Predictions of Missing Values in Given Digital Ad Spend By Course

Evan Woods

2024-01-24

```
# A tibble: 17 x 7
   `Course/Campaign`       `Ad Spend` `Paid Leads` `Paid App Submissions`
   <chr>                        <dbl>        <dbl> <chr>
 1 Branded                    1491524        50652 NA
 2 General                     974148        21842 NA
 3 CORe                        521368         8695 4720
 4 Finance                     376960         7957 1833
 5 Entrepreneurship            294894         3706 2909
 6 Business Analytics          276115         5232 2712
 7 Management Essentials       273244         4321 1596
 8 Negotiation                 252245         6967 2118
 9 Disruptive Strategy         169516         1433 2004
10 Financial Accounting        172689         2005 1408
11 Sustainable Business        160090         2210 1549
12 Leadership Principles       174408         1503 1914
13 Economics                   112047         1328 454
14 Alternative Investments     110574         1539 541
15 Global Business              61802          378 1017
16 Live Courses*               143452          845 364
17 TOTAL                      5565076       120613 25139
# i 3 more variables: `Paid Enrollment` <chr>, `Paid Revenue` <chr>, RAC <chr>
```

Results

Predicted Values of the Branded Course/Campaign Using Collaborative Filtering & Cosine Similarity:

Course with the Largest Cosine Similarity: Management Essentials

Cosine Similarity: 0.999835575

Branded Predicted Paid App Submissions: 8710

Branded Predicted Paid Enrollment: 3793

Branded Predicted Paid Revenue: 5856092

Branded Predicted RAC: 2144

Predicted Values of the General Course/Campaign Using Collaborative Filtering & Cosine Similarity:

Course with the Largest Cosine Similarity: Live Courses*

Cosine Similarity: 0.999863424

General Predicted Paid App Submissions: 2471

General Predicted Paid Enrollment: 631

```
General Predicted Paid Revenue: 1264266

General Predicted RAC: 10469
```

A tibble: 17 x 5

	`Course/Campaign`	`Ad Spend`	`Paid App Submissions`	`Paid Enrollment`
	<chr>	<dbl>	<int>	<int>
1	Branded	1491524	8710	3793
2	General	974148	2471	631
3	CORe	521368	4720	2655
4	Finance	376960	1833	697
5	Entrepreneurship	294894	2909	943
6	Business Analytics	276115	2712	942
7	Management Essentials	273244	1596	695
8	Negotiation	252245	2118	1056
9	Disruptive Strategy	169516	2004	1094
10	Financial Accounting	172689	1408	545
11	Sustainable Business	160090	1549	689
12	Leadership Principles	174408	1914	752
13	Economics	112047	454	98
14	Alternative Investments	110574	541	289
15	Global Business	61802	1017	371
16	Live Courses*	143452	364	93
17	TOTAL	5565076	25139	10919

i 1 more variable: `Paid Revenue` <int>

A tibble: 17 x 2

	`Course/Campaign`	RAC
	<chr>	<int>
1	Branded	2144
2	General	10469
3	CORe	196
4	Finance	540
5	Entrepreneurship	312
6	Business Analytics	293
7	Management Essentials	393
8	Negotiation	238
9	Disruptive Strategy	154
10	Financial Accounting	316
11	Sustainable Business	232
12	Leadership Principles	231
13	Economics	1143
14	Alternative Investments	382
15	Global Business	166
16	Live Courses*	1542
17	TOTAL	509

Discussion

The predicted values for the Branded and General courses are higher than other courses/campaigns, but the values do not seem to be obvious outliers except for the RAC. The predicted RACs are inordinately high with respect to the other values in the same column.

Predicting ROI

```
1/1 - 0s - loss: 15.6489 - accuracy: 0.0833 - 187ms/epoch - 187ms/step

      loss    accuracy
15.64890194  0.08333334

1/1 - 0s - 45ms/epoch - 45ms/step

Accuracy of the Neural Network Model: 0%.

1/1 - 0s - 40ms/epoch - 40ms/step

The resulting test predictions are:

5
5
```

Linear Model Prediction of ROI:

```
# A tibble: 17 x 25
   `Paid Click Source`    CORe Finance Entrepren `Business Analytics`
   <chr>                 <dbl>   <dbl>    <dbl>                <dbl>
 1 Branded                 774     173      360                  337
 2 General                 614     111      203                  205
 3 CORe                    714      38       64                   75
 4 Finance                 101     297       23                   32
 5 Entrepreneurship         39       7      189                    6
 6 Business Analytics       80       9       11                  196
 7 Management Essentials    40       6       14                   19
 8 Negotiation              16       9       16                   11
 9 Disruptive Strategy      22       7        9                   11
10 Financial Accounting     54      14        8                    6
11 Sustainable Business     14       4        5                    8
12 Leadership Principles   116      10       11                   16
13 Economics                42       4       10                   11
14 Alternative Investments  10       2        5                    0
15 Global Business           9       0        4                    2
16 LIVE Courses             12       6       12                    7
17 TOTAL                  2657     697      944                  942
# i 20 more variables: `Manegement Essentials` <dbl>, Negotiation <dbl>,
#   `Disruptive Strategy` <dbl>, `Financial Accounting` <dbl>,
#   `Sustainable Business` <dbl>, `Leadership Principles` <dbl>,
#   Economics <dbl>, `Alternative Investments` <dbl>, `Global Business` <dbl>,
#   `LIVE Courses` <dbl>, ...16 <dbl>, `Total Ad spend (Ex10)` <dbl>,
#   `non core enrolled` <dbl>, `core rev` <dbl>, `non-core rev` <dbl>,
#   `total rev` <dbl>, `ROI = (Rev-Ad)/Ad` <dbl>, `%Revenue` <dbl>, ...
```

The validation set predictions of ROI (no cross-over) are:

The predicted ROI (no cross-over) of Branded & General respectively from the linear model trained on the full dataset:

Branded: 20.712

General: 1.184

Discussion of Results

The predictions on the linear model are underperformant. I am hoping to improve

model performance through the application of material learned in the coming weeks of this course.