**Predicting Purchase Decision in Freemium Mobile Game**

1. **Goal**

Maximizing revenue

* Identify users who are likely to convert on their own.
* Identify true non-payers, failure to convert sufficient true non-payers may result in the not-yet-converted future payers taking advantage of the promotions instead of paying full price, thus hurt the revenue

1. **Dataset**

22K users who installed the game during 2016-03-01 to 2016-03-07

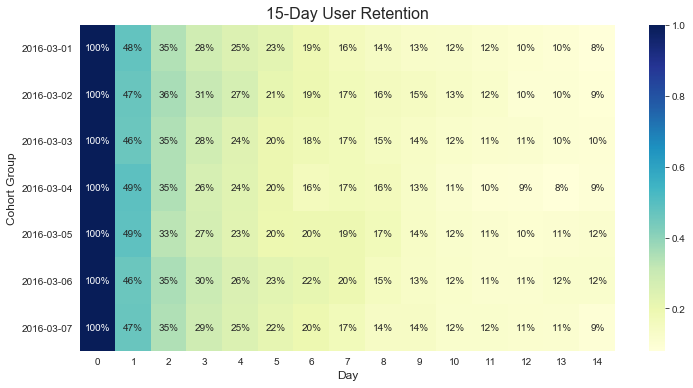
60 days

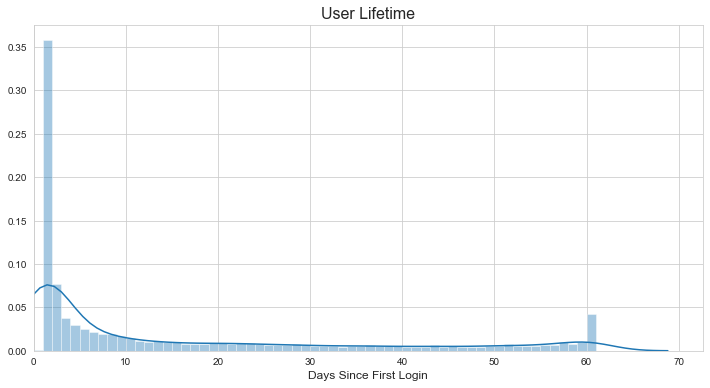
The data includes:

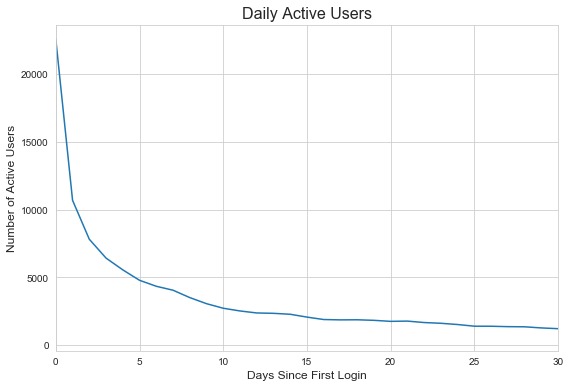
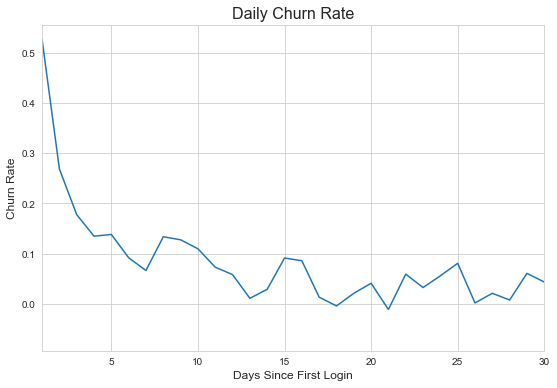
1. User data, including user ID and install date
2. Session history, including date and session number
3. Purchase history, including date and amount
4. Spending history, including date, currency, and amount
5. **EDA**

**3.1 User Retention & lifetime**

1. About 80% users finish/leave the game in 7 days, 90% users leave in 15 days
2. Highest churn rates at: T = 1, 2, 3, 5, 4, 8

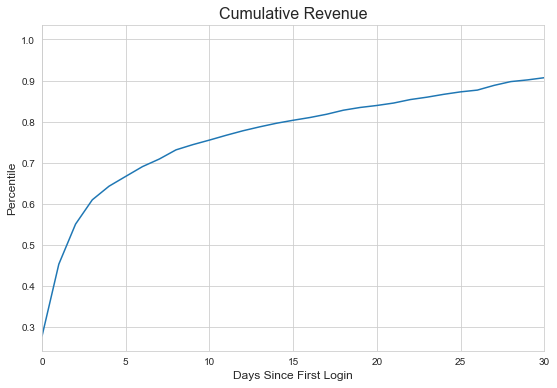


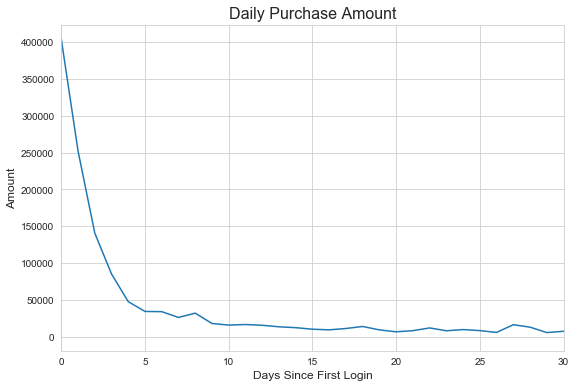
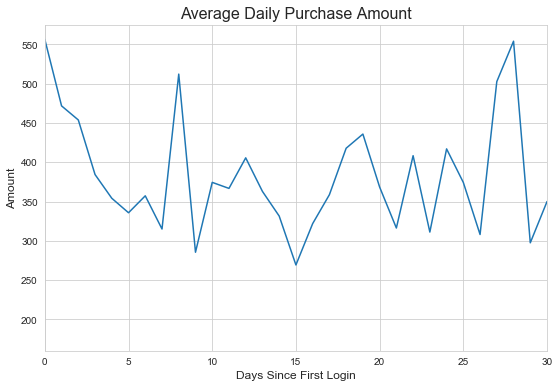


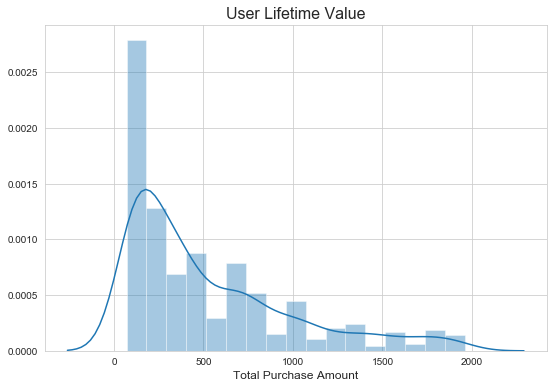
 

**3.2 Revenue**

1. 55% of revenue was gained in the first 3 days, 69% in 7 days, 80% in 15 days
2. Highest daily total revenue: T = 0, 1, 2, 3
3. Highest average revenue per paid user: T = 0, 8, 1, 2

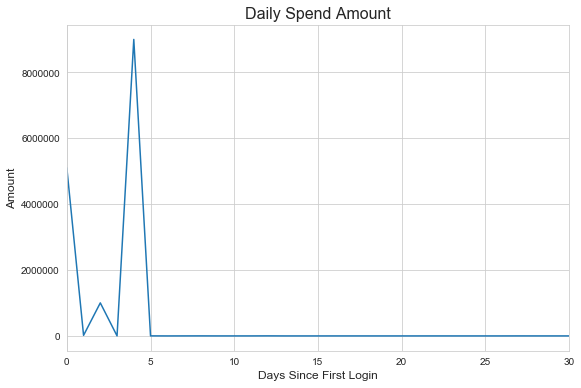
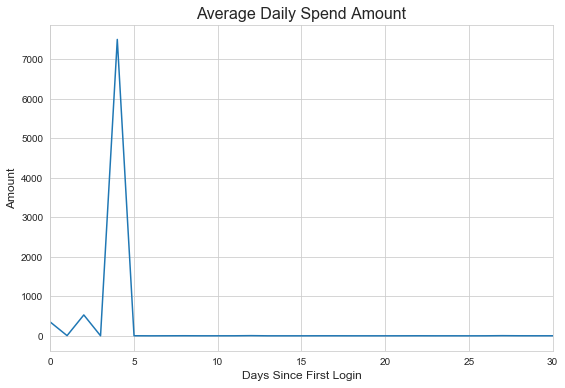




**3.3 Spend**

1. Highest total spend: T = 4, 0, 2
2. Highest average spend per user: T = 4, 2, 0

1. **Feature Engineering**

**Meta**: country, language, device

**Behavioral**: session count, active days, chapter count, days since last session,

purchase amount, purchase count, days since last purchase,

spend amount, spend count, days since last spend

**Composite**: average session count, median session count, average days between sessions, median days between sessions

1. **Model Evaluation**

LightGBM, random forest, neural network

**One-day Performance Window**

Start day: T = 0

Performance window: 1

Observation window: 1

Conversion rate: 2.37%

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Train score** | | | **Test score** | | |
| **Algorithm** | **auc** | **precision** | **recall** | **auc** | **precision** | **recall** |
| LGB | 0.8026 | 0.0522 | 0.6928 | 0.7838 | 0.0848 | 0.5955 |
| **RF** | 0.7942 | 0.0539 | 0.7040 | 0.7977 | 0.0525 | 0.7079 |
| NN | 0.7860 | 0.0504 | 0.6839 | 0.8004 | 0.0521 | 0.7079 |

**Two-day Performance Window**

Start day: T = 0

Performance window: 2

Observation window: 1

Conversion rate: 1.38%

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Train score** | | | **Test score** | | |
| **Algorithm** | **auc** | **precision** | **recall** | **auc** | **precision** | **recall** |
| **LGB** | 0.8948 | 0.0543 | 0.8062 | 0.9240 | 0.0629 | 0.8846 |
| RF | 0.8867 | 0.0515 | 0.7946 | 0.9184 | 0.0567 | 0.8654 |
| NN | 0.8781 | 0.0512 | 0.7946 | 0.9048 | 0.0505 | 0.8269 |

**Three-day Performance Window**

Start day: T = 0, 1, 2, 3

Performance window: 3

Observation window: 1

Conversion rate: 1.01%

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Train score** | | | **Test score** | | |
| **Algorithm** | **auc** | **precision** | **recall** | **auc** | **precision** | **recall** |
| **LGB** | 0.8944 | 0.0387 | 0.7995 | 0.8837 | 0.0535 | 0.7753 |
| RF | 0.8753 | 0.0346 | 0.7793 | 0.8769 | 0.0347 | 0.7865 |
| NN | 0.8776 | 0.0360 | 0.7860 | 0.8543 | 0.0318 | 0.7640 |

**Seven-day Performance Window**

Start day: T = 0

Performance window: 7

Observation window: T ∈ [7, 60]

Conversion rate: 2.37%

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Train score** | | | **Test score** | | |
| **Algorithm** | **auc** | **precision** | **recall** | **auc** | **precision** | **recall** |
| **LGB** | 0.9250 | 0.1085 | 0.8341 | 0.9331 | 0.1131 | 0.8427 |
| RF | 0.8857 | 0.0547 | 0.8062 | 0.9047 | 0.0558 | 0.8077 |
| NN | 0.8774 | 0.0853 | 0.7937 | 0.8980 | 0.0887 | 0.8202 |

SMOTE-NC tend to produce more FP, thus doesn’t improve test score.

1. **Recommendations**

**6.1. Sales Offers**

Offer 1

* 2nd – 7th day, multiple packages with decreasing discount, e.g., get 500 gems with 50% off, 1000 gems with 40% off…
* To users predicted to be unconverted on the next day given their recent 1-day/2-day/3-day behavior

Offer 2

* after 7 days
* at critical chapters (with highest spend)
* To users predicted to be unconverted on the rest of their lifetime given their recent 7-day behavior

**6.2 Test protocol**

Control group: send promotion offer to all users

Test group: send offer to predicted non-payers only

Duration: 20 days

Metrics:

1. Profit = revenue - cost
2. Conversion rate
3. Daily active users, paid users

Ref:

<https://www.aaai.org/ocs/index.php/AIIDE/AIIDE15/paper/view/11544/11359>

07/11/2020

05/25/2020

**N/A**

**MS Communication Data Science**