

Atkin attığı +92 kurbaya dök gelsin

rfm metrikleri → RFM skorları

keşfi kendi içinde ve birbirleri arasında kıyaslanabilir hale geldi

R	F	M
küçük iyi	büyük iyi	büyük iyi
en son alışveriş yaptığı gün ile bugün arasındaki fark	alışveriş yapma sıklığı	malzemenin harcandığı parametresi

RFM skorları

1	2	3	4	5	ile skorlandırılır
en kötü				en iyi	

5 5 5 → en iyi müşteri

1 1 -1 → en kötü müşteri

Frequency işlemler daha önemli  
monitörün önemi

R	F	bakılır
5	5	→ champion

RF ile Segmentler oluştur

df.shape

(gözlem birimi sayısı, değişken sayısı)

örn: (400, 8)

inplace = True → değişkenlerin kütlesi değişmemi sağlar

inplace = False → kütlesi C olarak geri döndürür

~ kütlesi değil aynı kütlesi

unique - number unique eşsiz olanın sayısı

tekrar eden şeyler gördüğümüz oluyor



qcut - quantile (dağılımlara göre) işlem yapar

qcut verilen değeri belirli büyüklüğe sınıflar

belirli büyüklükte label (katman) ayırır ve

o katmanlara vermiş olduğu skoru verir

örn: 5 label olur 100'den 0-20 20-40 40-60 60-80 80-100 olur

labels = [ilk 20, ikinci 20, , , son 20]

kuşakları büyüğe göre

labels = [5, 4, 3, 2, 1]

ilk 20 ye girenlere 5 skoru veririm

son 20 ye girenlere 1 skoru veririm.

index -> idlerini getirecek

- BG-NBD -

BG-NBD satın alma / işlem sayısını tahmin etmek için kullanılan model.

bu işlemi 2 aşamada ele alır: Buy Till you Die

Satın alma

(Buy)

• beta dağılır

Drop olma

(Till Die)

• gamma dağılır

X: bir müşterinin tekrar eden satış sayısı (yani en az 2 kere tekrar eden satış sayısı)

tx: recency, bir müşterinin ilk satın alması ile son satın alması arasında geçen süre (haftalık)

\* Bu RFM'deki recency değil!

Bu müşteri özelinde bir recency!

Yani Today date'e göre değil

Her bir müşterinin kendi ilk ve son aldığı arasındaki fark.

T: müşterinin ilk satın alması üzerinden geçen zaman

Yani müşterinin yaşı (haftalık)

X, tx, T -> elimizdeki her müşterinin kendi içinde değişen değerleri

r, a -> Gamma dağılımının parametre değerleri

a, b -> Drop rate, inaktif rate olasılığı hesaplayan parametreler

Bu parametreleri tahmin etmek için 'maksimum olasılık yöntemi' kullanılmaktadır.



$$E(Y(t)|X)$$

X rasal değişkeni belirli bir t periyodunda beklenen transaction sayısı

formül içindeki F → Gauss hiper-geometrik fonksiyonunun ifadesidir.

Gamma Gamma

X: frequency değeri, tekrardan satın sayısı yani en az 2. kez işlem yapmış alma sayısı

mx: monetary, yani gözetlenen transaction value'ları  
Total price / Toplam işlem sayısı

p, q, γ → dağılımdan gelen ifadeler

$$E(M | p, q, \gamma, mx, x)$$

kizi ve dağılım özellikleri girildiğinde monetary değerinin beklenen değerini verir

BG-NBD

Gamma Gamma

$$\left( \begin{array}{l} \text{Expected number of transaction} \\ \text{beklenen işlem sayısı} \end{array} \right) \times \left( \begin{array}{l} \text{Expected average profit} \\ \text{beklenen ortalama kâr} \end{array} \right) = CLTV$$

	RFM	CLTV	frequency = transaction
recency	(Today date - müşteri ilk satın aldığı tarih)	(müşteri son satın aldığı tarih - müşteri ilk satın aldığı tarih)	
(T) Frequency	haftalık	haftalık	
monetary	haftalık	aylık	