

A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light green. They are positioned diagonally, with the blue one partially covering the green one.

Online Shopping Intention

An analysis by Kai Tamashiro

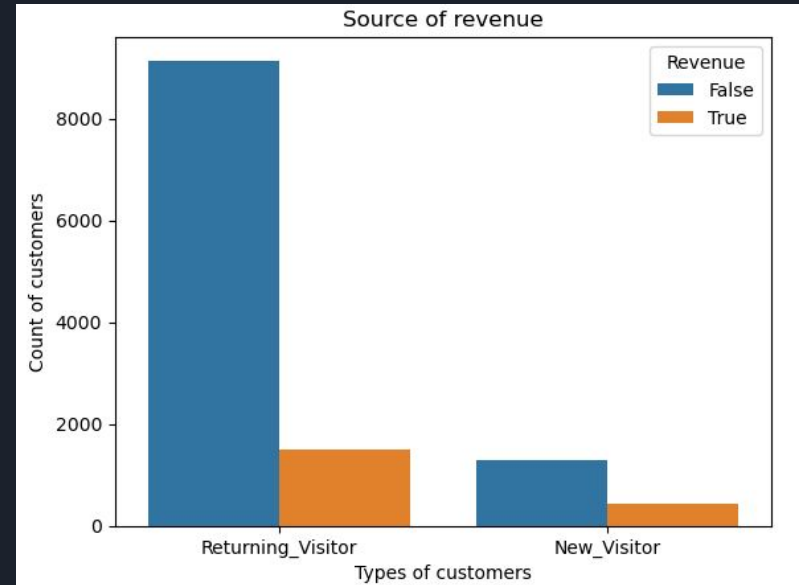
What can we do to
increase revenue?



Returning vs new visitors

Who tends to spend?

How can we nurture these sales?





Finding the best model

```
sm = SMOTE(random_state=42)
X_train05, y_train05 = sm.fit_resample(X_train, y_train)
Counter(y_train05)

Counter({False: 8307, True: 8307})
```

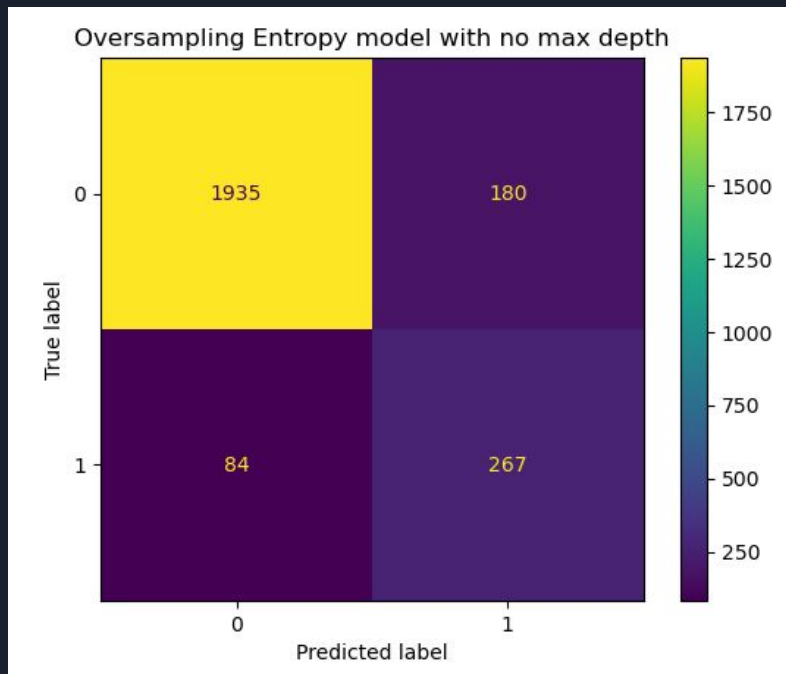
Over sampling to
create equal classes

Random Forest: Entropy

Avg precision: .91

Avg recall: .89

F1-score: .90

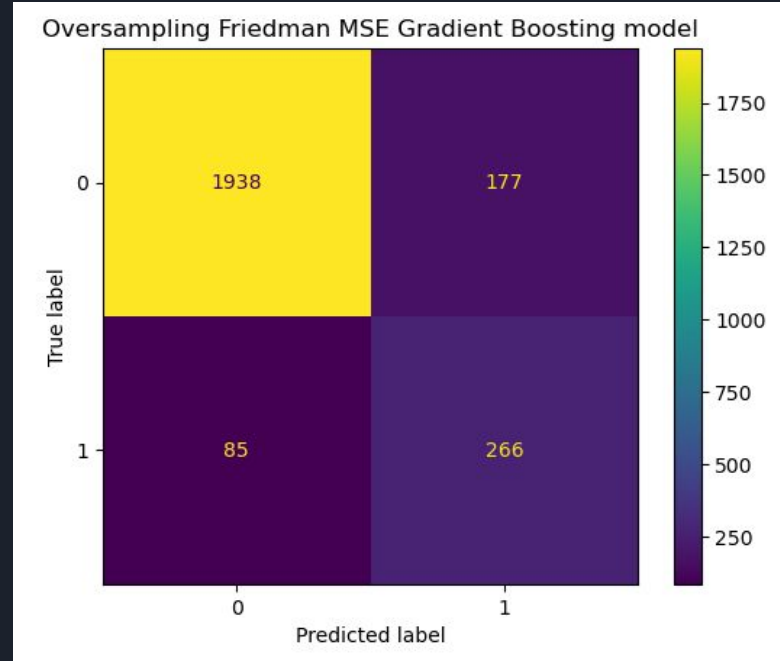



Gradient Boosting

Avg precision: .91

Avg recall: .89

F1-score: .90





Finding the best model, part 2

```
nm = NearMiss()  
X_trainUS, y_trainUS = nm.fit_resample(X_train, y_train)  
Counter(y_trainUS)  
  
Counter({False: 1557, True: 1557})
```

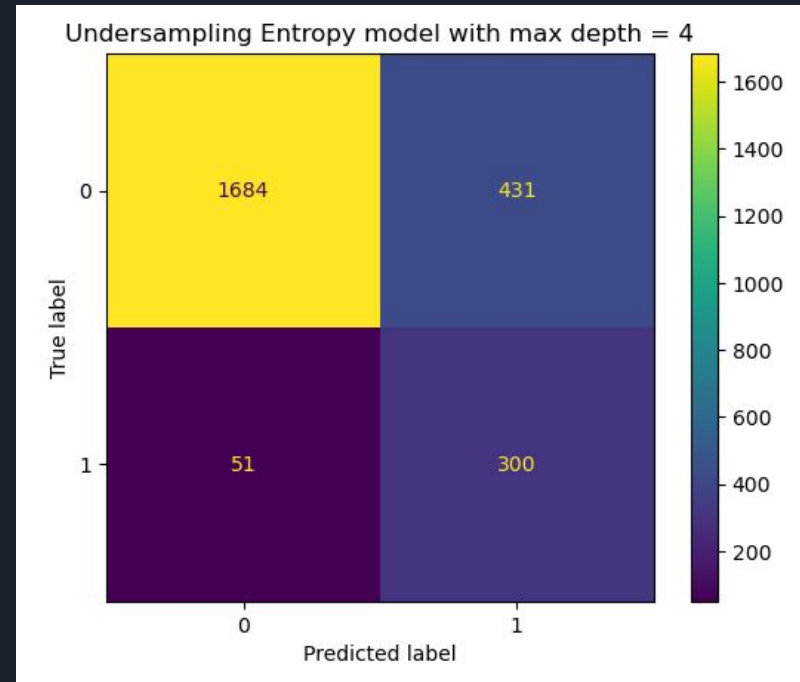
Under sampling to
create equal classes

Random Forest: Entropy with max depth

Avg precision: .89

Avg recall: .80

F1-score: .83

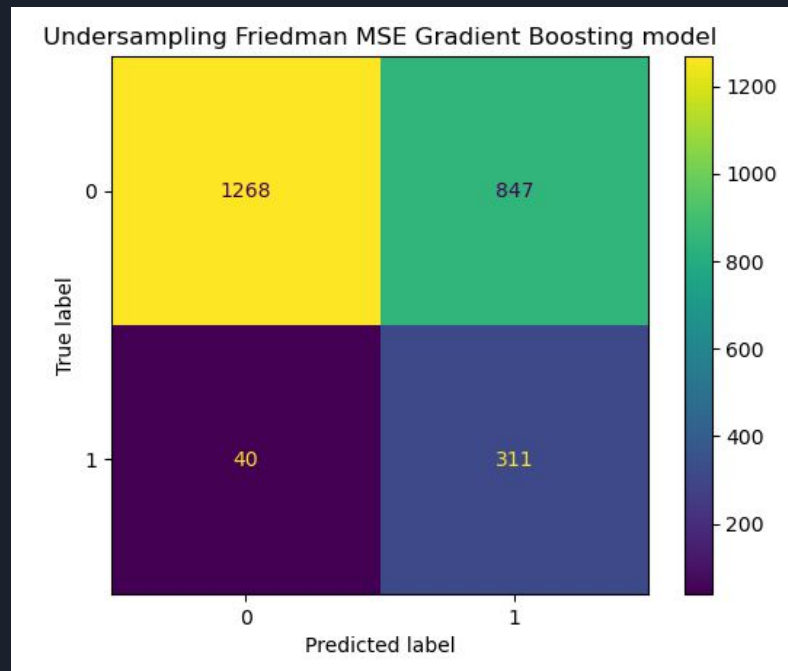


Gradient Boosting, part 2

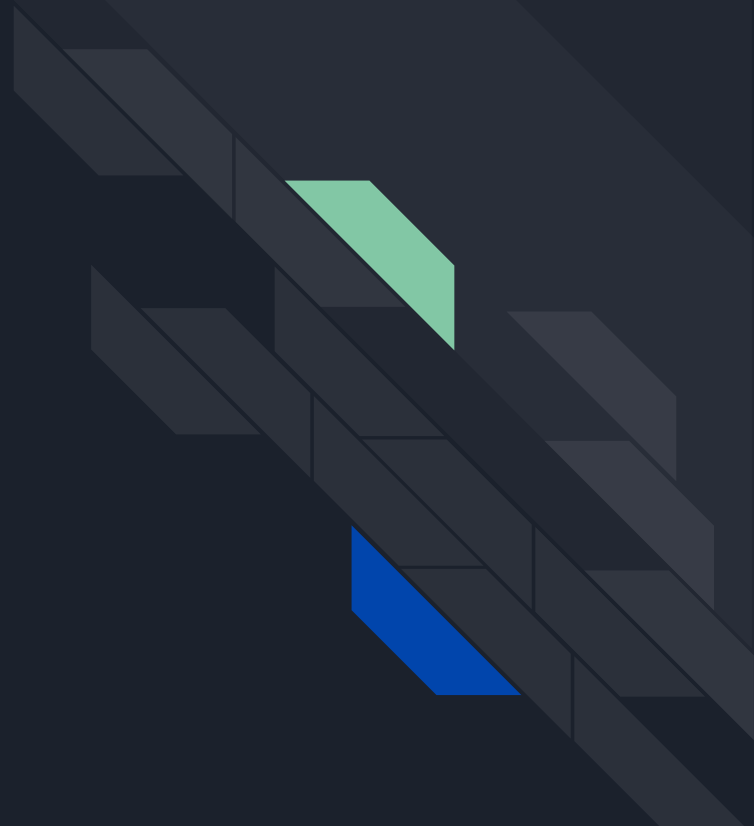
Avg precision: .87


Avg recall: .64

F1-score: .69



Next steps





Implement Entropy Random Forest Model

Find the customers that are spending money,
cater marketing towards them

Track those returning customers and reward
them for spending money

Most of the money comes from returning
visitors - how do we increase this number?



Psychology behind sales



Can website design affect how likely someone is to purchase something?

