Background

As smart phone penetration reaches the hundreds of millions mark, O2O (Online to Offline) requires businesses to have a strong presence both offline and online. APPs with O2O capabilities accumulate daily consumer behaviour and location data that require big data and commercial operations management. The competition at hand focuses on coupon redemption rates. Sending coupons is a general O2O marketing tool used to activate existing customers and attract new ones. While customers are happy to receive coupons that they want, they are frustrated when receiving coupons that they do not need. For merchants, sending unwanted coupons may erode brand equity and hinder marketing expense forecasting. Targeted marketing is an important technology to increase the coupon redemption rate, providing relevant discounts to customers and effective marketing tools to businesses. The competition provides participants with abundant O2O data in this field and expects contestants to predict whether the customer will use the coupon within a specified time frame.

Data

This competition provides real online and offline user consumption data from January 1, 2016 to June 30, 2016. The contestants are expected to predict the probability of customers redeeming a coupon within 15 days of receiving it.

Note: To protect the privacy of users and merchants, data is desensitized and under biased sampling.

Evaluation

The results are evaluated based on the average AUC value. That is, the AUC value is calculated for every coupon_id. The average of each AUC value is the evaluation score. More information on AUC value calculation method on wikipedia.

Table 1: Offline consumption & coupons

Field	Description
User_id	User ID
Merchant_id	Merchant ID
Coupon_id	Coupon ID, when coupon_id = null, this means that a coupon has not been redeemed. In such case, Discount_rate and Date_received don't matter.
Discount_rate	Discount rate, range in [0,1]

Distance	500x, the distance from the nearest shop around the user for locations in which a user is most active. x range in [0,10]; 0 – less than 500 meters; 10 – more than 5 kilometres.
Date_received	Date the coupon is received
Date	Purchase date, When Date=null & Coupon_id!= null, users receive coupon but don't redeem it; When Date!=null & Coupon_id= null, purchase happened but no coupon had been received; When Date!=null & Coupon_id!= null, 'Date' in which the coupon was used.

Table 2: Online click/consumption & coupons

Field	**Description**
User_id	User ID
Merchant_id	Merchant ID
Action	0 - click, 1 - buy, 2 - getcoupon
Coupon_id	Coupon ID, when coupon_id = null, this means that a coupon has not been redeemed. In such case, Discount_rate and Date_received don't matter. 'fixed' means Limited Time Offer.
Discount_rate	Discount rate, range in [0,1]. 'fixed' means Limited Time Offer.
Date_received	Date the coupon is received
Date	Purchase date, When Date=null & Coupon_id!= null, users receive coupon but don't redeem it; When Date!=null & Coupon_id= null, purchase happened but no coupon had been received; When Date!=null & Coupon_id!= null, 'Date' in which the coupon was used.

Table 3：Offline coupon redemption prediction

Field	Description
User_id	User ID
Merchant_id	Merchant ID
Coupon_id	Coupon ID
Discount_rate	Discount rate, range in [0,1].
Distance	500x, the distance from the nearest shop around the user for locations in which a user is most active. x range in [0,10]; 0 – less than 500 meters; 10 – more than 5 kilometres.
Date_received	Date the coupon is received.

Table 4：Submit file

Field	Description
User_id	User ID
Coupon_id	Coupon ID
Date_received	Date the coupon is received
Probability	Probability of using coupon in 15 days within it being received