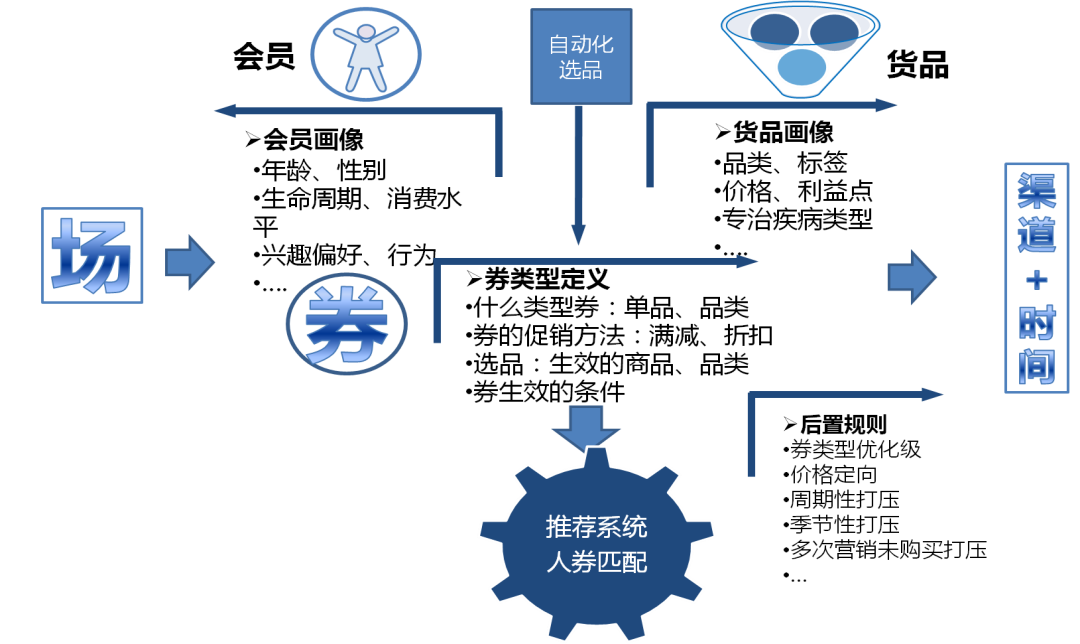
# 1精准营销概述

## 1.1基本思路

精准营销推荐算法在**特定的场景**下根据会员画像、会员行为、会员兴趣等，构建不同的会员群体模型。利用大数据，机器学习算法，捕获不同会员的对**商品的粘性、潜在需求**以及对**折扣的敏感性**，实现以**合适的承载方式**在**合适的时机**以**合适的渠道**实现**合适品种或品类**的推送，以此最大化平台收益。

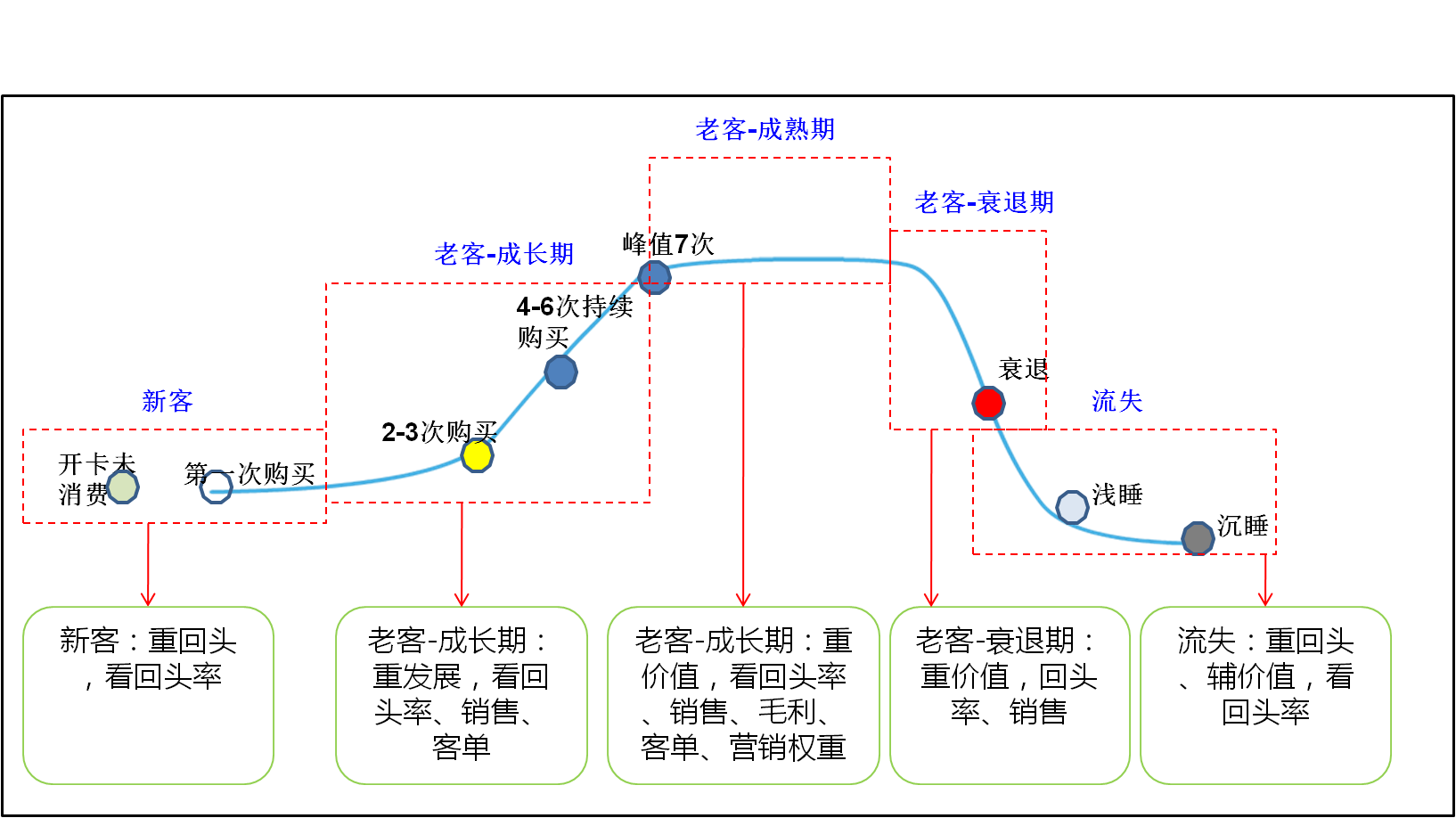


## 1.2会员生命周期分解

【V2-20180919】

### 1.21 会员生命周期

主流券精准营销以生命周期为主干预体系，如图所示：



各生命周期类型对应的口径、目的、策略及要求如下表所示：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **生命周期** | **口径** | **目的** | **策略** | **要求** |
| **新客** | 1、新客无消费：历史无消费会员  2、新客有消费：历史购买次数=1 | 侧重指标看**回头率** | 以细分维度找差异为主，尝试不同力度对回头率的影响 | 1、跟进实验进度、流程及实现过程中出现的问题，关注发券数的发送。  2、判断外部因素或内部因素导致出现的风险点  3、每天跟进效果情况，掌握每天回头的规律，判断数据异常  4、月度效果总结和分析  5、探索新的优化点（业务优化点及数据优化点） |
| **老客-成长期** | 近9个月有消费，同时购买次数为2~6次 | 侧重指标看**回头率、销售、客单** | 考虑成长期的特性：  1、探索会员（个体/群体）与商品间的‘桥梁’特征，评估会员对商品的偏好及需求程度。  2、探索差异性较大的特征，以此差异化运营策略  3、自动化选品的深度、宽度和准确度 |
| **老客-成熟期** | 近9个月有消费，同时购买次数为7次及以上 | 侧重指标看**回头率，销售，客单，毛利率，毛利额，营销权重** | 考虑成熟期的特性：  同‘老客-成长期’ |
| **老客-衰退期** | 近9个月未消费，购买次数>=2次 | 侧重指标看**回头率，销售** | 考虑衰退期的特性：  同‘老客-成长期’ |
| **流失** | 近12个月未消费，历史购买次数>=2次 | 侧重指标看**回头率** | 1、以细分维度找差异为主，尝试不同力度对回头率的影响  2、探索会员历史偏好及差异化特征对当前状态的影响，应用到当前的策略中  3、自动化选品的深度、宽度和准确度 |

### 1.22 特征加入标准

**标准：**后续各阶段的细分特征以策略干预形式加入，对比的方式 前期采用特征整体对比，可快速迭代优化方向的可行性，若有效果，后续再对该特征细化分析及优化。

**对比方式：**各生命周期优化策略全部采用ABtest对比方式， 新客和流失会员可结合回头率，分析数据时对比方式不限。

## 1.3券实现流程

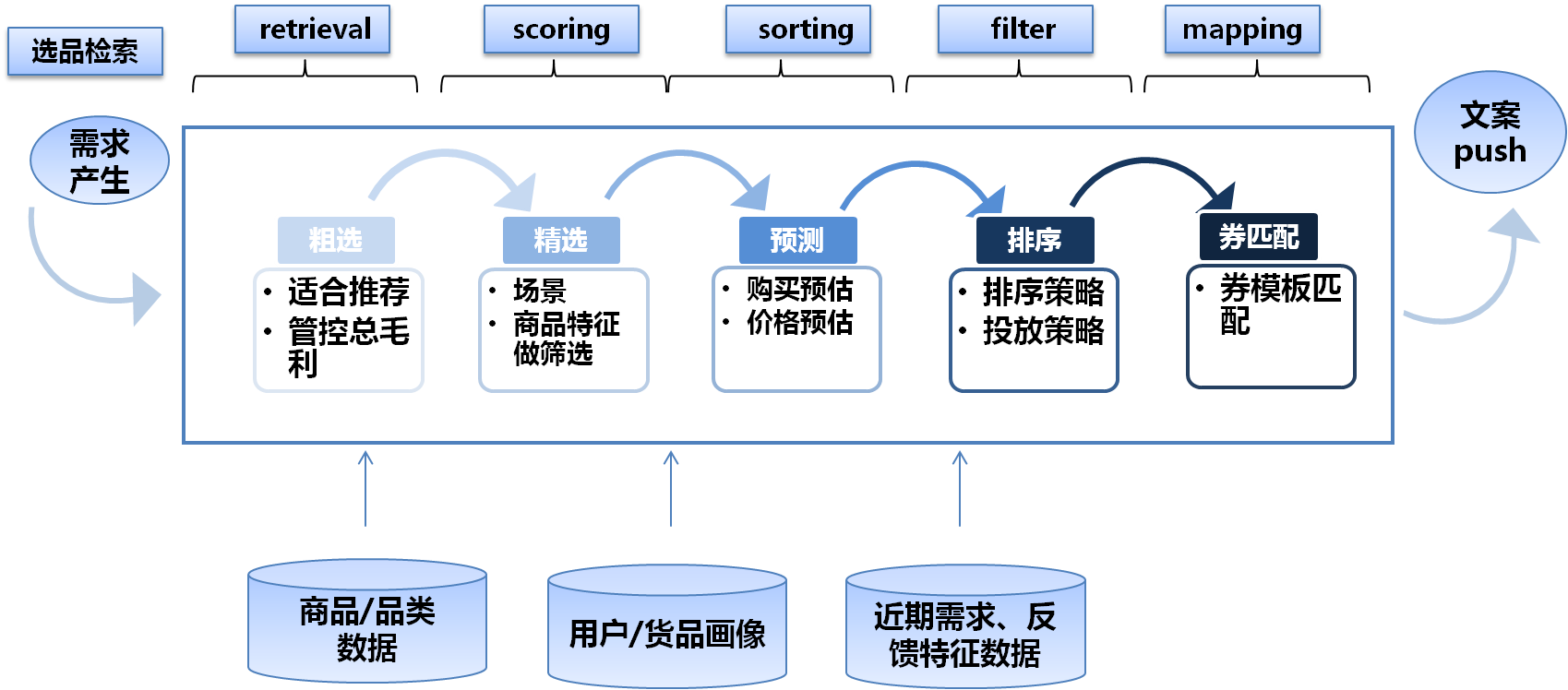
从需求的产生到实现及后续的迭代优化，流程及细分事项如下：



注： 红星表示核心项，需重点关注。

## 1.4业务逻辑开发框架

业务逻辑开发框架共分为5个模块， 分别是retrieval（选品）,scroing（精选）, sorting（精排）, filtering（规则），mapping（券匹配），如下图所示。



* retrievel 选品，根据场景需要根据商品/品类的毛利、销售等数据做选品；
* scroing：精选，基于商品特征做进一步筛选，比如毛利低的、无库存的，不符合参与活动的商品做剔除；
* sorting：精排，基于会员的推荐是否接受行为、购买行为、消费习惯等给于一系列商品的排序；
* filtering：规则，基于会员画像做筛选，剔除会员不适合的商品，比如短时间内购买过的，多次推荐没有购买的等商品剔除；
* mapping：券匹配，匹配券模板；

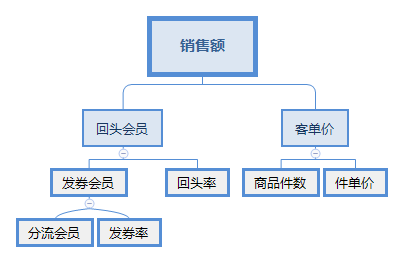
注: 上述模板并不一定全都需要，针对特定的应用，可省略其中某些模板。

## 1.5评估方式及优化目标

所有券方法的优化均采用ABtest分流方式进行评估

* 实验分流：
* ABT用MurmurHash算法， 对member\_id做稳定分流
* ABT以时间粒度，对实时场景，进行每天凌晨洗牌（保证流量的可对性）
* 每个分流人群结构相同，高等级比例相同
* 实验效果：

离线（dt+1）效果统计，准确评估策略好坏，评估指标采用单位发券会员金额和单位分流会员价值来评估，指标分解及计算公式如下所示：



* 单位发券会员金额 =销售额 /发券会员= 转化率（回头人数/发券人数） × 客单价(购买单价)（暂不考虑）
* 单位分流会员价值 =销售额 /分流会员= SCTR × SCVR × Price

即SCTR（ Send -Coupon-Through-Rate） = P(发券会员/ 分流分员)，表示会员是否满足触发券条件？

即SCVR（ Send -Coupon-Value-Rate ） = P(回头会员/发券会员) ，表示商品需要吗？价格实惠吗？

考虑目前分流会员的基数较大，而发券回头的基数较小，会导致效果的波动较大，因此对分流会员做了‘分流会员缩小版’处理，则SCTR则不考虑，处理后的分流会员都是满足发券触发条件的会员。

注： 前期为以转化率为主。

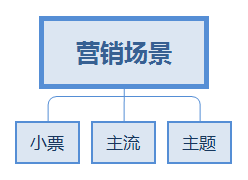
## 1.6 工具

1. HANA、mysql，sql 及数据模型的设计。
2. DS、kettle等工具的调度以etl处理
3. linux基本的命令，编写调度sh，传输数据等。
4. 常用的数据挖掘算法、推荐方法论以及适用场景；
5. 常用的分析方法及分析流程；
6. python、R 、SPSS
7. HADOOP、hive、spark、scala

# 2营销场景渠道分类

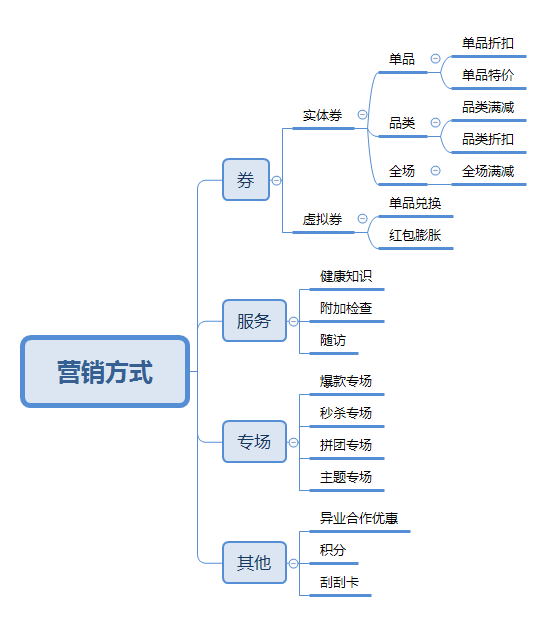
## 2.1 营销场景

从目前来看，营销场景大致分为三类，分别是小票、主流和主题。 小票场景是指通过线下门店的小票方式来进行营销；主流场景是指采用离线的方式来运营现有益丰会员；主题场景是指业务或运营人员确定某一主题临时性或间断性进行营销；



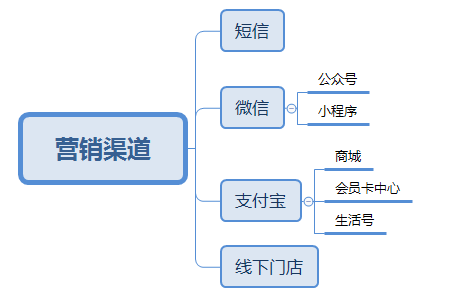
## 2.2 营销方式（承载形式）

针对不同的营销场景，会有不同的营销方式，有些营销方式是可以共享到不同的营销场景中，目前主要的营销方式是‘券’，具体分类如下图所示。 除些之外，还可以探索其他的营销方式，如积分、异业优惠、服务、专场（爆款、秒杀）、刮刮卡、虚拟券、红包膨胀等方式。



## 2.3 营销渠道

目前的渠道分为四类，分别是短信、微信、支付宝和线下门店，从会员的主动性来分，可以分为主动接收和被动接收，具体分类如下图所示

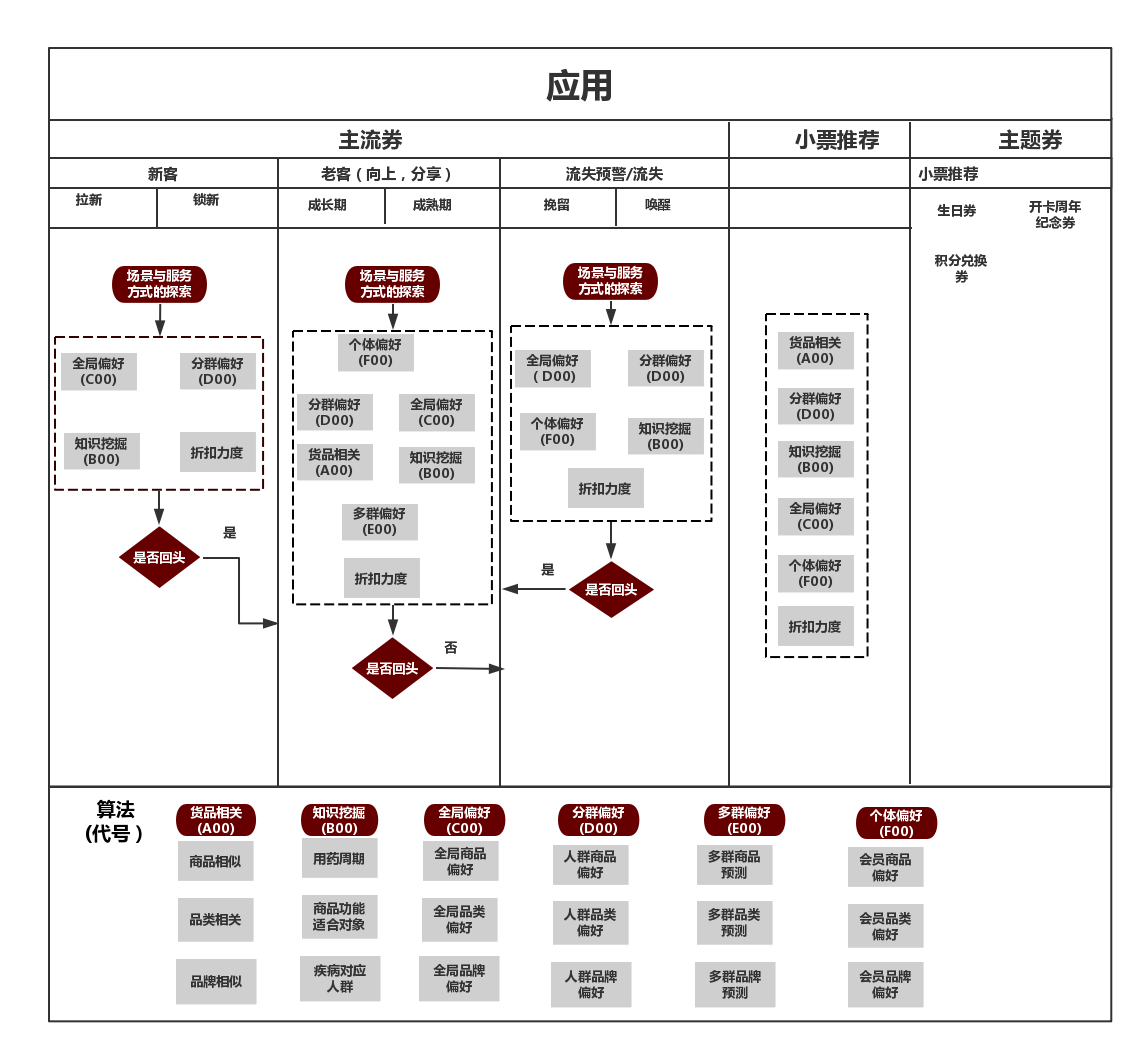


## 2.4 现有场景渠道

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **营销场景** | **营销方式** | **营销分类** | **营销渠道** | **会员主动/被动接收** |
| 主流 | 券 单品、品类、全场 | 千人千券 | 短信 | 被动接收 |
| 小票 | 券 单品、品类 | 千人千券 | 线下门店 | 被动接收 |
| 主题 | 生日券 全场 | 事件营销 | 短信 | 被动接收 |
| 主题 | 单品活动券 全场券 | 指定人群营销 | 短信 | 被动接收 |
| 主题 | 领券 全场券 | 主题活动营销 | 微信-小程序 | 主动接收 |
| 主题 | 新会员注册专享券 全场券 | 事件营销 | 微信 | 被动接收 |
| 主题 | 老会员绑定送券 | 事件营销 | 微信 | 被动接收 |
| 主题 | 门店优惠券 | 主题活动营销 | 微信-扫码 | 主动接收 |

# 2方法论

目前的应用主要分四种类型,分别主流券（生命周期的新客/拉新，老客，唤醒）、小票推荐（线下小票券）和主题券。从算法的思路来讲，分为货品相关、知识挖掘、全局偏好、人群偏好、多群偏好（预测模型）以及个体偏好。基本框架如下图所示：



* 货品相关：对商品间的潜在关系进行研究，包括商品相关性、品类相关性、疾病谱相关性以及品牌相关性等；基本思路从小粒度的相关性出发，可归到任意大粒度的相关性，可以避免大粒度相关性导致的不一致性问题。
* 知识挖掘：对医药知识的研究，包括用药疗程，商品的功能、疾病的特点等。
* 全局偏好：对公司、门店热门商品的研究，包括商品的畅销、平销、滞销状态，品类的热度以及品牌的热度等。
* 分群偏好：根据会员单维标签，构造多维人群，分析各个人群的特性以及偏好。
* 多群偏好：一是由于分群偏好由于受人工影响，受分群数量的限制，二是会员所属人群复杂，因此需要借助算法模型，通过会员画像、货品画像，构建会员的多维预测模型，实现交叉多群偏好预测；
* 个体偏好：针对会员自身的特点，分析其偏好及兴趣行为。后期也是可以通过模型融合多群偏好实现长期、短期、实时行为建模，实现千人千券。

## 2.1券模板定义与优化

### 2.11 选品

选品包装在券系统中很重要，直接关系到可以推荐的商品范围，而且医药行业对于商品推荐比较谨慎，需要有一个好的选品池，这个池子越大，可推荐的方式和商品就越多，成功转化的可能性越高。选品分为粗选和精选，粗选是整个精准营销的合适商品，即什么样的商品或品类适合推荐，例如从毛利率、会员感知以及易推广角度考虑；精选则考虑场景特点，进一步筛选。

1）选品的维度

从会员偏好、需求以及货品本身出发，选择适合的‘包装’形式根据会员的偏好、需求推送给会员，主要从以下几个方面着手，从粒度上可以分为1）单品；2）单品集合；3）通用名；4）通用名集合；5）品类 ；6）品类集合



2）选品的指标规则

满足什么样的条件的商品可以放入粗选池？ 单品、品类、通用名、疾病等 ，不同商品粒度所对应的条件不同。

* 单品：标签包括销售、转化率、毛利率、营销等级、价格，采取的选品方法有：

1）借助融合公式将各个因素融合得到权重值，通过权重值的高低每次选取合适的商品加入；2）通过不同标签的业务组合取值选取合适的商品加入；在无特殊情况下，建议采用第1）种方法，若受主题本身影响，则选用第2）种方法

* 通用名/品类：推出的通用名或者品类要确保盈利，，所以首先要保证总体是盈利的，同时确保店员更愿意配合推广，则采用以下方法：

a.计算通用名或者品类的总体毛利额是大于0的

b.计算通用名或者品类的总体毛利率是大于20%的

* 总体是盈利的情况下要确保店员对推出的通用名或者品类是愿意推广的：

|  |  |  |  |
| --- | --- | --- | --- |
| 分类 | 购买频次 | 毛利率（营销等级） | 方便推广程度 |
| 1 | 高 | 高 | 高 |
| 2 | 低 | 高 | 高 |
| 3 | 高 | 低或负 | 低 |
| 4 | 低 | 低或负 | 低 |

购买频次和毛利率的高低判断公式和阈值：

公式中K表示毛利率高的阈值，M表示购买频次高的阈值。

暂定K=50%，M=50%

由此选出的是通用名或者品类里面大部分被购买的商品是高毛利率的商品，保证选出的是方便推广程度高的通用名或品类

* 对于置信度分布进行分析，选定合适的阈值过滤（参考值0.016）
* 对于药师先验知识不符合用药规范的过滤
* 对于通用名或者品类不方便会员理解的进行包装或者过滤

### 2.12折扣力度

1）折扣力度分析

分为通用折扣力度的分析和个性化的折扣力度分析。通用折扣力度需要考虑商品本身毛利率以及历史的活动影响；个性化折扣力度，需要分析影响会员价格敏感的因素，例如购买的次数及购买的数量以及商品的价格及客单情况。

2）折扣力度制定

折扣力度的优化包括门槛和力度优化，为了满足这两个方面，同时考虑到券模板上线的复杂度，拟采用门槛和力度分段设定的方式，门槛的设定更多的倾向会员的行为差异性，所以建议通过对某个细分维度对应的客单价百分比分布进行优化，可以按5%、10%或更细粒度设定。而力度的设定一是考虑会员的行为，另外更多考虑业务上能承受的下限，更多的是验证不同力度所带来的效果变化，因此建议按中高低等级进行优化，初期由业务人员定义，后期可根据数据分析得到。

### 2.13生成券模板

参考 ‘新增模板说明’文档

拉新

分为线下和线上，数据的发力点主要在线上，通过异业合作，结合目前已有会员的特点，匹配最佳的合作方，采用适合的方式进行拉新。

## 2.2新客

新客券（A类会员）的评估指标着重关注回头率，引导会员产生第一次购买。

### 2.21流程（步骤）

新客券的实现流程如第1.3节所示。

### 2.22方法

新客/ 拉新由于会员的个体行为较少，个体的挖掘工作较少，工作重心在场景与渠道的探索、群体兴趣的挖掘、商品本身标签的挖掘以及折扣吸引力的设计。

* **场景与渠道的探索**

挖掘可利用的场景，尽量获取曝光机会和流量。除了券的方式之外，还可以探索其他适合该场景的营销方式，例如异业合作的优惠形式、买赠形式、专场形式。

* **排序方法（推荐商品）**
* **全局兴趣**

公司热门、畅销商品\品类等。

* **群体兴趣**

a) 从会员基本信息（标签）出发，寻找有差异的特征进行分群；借助具有同样特征的有行为的人群，计算其人群热度商品，从而捕获会员可能感兴趣的商品或品类，解决会员冷启动问题。例如年龄性别、地域等

b)根据相似人群分析会员行为及特点（消费客单价类型，消费毛利率类型，支付方式，惯性分布，时间间距等）。

c) 目标阶段的特点，例如第1次购买会员的特点，第2次购买会员的特点等。

* **知识挖掘：**从商品的功能、商品适用人群等出发，进行商品的定向人群投放。
* **折扣力度（差异化运营）**
* **全局最优力度**

探索全局特征，看其客单分布，通过调整不同百分位的客单定义门槛与力度来优化，优化标准见2.12节‘折扣力度制定’

* **细分群体最优力度**

通过分析不同人群间的特性及质量、给不同的促销方式或优惠力度，例如注册渠道（实体卡更难回头）、注册时长（新客无消费）、最后一次消费间隔时长（新客1次消费）、人群特性（基础特征和行为特征）等。若考虑运营成本，通过细分群体的方式可以对低质量会员不运营。

* **券类型**：尽量不发整单券，首要考虑单品及品类券；

新客无消费/新客有消费在策略逻辑上会有一定的差别，基本方法思路可复用。

### 2.23现状

目前全部采用全场券，全部会员统一折扣，只实现了全局最优力度对比，没有进行分群体实验。之前有分析过年龄、性别、分公司的客单情况，但未有实验。

### 2.24预期成果

1）新客无消费历史发券数、回头率、用券率



全量会员回头率在0.5~0.6左右， 单位发券金额0.4~0.5左右，近2月效果有明显下降，是否受策略影响需要跟进。

2）新客有消费

### 2.25注意事项

负责人不仅要关注自身工作，需跟进整个流程闭环，了解场景的会员特点、场景特点；具体的要求见1.21节的表格中的‘要求’。

## 2.3老客

老客的目的在稳定会员的基础上提升回头率、客单价、贡献力、新品的覆盖以及对品牌产生的影响力。老客按稳定性分为成长期、成熟期和衰退期，各自对应不用的策略。

**1）流程（步骤）**

券的实现流程如第1.3节所示。

**2）方法**

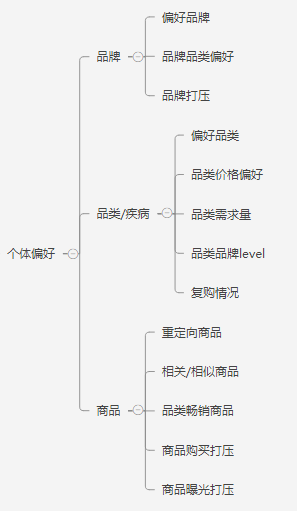
* **排序方法（推荐商品）**
* **全局兴趣：**畅销商品等。
* **群体兴趣**

a) 从会员基本信息（标签）出发，寻找有差异的特征进行分群；借助具有同样特征的有行为的人群，计算其人群热度商品，从而捕获会员可能感兴趣的商品或品类，解决会员冷启动问题。例如年龄性别、地域、消费等级、家庭情况、疾病谱等

b)计算相似人群，例如通过商品偏好、品类偏好聚类相似人群。如LDA，userCF等

* **个体偏好**

个体偏好分为品牌、品类/疾病和商品三大类，包括正向行为与反向行为，示例如图所示，后续不断完善。



优先计算会员在慢病商品、慢病品类及慢病品牌上的依赖度。

* **知识挖掘（优先3）**

从商品的功能、商品适用人群等出发，结合用药周期进行商品定向投放；

* **货品相关（优先4）**

对会员的偏好、兴趣进行扩展，用于新品的推荐。

a) 相似商品：

* **推送时间（差异化运营）**
* **折扣力度（差异化运营）**

单品券根据购买数量给予折扣，品类券根据会员客单定义满减力度。

* **券类型**：首要考虑单品券，其次品类券；

3）预期成果

## 2.5流失挽留

流失预警会员，评估指标着重关注转化率，强挽留。

**1）流程（步骤）**

券的实现流程如第1.3节所示。

**2）方法**

* **排序方法（推荐商品）**
* **全局兴趣（优先3）**

公司/门店热门、畅销商品等。

* **群体兴趣（优先2）**

**a**) 从会员基本信息（标签）出发，寻找有差异的特征进行分群；借助具有同样特征的有行为的人群，计算其人群热度商品，从而捕获会员可能感兴趣的商品或品类，解决会员冷启动问题。例如年龄性别、地域等

b)根据相似人群分析会员行为及特点（消费客单价类型，消费毛利率类型，支付方式，惯性分布，时间间距等）。

* **个体偏好（优先1）**

见第4节‘老客’。

* **知识挖掘：**从商品的功能、商品适用人群等出发，进行商品的定向人群投放；
* **折扣力度（差异化运营）**

结合历史毛利贡献

* **券类型**：针对强唤醒会员，可采用整单券，针对流失预警会员，尽量不发整单券，首要考虑单品及品类券；

3）预期成果

## 2.6流失唤醒

流失会员，评估指标着重关注转化率，强唤醒。

### 2.61 流程（步骤）

券的实现流程如第1.3节所示。

### 2.62方法

* **排序方法（推荐商品）**
* **全局兴趣（优先2）**

公司/门店热门、畅销商品等。

* **群体兴趣（优先1）**

**a**) 从会员基本信息（标签）出发，寻找有差异的特征进行分群；借助具有同样特征的有行为的人群，计算其人群热度商品，从而捕获会员可能感兴趣的商品或品类，解决会员冷启动问题。例如年龄性别、地域等

b)根据相似人群分析会员行为及特点（消费客单价类型，消费毛利率类型，支付方式，惯性分布，时间间距等）。

* **个体偏好（优先4）**

个体偏好较弱，针对强行为可进行干预

* **知识挖掘（优先3）：**从商品的功能、商品适用人群等出发，进行商品的定向人群投放；
* **折扣力度（差异化运营）**
* **全局最优力度**

探索全局特征，看其客单分布，通过调整不同百分位的客单定义门槛与力度来优化，优化标准见2.12节‘折扣力度制定’

* **细分群体最优力度**

通过分析不同人群间的特性及质量、给不同的促销方式或优惠力度，例如流失时长、流失前会员类型（新客、老客）、人群特性（基础特征和行为特征）等。若考虑运营成本，通过细分群体的方式可以对低质量会员不运营。

* **个体最优力度**

针对行为较强的会员，采用历史有差异度的特征针对性给予不同力度，如历史毛利贡献、客单价、购买次数； 若会员行为不强，则不建议采用该方法。

* **券类型**：针对强唤醒会员，可采用整单券，针对流失预警会员，尽量不发整单券，首要考虑单品及品类券；

### 2.63现状

目前两组流量采用分公司细分群体，将7个分公司分为两个折扣力度实现（细分群体最优力度）。另外两组流量在考虑分公司细分的基础上，对历史对慢病品类有强行为的会员推送统一品类券（部分采用个体偏好）。

### 2.64预期成果

数据看会员基数、前几月发券数、回头率、客单价等情况

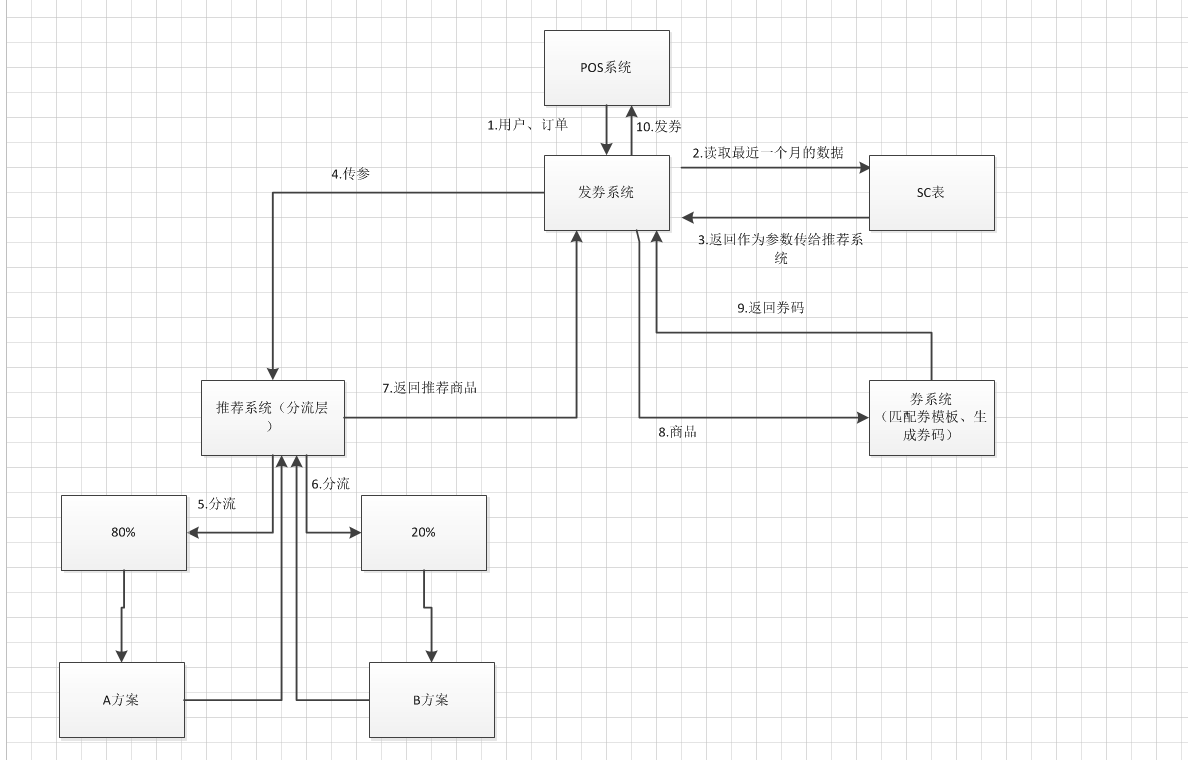
### 2.65注意事项

负责人不仅要关注自身工作，需跟进整个流程闭环，了解场景的会员特点、场景特点；具体的要求见1.21节的表格中的‘要求’

## 2.7小票推荐（实时）

### 2.71 流程（步骤）

券的实现流程



### 2.72方法

* **排序方法（推荐商品）**
* **全局兴趣**

品牌因素和热度商品都是属于商品层的优化，因为品牌商品和热度商品能尽可能的保证用户用券率，可以用于提高转化率的一种途径。

目前因为基础数据中没有品牌维度，所以品牌目前做起来比较困难。热度商品可以根据销量做一个基础数据进行商品优化，后续优化：

1）主要以热度商品为主，可以先计算出来优化商品

2）品牌维度想办法维护起来

* **群体兴趣**

a)见第4节 ‘老客’。

b)相似会员

根据用户购买商品寻找相似用户，基于相似用户推荐商品，之前已经做过相似用户计算，因为数据量比较大，而且是稀疏矩阵，所以计算方式并不完善，暂时没有用起来。后续优化：

1）可以在算法层做优化；2）算法的训练样例上可以先做一些预处理

* **个体偏好**

见第4节‘老客’。

* **知识挖掘**

从商品的功能、商品适用人群等出发，结合用药周期进行商品定向投放。

* **货品相关（优先1）**

**a) 根据当前订单推荐相关商品（相关商品计算）**

类似于关联规则算法计算得到的商品组合，只是算法不同，用余弦定理itemCF算法计算。计算所有用户订单中用户同时购买的商品，找到商品组合。根据itemCF计算商品与商品间的相关性，从而向上汇总到品类、品牌等大粒度的相关性，相关数据如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **商品** | **通用名** | **品类** | **疾病** |
| **商品** | 1）商品->商品 | 2）商品->通用名 | 3）商品->品类 | 4）商品->疾病 |
| **通用名** | 5）通用名->商品 | 6）通用名->通用名 | 7）通用名->品类 | 8）通用名->疾病 |
| **品类** | 9）品类-> 商品 | 10）品类-> 通用名 | 11）品类->品类 | 12）品类- >疾病 |
| **疾病** | 13）疾病->商品 | 14）疾病->通用名 | 15）疾病->品类 | 16）疾病->疾病 |

上述的1、2、3、6目前已计算得到的数据；1、3已经在使用。

**b) 根据当前订单推荐相关商品（旧疾病分类-人工给予用药方案）**

旧的疾病分类是之前做小票促销时，药师根据益丰的商品分类进行的对应疾病人工分类，整理出购买了某个商品应该所属的疾病，然后人工给出疾病用药方案。目前已做为商品组合的填充逻辑已上线，但疾病分类对应的商品目前没有重新整理，有优化空间。新的疾病分类目前正在整理中。

* **折扣力度**

单品券根据购买数量给予折扣，例用户经常购买的商品件数是N件，则发放N+1件可使用的券，或者用户经常购买X商品可以推出买Y送X或者买X+Y打折等形式。品类券根据会员客单定义满减力度。折扣力度优化可参考老客‘折扣力度’优化。

* **券类型**：

小票目前有4种形式：

1）单品折扣

2）单品特价

3）品类折扣（因技术原因未上线）

4）品类特价

短期内准备尝试通用名满减，以4级品类的形式实现，后续可完善的有

1）通用名折扣

2）自定义品类折扣、满减

3）自定义商品折扣、满减

### 2.73现状

目前只应用了货品相关的逻辑，会员个体偏好数据未引入实验。各数据应用情况如下：

**1、商品组合**

最主要的策略，基于用户购买的订单商品与相似商品组合及衍生组合，实时计算可推荐的商品。目前状况：已经使用过的有a)商品--> 商品 、b)商品-->4级品类两种

**a)商品--> 商品**

目前状况：（已实现）

1） 650组全公司级的折扣商品组合（毛利率>50%的）

2） 72组全公司级的特价商品组合（毛利率>30% and <50%的）

3） 650组全公司级的特价商品组合（毛利率>50%的）

可扩展的：

目前总池子还剩下5000个商品组合（毛利率>20%）

接下来准备扩展的：

全公司级的特价商品组合（毛利率>20% and <50%的）

**b)商品--> 通用名**

目前已经得到了推通用名的组合以及通用名对应包含的4级品类表

目前通用名核销两种方案：

1）自定义商品券

优点：最精确

缺点：可能会把POS撑爆，而且目前没有现成的方式上券模板，只能手工上

2）四级品类券（目前准备采用这种）

优点：实现方便

缺点：超出范围

**c)商品--> 品类**

目前已经上了7个4级品类商品，做了高门槛及低门槛的满减品类券，但是因为转换率都只有0.1%，所以都下线了。之前上线分析来看高门槛的可以提高客单，而低门槛的会拉低客单，但是转化率提升并不高。所以目前来看效果不好

**d) 待完善类型**

商品推通用名近期可以尝试，其他的选品较难。

a)通用名-->通用名

b)通用名-->品类

c)品类--> 品类

**2小票促销（后续填充规则）**

如果商品组合找不到可推荐的商品，基于购买订单商品推出用户所属疾病分类，随机推荐对应保健品商品。

目前状况：

之前检查发现疾病分类以及疾病推荐保健品都存在问题：

1. 品分类也未更新同步 -- 已同步至最新的商品分类。
2. 商品对应疾病表数据有误 -- 已经将慢病的全部调整成正确的分类，
3. 商并且之前所有购买商品都会有对应的疾病分类 -- 现已将非慢病的疾病分类全部下线；
4. 部分慢病转换率较低，不到0.1% -- 已将这部分慢病也下线

待完善：

1. 因为保健品自然转化率较低，只有0.1%，所以疾病分类对应的商品没有进行进一步的调整，这里可以进一步优化。
2. 现在对应疾病推荐的保健品是随机的，这里可以做策略优化
3. 现在对应保健品爆品已经可以实现，但还没有具体的爆品策略，这里可以完善
4. 保健品因为单价高、毛利高，所以如果能提高保健品的转换对小票整体销售、客单、毛利都是有提升作用的，可以想办法提高保健品的转化率。

**3重定向**

前面都是基于当前订单做的策略，属于实时场景，主要目的是让用户在购买完成之后立刻用券多买一些其他商品。重定向属于基于用户历史信息的，根据用户历史购买的单品和品类可以推单品和品类券。

目前状况：

因为用户偏好还没有接入推荐系统，所以目前暂未实现

待完善：

这里可以拓展小票的场景，但是目前实时场景用券占总体小票用券的95%以上，7天内回头用券数非常小，所以一直没有把这个点实现，后续可以尝试这种策略

**4疾病分类（新）**

基于药师新做的疾病分类，与疾病分类用药，做商品、品类重定向

目前状况：

药师还在做数据

待完善：

待使用

**5人群分类**

找出相似人群，用人群的热卖商品进行推荐

目前状况：

暂未开始

待完善：

**6地域划分**

基于地域推荐。结合具体区域特点，按分公司特色推荐商品

目前状况：

目前暂未考虑地区因素，相似商品也是基于全公司数据计算的

待完善：

**7相似用户**

可以基于相似用户推荐商品

目前状况：

目前暂未计算

待完善：

**8品牌**

优化商品组合，结合品牌或者热卖的商品，将转化率低的商品用品牌或者热卖的商品替换。

目前状况：

因为用户购买订单及推荐商品的效果没有统计出来，难以分析

待优化：

先将效果报表做出来之后可以进行优化

9**选品**

目前只做了单品、品类、品类集合（主流券），因为小票场景受POS机系统影响，无法自由推荐2、3、4选品结果，所以选品进度缓慢，目前主要的推荐是单品，以及部分4级品类做了选品包装，但可选性并不大，影响较小。

### 2.74 预期成果

数据看会员基数、前几月发券数、用券率、用券客单等情况

### 2.7.5注意事项

负责人不仅要关注自身工作，需跟进整个流程闭环，了解场景的会员特点、场景特点；具体的要求见1.21节的表格中的‘要求’。

## 2.8 Filtering后置规则

根据会员的特性提前或打压某些商品或品类，对sorting的结果进行调整，主要包括以下几点：

1. 推送频率过滤（scoring计算）
2. 同一个商品推送间隔过滤
3. 个人所在门店库存过滤
4. 商品/品类购买打压
5. 曝光商品/品类打压
6. 政策打压，如处方药打压等
7. 结合个人或群体特征将商品/品类前置或后置
8. 其他业务规则

## 2.9券匹配

根据会员不同特征，采用不同的策略匹配券模板，会员的特征分为两类，一类是决定如何调整商品/品类偏好排序，另一类决定折扣力度。

# 3各类数据逻辑

## 3.1用户相似商品

通过计算得到用户的上次购买单品或者历史最偏好单品的相似商品。

参考交叉券相似商品逻辑，使用关联规则算法计算得到商品组合，用余弦定理itemCF算法计算。计算所有用户订单中用户同时购买的商品，找到商品组合。利用用户实时行为或历史偏好商品计算出的商品组合进行用户相似商品推荐。

由商品组合可以得到以下数据（从而向上汇总到品类、品牌等大粒度的相关性）：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **商品** | **通用名** | **品类** | **疾病** |
| **商品** | 1）商品->商品 | 2）商品->通用名 | 3）商品->品类 | 4）商品->疾病 |
| **通用名** | 5）通用名->商品 | 6）通用名->通用名 | 7）通用名->品类 | 8）通用名->疾病 |
| **品类** | 9）品类-> 商品 | 10）品类-> 通用名 | 11）品类->品类 | 12）品类- >疾病 |
| **疾病** | 13）疾病->商品 | 14）疾病->通用名 | 15）疾病->品类 | 16）疾病->疾病 |

目前只应用了货品相关的逻辑，会员个体偏好数据未引入实验。各数据应用情况如下：

**1、商品组合**

最主要的策略，基于用户购买的订单商品与相似商品组合及衍生组合，实时计算可推荐的商品。目前状况：已经使用过的有a)商品--> 商品 、b)商品-->4级品类两种

**a)商品--> 商品**

目前状况：（已实现）

1） 650组全公司级的折扣商品组合（毛利率>50%的）

2） 72组全公司级的特价商品组合（毛利率>30% and <50%的）

3） 650组全公司级的特价商品组合（毛利率>50%的）

可扩展的：

目前总池子还剩下5000个商品组合（毛利率>20%）

接下来准备扩展的：

全公司级的特价商品组合（毛利率>20% and <50%的）

**b)商品--> 通用名**

目前已经得到了推通用名的组合以及通用名对应包含的4级品类表

目前通用名核销两种方案：

1）自定义商品券

优点：最精确

缺点：可能会把POS撑爆，而且目前没有现成的方式上券模板，只能手工上

2）四级品类券（目前准备采用这种）

优点：实现方便

缺点：超出范围

**c)商品--> 品类**

目前已经上了7个4级品类商品，做了高门槛及低门槛的满减品类券，但是因为转换率都只有0.1%，所以都下线了。之前上线分析来看高门槛的可以提高客单，而低门槛的会拉低客单，但是转化率提升并不高。所以目前来看效果不好

**d) 待完善类型**

商品推通用名近期可以尝试，其他的选品较难。

a)通用名-->通用名

b)通用名-->品类

c)品类--> 品类

## 3.2 用户偏好商品

老客目前选品逻辑优先级最高的是用户重定向偏好商品。在大体上推荐策略优先级为：单品>品类>全场。当前用户偏好商品的计算决定了用户可推荐单品的池子有多大及商品在用户偏好单品中的优先级。而用户偏好品类的计算决定了用户可推品类的池子有多大及品类在用户偏好品类中的优先级。

目前状况：

已实现部分：

* 用户偏好商品计算完成

1. 用户偏好商品在一年内有购买并且用户的偏好随着用户对该商品的购买时间每3个月衰减一定分数
2. 把用户对商品的购买次数乘以时间对应的衰减系数再全部相加，得到用户对该商品的分数
3. 取得分大于0.8的用户单品

* 用户偏好品类计算完成

1. 通过用户偏好商品关联表得到对应品类的得分
2. 把用户品类得分汇总得到用户对该品类总得分
3. 取品类总得分第一的用户品类

* 用户匹配单品逻辑完成

1. 如果上次用户购买2件，原则上推2+1件
2. 如果券模板库只有2件或者1件，则推2件或者1件
3. 如果匹配到多个券模板，按照件数取最低的推荐

待优化部分：

* 用户偏好商品及偏好品类口径的调优
* 券模板匹配调优

待开发部分：

* 用户偏好商品及偏好品类算法评估机制

## 3.3用户偏好时间

用户偏好时间关系到给用户推送的具体时机，对于有很多消费行为的老客，从偏好月、周、日、天气、购药周期、节假日、会员日、促销日等方面进行精准发放。

目前状况：

已实现部分：

* 星期偏好计算完成

1. 星期偏好的条件为总消费次数大于4并且周中某天消费次数大于总次数40%则认为该用户在这一天有用药偏好
2. 历史数据分析得到一般发券第二天和发券第六天用券最好
3. 让用药星期偏好日子等于发券第二天或者发券第六天

* 购药周期（慢病2-6次和慢病7次）

1. 计算单品一盒的最大使用周期
2. 计算品类
3. 消费间隔=上次消费数量X该单品（品类）最大使用周期 （上次消费数量没有则默认1，单品最大使用周期没有则默认7天，品类最大使用周期没有则默认15天，品类为偏好单品的品类，周期为单品的周期）
4. 根据消费间隔和上次消费时间计算下次消费时间
5. 如果当前日期距离下次消费时间在两天内，则进入推送池子

待优化部分：

* 星期偏好的口径调优，做放开或者缩小尝试
* 购药周期增加选品并调优
* 单品消费时间间隔的计算方式调整

待开发部分：

* 引入用户偏好月时间，对用户发券在月内做均衡
* 引入用户偏好日时间，结合用户常购买日时间段发券
* 引入用户偏好天气，在适当天气发券，控制不好天气的发券量
* 引入用户偏好节假日，对于用户节假日与工作日购买相差巨大的选择在节假日推送
* 引入用户偏好会员日、促销日，在公司有促销活动时将该部分用户拉出来做策略

## 3.4特定人群提取

针对某些特定人群，需要做特定的策略，具体来说有以下几类群体：1、毛利率划分；2、年龄、性别、地域等属性群体划分；3、客单价划分；4、流失预警

目前状况：

已实现部分：

* 毛利率划分群体
* 客单价划分群体

待优化部分：

* 毛利率划分群体具体策略调优
* 客单价提升调优

待开发部分：

* 年龄、性别、地域等属性群体探索与划分
* 新的群体探索
* 流失预警人群开发

此处给出整个系统中所采用的关键技术、最新技术、体现本系统技术特色等的内容。

# 3工作分配

|  |  |  |
| --- | --- | --- |
| **工作类型** | **相关工作** | **人员安排** |
| 拉新 | 跟进整个流程 | 王婷 |
| 新客 | 跟进整个流程 |
| 老客-成长期 | 跟进整个流程 | 姚泊彰 |
| 老客-成熟期 | 跟进整个流程 |  |
| 流失挽留 | 跟进整个流程 |  |
| 流失唤醒 | 跟进整个流程 |
| 小票推荐（实时） | 1、跟进整个流程  2、保证系统稳定性 | 刘偲 |
| 主题券 |  |  |
| 券模板生成 | 1、所有场景涉及到的券生成  2、底层框架的建立与维护 | 主要人员:姚泊彰、王路; 所有券相关人员都会参与 |
| 会员画像 | 涉及会员角度的所有特征，具体见‘会员画像’说明 | 主要人员：吴海艺，所有券相关人员都会参与 |
| 货品画像 | 涉及货品角度的所有特征，具体见‘货品画像’说明 |
| 效果评估 | 负责所有券的效果评估及kpi的考核指标 | 彭卫 |
| 券优化分析和探索 | 负责优化 | 所有券相关人员都会参与 |

按工种分类型：

1）数据

2）策略开发

3）策略分析

2）效果开发

4）算法

5）系统

4、技术框架