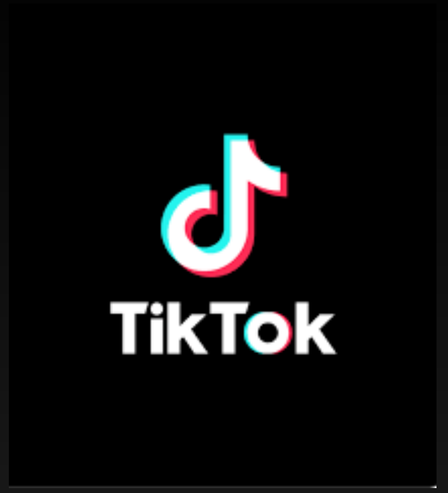
TikTok Ad Targeting Analysis

By Hermia Ren & Jordan Kane

Topic Choice



- TikTok advertisement targeting
- Interesting topic related to online social media
- Trending and popular platform that is used daily worldwide
- Accessible API
- Large dataset of users and ads

Data Preparation


```
{
  "ad": {
    "id": 1776307088392242,
    "image_urls": [
      "https://p21-ad-sg.ibyteimg.com/origin/tos-alisg-p-0051c001-sg/oEu8nbPfGDLgQip2yDJ
    ],
    "last_shown_date": "20231117",
    "reach": {
      "unique_users_seen": "10M-20M"
    },
    "status": "active",
    "status_statement": "N/A",
    "videos": [
      {
        "url": "https://v77.tiktokcdn.com/845faa77aa075ea6615c80ae7fe92ffb/655f6126/vide
        "cover_image_url": "https://p16-sign-sg.tiktokcdn.com/tos-alisg-p-0051c001-sg/oE
      }
    ],
    "first_shown_date": "20230906"
  },
  "advertiser": {
    "business_id": 6979881690004980482,
    "business_name": "International Federation of Red Cross and Red Crescent Societies",
    "paid_for_by": "SOCIAL SOCIAL GmbH"
  }
...
```

```
{'data': {'ad': {'id': 1776307088392242,
                 'reach': {'unique_users_seen': '10M-20M',
                           'unique_users_seen_by_country': {'AT': '166K',
                                                            'BE': '490K',
                                                            'CZ': '1.3M',
                                                            'DK': '38K',
                                                            'ES': '6.9M',
                                                            'FI': '478K',
                                                            'IE': '121K',
                                                            'IT': '8.0M',
                                                            'NL': '165K',
                                                            'SE': '60K'}}},
          'ad_group': {'targeting_info': {'age': {'13-17': True,
                                                  '18-24': True,
                                                  '25-34': True,
                                                  '35-44': True,
                                                  '45-54': True,
                                                  '55+': True},
                                          'audience_targeting': 'No',
                                          'country': ['BE',
                                                      'AT',
                                                      'DK',
                                                      'CZ',
                                                      'IE',
                                                      'IT',
...
                                          'video_interactions': ''}}},
 'error': {'code': 'ok',
           'log_id': '20231123091619812CC42FC97D2E24DEFC',
           'message': ''}}
```

Machine Learning

Data Preparing

- Split multi-values into single values and add on more columns to store single values
- Do some essential calculations and make my own target value: “successfully_reached”(when actual reaching amount >= min_targeting_amount)
- Convert strings into numerical values
- Drop unnecessary rows

ad_group.targeting_info.country
['BE', 'AT', 'DK', 'CZ', 'IE', 'IT', 'FI', 'ES...
['NL', 'CZ', 'IT', 'GR', 'FI', 'RO', 'DK', 'IE...
['ES', 'GB', 'DE', 'IT']
['BE', 'AT', 'DK', 'CZ', 'IE', 'IT', 'FI', 'ES...
['IT', 'GR', 'AT', 'HU', 'IE', 'PT', 'CZ', 'FI...

...	SK	SI	ES	SE	NO	IS	LI	GB	CH	successfully_reached
...	0	0	0	1	1	0	0	1	1	0
...	0	0	1	0	0	0	0	1	0	0
...	0	0	0	0	0	0	0	0	0	1
...	0	0	1	1	1	0	0	1	1	0
...	0	0	0	0	0	0	0	1	0	0

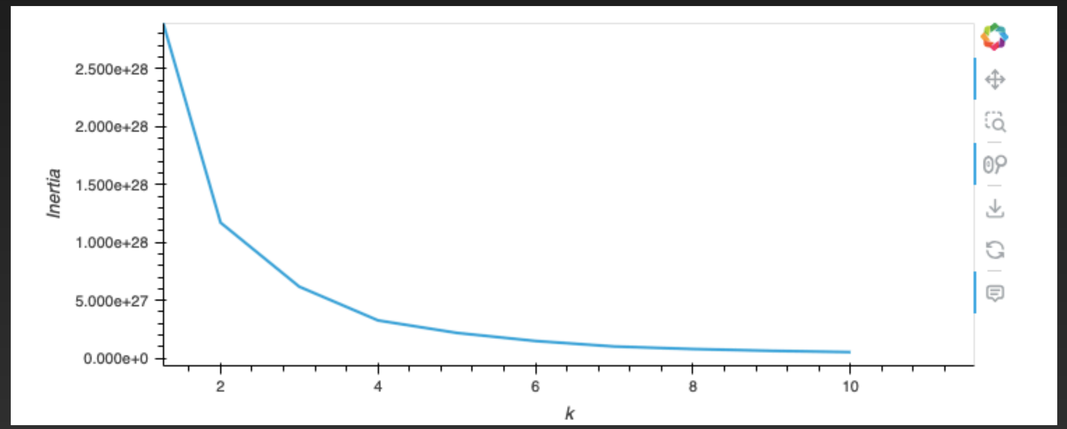
ad_group.targeting_info.number_of_users_targeted
NaN
200.1M-244.6M
27.8M-34.0M
NaN
NaN

ad.reach.unique_users_seen_x
10M-20M
10M-20M
10M-20M
10M-20M
10M-20M

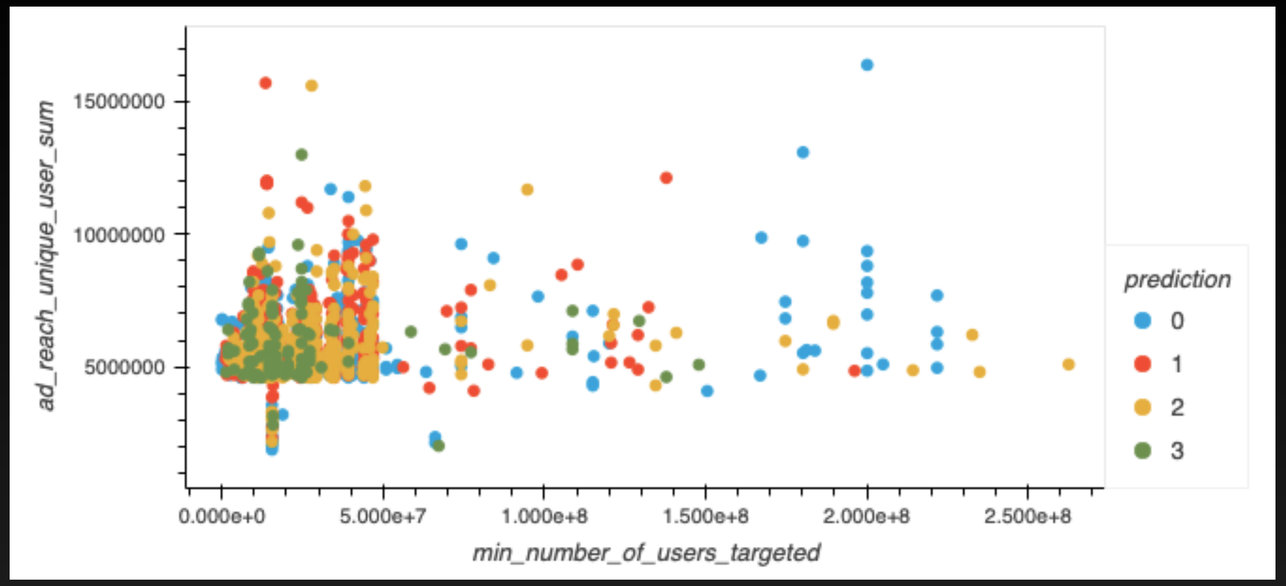
min_number_of_users_targeted	max_number_of_users_targeted	ad_reach_unique_user_sum
200100000.0	244600000.0	16384000
27800000.0	34000000.0	15600000
13500000.0	16500000.0	15700000
180200000.0	220300000.0	13087000
24700000.0	30100000.0	13000000
...	...	...
27700000.0	33800000.0	4600000
39200000.0	47900000.0	4600000
11700000.0	14300000.0	4600000
34800000.0	42500000.0	4600000
14000000.0	17100000.0	4600000

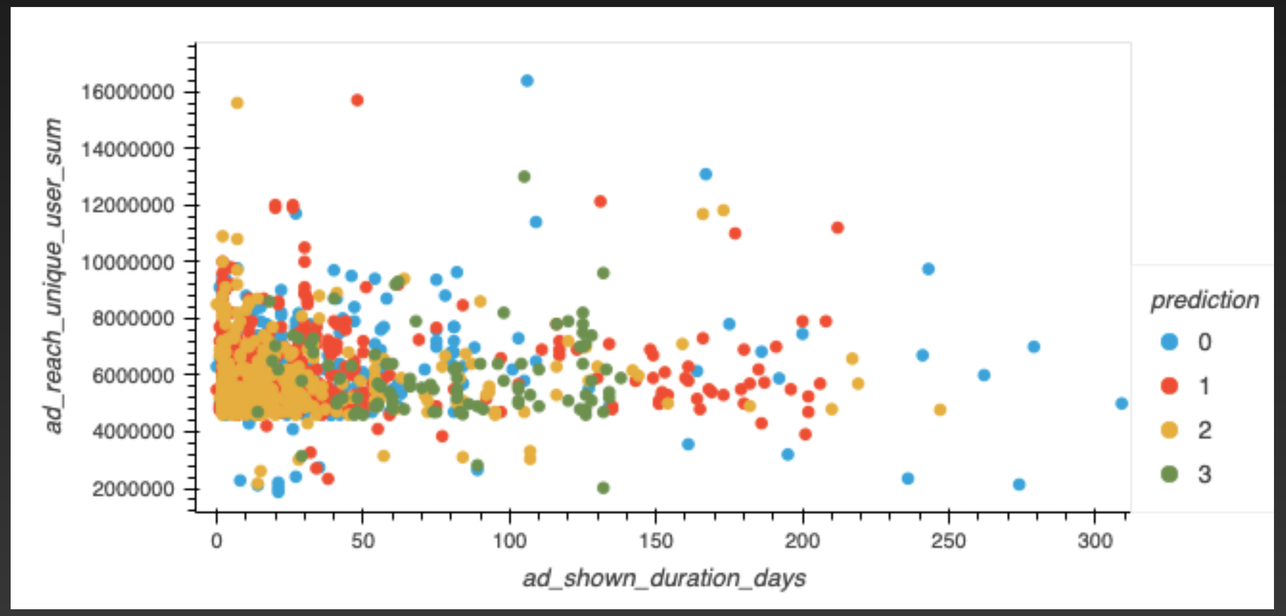
Cluster Model

- Doesn’t work well for this case.
- Chaos.



Elbow chart for finding the k value





Logistic Regression

- Terrible performance when predicting for Class 1: 0 precision and 0 recall.
- Tend to predict all Class 1 as Class 0.
- Overall performance is lower than 50%.

	precision	recall	f1-score	support
0	0.96	1.00	0.98	847
1	0.00	0.00	0.00	35
accuracy			0.96	882
macro avg	0.48	0.50	0.49	882
weighted avg	0.92	0.96	0.94	882

Decision Tree Model

- Outstanding performance: achieving perfect accuracy and precision, recall, and f1-score for both classes.
- Correctly predicted all instances of both Class 0 and Class 1.
- The decision tree model is well-suited to this data and has learned the patterns effectively.

confusion matrix

	Predicted 0	Predicted 1
Actual 0	288	0
Actual 1	0	6

Accuracy Score: 1.0
Classification report

	precision	recall	f1-score	support
0	1.00	1.00	1.00	288
1	1.00	1.00	1.00	6
accuracy			1.00	294
macro avg	1.00	1.00	1.00	294
weighted avg	1.00	1.00	1.00	294

TensorFlow Model

10/10 - 0s - loss: 0.1027 - accuracy: 0.9796 - 367ms/epoch - 37ms/step
Loss: 0.10269585251808167, Accuracy: 0.9795918464660645

- **Loss:** 0.1027 - This is the value that the model is attempting to minimize during training. A lower loss indicates better convergence.
- **Accuracy:** 97.96% - The proportion of correctly classified instances. It indicates that the model is accurate in its predictions for approximately 98% of the cases.

Data Visualisations

Tableau link: https://public.tableau.com/app/profile/jordan.kane/viz/TikTokAdTargetingAnalysis_17016755728710/CategoryPerformance?publish=yes

Limitations

- Could only access data from Europe. Australia was unavailable.
- TikTok data needed a lot of cleaning as was unorganised.
- Data was mostly string. Had to convert to correct format.
- API was challenging to access. Needed to apply and await approval.
- Can be inaccurate as under age kids can lie about their age when making a TikTok account.

Conclusion

- Tree Decision Model performs the best for this data set while Cluster Model and Logistic Regression Model are not ideal.
- Most successful demographic to target are people in their 20s who are interested in the categories of Entertainment, Gaming and Pop Culture and live in the United Kingdom.
- Gender and ad duration provided negligible results.

Thank You!