1. **List of codes**

The code contains 11 scripts in Python or R divided into construction (scripts 1-5), evaluation (scripts 6-7), and interpretation (scripts 8-11) of the algorithm. Each script is named after the task it performs. The main results reported in the paper and online appendices can be reproduced from scripts 6, 7, and 11.

1. **List of data frames in tiktok.RData**

| **Name** | **Description** |
| --- | --- |
| activity | Activities detected by the I3D algorithm and their association with engagement scores |
| e\_space | Average pixel-level engagement scores over space |
| e\_time | Average pixel-level engagement scores over time |
| p\_space | Average pixel-level product scores over space |
| p\_time | Average pixel-level product scores over time |
| pe\_space | Average pixel-level PE-scores over space |
| pe\_time | Average pixel-level PE-scores over time |
| emotion | Scores for different emotions detected by the FER algorithm and engagement scores |
| loss | Training and validation losses for the 3D CNN |
| model\_construction | Video engagement data for constructing the 3D CNN |
| model\_evaluation | Sales panel |
| object | Objects and number of instances detected by YOLO algorithm |
| own | Influencer and product features for videos in the sales panel and videos that advertise the influencers’ own products |
| search | Baidu search index |

1. **Variable dictionary**

**“activity”**

Each observation is at activity level.

| **Variable** | **Description** |
| --- | --- |
| Estimate | The association between the probability of an activity with engagement score at video-segment (15 seconds) level |
| Std..Error | Standard error |
| t.value | T value |
| Pr...t.. | P value |
| act | Name of the activity |

**“e\_space”**

A 224 x 224 matrix of engagement scores at each pixel location averaged over all videos.

**“e\_time”**

Each observation is at video-second level.

| **Variable** | **Description** |
| --- | --- |
| time | Time in second |
| value | Average engagement score at that time point |

**“p\_space”**

A 224 x 224 matrix of product scores at each pixel location averaged over all videos.

**“p\_time”**

Each observation is at video-second level.

| **Variable** | **Description** |
| --- | --- |
| time | Time in second |
| value | Average product score at that time point |

**“pe\_space”**

A 224 x 224 matrix of product engagement scores at each pixel location averaged over all videos.

**“pe\_time”**

Each observation in this data file is at video-second level.

| **Variable** | **Description** |
| --- | --- |
| time | Time in second |
| value | Average product engagement score at that time point |

**“emotion”**

Each observation in this data file is at video-frame level.

| **Variable** | **Description** |
| --- | --- |
| video\_id | Video ID |
| time | Time in second |
| angry | Anger score |
| disgust | Disgust score |
| fear | Fear score |
| happy | Happiness score |
| sad | Sadness score |
| surprise | Surprise score |
| neutral | Neutral score |
| enga\_score | Average pixel-level engagement score |

**“loss”**

Each observation is at epoch level.

| **Variable** | **Description** |
| --- | --- |
| epoch | Epoch number |
| Training Loss | Training loss (mean absolute percentage error) |
| Validation Loss | Validation loss (mean absolute percentage error) |

**“model\_construction”**

Each observation is at video level.

| **Variable** | **Description** |
| --- | --- |
| like | Number of likes received by the video |
| comment | Number of comments received by the video |
| share | Number of shares received by the video |

**“model\_evaluation”**

Each observation is at product/video-day level.

| **Variable** | **Description** |
| --- | --- |
| video\_id | Video ID |
| updated\_time | Date |
| updated\_time\_rel | Number of days since the first day in the sales panel for each product |
| video\_posted\_time | Posting date of the video |
| treated | Whether the video ad has been posted for a product (1 for posted, 0 for not yet) |
| like | Number of likes received by the video |
| comment | Number of comments received by the video |
| share | Number of shares received by the video |
| len | Length of the video in second |
| e\_score | Engagement score constructed by the number of shares (normalized to the interval of [0, 1]) |
| e\_score\_like | Engagement score constructed by the number of likes |
| e\_score\_comment | Engagement score constructed by the number of comments |
| e\_score\_unsup | Engagement score constructed using an unsupervised approach |
| p\_score | Product score of the video (normalized to the interval of [0, 1]) |
| pe\_score | PE-score (normalized to the interval of [0, 1]) |
| pe\_score\_like | PE-score constructed by the number of likes (normalized to the interval of [0, 1]) |
| pe\_score\_comment | PE-score constructed by the number of comments (normalized to the interval of [0, 1]) |
| pe\_score\_unsup | PE-score constructed using an unsupervised approach (normalized to the interval of [0, 1]) |
| influencer\_id | Influencer ID |
| gender | Gender of the influencer (0 for female, 1 otherwise) |
| fans | Number of followers of the influencer |
| order\_cnt | Number of video ads the influencer has posted |
| avg\_play | Average play for the influencer |
| influencer\_price | Price per video ad for the influencer (in RMB) |
| expected\_cpm | Expected CPM for the influencer (in RMB) |
| taobao\_id | Product ID |
| rev | Previous 30-day sales revenue (in RMB) |
| search | Baidu search index for the product |
| avg\_search | Average Baidu search index over time |
| price | Product price (in RMB) |
| discount | Product price discount (in RMB) |
| category | Category of the product |

**“object”**

Each observation is at object level.

| **Variable** | **Description** |
| --- | --- |
| object | Object detected by the YOLO algorithm |
| count\_high | Number of object instances detected in high-engagement pixels |
| count\_low | Number of object instances detected in low-engagement pixels |
| count\_diff | Difference between the number of object instances detected in high-engagement pixels and low-engagement pixels |

**“own”**

Each observation is at video level.

| **Variable** | **Description** |
| --- | --- |
| video\_id | Video ID |
| pe\_score | PE-score of the video |
| own | Whether the video advertises the influencer’s own product |
| category | Category of the product |
| fans | Number of followers for the influencer |
| gender | Gender of the influencer (0 for female, 1 otherwise) |
| order\_cnt | Number of video ads the influencer has posted |
| price | Product price (in RMB) |
| discount | Product price discount (in RMB) |

**“search”**

Each observation is at product-day level.

| **Variable** | **Description** |
| --- | --- |
| taobao\_id | Product ID |
| updated\_time | Date |
| search | Baidu search index |