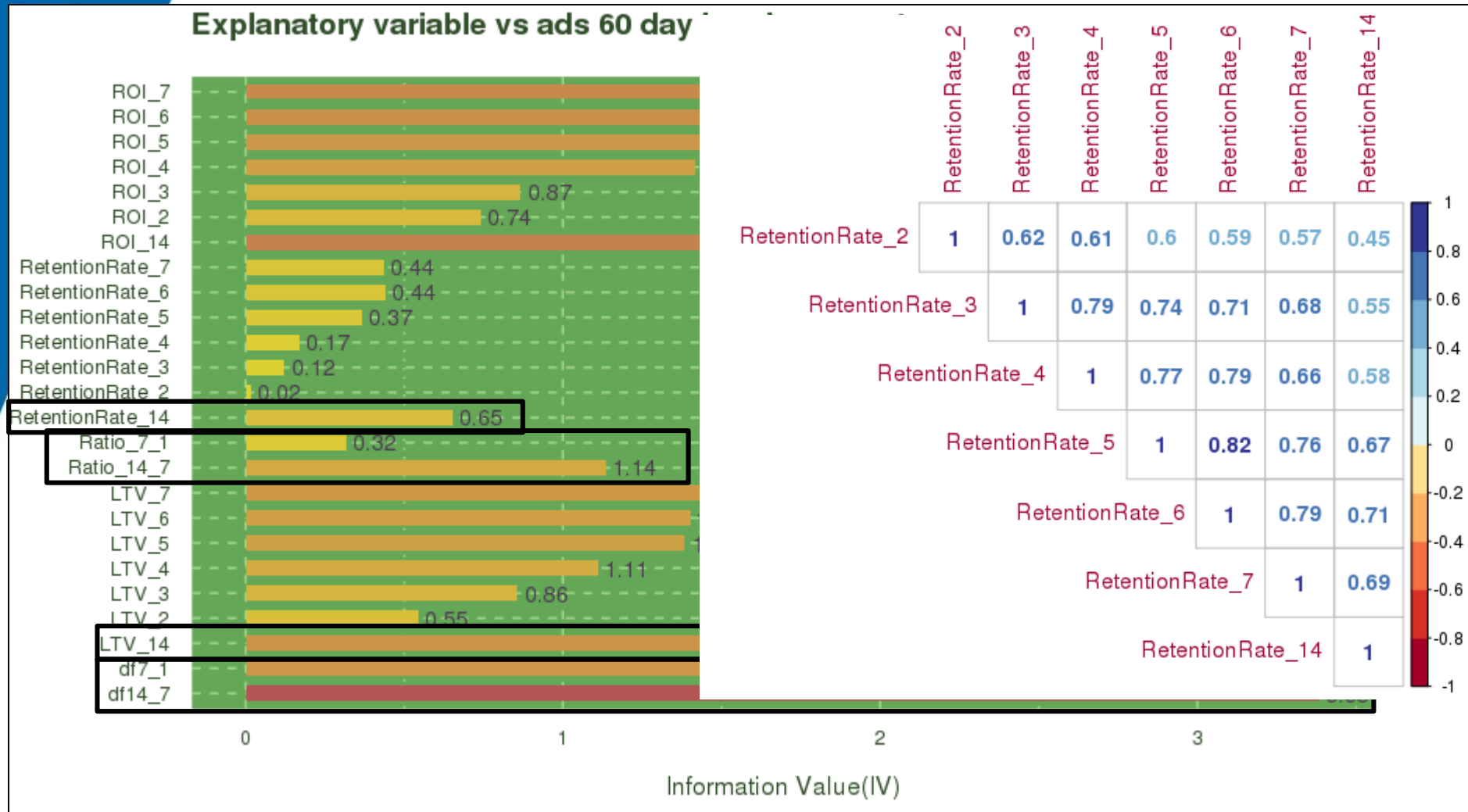
回本其餘30%的資料，共 $192 - 134 = 58$ 筆未回本其餘30%的資料，共 $810 - 567 = 243$ 筆

301筆資料

解釋變數- IV(Information Value) :



解釋變數- IV(Information Value)：留存率-相關係數表：



廣告累積60日，是否能回本之機率

Logistic Regression :

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 * impressions + \beta_2 * clicks + \beta_3 * fb_install +$$

$$\beta_4 * CPI + \beta_5 * CR + \beta_6 * DateDay + \beta_7 * \cancel{\text{留存}} + \beta_8 * \cancel{LTV} +$$

$$\beta_9 * Diff_{7_1} + \beta_{10} * Diff_{14_7} + \beta_{11} * Ratio_{7_1} + \beta_{11} * Ratio_{14_7} +$$

$$\beta_{12} * Charge_{CPA} + \beta_{13} * Charge_{CPM} + \beta_{14} * Channel_{VO} +$$

$$\beta_{15} * Gender_{IOS} + \beta_{16} * Gender_{男} + \beta_{yi} * Location_{yi} +$$

$$\beta_{17} * 14\text{日留存} + \beta_{18} * 14\text{日LTV} + e$$

虛擬變數-基準類：

投放方式：AEO

系統：安卓

受眾性別：全

國別：HK & MO

解釋變數之間是否存在高度共線性(Variance Inflation Vector , VIF) :

Explanatory variable	VIF
impressions	3.88
clicks	3.16
FB_install	3.46
CR	1.59
CPI	4.29
DateDay	2.00
Ratio_7_1	1.14
Ratio_14_7	1.08
df7_1	18.59
df14_7	3.83
LTV_14	26.31
RetentionRate_14	1.38
TA_投放方式CPA	4.25
TA_投放方式OCPM	2.15
TA_投放方式VO	1.61
TA_系統IOS	1.16
TA_地區_HK & TW & MO	1.24
TA_地區_HK & TW & SG	2.36
TA_地區_MY	1.53
TA_地區_SG	1.15
TA_地區_TW	2.58
TA_地區_US	1.24
TA_地區_worldwide	3.41
TA_性別_男	1.45



由結果顯示，df7_1
與LTV_14兩者間存在
高度共線性。



廣告累積60日，是否能回本之機率

Logistic Regression :

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 * \cancel{impressions} + \beta_2 * \cancel{clicks} + \beta_3 * fb_install + \\ \beta_4 * CPI + \beta_5 * CR + \beta_6 * DateDay + \beta_7 * \cancel{次日留} + \beta_8 * \cancel{次日LTV} + \\ \beta_9 * \cancel{Diff}_{7_1} + \beta_{10} * \cancel{Diff}_{14_7} + \beta_{11} * \cancel{Ratio}_{7_1} + \beta_{11} * \cancel{Ratio}_{14_7} + \\ \beta_{12} * Charge_{CPA} + \beta_{13} * Charge_{CPM} + \beta_{14} * Channel_{VO} + \\ \beta_{15} * Gender_{IOS} + \beta_{16} * Gender_{男} + \beta_{yi} * Location_{yi} + \\ \beta_{17} * 14日留存 + \beta_{18} * 14日LTV + e$$

虛擬變數-基準類：

投放方式：AEO

系統：安卓

受眾性別：全

國別：HK & MO

解釋變數之間是否存在高度共線性(Variance Inflation Vector , VIF) :

Explanatory variable	VIF
FB_install	1.14
CR	1.30
CPI	4.15
DateDay_7	1.90
RetentionRate_14	1.35
Ratio_7_1_10	1.11
Ratio_14_7_10	1.08
df14_7_10	1.55
LTV_14_10	3.30
TA_投放方式CPA	4.11
TA_投放方式OCPM	2.09
TA_投放方式VO	1.57
TA_系統IOS	1.14
TA_地區_HK & TW & MO	1.23
TA_地區_HK & TW & SG	2.35
TA_地區_MY	1.52
TA_地區_SG	1.15
TA_地區_TW	2.57
TA_地區_US	1.24
TA_地區_worldwide	3.10
TA_性別_男	1.42



解釋變數之間不存在
高度共線性問題。



Confusion Matrix

	實際沒回本	實際回本	
預測沒回本	231	20	251
預測回本	12	38	50
	243	58	

$$\text{正確率(Accuracy)} = \frac{\# \text{預測回本且實際回本} + \# \text{預測沒回本且實際沒回本}}{\# \text{全樣本}} = \frac{231+38}{301} \approx 0.89$$

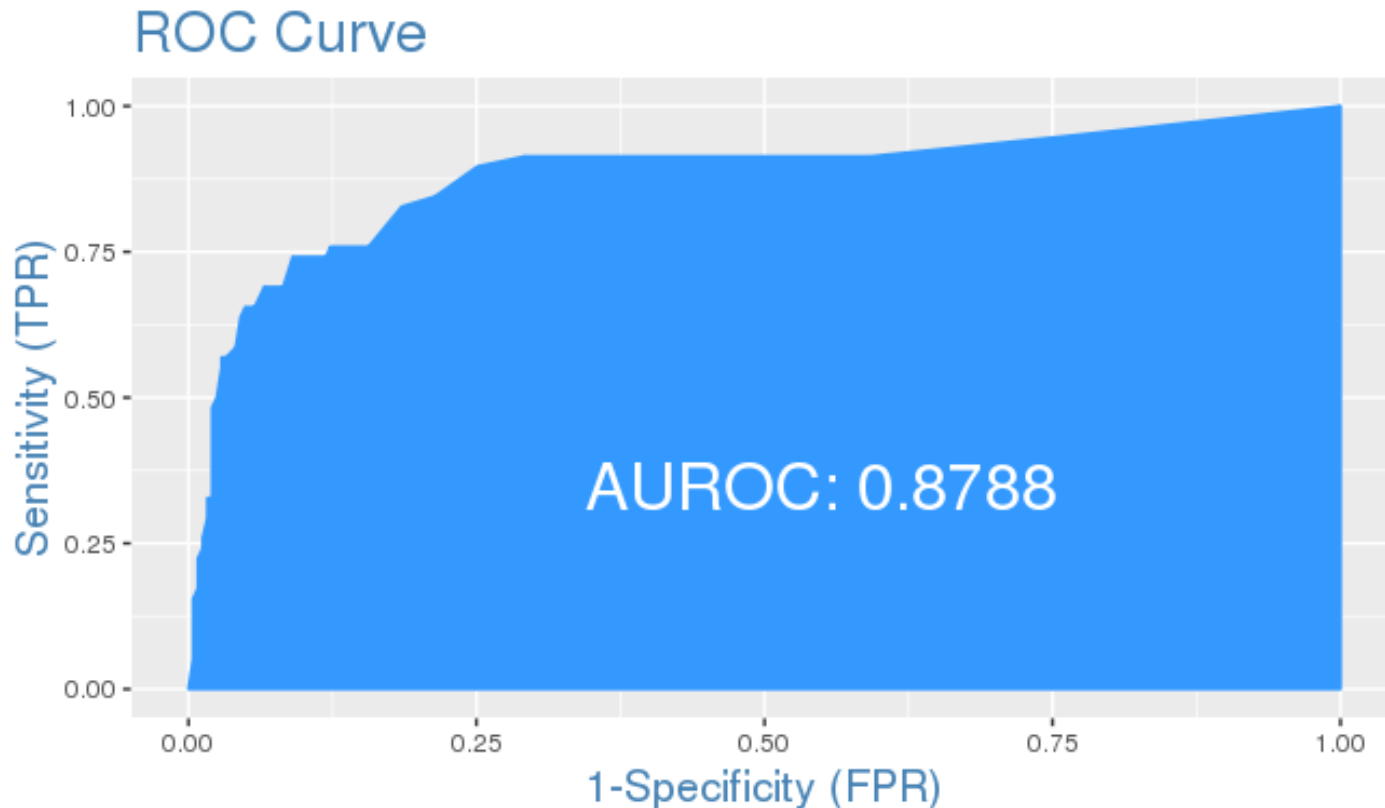
$$\text{模型捕捉率(Sensitivity)} = \frac{\# \text{預測回本且實際回本}}{\# \text{實際回本}} = \frac{38}{58} \approx 0.66$$

$$\text{特異性(Specificity)} = \frac{\# \text{預測沒回本且實際沒回本}}{\# \text{實際沒回本}} = \frac{231}{243} \approx 0.95$$

模 型 建 立 介 紹

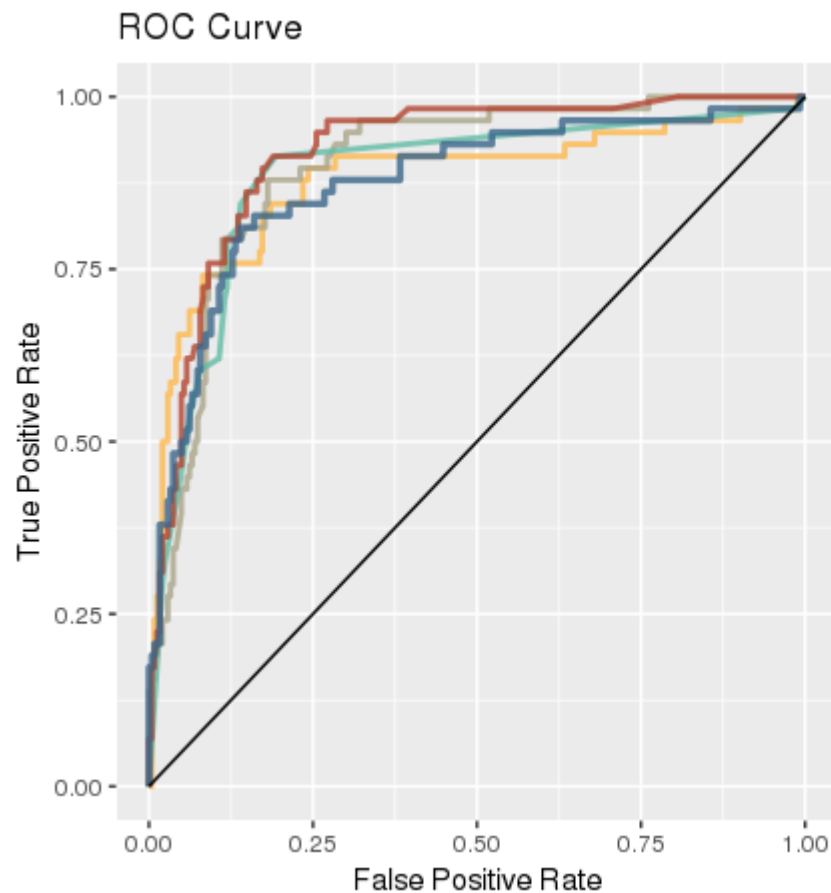
$$\text{模型捕捉率(Sensitivity)} = \frac{\# \text{預測回本且實際回本}}{\# \text{實際回本}} = \frac{38}{58} \approx 0.66$$

$$\text{特異性(Specificity)} = \frac{\# \text{預測沒回本且實際沒回本}}{\# \text{實際沒回本}} = \frac{231}{243} \approx 0.95$$



不同演算法-模型結果比較：

Test：



	Area.Under.ROC.Curve
CART	0.8847
Logistic Regression	0.8790
Neural Networks	0.8997
Random Forest	0.9177
Support Vector Machine	0.8751

不同演算法-模型結果比較：

Training：

演算法	Accuracy	MisClassError	Precision	Sensitivity	Specificity
logistic	89.16%	10.84%	76.27%	67.16%	95.06%
NN	91.16%	8.84%	75.84%	84.33%	93.65%
Decision_Tree	92.87%	7.13%	80.43%	82.84%	95.24%
RF	100.00%	0.00%	100.00%	100.00%	100.00%
SVM	93.30%	6.70%	88.14%	77.61%	97.53%

Test：

演算法	Accuracy	MisClassError	Precision	Sensitivity	Specificity
logistic	89.37%	10.63%	76.00%	65.52%	95.06%
NN	85.05%	14.95%	66.15%	74.14%	90.95%
Decision_Tree	84.05%	15.95%	61.33%	79.31%	88.07%
RF	87.38%	12.62%	66.67%	75.86%	90.95%
SVM	86.38%	13.62%	75.68%	48.28%	96.30%

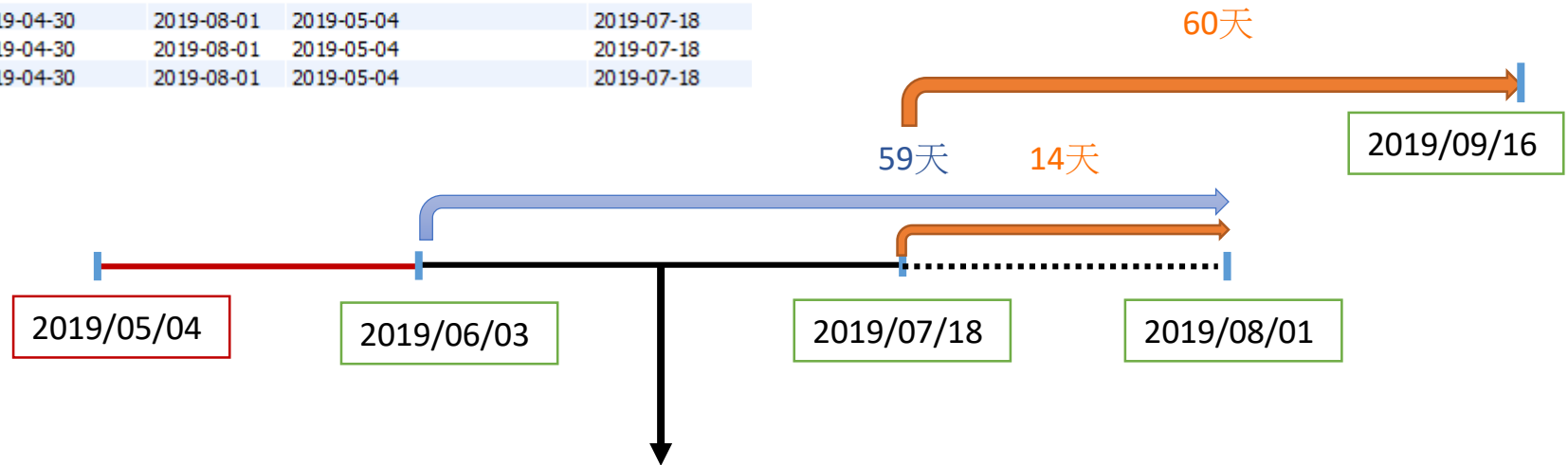
今日執行時間：2019/08/01



model_end_date	data_date	one_month_data_begin_date	data_end_date
2019-04-30	2019-08-01	2019-05-04	2019-07-18
2019-04-30	2019-08-01	2019-05-04	2019-07-18
2019-04-30	2019-08-01	2019-05-04	2019-07-18
2019-04-30	2019-08-01	2019-05-04	2019-07-18

ROI_2	ROI_7	ROI_14	ROI_60
-88.63%	-58.62%	-43.90%	NULL
-96.64%	-62.95%	-28.98%	NULL
-91.30%	-62.34%	-43.54%	NULL
-75.44%	-62.36%	-45.39%	NULL

資料時段



1. 6/3至7/18投放的廣告，數據計算範圍是：2019/05/04-2019/07/18

2. 2019/08/01跑的資料，要到2019/09/16才能驗證預測準確性

總結

1. 2019/08/01跑的資料 ↓
廣告投放時間：2019/06/03-2019/07/18
數據計算範圍：2019/05/04-2019/07/18
2. 2019/08/01跑的資料，要到2019/09/16才能驗證數據。
3. R程式排程-會執行當天、可驗證日(數據更新)。
4. R程式排程-數據驗證統計 ↓
每天都會有一個準確度驗證(整體、回本、不回本準確度)。

建立模型相關資訊

廣告60日回本預測-機器學習模型應用

篩選條件

投放方式 篩選條件

不篩選

FB 篩選條件

不篩選

系統 篩選條件

不篩選

AF 篩選條件

不篩選

2021-12-17-預測資料
(09:25AM 更新資料)

遊戲：將膽

廣告投放期間：2021-10-19~2021-12-03

數據計算範圍：2021-09-19~2021-12-03

模型訓練範圍：2018-04-01~2019-04-30

目前採用的演算法：Neural Network

查詢

建 模 資 訊

廣告樣本：

- 1、排除安裝數等於「0」。
- 2、排除impression小於「1,000」的廣告。
- 3、排除次日ROI = 「-100%」的廣告。
- 4、排除14日ROI = 「-100%」的廣告。
- 5、排除14日ROI大於「0%」的廣告。

註解：詳細廣告樣本篩選流程及統計方法，請參考第一次簡報

模型選取特徵值：

- 1、投放方式、系統、性別、年齡、地區。
- 2、安裝、CR、CPI。
- 3、14日留存、14日LTV。
- 4、投放天數、Diff_1-7、Ratio_1-7、Ratio_7-14。

預測變數(Y)：廣告60日ROI是否能回本 (是：1、否：0)。

註解：模型特徵值挑選準則，請參考第一次簡報

建模樣本資訊：

- 1、Data：1002筆樣本《回本：不回本 = 19.2% (192)：80.8% (810)》。
- 2、TrainingData：701筆樣本《回本：不回本 = 19.2% (134)：80.8% (567)》。
- 3、TestData：301筆樣本《回本：不回本 = 19.2% (58)：80.8% (243)》。

Sheet Options

廣告60日回本預測-機器學習模型應用

廣告60日預測-資料準確度驗證

建立模型相關資訊

User 使用版面

廣告60日回本預測-機器學習模型應用

顯示 50 項結果

file_path	TA_name	data_date	投放天數	impressions	AF_install	FB_spend	次日ROI	七日ROI	十四日ROI	六十日ROI	廣告60日回本預測
	「OCPM_IOS_HK & MO_男_18-57」: 0415_將膽_港澳_再行銷名單_類似5%@	2020-06-23	2	1,004	1	12	-57.08%	-57.08%	-57.08%	-	X
	「OCPM_IOS_NA_全_18-57」: 0415_將膽_再行銷名單@	2020-06-23	10	15,749	34	339	-99.42%	-99.42%	-99.42%	-	X
	「OCPM_Android_NA_全_18-57」: 0415_將膽_再行銷名單@	2020-06-23	2	2,285	6	45	-95.65%	-95.65%	-95.65%	-	X
	「CPA_IOS_HK & MO_男_18-57」: 將膽_港澳_購買5%_18-54_男@	2020-06-23	3	13,006	16	215	-99.08%	-99.08%	-99.08%	-	X
	「AEO_Android_TW_男_18-57」: 將膽_AEO_台灣_18-54_男@	2020-06-23	7	10,308	13	551	-92.75%	-84.72%	-84.72%	-	X

顯示第 1 至 16 項結果，共 16 項

Download Excel

搜尋:
 六十日ROI 廣告60日回本預測
 上一頁 1 下一頁

下載預測數據

提供預測回本資訊

Sheet Options

廣告60日回本預測-機器學習模型應用

廣告60日預測-資料準確度驗證

廣告60日預測-準確度評估

請選擇 資料時間

2019-06-01

to

2021-11-01

挑選 模型區間

建模範圍：2018/04/01 ~ 2019/04/30

查詢

請選擇 穩定度指標

整體-準確率

整體-準確率

回本-準確率

不回本-準確率

準確度 (Accuracy) 評估

請選擇 穩定度指標

整體-準確率



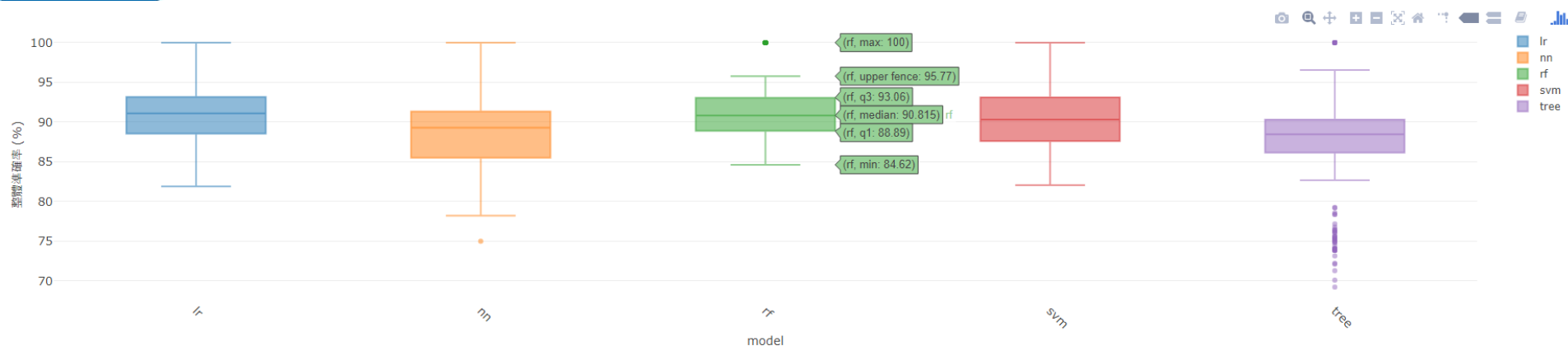
顯示 5 項結果

Model	模型迄日	資料時間	整體準確率	回本準確率	不回本準確率	實際不回本且預測不回本	實際不回本且預測回本	實際回本且預測不回本	實際回本且預測回本	廣告數
lr	2019-04-30	2019-06-01	91.86%	50.00%	96.65%	73	1	6	6	86
nn	2019-04-30	2019-06-01	89.53%	75.00%	91.89%	68	6	3	9	86
rf	2019-04-30	2019-06-01	90.70%	66.67%	94.59%	70	4	4	8	86
svm	2019-04-30	2019-06-01	91.86%	58.33%	97.30%	72	2	5	7	86
tree	2019-04-30	2019-06-01	90.70%	66.67%	94.59%	70	4	4	8	86

顯示第 1,871 至 1,875 項結果，共 1,875 項

上一頁 1 ... 371 372 373 374 375 下一頁

準確度 (Accuracy) 評估



顯示 5 項結果

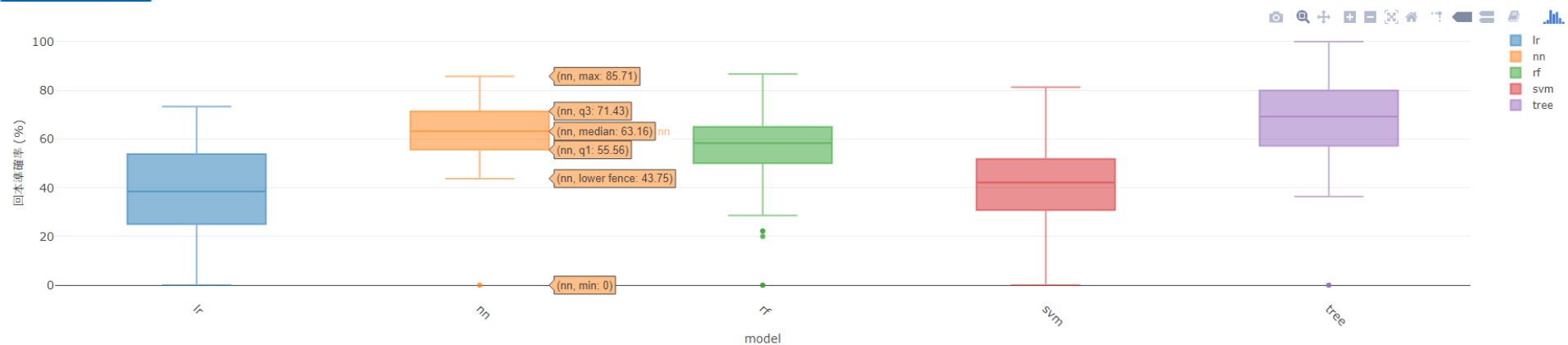
搜尋:

Model	模型迄日	資料時間	整體準確率	回本準確率	不回本準確率	實際不回本且預測不回本	實際不回本且預測回本	實際回本且預測不回本	實際回本且預測回本	廣告數
lr	2019-04-30	2019-06-01	91.86%	50.00%	98.65%	73	1	6	6	86
nn	2019-04-30	2019-06-01	89.53%	75.00%	91.89%	68	6	3	9	86
rf	2019-04-30	2019-06-01	90.70%	66.67%	94.59%	70	4	4	8	86
svm	2019-04-30	2019-06-01	91.86%	58.33%	97.30%	72	2	5	7	86
tree	2019-04-30	2019-06-01	90.70%	66.67%	94.59%	70	4	4	8	86

顯示第 1,871 至 1,875 項結果，共 1,875 項

上一頁 1 ... 371 372 373 374 375 下一頁

回本準確度 (Recall) 評估



顯示 5 項結果

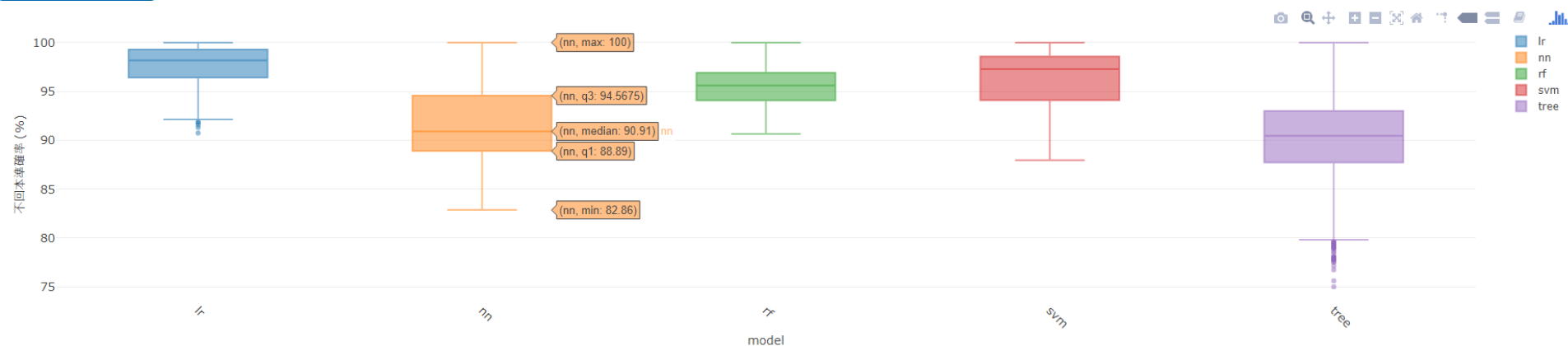
搜尋:

Model	模型迄日	資料時間	整體準確率	回本準確率	不回本準確率	實際不回本且預測不回本	實際不回本且預測回本	實際回本且預測不回本	實際回本且預測回本	廣告數
lr	2019-04-30	2019-06-01	91.86%	50.00%	98.65%	73	1	6	6	86
nn	2019-04-30	2019-06-01	89.53%	75.00%	91.89%	68	6	3	9	86
rf	2019-04-30	2019-06-01	90.70%	66.67%	94.59%	70	4	4	8	86
svm	2019-04-30	2019-06-01	91.86%	58.33%	97.30%	72	2	5	7	86
tree	2019-04-30	2019-06-01	90.70%	66.67%	94.59%	70	4	4	8	86

顯示第 1,871 至 1,875 項結果, 共 1,875 項

上一頁 1 ... 371 372 373 374 375 下一頁

不回本準確度(Specificity)評估



顯示 5 項結果

搜尋:

Model	模型迄日	資料時間	整體準確率	回本準確率	不回本準確率	實際不回本且預測不回本	實際不回本且預測回本	實際回本且預測不回本	實際回本且預測回本	廣告數
lr	2019-04-30	2019-06-01	91.86%	50.00%	98.65%	73	1	6	6	86
nn	2019-04-30	2019-06-01	89.53%	75.00%	91.89%	68	6	3	9	86
rf	2019-04-30	2019-06-01	90.70%	66.67%	94.59%	70	4	4	8	86
svm	2019-04-30	2019-06-01	91.86%	58.33%	97.30%	72	2	5	7	86
tree	2019-04-30	2019-06-01	90.70%	66.67%	94.59%	70	4	4	8	86

顯示第 1,871 至 1,875 項結果，共 1,875 項

上一頁 1 ... 371 372 373 374 375 下一頁