# Coupon Challenge

Raul Sanchez-Vazquez

25 June 2018

# The Problem

The problem:

- Create a recommender system for coupons.

This would be done in two stages:

- First, learn a model for classic top-n recommendation.
- Second, learn a binary classification model where the positive class correspond to actual purchase of the item.

Produce recommendation as:

- Produce candidate items (RecSys Model)
- Order by purchase probability (Positive probability score)

# Analysis

We first built data profiling documents for:

- The item catalog.
- The user catalog.
- The user session log with purchase information.

We first noticed that the binary classification model would have to deal with class unbalance as shown on the plot below:

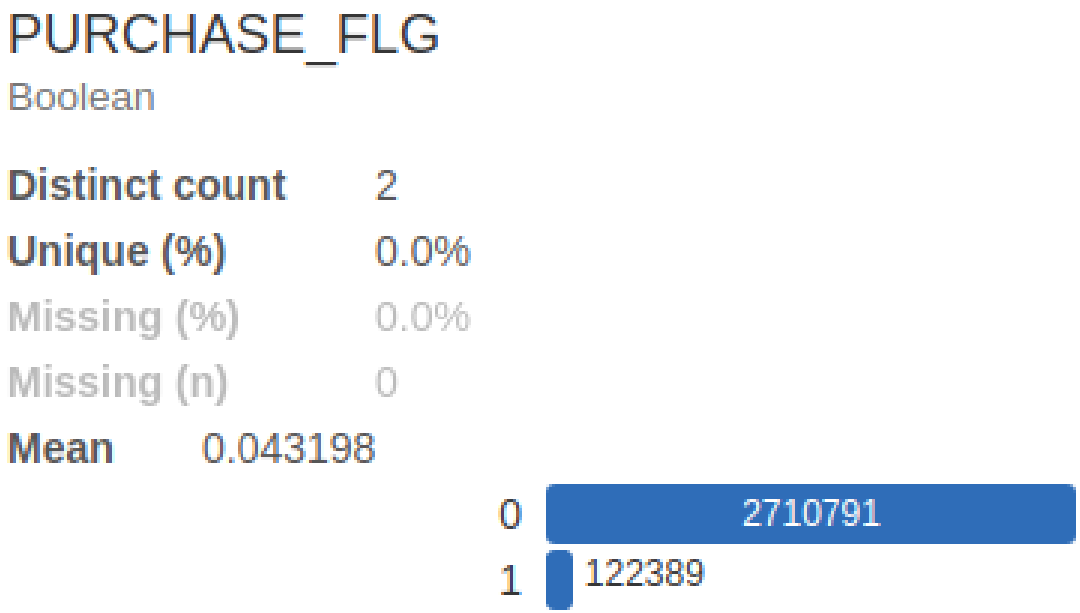



Figure 1

# Features

## Item features:

- item age promo
- item n views
- item n purchases
- item area
- item ken
- item small area
- item ken most buy
- item ken most buy2
- item large area most buy
- item large area most buy2
- item small area most buy
- item small area most buy2
- item buyers age mean
- item buyers age median
- item buyers age std
- item DISCOUNT PRICE percentage
- item price
- item purchase sex f
- item purchases sex m

# Features

## User features:

- user same item n views
- user same item n purchases
- user last item large area name
- user last item ken name
- user last item small area name
- user purchases same item area
- user purchases same item ken name
- user purchases same item small area
- user AGE
- user SEX ID

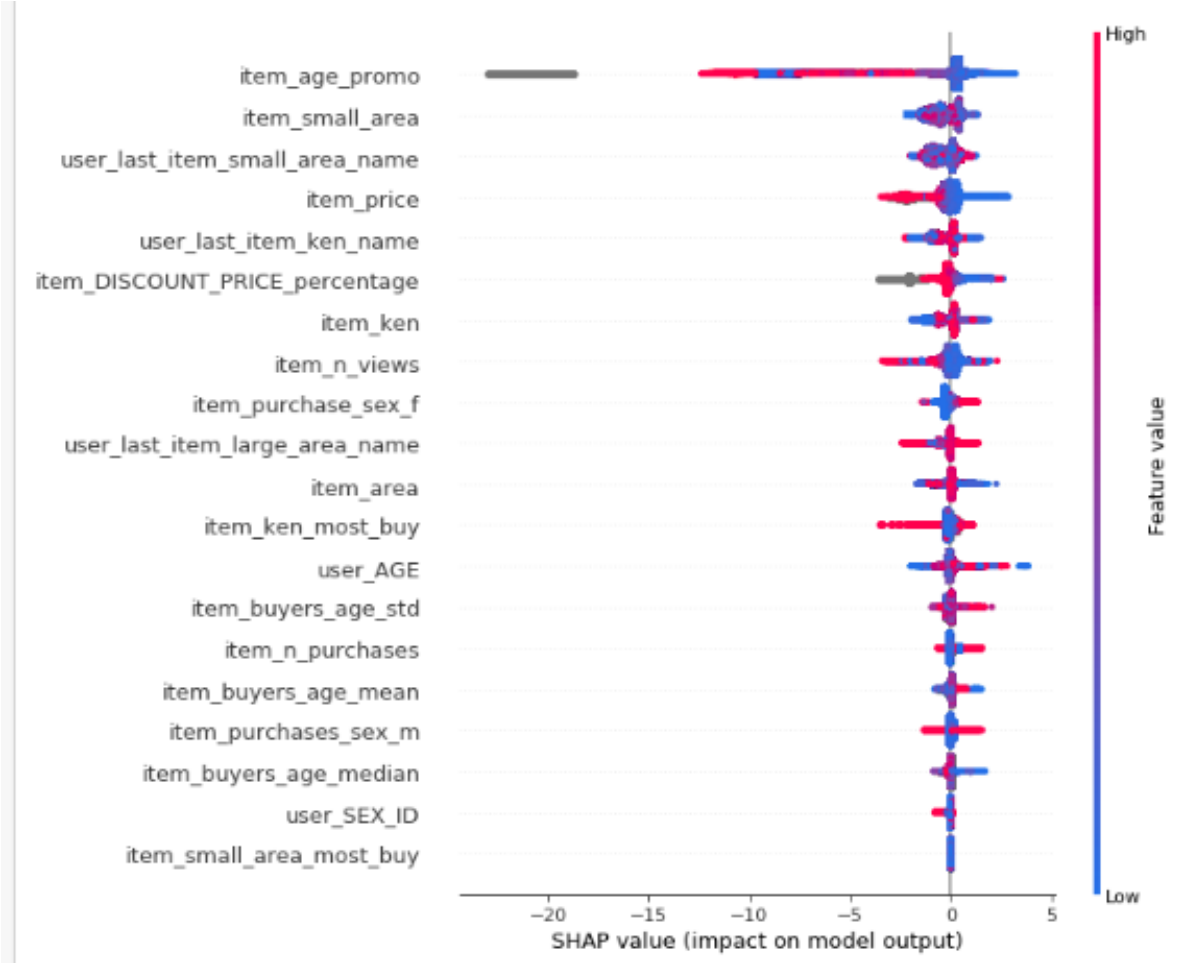
# Model Performance

Recsys on Test:

- P@10: 0.0018949648

Binary Classification on Test:

- AUC:0.900254



# Model Performace

Hybrid recommender:

- P@10: 0.029458179905312992