Napster 2.0 Churn

Predicting customer churn rate for the next month

Junhai Ma, Cinny Lin

The problem

Company	Context	Problem statement
Napster 2.0 is a music streaming service.	Subscribers pay **$10** for access to the service monthly.	Customers have an average lifetime of only **1.75 months**.

Executive Summary

Revenue impact

Offering a 30% discount to potential churners can **improve our revenue by 7%** under 30% acceptance rate.

Scope of Offering

We should offer discount to those with **more than 60% probability of churning** as predicted by our model
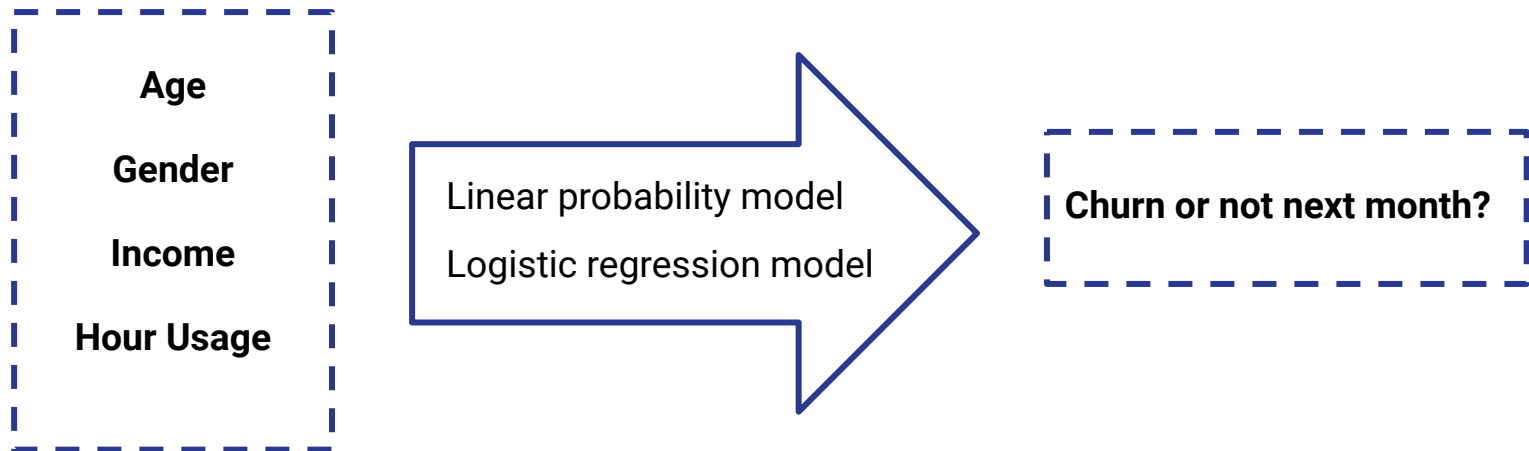
Solution

Offer discount to churners

According to our models and projected revenues, we can see that offering discounts to potential churners will extend customer lifetime value and increase our projected revenue.

Implementation

Solution

Predictive modeling to identify those who are likely to churn

Age

Gender

Income

Hour Usage

Linear probability model

Logistic regression model

Churn or not next month?

Linear probability model

```
                           OLS Regression Results
==============================================================================
Dep. Variable:               churn_TF   R-squared:                       0.122
Model:                            OLS   Adj. R-squared:                  0.122
Method:                 Least Squares   F-statistic:                     244.1
Date:                Tue, 17 Nov 2020   Prob (F-statistic):           1.46e-196
Time:                        10:09:22   Log-Likelihood:                 -4546.0
No. Observations:                7000   AIC:                             9102.
Df Residuals:                    6995   BIC:                             9136.
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
age            0.1309      0.006     23.618      0.000       0.120       0.142
income        -0.0903      0.006    -16.391      0.000      -0.101      -0.080
hours         -0.0636      0.006    -11.531      0.000      -0.074      -0.053
female         0.0590      0.011      5.322      0.000       0.037       0.081
const          0.5465      0.008     70.410      0.000       0.531       0.562
==============================================================================
```

Logistic regression model

```
                      Logit Regression Results
=================================================================================
Dep. Variable:                 churn_TF    No. Observations:               7000
Model:                            Logit    Df Residuals:                   6995
Method:                             MLE    Df Model:                          4
Date:                Tue, 17 Nov 2020    Pseudo R-squ.:                 0.09580
Time:                          08:34:50    Log-Likelihood:               -4317.9
converged:                         True    LL-Null:                      -4775.4
Covariance Type:              nonrobust    LLR p-value:                9.545e-197
=================================================================================
                 coef     std err         z      P>|z|     [0.025      0.975]
---------------------------------------------------------------------------------
age            0.6109       0.028    21.887      0.000      0.556       0.666
income        -0.4219       0.027   -15.818      0.000     -0.474      -0.370
hours         -0.2967       0.026   -11.315      0.000     -0.348      -0.245
female         0.2755       0.052     5.321      0.000      0.174       0.377
const          0.2113       0.036     5.876      0.000      0.141       0.282
=================================================================================
```
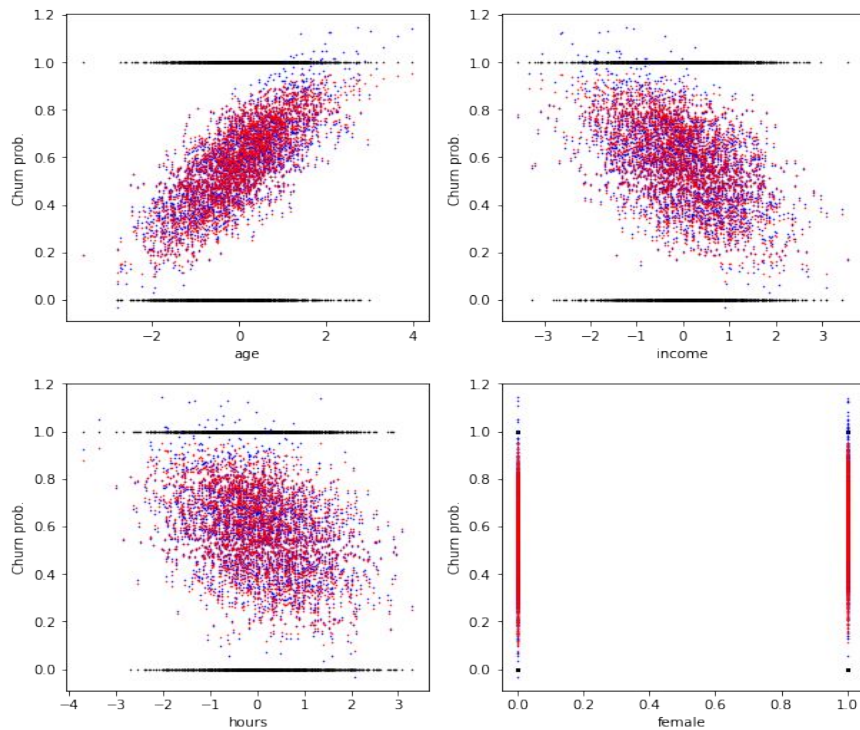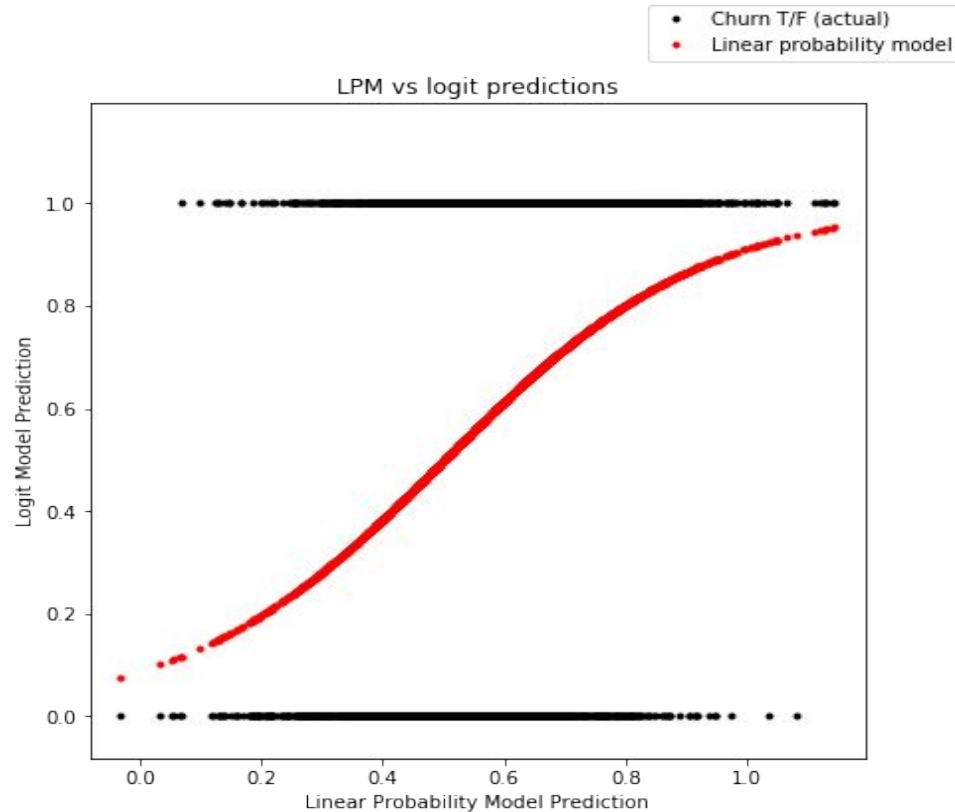
Feature importance

Single factor relationship with churn

Model consistency

ROC curve
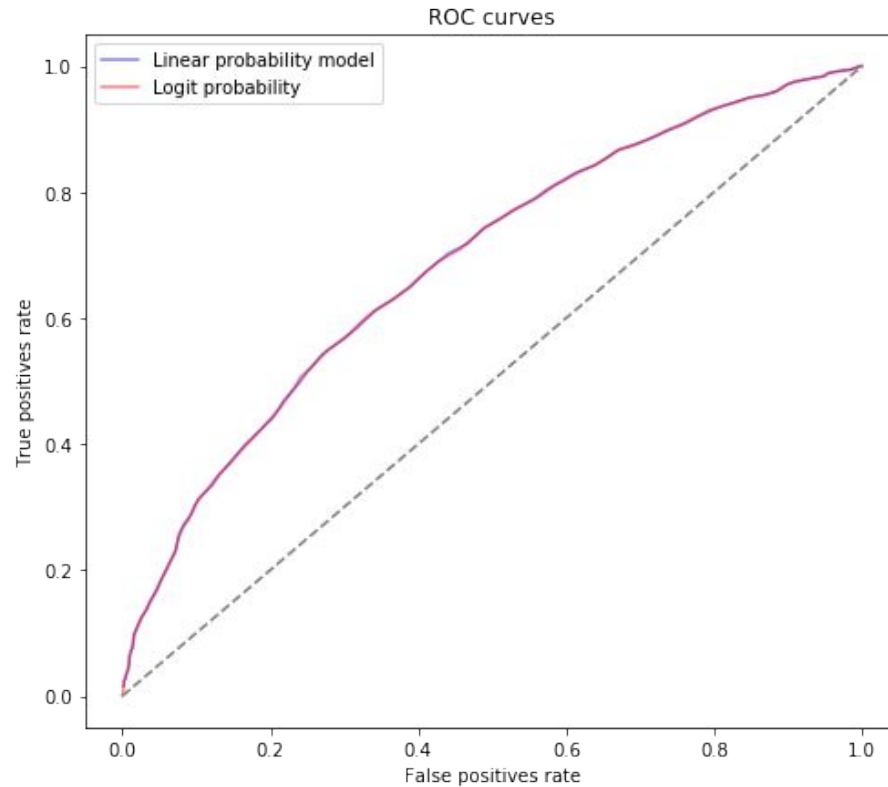
Impact of probability threshold on revenue

Lift curves

- Linear probability model
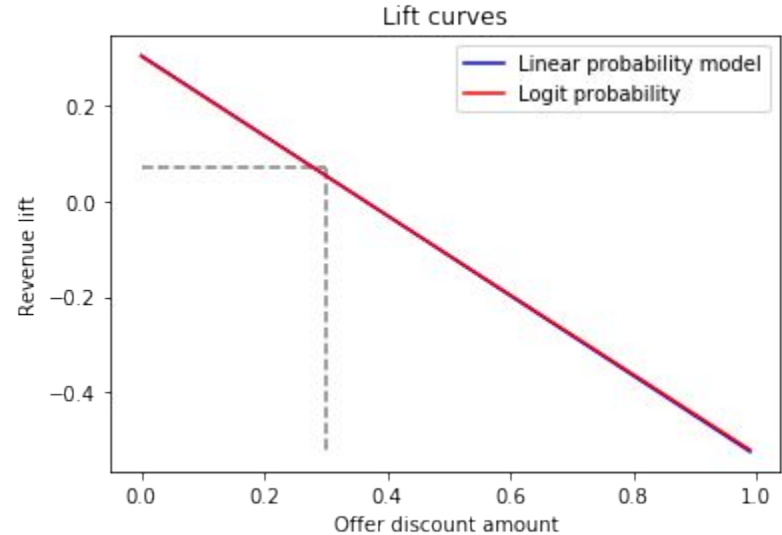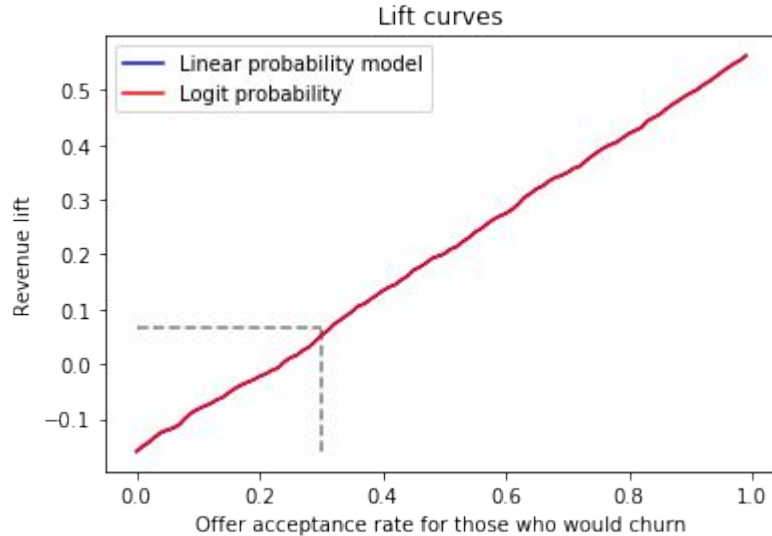- Logit probability
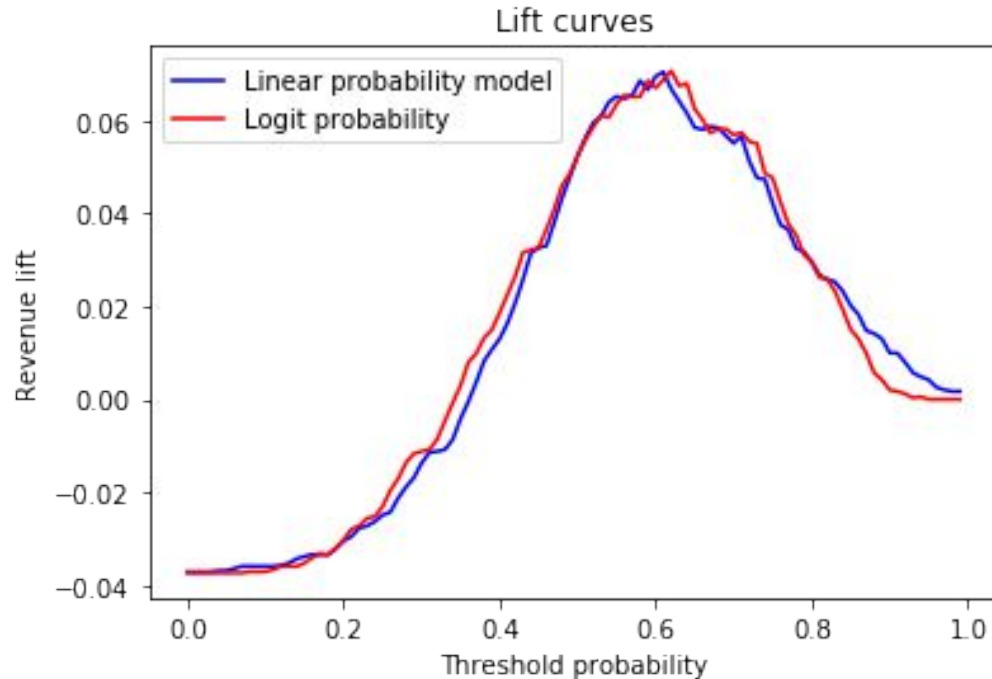
Under assumptions:

1. Discount amount 30%
2. Acceptance rate 30%

Sensitivity of our assumptions

Under assumptions: probability threshold 50%

Other issues

1. **How much should the discount be?**

    It depends on the relationship between discount rate and acceptance rate.

2. **What are channels to offer discount to potential churners?**

    Email could have a higher acceptance rate than phone call

3. **Offering discount is expensive, how to cut more cost?**

    Call or send emails checking customer satisfaction before making the offer

Thank you!

Any question?