<Buy it again 흐름>

**\*step1 - 해당 사용자의 구매 이력 조회**

|  |  |
| --- | --- |
| 1.구매 이력 존재 경우 | 상품 Category별 RCP 계산 진행 -> 추천 Score 제공 가능! |
| 2.구매 이력 미존재 경우 | *RCP*점수자체가추천 *Score*로제공*!* |

**\*step2 - time-interval 계산 & Gamma fitting시 주의할 점**

: 카테고리 내에서 고객들마다의 time-interval을 구해야한다!

예를 들어 100명의 ‘기저귀’ 구매 고객이 있고,

11월 3일 일자에 ‘기저귀’ 를 구매한 A고객의 구매 이력이 오전 10시, 오전 11시 총 2회이고, 11월 5일 일자에 구매 이력은 오후 10시 1회라면, A고객의 구매 주기는 |5일-3일|=2일, 구매 횟수 k=3으로 계산된다. 이와 동일하게 나머지 99명의 구매 주기도 구해져야 한다.

: (주의) 구매 이력이 1회인 고객의 경우 time-interval은 ’NaN’으로 계산되고, ’NaN’ 인 time-interval값은 Gamma fitting에 개입되지 않는다!

: (주의) 구매 이력이 k회 존재하지만, 모두 동일 일자에 구매하여 time-interval이 ‘0’으로 계산되는 경우 => ’NaN’의 경우와 마찬가지로 ‘0’의 time-interval은 Gamma fitting에 개입되지 않는다!

: 또한, Gamma fitting을 하는 값은 time-interval값이 아닌, 1/time-interval값이다 : lambda는 ‘반복 구매율’을 의미하며, 이를 추정하기 위한 적절한 값은 time-interval이 아닌 ‘1/time-interval’ 이다.(time이 길어지면, 반복구매율은 짧아진다.)

**\*step3 - lambda값 계산시 주의할 점**

|  |  |
| --- | --- |
| **구매 이력이 1회인 고객의 경우, 고객의 구매 주기 ‘평균’을 구할 수 없는 issue** | : time-interval 구할 수 없다 -> 모든 고객들의 time-interval들의 대표 값을 쓴다*.(*고객들의평균등구매주기*)* |
| **구매 이력 >=1 But 동일 일자에 구매가 이루어져서, 고객의 구매 주기 ‘평균’ 을 구할 수 없는 issue** | : 이 경우에도 고객의 구매 time-interval 존재하지 않는다. -> 모든 고객들의 평균 구매 주기 사용. |

**\*step4 - Buy it again 추천 score**

: 실제 추천이 이루어 질 경우, Category별 ‘RCP점수’ X ’시점 구매확률’ 로 Buy it again Score를 계산한다.

**\*cf. 모델 훈련과정 간 issue되는 점**

: RCP가 계산되는 정도는, 4계절을 모두 살필 수 있는 1년이상 되면 좋겠다.

: unit time-interval을 어떻게 정의하느냐에 따라 score가 상이하다! : lambda의 정의해 의해서 이런 issue가 생기는데, 적절한 unit time 정의가 필요하다. 기본은 ‘일주일’로 정했다.(category마다 가진 mean\_time을 unit으로 둔다면? -하나의 방법이긴 하지만, category마다 상이한 unit time을 가진다면 합리적인 score비교가 불가능 할 것이다.)

: lambda를 구하는 훈련기간에 따라, 추천 score가 상이하다!