Which customers are unhappy?

Nadya Agrawal

March 2023


METALLICA
I understand your frustration.

1. Overview & Business Problem
2. The Data
3. The Method
4. The Model
5. Findings
6. Recommendations
7. Next Steps

Overview & Business Problem

Our e-commerce company loses money when we lose customers.

How can we predict which customers will leave us before they do?

Through machine learning, we can determine the best predictors for churn.

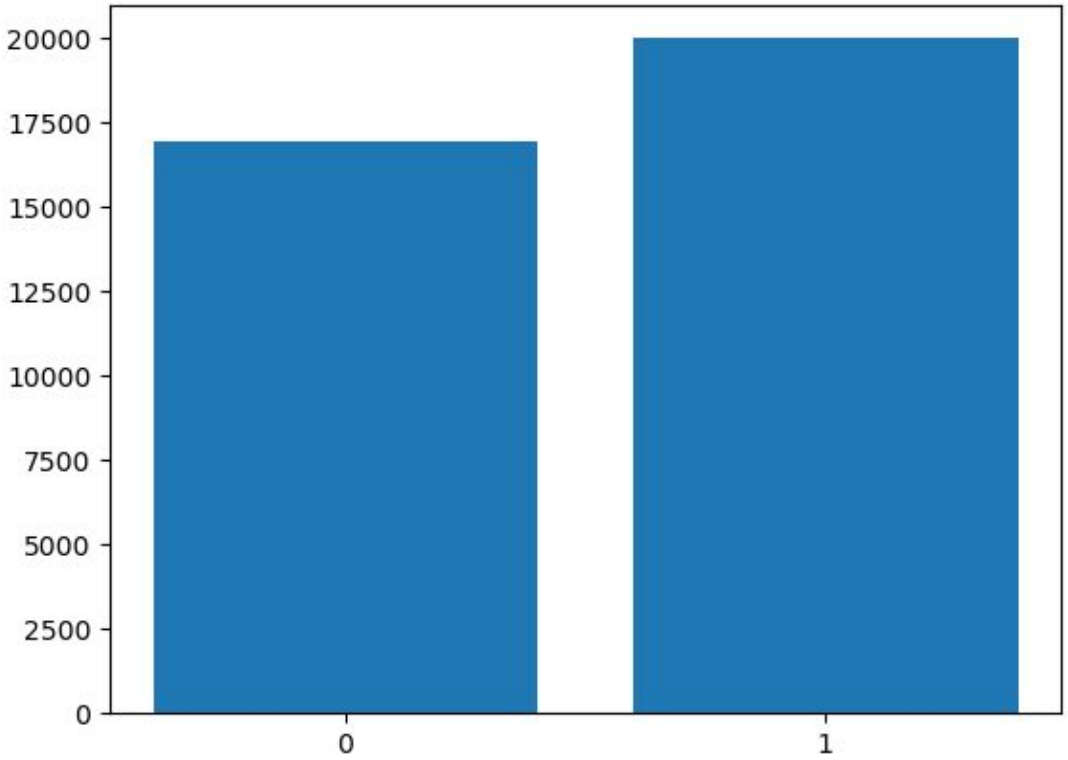
The Data

The data comes from the **“Customer Churn”** dataset, posted by Pawan Trivedi on Kaggle.

It contains data on membership, website use, transactions, and more.

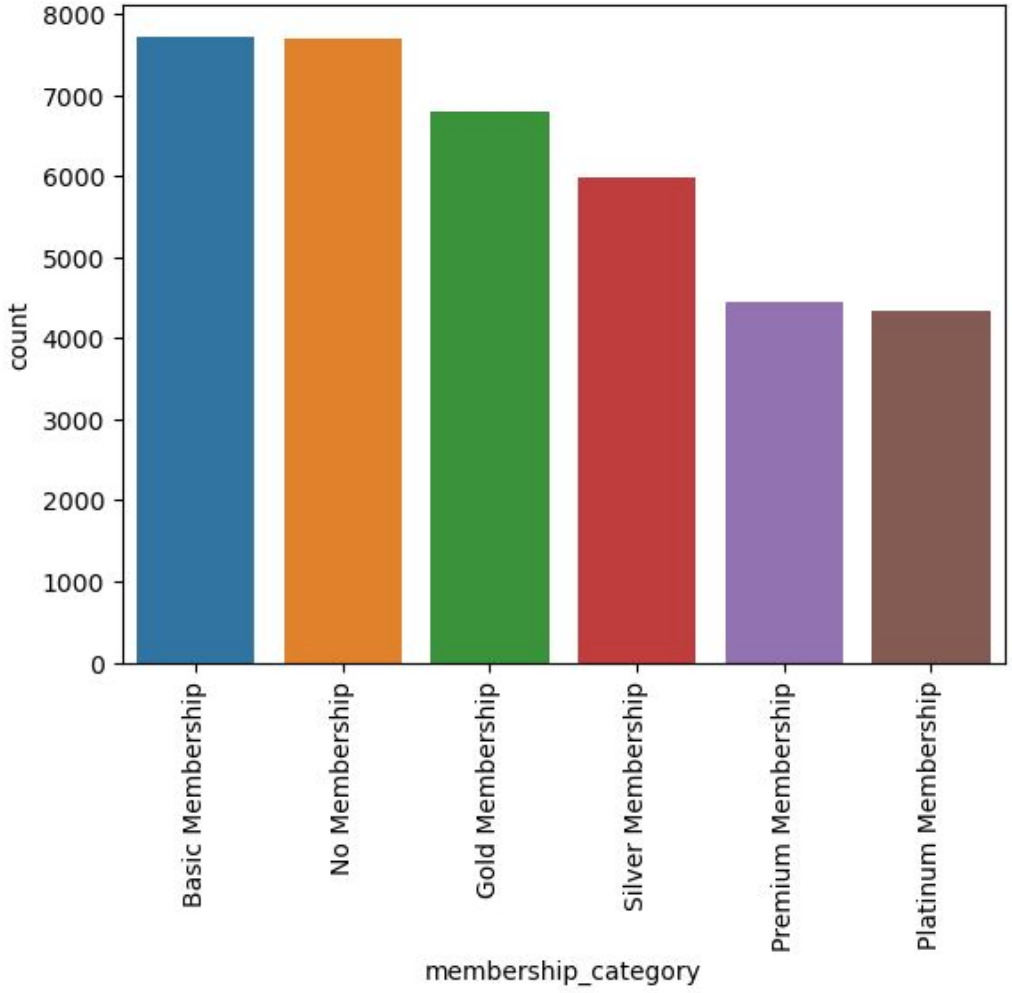
Data Split

Pretty even split between those at risk of churn and not



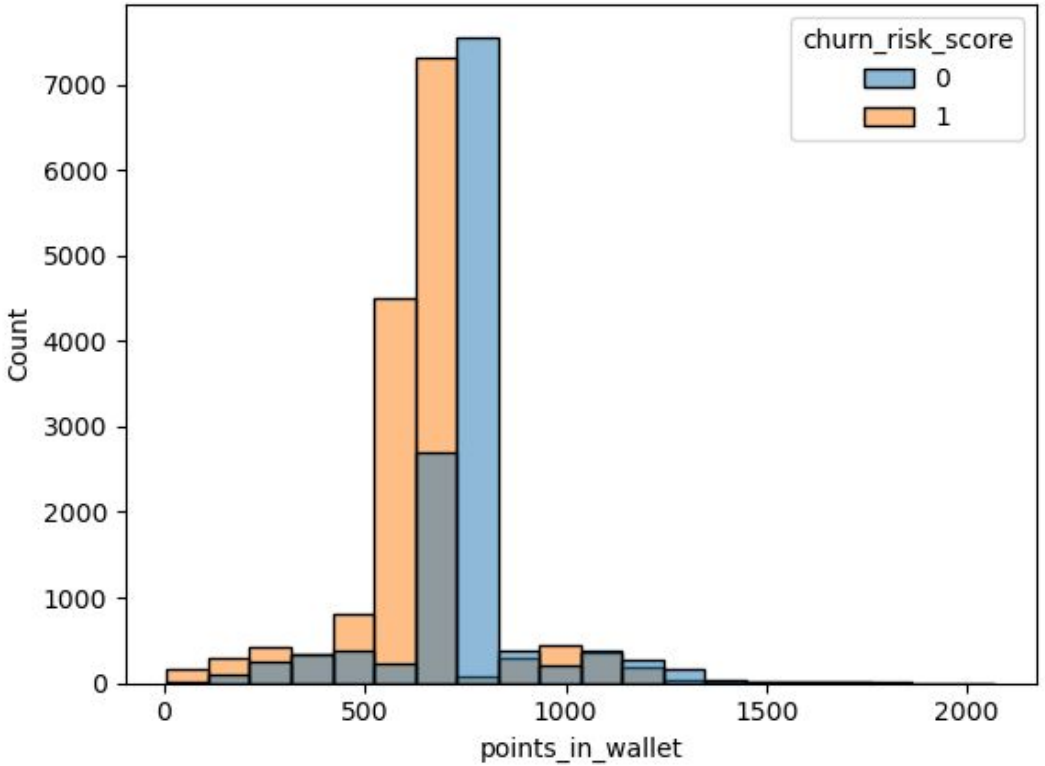
Membership

- Six membership types
- Most are in the **“basic” or “no”** membership categories, these are **most are at risk** of churn
- The data shows everyone with a “platinum” membership is not at risk of churn



Points

Some overlap, but key sections where churn and not churn are dominant

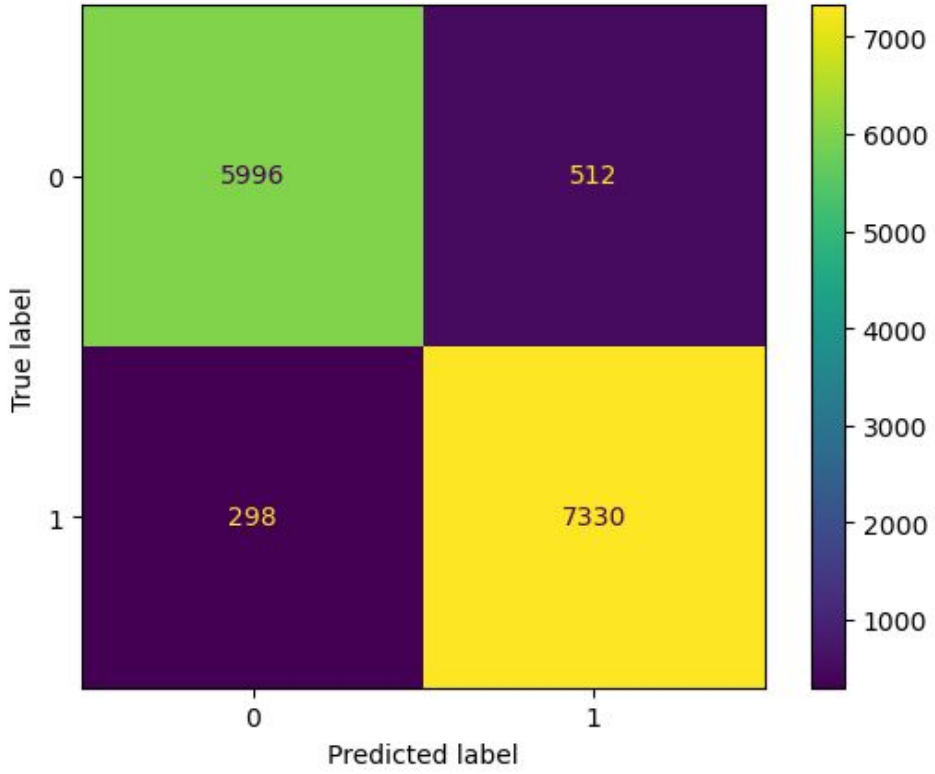


The Model

- Ran a range of models, found XGBoost Classifier to be the best
- Removed negative values since unreadable
- Wanted to minimize false negatives

Confusion Matrix

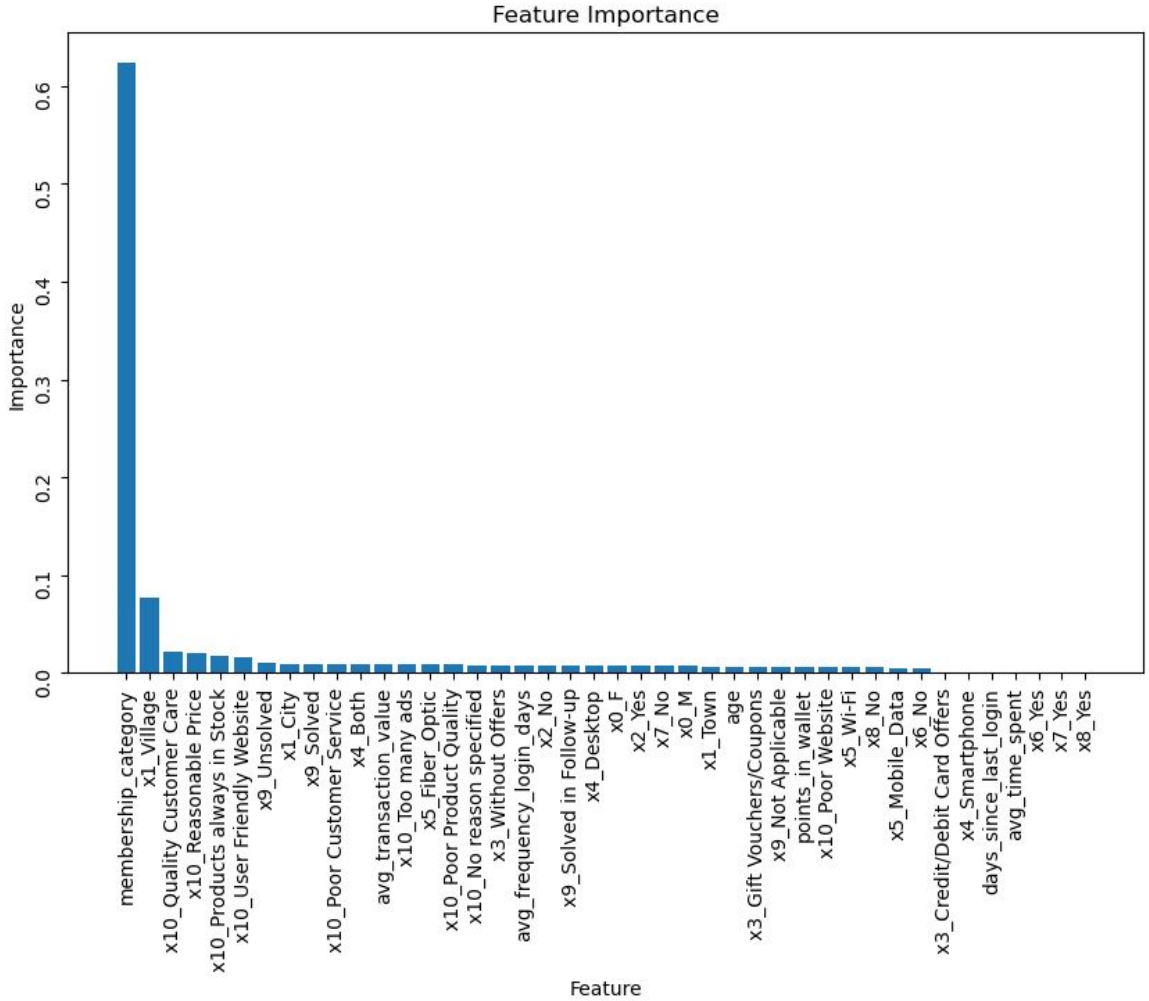
XGBoost Classifier **minimized false negatives** the best compared to other models



Features

Membership category is by far the most important feature in our model.

Followed by whether the customer is in a **village** and if they've complained about **customer care.**



Recommendations

- Use XGBoost Classifier to predict churn risk
- Focus marketing on customers that are in lower membership categories or do not hold a membership
- Re-target customers who live in villages
- Re-target customers who have given feedback on customer care quality

Next Steps

1. **Look at previous marketing to at-risk customers.**

2. **Pull more recent data.**

3. **Determine whether re-targeting helps or hurts.**

THANK YOU

Nadya Agrawal

/npagrawal

nadyaagrawal@gmail.com