

Bot or Human

- Akshay Goel, Hiroko Sayama, Abhik Ghosh, Daphne Chui



Format of Trainer Data (first 6 rows)

```
"bid_id" "bidder_id" "auction" "merchandise" "device" "time" "country" "ip" "url" "outcome"
"1" 89 "037d9f10da403d8d5f94b6e2957a3702f6x07" "cxcqe" "home goods" "phone71" "onepm" "za" "38.122.182.50" "q3ptdym7sr4agaj" 0
"2" 133 "0547dbd3ecb4a293913f138d60b586fdf4ane" "9e6f5" "sporting goods" "phone90" "onepm" "mx" "96.217.45.28" "vasstdc27m7nks3" 0
"3" 134 "037d9f10da403d8d5f94b6e2957a3702f6x07" "sclcr" "home goods" "phone33" "onepm" "th" "86.200.121.249" "ztwq32qvmby357" 0
"4" 381 "0565a780f2c5b46837bbe9739ef15ea92fdu0" "gjkng" "mobile" "phone177" "onepm" "in" "113.172.86.24" "q2xfkx8ike6ema4" 0
"5" 398 "055282b75717345ba6f53d40e9e6d6c56qrzp" "crsvb" "sporting goods" "phone181" "onepm" "it" "228.147.92.222" "9ysjpbdbvbmigsc" 0
"6" 500 "037d9f10da403d8d5f94b6e2957a3702f6x07" "k3qaa" "home goods" "phone208" "onepm" "in" "243.69.78.218" "tpvyo4b9pb5grhu" 0
```

Format of prepared data using factors (first 6 rows)

```
"bid_id" "bidder_id" "auction" "merchandise" "device" "time" "country" "ip" "url" "outcome"
"1" 89 28 93 2 151 3 54 287 135 0
"2" 133 42 72 6 165 3 32 378 162 0
"3" 134 28 215 2 99 3 48 358 177 0
"4" 381 46 120 4 52 3 23 29 134 0
"5" 398 44 92 6 55 3 25 220 47 0
"6" 500 28 156 2 62 3 23 249 151 0
```

Trainer Data

Descriptive Statistics for Trainer Data

| Statistic | bid_id | bidder_id | auction | merchandise | device | time | country | ip | url | outcome |
|-----------|-------------|-----------|---------|-------------|--------|------|---------|-------|-------|---------|
| N | 488 | 488 | 488 | 488 | 488 | 488 | 487 | 488 | 488 | 488 |
| Mean | 1,815,737.0 | 30.6 | 145.5 | 3.8 | 93.6 | 3.5 | 30.2 | 201.4 | 119.0 | 0.2 |
| St. Dev. | 1,584,536.0 | 17.0 | 75.1 | 1.3 | 46.7 | 1.7 | 15.7 | 107.9 | 54.9 | 0.4 |
| Min | 89 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Max | 7,651,994 | 58 | 286 | 6 | 171 | 6 | 55 | 379 | 178 | 1 |

Test Data

Descriptive Statistics for Test Data

| Statistic | bid_id | bidder_id | auction | merchandise | device | time | country | ip | url | outcome |
|-----------|--------|-----------|---------|-------------|--------|-------|---------|-------|-------|---------|
| N | 1,043 | 1,043 | 1,043 | 1,043 | 1,043 | 1,043 | 1,043 | 1,043 | 1,043 | 1,043 |
| Mean | 521.0 | 169.0 | 207.0 | 5.6 | 177.0 | 6.4 | 44.4 | 439.9 | 362.0 | 0.1 |
| St. Dev. | 301.2 | 82.0 | 111.0 | 1.4 | 92.5 | 3.3 | 19.0 | 247.0 | 155.7 | 0.3 |
| Min | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| Max | 1,042 | 303 | 402 | 8 | 305 | 12 | 82 | 866 | 556 | 1 |

Regression Models

Descriptive Statistics for Trainer Data*Dependent variable:*

| | outcome | |
|--------------|-------------------------|-----------------------|
| | (1) | (2) |
| bid_id | 0.00000*** (0.00000) | |
| bidder_id | 0.281*** (0.039) | 0.199*** (0.025) |
| auction | 0.003 (0.003) | 0.005* (0.002) |
| merchandise | -0.913*** (0.247) | -0.577*** (0.177) |
| device | 0.010** (0.005) | 0.010** (0.004) |
| time | 0.605*** (0.135) | 0.616*** (0.125) |
| country | 0.072*** (0.016) | 0.059*** (0.013) |
| ip | 0.006*** (0.002) | 0.006*** (0.002) |
| url | -0.003 (0.004) | -0.001 (0.003) |
| Constant | -17.674*** (2.242) | -13.895*** (1.741) |
| Observations | 487 | 487 |

Predicted Results using Model 1**Number of Humans - depicted by 0; Number of Bots - depicted by 1**

"x"
 "0" 945
 "1" 98

Predicted Results using Model 2**Number of Humans - depicted by 0; Number of Bots - depicted by 1**

"x"
 "0" 976
 "1" 67