**精准营销项目介绍**

——BI团队

# 1概述

## 1.1 核心思路

精准营销推荐在**特定的场景**下根据会员基本信息、会员行为、会员兴趣等，构建不同的会员个体/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**，实现给**合适的会员**在**合适的时机**以**合适的渠道**实现**合适商品**承载**合适的券**的推送，以此最大化平台收益。核心思路图见图1.1所示。

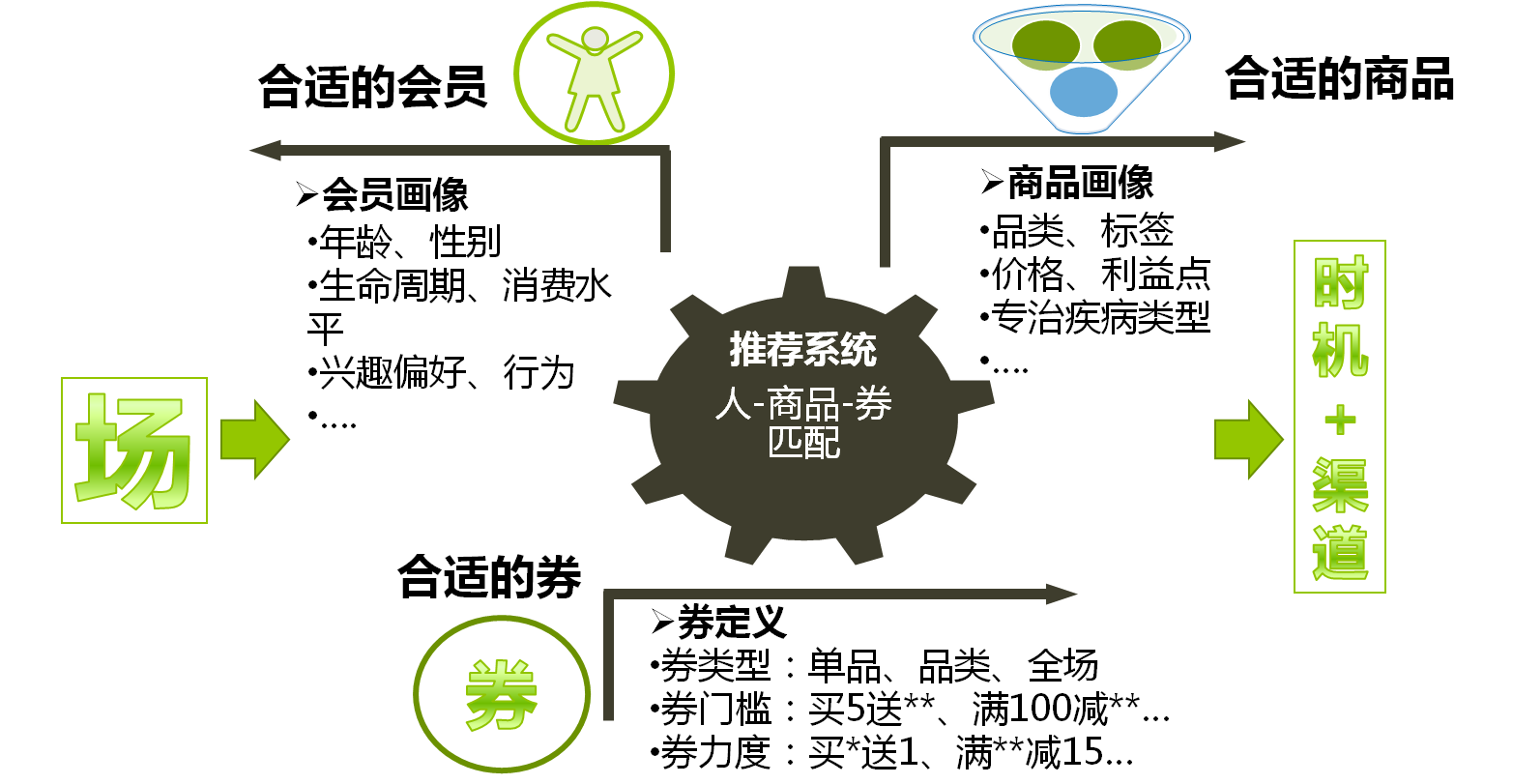


图1.1 精准营销实现核心思路图

## 1.2 业务场景

精准营销的推荐主要应用在以下五个场景，分别是主流场景、小票场景、小程序、微信公众号和支付宝。主流场景主要是通过“短信”的渠道推送；小票场景通过购物小票推送；小程序、微信公众号和支付宝由会员或粉丝主动领取。

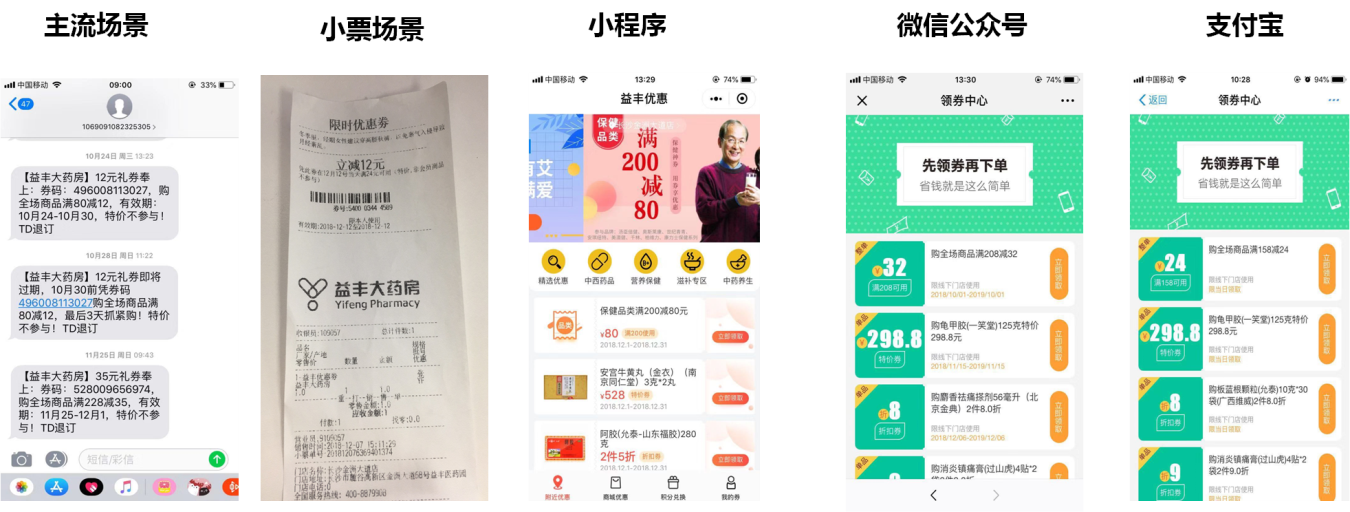


图1.2 业务场景示例

## 1.3业务实现逻辑框架

精准营销的业务需求和策略，从技术实现上，BI团队构建了一套业务实现的逻辑框架，见图1.3所示。围绕着选择“合适的会员”、“合适的商品”、“合适的券”的业务综旨，从"需求产生"到“文案推送"，经过了四个阶段，分别是"初选"、"精选"、"规则调整"和"券匹配"。在每个阶段中，分别包含了不同机制，定义了不同机制的功能和界限范畴，不同的业务需求可在不同的阶段和机制中实现。

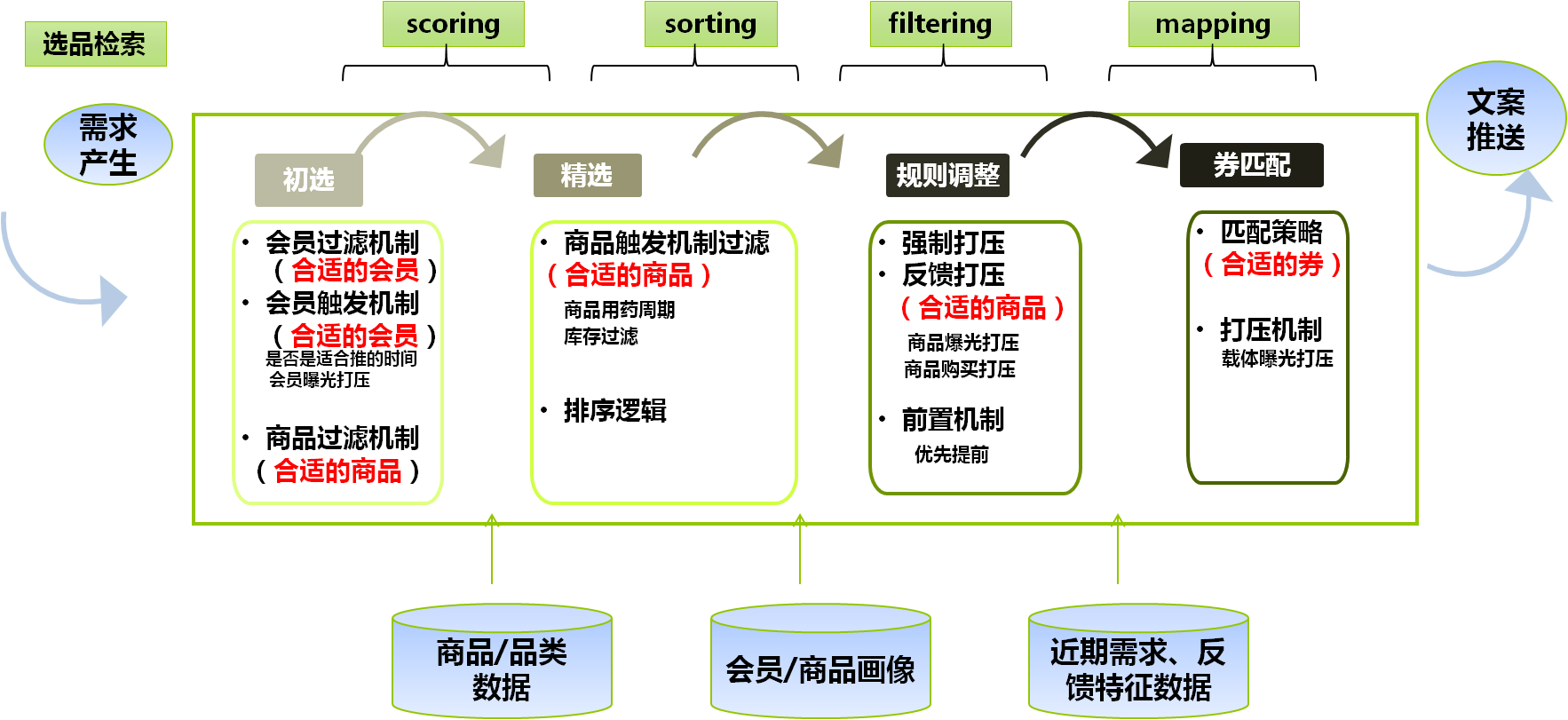


图1.3 业务实现逻辑框架

# 2、各模块设计原理

精准营销个性化推荐以会员生命周期标签为主线，按照会员的特点把会员切割成新客、成长期、成熟期、衰退期、流失期五个状态，具体口径见图2.1。通过会员生命周期动线图定位会员生命周期状态，再结合会员画像、商品画像给予会员不同的权益、服务、体验与合适的利益点，缩短新客的成长周期，延缓衰退与流失，促使会员向成熟期定向转化，达到会员质量（忠诚度）整体提升，从而实现商业价值。表2.1描述了各生命周期要达成的的目的、策略方向和渠道载体。下面从各个会员生命周期状态来阐述不同业务模块的设计原理。

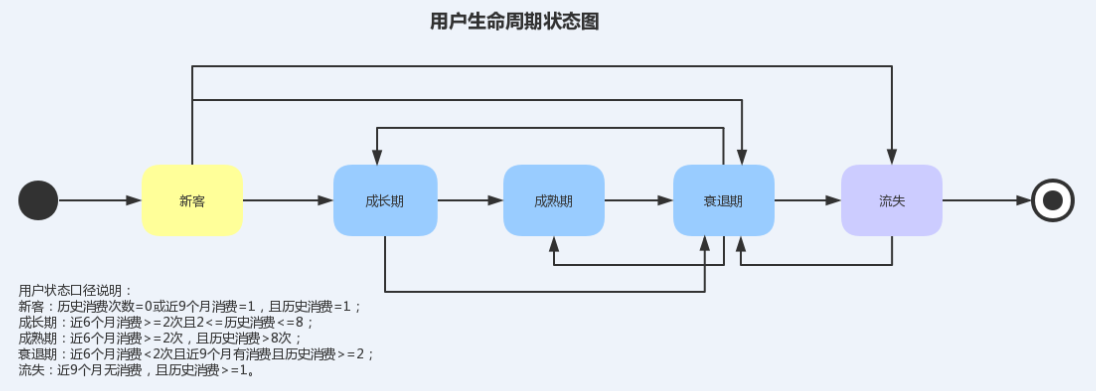


图2.1 会员生命周期状态图

表2.1 各生命周期状态目的策略表

|  |  |  |  |
| --- | --- | --- | --- |
| **生命周期** | **目的** | **策略方向** | **渠道载体** |
| **新客** | 促使会员干预下回头  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 | 短信发券干预：全场券  自领券干预：全场券；单品券  等 |
| **成长期** | 促使会员消费更多次数，逐步提升消费金额  侧重指标看**回头率、销售、客单，衰退率** | 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 | 短信发券干预：全场券；品类券；单品券  自领券干预：全场券；品类券；单品券  等 |
| **成熟期** | 影响会员消费习惯，促使会员为公司创造更大价值  侧重指标看**回头率，销售，客单，毛利率，毛利额，营销权重，衰退率** | 考虑成熟期的特性：  同‘老客-成长期’ | 短信发券干预：全场券；品类券；单品券  自领券干预：全场券；品类券；单品券  等 |
| **衰退期** | 有效挽回会员，防止会员流失  侧重指标看**回头率，销售，流失率** | 考虑衰退期的特性：  同‘老客-成长期’ | 短信发券干预：全场券；品类券；单品券  自领券干预：全场券；品类券；单品券  等 |
| **流失** | 促使会员干预下回头  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 | 短信发券干预：全场券；品类券  自领券干预：全场券；品类券  等 |

## 2.1合适的会员

合适的会员是指筛选适合推券的会员，目前主要从政策约束、硬性约束和业务约束上进行初步过滤筛选。具体操作见表2.2所示。

表2.2 各约束类型对应操作

|  |  |
| --- | --- |
| **约束类型** | **具体操作** |
| 政策约束 | 徐州、南京门店，灌云社区、泰州门店，粤海门店会员不能发券 |
| 硬性约束 | 短信不可打扰会员、手机号码异常会员、状态不正常会员不能发券 |
| 业务约束 | 近一年毛利为负会员、近三年未消费会员不发券、收购会员导入后一个月不干预 |

## 2.2合适的时机

选择了合适的会员，需要进一步为这些会员选择合适推送的时机。目前主要从“主动”和“被动”两个角度进行选择。具体操作如下：

1. **主动（会员购药时间习惯偏好）**

基于会员在益丰的历史行为，能够捕获到会员购药时间偏好，比如会员一般周几来购药？购药的时间点是上午、中午、下午还是晚上？ 会员的购药周期一般是间隔多少天？等等。从这些角度，我们可以找到推送的最佳时机。目前我们采用了会员购药星期偏好模型，具体模型原理见3.1节。

1. **被动（会员反馈行为）**

针对会员在若干次被干预后表现极差的，考虑会员是否暂时无需求或是对此不感兴趣？会进行一定的打压（惩罚）操作；而针对那些被干预后表现良好的，若会员拉到正常轨迹，会通过改变会员推送间隔时间改变推送时机。

## 2.3合适的商品

在找到合适的会员，确定合适的时机后，需要为这些会员找到合适他们的商品。目前主要从公司价值、会员立场以及商品可推性三个角度考虑，下面具体阐述。

**1、公司价值**

从公司价值来讲，需要考虑可推商品的毛利情况，特别是针对成熟期会员，毛利率是首要保证的，而对于新客和流失会员，让利空间可大些。因此，在选择可推商品池子时，初步把不利于公司价值的商品进行剔除。

**2、会员立场**

从会员角度来讲，第一步，需要考虑会员真正偏好、需要或潜在需要的东西。目前会员商品偏好模型主要从会员历史购买的商品（简称重定向）、相似人群偏好商品、相似相关商品组合等进行计算，寻找会员偏好的单品和品类，具体模型介绍见3.2节。第二步，需要找到会员推送该商品合适的时机，主要从会员的购药周期或者近期购买的情况等角度考虑。下面具体介绍如何从会员立场选择合适的单品和品类。

1）单品

* 针对新客：个体行为缺乏，难以从个体行为中得到单品偏好，目前主要从群体上获取人群热度商品（是指具有同一特性的群体，共同喜欢的商品）。
* 针对老客（成长期、成熟期、衰退期）：个体行为较丰富，且部分商品具有明显的购药周期，则首先通过“会员单品偏好模型”得到会员偏好的单品，然后结合会员上次购买的时间和数量以及单品的用药周期，决定该单品是否可推。
* 针对流失会员：个体消费信息时效性太低，自身的行为可靠性偏弱，同新客一样，从群体上获取人群热度商品。

2）品类

品类的定义不仅仅指当前益丰的品类结构，可以是由“单品”组成的任何“类”，从会员的感触上可以更加丰富。

* 针对新客：同单品一样，难以从个体得到品类偏好，由于目前受“类”的可推性限制，当前新客是没有品类偏好的。
* 针对老客（成长期、成熟期、衰退期）：目前的“类”采用的是慢病品类，因此第一步基于“会员品类偏好模型”计算出会员偏好的慢病品类，然后结合会员上次购买的时间和数量以及单品的用药周期，决定该单品是否可推。
* 针对流失会员：借助老客部分逻辑，只基于会员历史购买的行为计算会员偏好的品类。

**3、商品可推性**

从公司价值、会员角度都具备可推，从商品自身考虑，主要有两方面：

1）由于该商品具有行业特殊性，有些商品若没推好，则会引发会员的反感及投诉，针对风险品种，直接不推。例如政策类商品、隐私类商品。

2）从单品组成的部分“类”，由于大众的普识度限制，目前没有探索出可展现给会员的方式。

## 2.4合适的券

在得到合适的会员在合适的时机推送的合适的商品后，需要找到一个合适的承载方式（券）把商品呈现给会员。具体来说就是要选择合适的券类型、合适的券门槛、合适的券折扣力度，实现差异化、个性化推送，从而打动会员。目前支持的券类型分类单品券、品类券和全场券，下面具体阐述。

1. **单品券匹配**

1）针对新客：单品的券门槛和力度通过数据分析得到历史最多的会员购买的件数和享受到的力度，再由业务人员修正，采用通用折扣力度。

2）针对老客（成长期、成熟期、衰退期）：采用个性化折扣力度，券门槛和力度的定制如下：

* 券门槛：根据会员历史购买的单品数量，计算出会员可能购买的数量，在此基础上再做适当提升，目前是在原基础上做1件提升，例如平时该会员购买2件，则这次券门槛设为3件。
* 券力度：通过数据探索，计算出单品不同购买件数的让利的空间，再由业务人员修正，定义不同券门槛对应的力度。

3）针对流失会员：同理新客。

1. **品类券匹配**

1）针对新客：暂无品类券

2）针对老客（成长期、成熟期、衰退期）：采用个性化折扣力度，券门槛和力度的定制如下：

* 券门槛：根据会员历史行为计算出会员品类客单价和品类毛利率，通过对会员品类客单价和品类毛利率的探索，目前是采用会员品类毛利率的高低来控制客单价的提升系数，算法原理是会员品类客单价的提升系数会随着毛利率的增大而增大，具体模型见3.4节品类匹配模型。最后优惠的券门槛为会员的客单价与客单价提升系数相乘。例如会员平时的品类客单价为60元，品类毛利率为50%，则这次券门槛设置满100元减\*\*。
* 券力度：采用会员品类毛利率来衡量，若品类毛利低于10%，则不推荐该品类；否则，随着会员毛利率的升高，定制的力度就越大，因此按照毛利分段排序，得到每段能给到的折扣力度。

目前的品类只做了慢病品类，而偏好该品类的都是慢病会员，他们具有的特点是高忠诚度，偏低品类毛利，而公司很多的慢病药品毛利较低。因此，在会员品类毛利极低时，不发品类券；品类毛利较低时，品类券需要以低门槛低折扣保证会员回头同时损失较少毛利；随着会员品类毛利的升高，需要逐渐提高门槛并在适当的时候给予会员较大力度的奖励。

3）针对流失会员：对有强偏好的品类的会员，采取的方法与老客一致。

1. **全场券匹配**

1）针对新客：从“全局最优力度”到当前的“细分群体最优力度”进行转变，目前“细分群体最优力度”采用会员“开卡时长”+“开卡来源”特征进行群体细分，计算每个群体的购买率和客单价，初始券门槛和力度的定制如下。

* 券门槛：采用会员所属群体的客单价做为优惠券门槛。
* 券力度：由于每个群体的质量（购买率）有较大差异，因此根据不同群体的质量设置不同的力度，会员质量越差，享受的力度就越大，而对于质量稍好的，则力度就偏小。

针对新客，后续将采用了“优惠力度阶梯化”干预，即对会员进行干预却没有回来的会员，随着干预次数的增多，会采用比上次更低的优惠力度，若还不回来，则强制停止干预。

2）针对老客（成长期、成熟期、衰退期）：与老客的品类券思路一样。具体模型见3.4节全场匹配模型。

* 券门槛：算法原理是毛利率在50%以下的会员，其客单提升数系会随着毛利率的升高而降低，而针对毛利率50%以上的会员，其客单提升系数会随着毛利的升高而升高。具体算法见3.3节。
* 券力度：同老客品类券。

全场券的情况略有不同，对于会员消费毛利极低时，我们不发全场券；全场毛利较低时，说明他们对价格的敏感度较高，同时又不想公司亏损，则会匹配高门槛低折扣力度的全场券；随着会员全场毛利的升高，消费习惯变好，为了培养会员消费习惯，门槛会逐渐降低，折扣力度有所提升；随着会员全场毛利率上升到一定程度，需要逐渐提高门槛并给予会员较大力度奖励。

3）针对流失会员：原理同新客，流失会员采用的细分群体的特征为会员“流失月”+“消费次数”。券门槛与力度的定义同新客。

**4、券类型的优先级**

目前会员推送的券类型优先级为：单品券、品类券、全场券。

# 3 算法介绍

## 3.1会员购药星期偏好模型

会员购药星期偏好模型的具体做法为：

**步骤1**：分别统计每个会员所有消费中周一~周日消费次数

**步骤2**：得到每个会员周一~周日消费次数占比=当天消费次数/总消费次数

**步骤3**：定义模型参数为会员消费次数下限及消费次数占比下限

**步骤4**：通过选择合适的参数得到会员偏好星期时间

## 3.2会员单品偏好模型

会员单品偏好有多种数据来源，当前已实现的有重定向会员单品偏好模型和商品组合单品偏好模型。

**1、会员历史购买（重定向）单品偏好模型**

**步骤1**：分别统计每个会员订单，以三个月为衰减单位进行衰减处理，即三个月内为3，三到六个月为6，以此类推，定义为变量x

**步骤2**：根据公式算出一年以内时间衰减权重weight，根据公式算出一年以上时间衰减权重weight。

**步骤3**：会员购买商品数量\*weight，再统计求和即为会员对商品的偏好得分。

**2、相似相关商品组合单品偏好模型**

**步骤1**：设置最小支持度minsupport：依据商品组合出现概率所定的阈值，高于该值，商品组合将被保存在候选集Ck中，否则将被剔除出候选集。此处依据假设：频繁项的所有子集均为频繁项。

minsupport=a/len(D) (1-1)

a为可调参数（经验值、静态参数），len(D)为每个会员的订单数。

**步骤**2：计算商品组合支持度supportdata：每个商品组合出现的概率，

supportdata=商品组合在数据集出现的行数/数据集总行数 (1-2)

**步骤**3：计算商品组合置信度

conf（A to B）= supportData[AB]/supportData[AB - B]

=supportData[AB]/supportData[A] (1-3)

**步骤**4：把每个商品存在的商品组合按照商品组合置信度按降序排列。

**步骤**5：得到最新会员购买单品信息，找到会员购买商品的商品组合，得到会员偏好单品。

**3、会员单品偏好融合模型**

得到不同方式计算出来的会员偏好单品后，融合的具体做法为：

**步骤**1：将所有会员偏好单品按照来源定义优先级，放在一起

**步骤**2：如果有会员具有不是一个来源的同一个商品，选择优先级最高的作为单品偏好得分

**步骤**3：最终得到的单品偏好数据即是会员单品偏好数据

## 3.3会员价值模型

会员价值模型是使用RFM模型进行会员的价值评估，其中，R(Recency)表示客户购买的时间有多远，F(Frequency)表示客户在时间内购买的次数，M (Monetary)表示客户在时间内购买的金额。在具体应用中，我们计算了以下三个指标：

* 当前会员价值

会员近一年时间展现出来的价值

* 历史会员价值

会员历史上展现出来的价值

* 忠诚度阶梯划分

RFM按照三个维度划分八个象限得到不同意义的人群

计算会员价值的公式为：



其中，M取毛利额，L为会员第一次消费与最后一次消费时间差，用来展现会员的生命长度。

忠诚度阶梯划分是RFM在三个指标上分别取中位值可以得到八个象限，如图3.1所示。

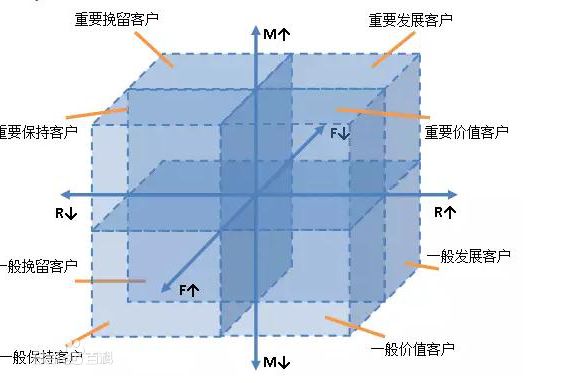


图3.1 RFM忠诚度阶梯划分

## 3.4品类匹配模型与全场匹配模型

**1、品类匹配模型**

**步骤**1：根据品类需求，确定需求为“品类毛利10%以下不发品类券，品类毛利10%到50%时发低折扣券，品类毛利为50%以上时发高折扣券，客单提升随着品类毛利的升高而升高，并且在客单提升幅度逐渐加大，最低客单提升系数要求1.4”。因此，我们选择客单提升计算模型的数学形式为抛物线，公式为y=k\*(x-t)^2+b

**步骤**2：根据品类的需求数据结点，代入公式中求解系数，得到k=1.875，t=0.1，b=1.4。

**步骤**3：得到品类的客单提升系数公式为y=1.875\*(x-0.1)^2+1.4，其中x为品类毛利率。

**步骤**4：根据生命周期确定可用折扣力度，折扣力度按从小到大排序，将品类毛利分段 分别匹配对应折扣力度。

**2、全场匹配模型**

**步骤**1：根据全场需求，确定需求为“全场毛利0%以下不发全场券，全场毛利0%到10%时发低折扣券，全场毛利10%到50%时发中等折扣券，全场毛利50%以上时发高折扣券，客单提升在50%以下随着全场毛利的升高而降低，在50%以上随着全场毛利的升高而升高，并且在客单提升幅度在两头迅速提升，在50%附近平稳过渡，最低客单提升系数为1.4”。因此，我们选择客单提升计算模型的数学形式为抛物线，公式为y=k\*(x-t)^2+b。

**步骤**2：根据全场的需求数据结点，代入公式中求解系数，得到k=7.2，t=0.5，b=1.4。

**步骤**3：得到品类的客单提升系数公式为y=7.2\*(x-0.5)^2+1.4，其中x为品类毛利率。

**步骤**4：根据生命周期确定可用折扣力度，折扣力度按从小到大排序，将品类毛利分段 分别匹配对应折扣力度。

# 4 效果评估

## 4.1 效果评估方式

精准营销个性化推荐效果的评估方式均采用A/B测试（简称ABT）方式进行评估。ABT的原理是将益丰所有会员通过算法合理地分成两部分（简称分流），一部分称为“实验组”，另一部分称为“对照组”。 “实验组”进行发券操作，“对照组”不进行发券操作，两者的差异只体现在是否有发券操作，其他运营等受影响的因素均匀，可忽略。最后通过对比“实验组”和“对照组”的评估指标的差值来计算最终的提升金额和毛利。ABT分流的具体实现见4.3节， 评估指标的定义见 4.2节。

## 4.2 效果评估指标

“实验组”和“对照组”的会员基数不同，为保证两者间具有可比性，效果评估指标采用“人均分流会员金额”和“人均分流会员毛利额” （“实验组”和“对照组”的会员基数统称为分流会员）。

从核心指标销售额进行分解，见图4.1所示：

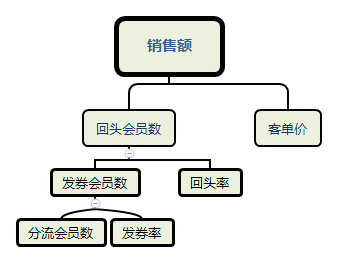


图4.1 销售额指标分解

从上图可知：

得：

即： **SCTR（ Send -Coupon-Through-Rate）** = 发券会员数/ 分流分员数

* 表示是否满足触发券条件？

**SCVR（ Send -Coupon-Value-Rate ）** = 回头会员/发券会员

* 表示商品需要吗？
* 表示价格实惠吗？

从指标的分解来看，若想让“人均分流会员金额”提升越高，则要提升发券率、回头率 和客单价。 发券率受会员基数以及“合适发券会员”的影响，提升空间有限， 而回头率和客单价的提升是整个精准营销策略方法的优化目标。

## 4.2 效果评估实现

**1、ABT分流原则**

* 分流算法：ABT用MurmurHash算法， 对会员id做稳定分流。
* 会流方式：ABT以时间粒度分流，在小票场景，每天凌晨对会员洗牌，重新分流（保证流量的可对性）；在主流场景，所有会员从第一次进入进行分流，后每月1号对会员洗牌重新分流。
* 会员影响：“实验组”和“对照组”会员结构相同，各生命周期会员人数一致，高等级会员一致。
* 实验效果：离线（当天对前一天）效果统计，准确评估策略好坏。

**2、ABT实际分流情况**

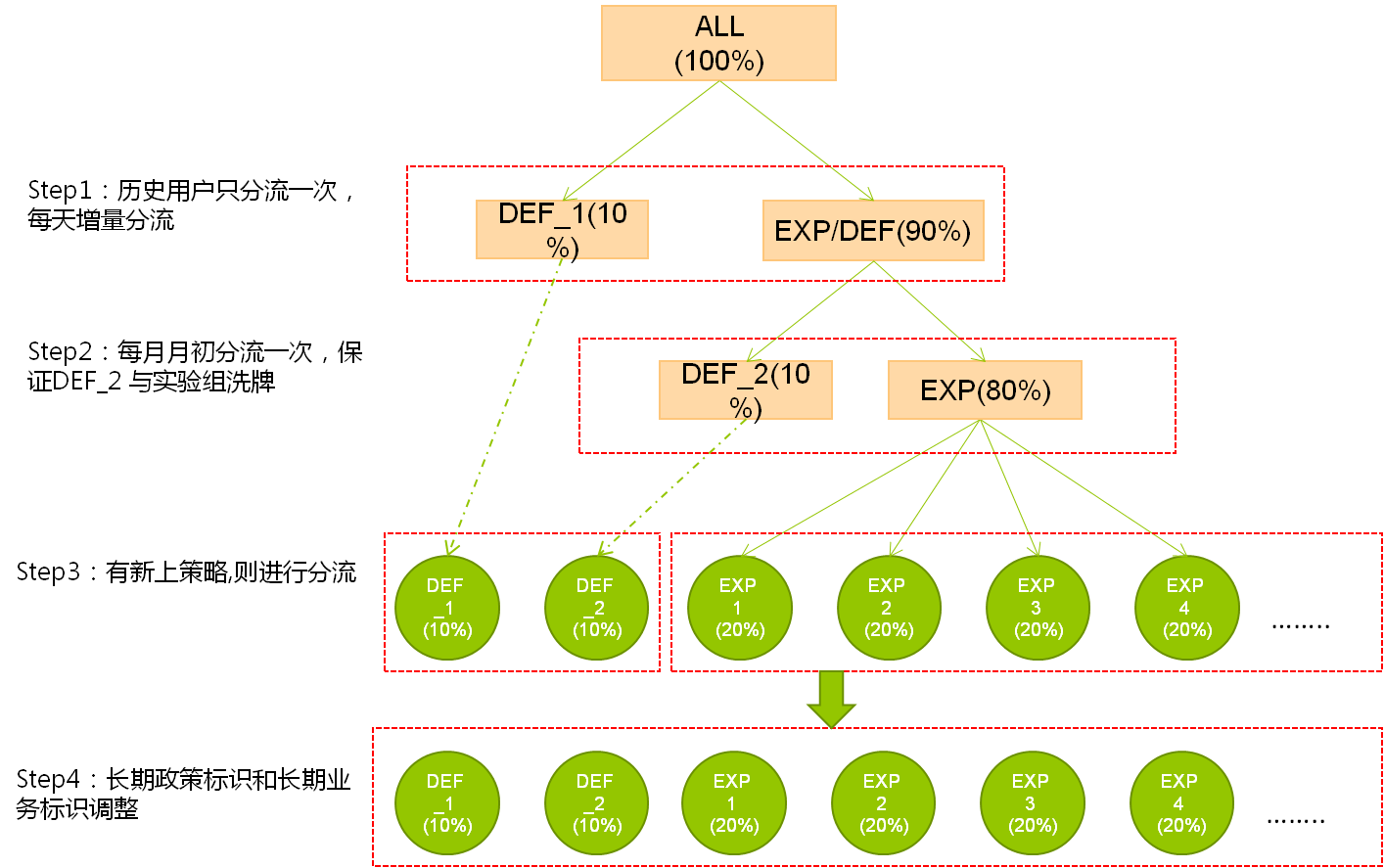
实际分流情况如图4.2 所示，“实验组”（EXP）会员占所有会员80%，“对照组”（DEF）占比20%， 实验组成为4组，各自占比20%，分别是EXP1~EXP4，主要是用于各个不同策略的对比；对照组分成2组，分别是DEF\_1 和DEF\_2，DEF\_1（10%）始终不做发券操作，DEF\_2每月1号重新分流，不做发券操作。

图4.2 ABT实际会员分流情况

# 5 总结

## 5.1成果

**1、显性价值**

从精准营销总体带来的kpi价值和个性化推荐部分带来的kpi价值两块进行总结，图 5.1展示2018年度精准营销效果，上图是销售额，下图是毛利额。 从销售额来看，个性化推荐部分12份占比70%，且随着方法策略的完善，每月其本呈上升趋势（11月受双11影响较大）；从毛利额来看，个性化推荐部分12月份占比71 %， 每月基本呈上升趋势。个性化推荐部分12月份毛利率达38%。

图5.1 2018年度精准营销效果

**2、隐性价值**

精准营销经过一年多的实际探索，形成了一套独有的推荐引擎，并形成了一套趋近完整的方法论，具有可复制性的扩展性。具体体现在如下：

* 在实现上：具有一套精准营销实现系统，以及业务实现的逻辑框架。
* 在方法策略上：形成一套推荐业务思路以及券的券门槛力度的制定与匹配的方法。
* 在效果评估上：引入互联网ABT的思路，并结合益丰的情况得以合理实现。
* 在底层模型上：借助精准营销项目，形成了益丰特有的会员画像、商品画像，目前会员画像已有近90个可用标签，商品画像已有近40个可用标签。

从上述的描述中，精准营销项目具备了“产品”的能力 ，可以向外孵化新应用的能力。对于益丰数据而言，其底层数据、底层模型具有可共享性，可以为新应用赋能，同时新项目亦可为数据赋能，形成业务数据闭环。

## 5.2不足

**1、业务问题**

* 目前精准营销（券系统）场景受限、单次曝光机会少、渠道单一
* 选品较难突破，可推范围受限（单品、品类推送比例不到30%）
* 载体太少，未突破券

**2、系统问题**

* 券模板的设计存在不合理性，难以做到同一个券在不同场景复用，由BI 底层脚本生成，维护成本较高
* BI与发券系统对接 路径太长；同时梦网数据没有返回，没有形成闭环
* 业务变化较大，各个模块初期未考虑其因素，导致各系统、各模块修修补补，扩展性较差
* 券的来源较多，没有统一出口进行管控

# 6规划

精准营销的规划从横向和纵向分别阐述，横向上有望突破“券”，可以横向扩展多场景、多渠道以及营销方式及载体的多样化，增加曝光机会，最大化平台收益；纵向上让精准营销推荐引擎可以根据不同应用封装不同服务接口，底层数据模型可共享，最大化数据应用，同时让精准营销方法论更完善。

## 6.1横向业务扩展

对于当前的“券”而言，“精准营销”是它的方法论； 对于“精准营销”而言，“券”只是它的其中一个承载方式。 因此，精准营销除了支持“券”之外，还可以支持其他的承载形式，例如服务、运营、积分、会员权益等。精准营销与应用关系图如图6.1所示。

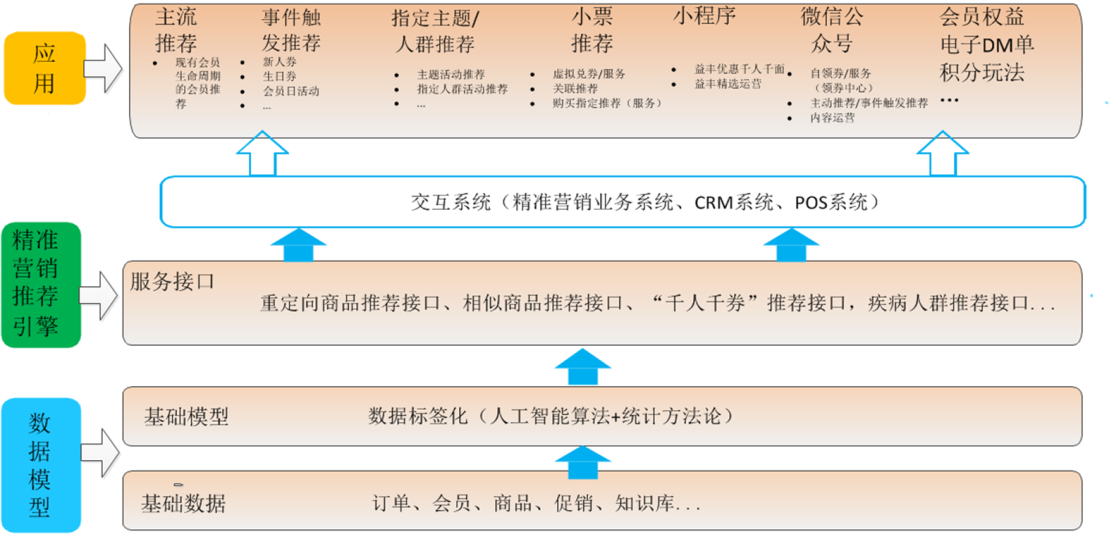


图6.1 精准营销与应用关系图

## 6.2纵向技术深入

借助“数据中台”的概念，希望BI底层数据模型具有可共享性，前端应用能快速、灵活支持。因此，将精准营销底层能力进行模块化、接口化、服务化。将业务可变和业务不可变模块进行解藕、数据与业务进行解耦设计。精准营销模块化设计架构图见图6.2。

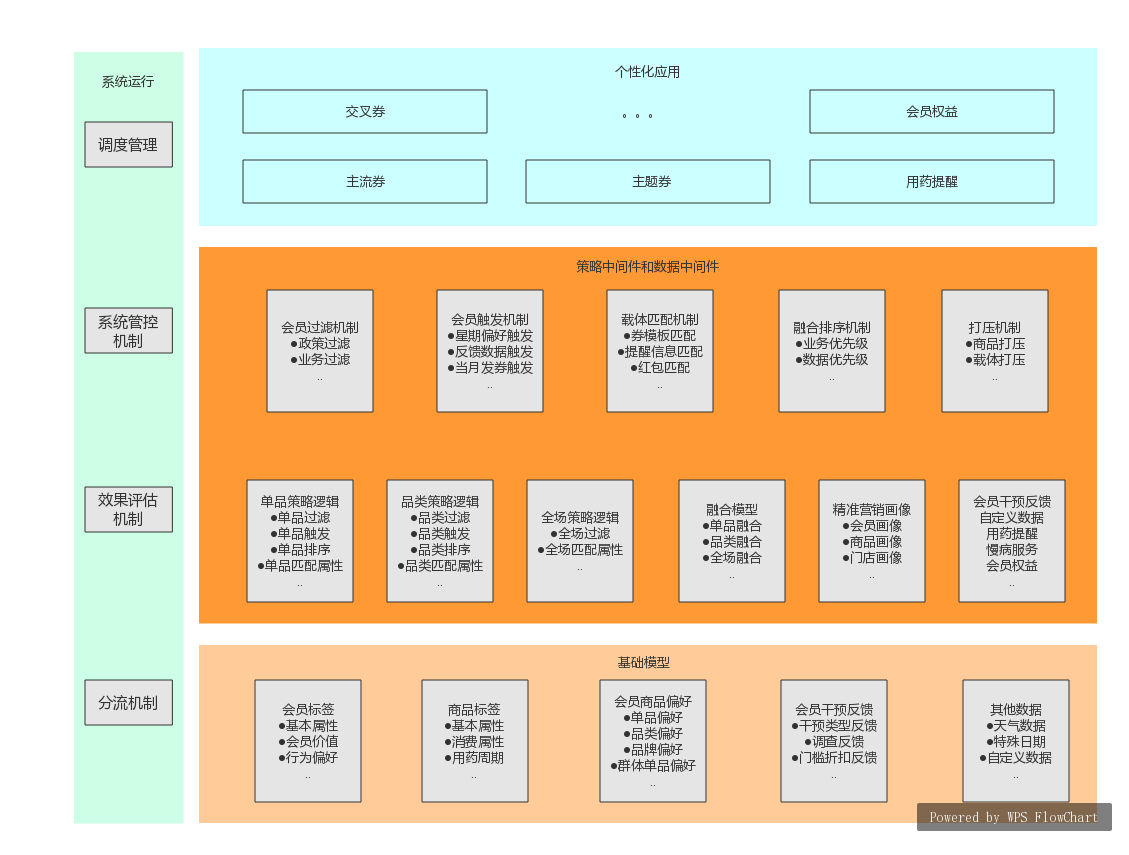


图6.2 精准营销模块化设计架构图

架构原则如下：

1. 服务模块分为数据服务和策略服务，两者之间及其内部均解耦设计
2. 数据服务实现数据的模块化，是对数据中心的整合，不提供动态修改参数服务，即其数据内容只与其依赖数据源有关；
3. 策略服务层实现业务的模块化，提供动态修改参数服务，即其数据内容与场景配置有关；
4. 所有参数由系统管控机制统一管理。