

# 美团外卖推荐业务中的数据处理流程以及性能优化

191250059 蒋承霖

191250172 颜永健

191250148 吴嘉豪

191250146 文梟侠

191250078 李智强

## 引言

美团外卖是中国生活服务网站美团网旗下的互联网外卖订餐平台，由北京三快在线科技有限公司运营，创立于 2013 年，是目前中国最大的网上订餐外卖平台。美团外卖自诞生之初便以全面的商家信息整合、合理的配送业务分发、高水平的个性化商品推荐以及搜索推荐等引人注目并得到用户和市场的好评，而这些业务离不开美团对于其业务运营原始数据的智能化应用。本文试从美团外卖推荐业务出发，分析美团外卖推荐业务中的原始业务数据是如何进行处理转化成业务需要的数据，为智能化业务服务的，并达到学习互联网企业数据的智能应用的技术栈和技术逻辑的目的。

## 一 应用场景概述

本文试从“消费者在美团外卖上下单购买外卖”这一具体消费场景作为引子，重点分析该场景下的推荐相关的细分业务（如：个性化推荐，搜索结果排序），分析这些业务的处理思路以及其处理思路下的数据转换逻辑，以达到学习美团的数据视角的智能业务处理逻辑的目的。

需要指出的是,推荐业务一般情况下属于机器学习的业务范畴,更多注重于建模、模型的训练以及数据的处理方法,而本次调研的主要目的是学习先进互联网企业对于数据的智能化应用的处理思路,以及数据在智能应用的过程中的转换、转移等的实现方法以及对应的优化方式,因而对于推荐业务的部分核心工作的分析会避免聚焦于技术细节以及数据细节处理,而更关注整个推荐业务所涉及的技术栈以及基础设施,以及数据处理后会保留或者凸显什么特征。此外本报告更加关注于分层的数据处理,而不同层次的数据处理在某一推荐业务执行的子环节通常是连续的,因而会导致本文不可避免的在具体流程的描述上存在割裂,以及对不同层次的数据处理环节的介绍中存在重复的处理方法介绍,还望理解。

## 1.1 业务特征分析与业务目标确定

在进入对于推荐业务流程的具体分析之前,我们需要处理两个问题:外卖业务具有什么样的特征?在外卖业务的特征的基础上,我们的推荐系统所需要最终实现的目标是什么?如果用户在外卖业务中的消费特征(消费习惯)倾向于稳定的话,我们在外卖业务应用推荐系统就有业务价值,值得为其打造并且维护一套推荐系统,并且会影响后续的数据源选取、数据特征设定的导向;后者则会影响推荐系统的输出的价值导向,例如,我们推荐的目的是要维护用户对于个性化推荐系统的满意度,还是构建 O2O 的推广渠道。

美团外卖团队方面对于外卖业务的特征定性如下:

1. 多样性,快速发展:这意味很多用户对平台上的新品类缺乏了解,对自身的**需求也没有充分意识**。平台需要去发现用户的消费意愿,以便对用户的消费进行引导。该特性也决定了个性化推荐业务在美团外卖业务中举足轻重的地位。
2. 高频:外卖是个典型的高频 O2O 应用。一方面**消费频次高**,从用户的消费频次和消费周期能够较好地看出其对外卖业务的依赖性,相应的能较好的预判用户对本平台的粘性;另一方面**消费单价较低**,用户**决策时间短、随意性大**。
3. 场景驱动:场景是特定的时间、地点和人物的组合下的特定的消费意图。不同的时间、地点,不同类型的用户的消费意图会有差异。例如白领在写字楼中午的订单一般是工作餐,通常在营养、品质上有一定的要求,且单价不能太高;而到了周末晚上的订单大多是夜宵,追求口味且价格弹性较大。场景**辨识越细致**,越能了解用户的消费意图,运营效果就越好。

4. 用户消费的地理位置相对固定，结合地理位置判断用户的消费意图是外卖的一个特点。

在对于以上的对于外卖业务的定性的基础上，美团团队方面确定了以下的业务目标：

1. 以个性化推荐业务作为吸引新用户使用外卖业务的核心手段。
2. 提升用户价值。具体而言就是提升用户的单均价和消费频次，从而提升用户的LTV (life-time value)。
3. 保障用户的忠诚度，通过提升用户总体体验将用户留在外卖平台。

## 二 抽象数据的标准化构建

本章节试描述在具体的推荐环节中所需要的数据，以及美团外卖如何从客户的原始交互数据中进行集成、提炼与处理，最终得到推荐业务所需的、可用的、有价值的输入数据。

本章节中所描述的“原始数据”可大致分为以下两种：

1. 用户（包括客户和商家）在进行相关业务执行的交互时，对美团方面提交的原始数据。
2. 美团方面通过社会调研以及统计所获取的，客户端外的“软数据”。

推荐业务的核心要点在于“猜你喜欢”，即系统根据用户在美团外卖平台上表现出来的喜好，按照一定的顺位，推荐可能符合其喜好的产品。在这一套业务逻辑下，我们有两个“是什么”的问题需要处理以保障业务的可实现性：

1. 在系统眼里的“你”是什么样的？要怎么构建出“你”，以及如何知道“你”喜欢什么？
2. 要如何知道哪些商品具有什么样的特性，是否满足“你”的喜好要求？

其中，前者对应的具体业务为用户画像的构建，后者对应的具体业务为外卖商品的特征提取以及商品标准构建。

### 2.1 用户画像构建

用户画像是对于本业务用户的特征抽象，用以记录并且维护用户在使用美团外卖业务时表现出来的特征（比如商品喜好，下单时间，订单浏览时间等）。

用户画像的构建直接关系到推荐系统对每个用户消费习惯和口味的把握，直接影响推荐的准确度。通过对收集到的用户（包括商家和消费者）在平台上产生的数据进行系统性的分析，可以得到统计学视角的美团整体用户的用户画像以帮助平台进行商业决策，也可以针对每一个不同的用户进行细致的用户画像构建，以达到精准的推荐。

美团采取了两种方式完成用户画像的构建。第一种是结合美团外卖业务产生的特征数据以及对美团外卖用户的市场调研数据，从统计学的视角构建美团外卖用户整体的用户画像；第二种是对个体用户在使用美团外卖的过程中产生的业务数据进行建模，所构建出来的具体用户的基于计算数学建模的用户特征画像。前者主要服务于美团外卖宏观服务的优化，后者主要服务于具体用户在美团外卖业务中的个性化推荐服务。

### **2.1.1 统计学意义用户画像建模**

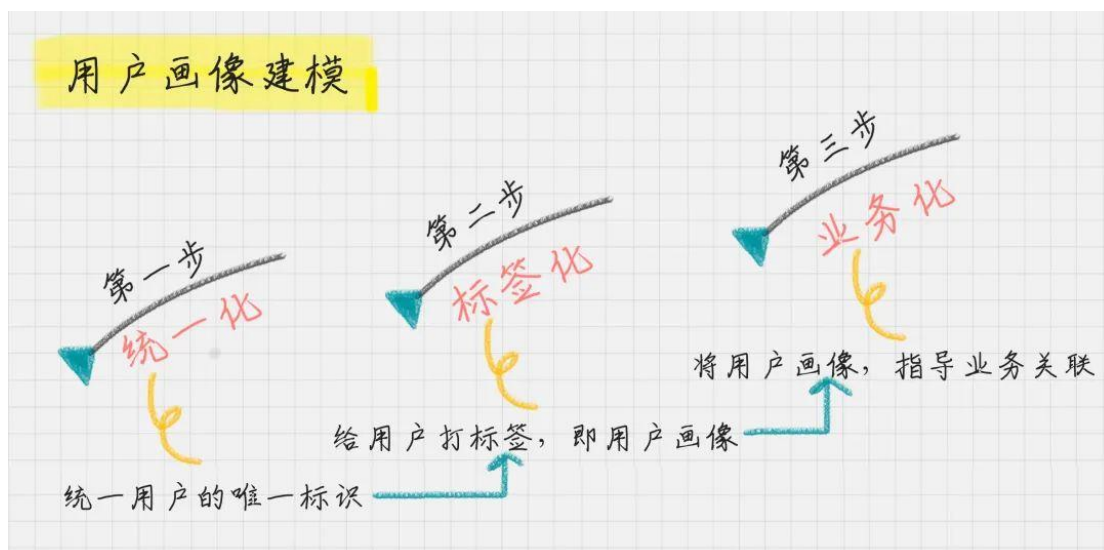
为了方便进行宏观的推荐业务的定性以及调整，需要先通过统计学的方式，对于用户的身份信息（性别、年龄、职业、活跃区域等）进行统计提炼，以构建美团外卖业务的宏观的用户画像。宏观的用户画像将作为美团外卖业务战略定性的关键指标（本文不涉及），并指示推荐业务的可选的研究倾向。



@逸宏

### 2.1.2 个人用户画像建模

对于个人用户画像的建模，主要分为三个步骤：统一维护唯一标识、设定画像特征类型并对用户进行各特征的标签化处理，将初步建模的用户画像和外卖业务构建关联。

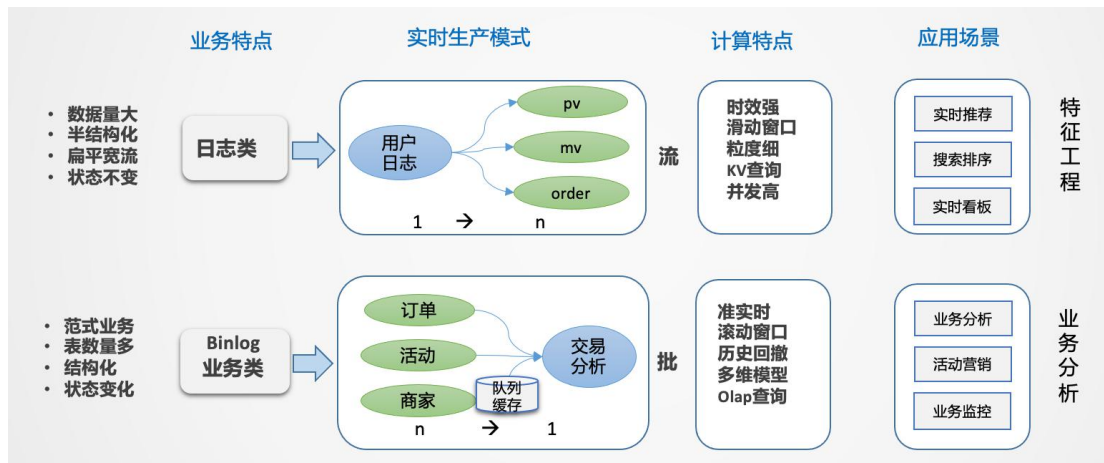


## 1. 用户身份标识确定以及画像建模数据源处理

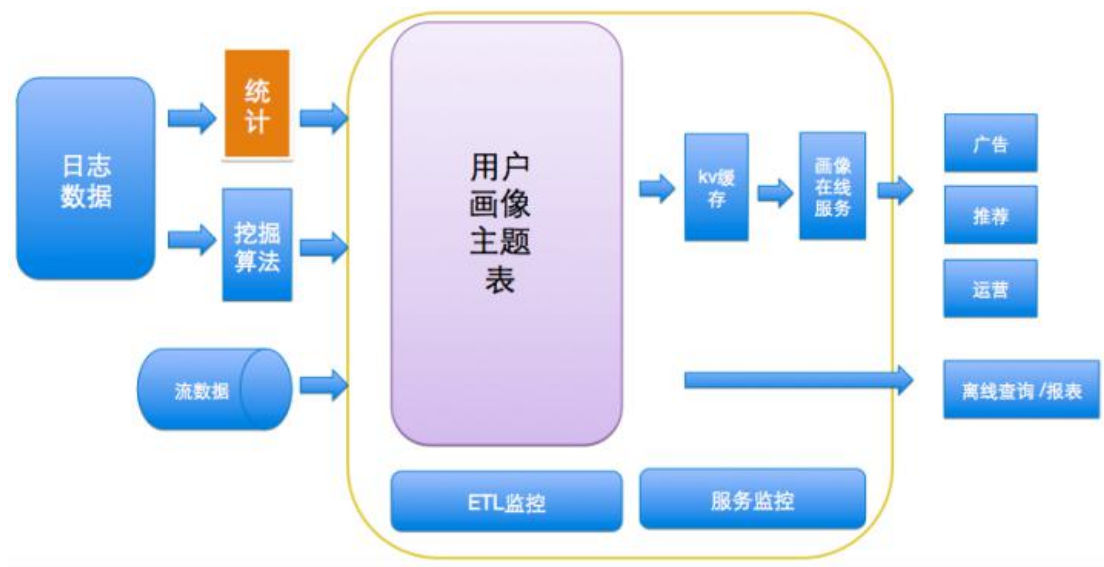
构建用户画像首先要对用户建立统一化的标识 id，以便追踪具体用户的具体行为。美团主要以手机号作为用户的标识 id。

美团画像服务的架构中，数据源包括基础日志、商家数据和订单数据。基础日志一般指用户业务流程中所产生的符合日志规范的记录性文件，这里根据需求可以狭义解释为记录了用户进行外卖业务时所产生的数据（比如用户进入某一店铺时发起的网络请求）的业务日志，主要通过记录用户的业务执行信息，用以反映用户在外卖业务的行为特征；商家数据则记录了商家及其所售卖商品的相关数据（详见 2.3），用户和商家相关的请求结合商家数据可以一定程度上反映用户对于商家的好恶，进而还可以从商家数据所记录的特征信息（例：商家主要售卖的商品类型）进而反应用户对于外卖产品类型的好恶倾向；订单数据则记录了用户的下单时间、下单地点、商家以及其产品、单价、优惠使用情况、备注信息等数据，这份业务数据中最为核心的数据除了能够用作反映上文所提及的外卖店家与产品的好恶信息，更能反映用户的消费习惯、消费倾向等其他数据难以有效反映的用户特征信息，同时也可作为外卖业务场景建模的输入数据（见 2.2）。

用户画像处理引擎可以从存储中获取这些数据源作为用户画像建模的输入，而在业务密集的环境下为了提高处理效率，相关数据往往还会整理成流作为用户画像建模流程的输入。



数据收集完成之后，存放于用户画像的主题表中，之后再进行 kv 存储和为画像在线服务提供支持。美团会对业务流程进行监控，包括利用数据治理平台对数据处理流程进行监控以及对下游服务（广告、推荐、排序等）进行监控。下图展示了美团画像服务的整个处理逻辑。



## 2. 用户画像具体建模

用户画像具体建模对应了上文所指的“打标签”，本质上是对用户进行特征提取，对于特征进行建模，并且再结合数据源对于建模的特征进行调整，以构建或者完善用户的用户画像模型的流程。

在文中，“特征”一词有两种含义：一表示可以和事物的某一特性或者行为倾向构建映射，进而以具体的数据反映事务模糊的特性或者行为的倾向以及倾向程度的数据，二表示机器学习理论中的特征，也即“被观测对象的可测量性能或特性”。业务逻辑视角下的特征取前一含义，具体的数据处理逻辑分析时的特征一

词取后者含义。而“打标签”即为“明确特征需求”的过程，即“选取既能和用户的外卖业务构建深度关联，又方便获取以及处理相关数据的特性”。

在对用户打标签的阶段，需要着重考虑的有：用户标签，消费标签，行为标签，内容分析。

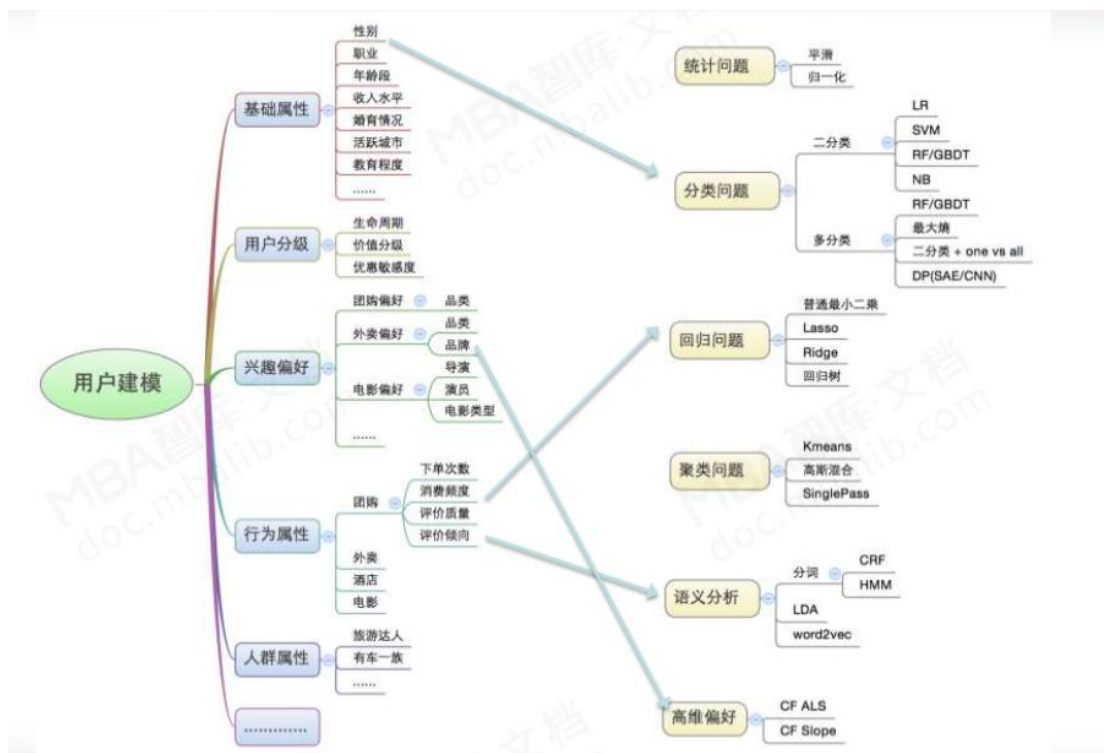
- 用户标签为用户的主要信息，如性别、年龄、地区、家庭婚姻状况等等，通过对用户数据的分析来描述一个用户的个人情况。
- 消费标签则重点在于用户的消费状况。如餐饮口味，消费均价，团购等级，预定使用等级，排队使用等级等等。
- 行为标签则是对用户在平台中产生的行为进行分析，如点外卖的时间段，使用频次，平均点餐用时，访问路径等。
- 内容分析则是对用户平常浏览的内容进行统计以描述用户喜好。

由于数据量是巨大的，使用基于 hadoop 生态衍生出的大数据工具对数据进行存储（HDFS，MongoDB 等）和计算（MapReduce，Spark 等）。针对海量的数据，需要开发自动化的处理算法，以提高分析效率、降低人工成本。

### 3. 用户画像处理的算法流程

先收集到用户的特征值信息，然后对用户特征进行处理，包括数据清洗去除异常数据，对特征进行值处理、选择、组合和降维，然后生成有价值的商业数据。具体流程如下图。





## 2.2 用户场景建模

除了用户画像的构建, 还需结合业务发生时的其他数据对于用户请求这一业务的场景进行特征提炼以及建模。场景可以分为时间, 地点(如商业圈), 周围环境

（比如天气）等，然后对场景进行初步的分析，比如对于地点，通过周边环境、用户处于该地点得时间段，判断该地点是否为用户得常驻点（家、学校等），然后结合用户画像分析其行为动机。再获得用户场景的基础上，还要深入挖掘用户生活中的真实场景，将用户生活中的一个个场景点串成线形成面，最终形成用户一天的场景转换动态，从而更准确的匹配到具体的功能点。

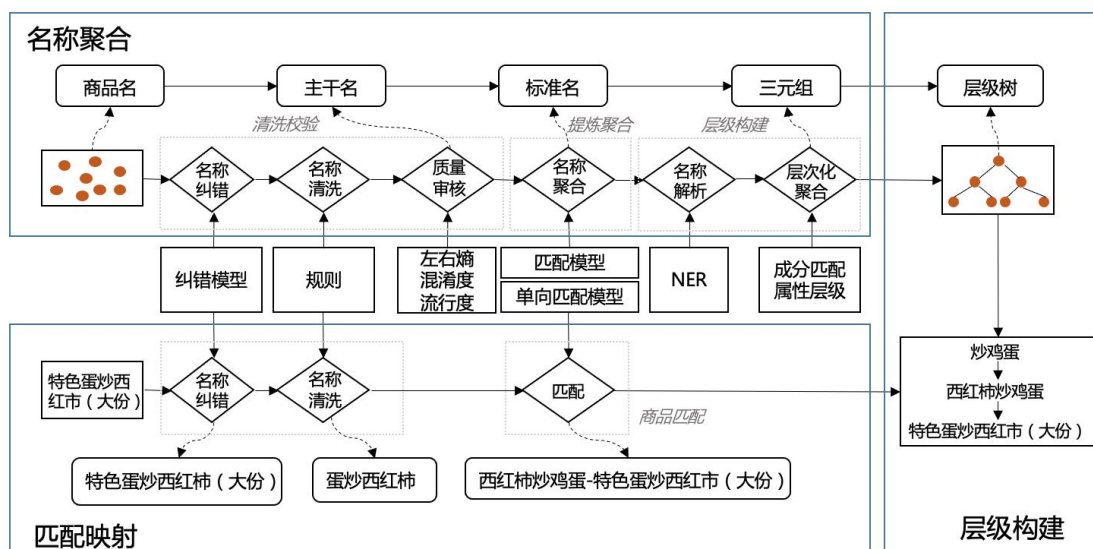


## 2.3 商品信息标准化

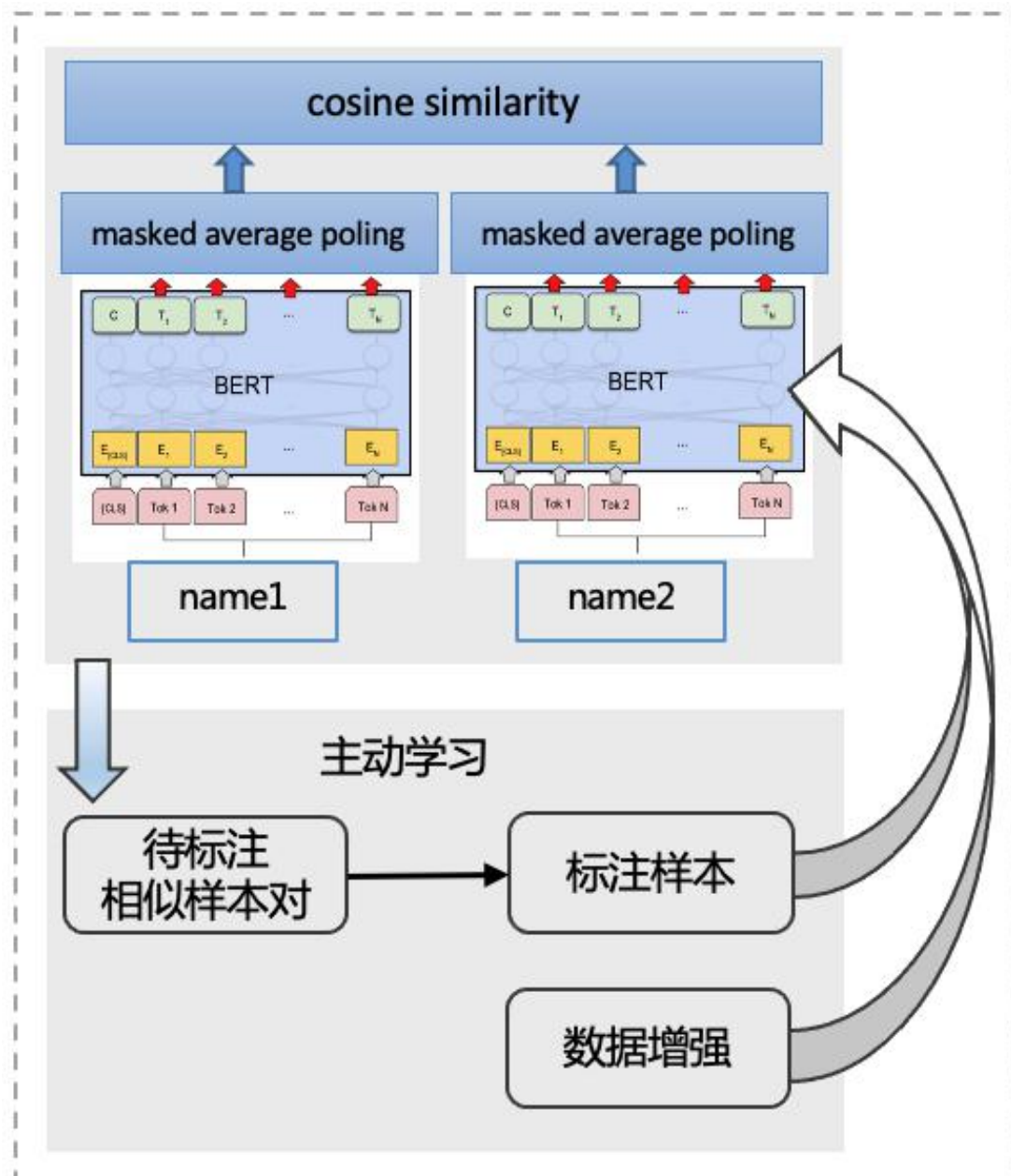
推荐的关键环节，是根据用户的标签提供匹配的菜品种。为此，将菜品标签化，标准化，将同类商品聚合从而构建美食知识图谱是重中之重。其建设可以参考淘宝的SPU建设。SPU建设实际上是对属性的建设，比如品牌、类目、型号等将商品标准化，唯一化。然而与淘宝商品不同，菜品的标准化面对更大的困难。以下是几个主要问题：

1. 菜品的命名具有很高的自由度，个性化程度较高。对同一个菜品的命名，不同商家的不同语言习惯会带来不同的命名方式，甚至可能为了提高销量为菜品取“昵称”，导致菜品没法像手机等其他商品那样通过关键字进行聚合。在某些情况下，商品名称甚至无法明确菜品的类型。
2. 如果按照食材、口味、做法等菜品详情进行打标签和细分，则会大大增加商家的录入成本；但由于菜品属性没有标准化，有无法直接将菜品和属性自动关联起来。
3. 标准化的粒度难以把控，容易出现粒度过粗或者过细。比如将香辣鸡翅标准化为鸡翅会导致粒度过粗，传统黄焖鸡【大碗】”本身粒度过细，需要提炼到“黄焖鸡”。

美团的商品标准化方案分为一下几个步骤：



1. **清洗：获得主干名**
2. **名称聚合：合并同义主干名。** 先后采用规则匹配和语义匹配。规则匹配利用 NER 模型对主干名进行成分识别，结合知识图谱构建的属性同义词表，判别两个主干名是否是同义关系。语义匹配弥补了规则匹配中同义词有限的问题。美团采用了 BERT+DSSM 的语义匹配模型进行同义关系的扩覆盖。



主动学习的方式是先利用基础模型，圈定一批待标注的相似样本，交与外包标注，将标注正确的样本补充至已有同义词中，标注错误的样本作为负例加入训练集，用于模型的优化迭代。对于字面相似度高但核心成分不同的匹配错误，先根据字面距离圈定了一批字面相似度高的样本，再利用名称解析模型对它们进行成份识别，找出其中的负例。通过这种方式，在不增加标注成本的情况下，自动补充了十万级样本，进一步提升了模型准确率。

### 3. 匹配映射。

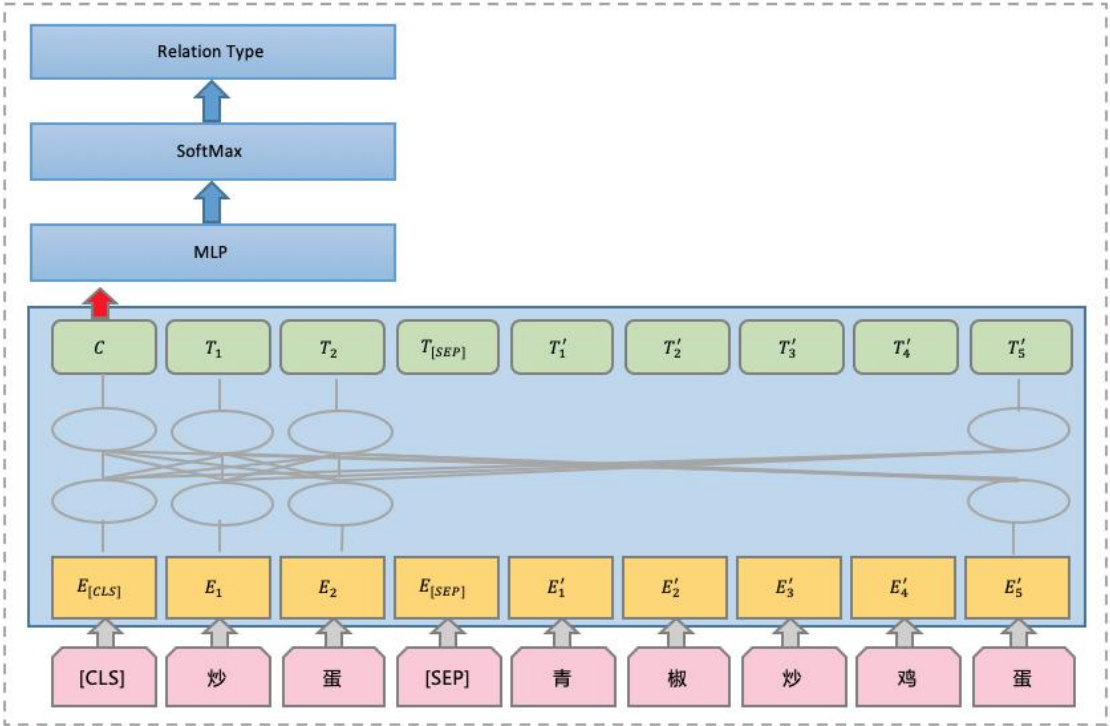
1. **文本匹配：**整体上包括召回、排序两个阶段。召回阶段对商品名称中的描述性信息进行清洗并进行切片，通过关联相同切片的方式召回待匹配的标准名；基于召回的标准名，通过计算 Jaccard 距离，保留其中 Top 20 的标准名；在此

基础上，生成商品名和标准名的向量表示，通过计算 Jaccard 字面距离以及 Cosine 向量相似度，获取其中综合得分最高的标准名。

**2. 图像匹配：**由于菜品名称的高自由度，会导致从菜品名称获得的信息有限，因此通过引入商品图片信息提升匹配的准确度。

**4. 层级构建**

针对标签粒度问题，对粒度进行合理的聚合，建立层次化的商品体系。构造方法包括规则匹配和模型识别。规则匹配是基于已有的 NER 模型和属性词表，通过结构化匹配的方法，遍历 1 生成标准名层级树。判别模型则进一步提高泛化性。基于 BERT 的关系分类模型如下图所示，对待分类一对标准名用[SEP]进行拼接，并在开头增加[CLS]标识符；将拼接结果编码后，传入 BERT 模型，取出[CLS]位的 Embedding；再接一个全联接层和 Softmax 层，输出关系分类结果。标准名关系包括：同义、上级、下级、无关系，一共四个类别。



经过以上处理后，模糊的，高度多样的商品信息便转换成了可以精确定位与分类聚类的商品信息知识图谱，再结合同样经过特征化处理构建的用户画像，我们便可通过捕获用户画像模型和商品之间的好恶关联来获取用户对于具体商品的，置信度较高的好恶倾向，便可进一步形成推荐业务。

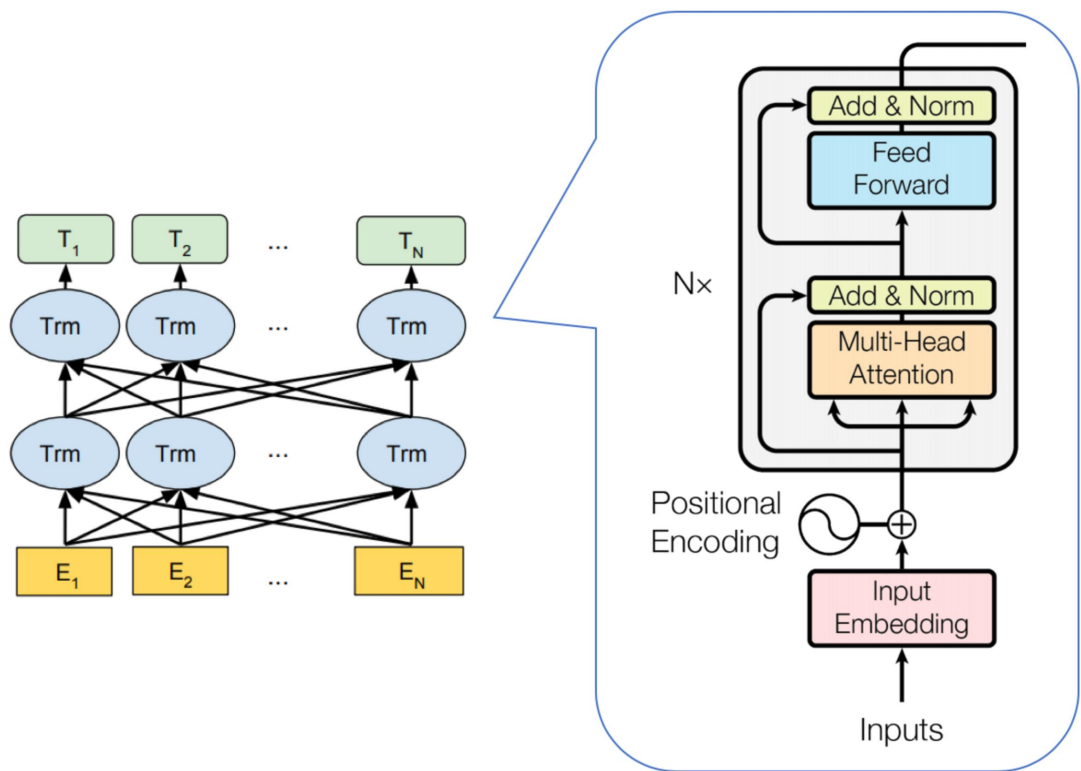
**2.4 模糊业务数据的特征提炼**

在具体外卖业务执行的过程中，可能会产生一些含义并不显然的，模糊的业务数据，比如用户对于外卖产品的文字评价；同时，这些表述模糊的数据一方面承载



了更多的业务信息，相比于明确的原始业务数据可能更有业务分析价值，另一方面由于表达的自由度，也更加贴近于业务的核心环节（比如业务的完成质量评估等）。因而，从模糊的数据中提取出计算机可以理解的数据显得尤为关键。

以客户产生的评价数据为例，在美团的外卖推荐系统中，除了好评差评这种比较模糊的界定之外，美团还采用了 BERT 自然语言处理系统，用以处理用户在外卖交易过程中产生的评论。BERT 是基于 Transformer 的深度双向语言表征模型，基本结构如下图所示，本质上是利用 Transformer 结构构造了一个多层双向的 Encoder 网络。



### 三 数据清洗和特征处理

经过以上操作，美团将抽象的统计学数据以及原始的业务数据转化成了标准化的、特征化的、可以作为推荐系统的输入的数据。随着美团交易规模增大，积累下的业务数据和交易数据越来越多。对这些数据进行分析 and 挖掘，可以为业务方向提供决策支持。目前美团的业务系统中大量地应用到了机器学习和数据挖掘技术，本章节将介绍如何进行数据清洗和特征处理。

### 3.1 数据选取

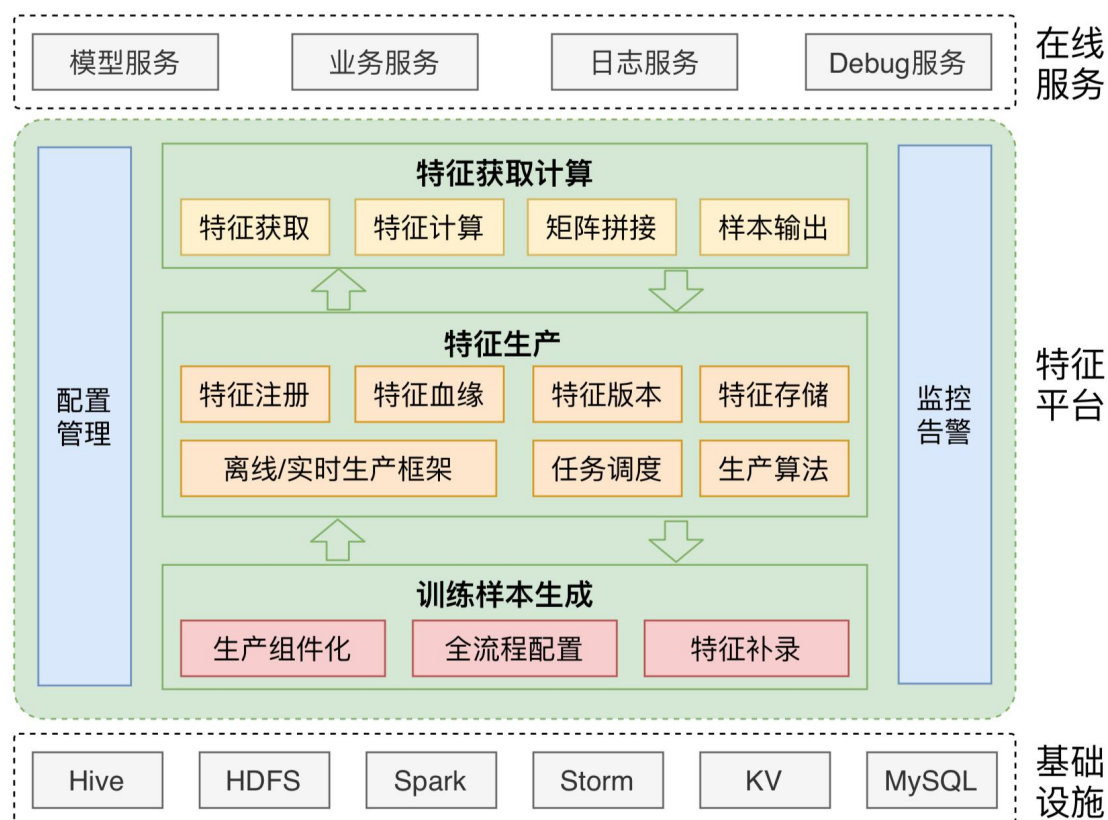
在确定业务目标后，需要确定使用哪些数据可以达到目标——即事先梳理可能与用户是否点击下单有关的数据。一般来说，可以借鉴一些业务经验，同时运用特征选择、分析的方法辅助选择数据。

从业务经验来判断，可能影响用户是否下单的因素有距离、用户历史行为、用户实时兴趣、是否热门、评价人数等等。在确定数据范围后，需要对数据的可用性进行评估，包括数据的获取难度、数据规模、数据准确率、数据覆盖率等。

在选定好要用的特征后，我们需要可靠的特征获取方案，一般有如下两种方案：

- 离线特征获取方案：可以离线使用海量的数据，借助于分布式文件存储平台，例如 HDFS 等，使用例如 MapReduce，Spark 等处理工具来处理海量的数据等。
- 在线特征获取方案：在线特征比较注重获取数据的延时，由于是在线服务，需要在非常短的时间内获取到相应的数据，对查找性能要求非常高，可以将数据存储索引、KV 存储等。

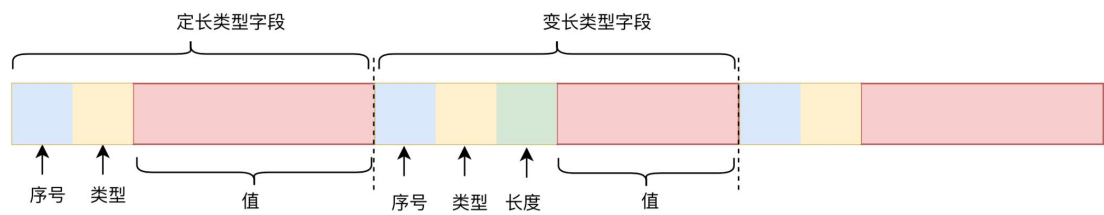
美团外卖部门已建立如下图所示的特征获取平台，该平台框架由三部分组成：训练样本生成（离线）、特征生产（近线）以及特征获取计算（在线）。



### 3.2 数据清洗

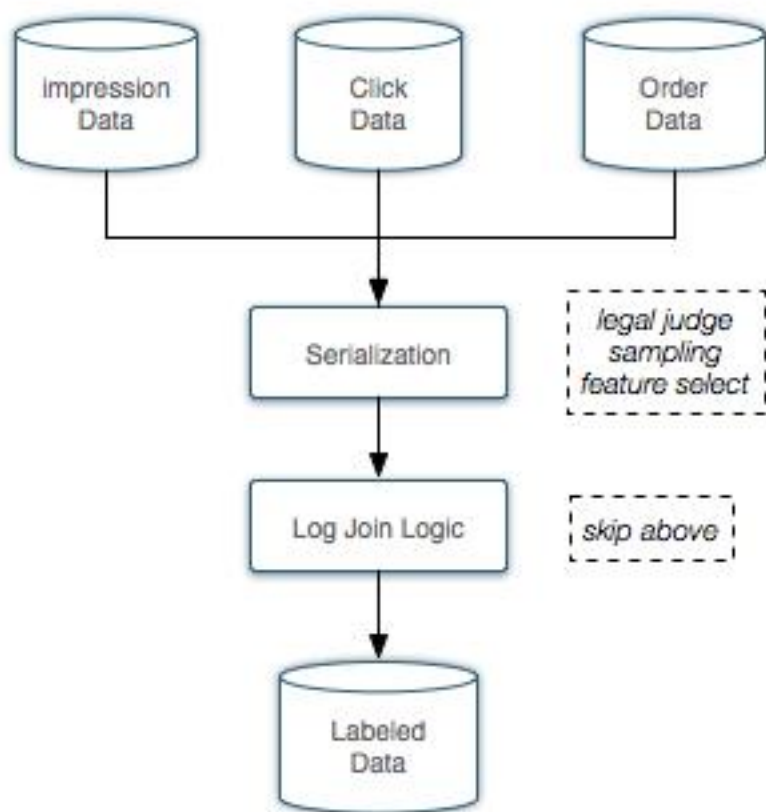
在了解特征数据源与获取方法后，就需要处理数据，首先进行的是数据清洗。数据清洗的主要工作是**为离线模型训练准备标注数据，同时洗掉不合法数据**。数据清洗的大致流程如下：

- 1. 序列化：特征数据通过聚合处理后，需存储到 HDFS/KV 系统中，用于后续服务的使用。序列化的过程中，如果日志字段不合法或者单一用户曝光、点击或下单超出设定的阈值，相关日志都会被清洗掉，避免数据对模型训练造成影响。



- 2. 数据标注：数据序列化之后在 HDFS 上保存多份文本文件，分别是曝光、点击和下单等数据。数据标注模块根据 `globalId`（一次搜索的全局唯一标示）和相应的团购 `id` 为 `key`，将相关数据关联起来，最终生成一份标注好是否被点击、下单、支付等等的标注数据。同时这份标注数据携带了本次展现的详细特征信息。





### 3.3 特征分类

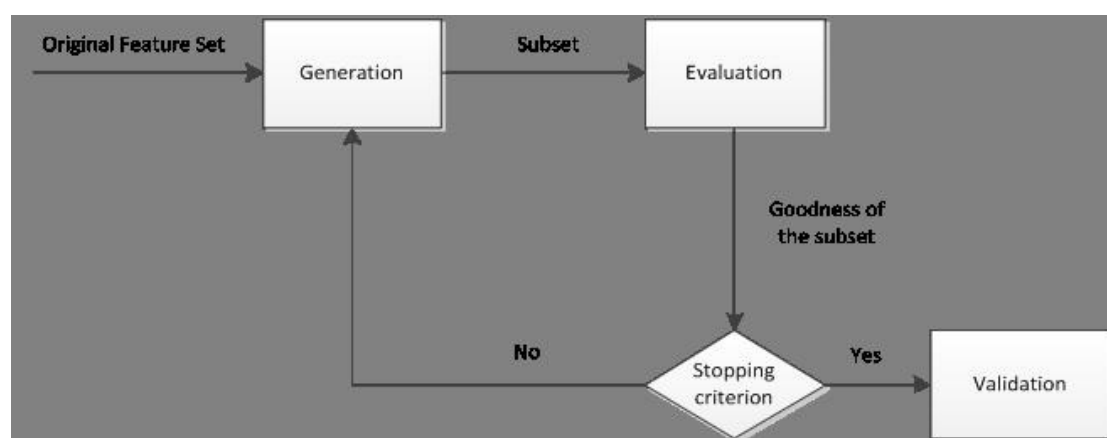
在清洗及标注好数据后，应当对特征进行分类，对于不同的特征应该有不同的处理方法。常用的特征分类方法有如下三种：

1. Low Level 特征和 High Level 特征。Low level 特征是较低级别的特征，主要是原始特征，不需要或者需要非常少的人工处理和干预，例如文本特征中的词向量特征、图像特征中的像素点、用户 id、商品 id 等；High level 特征是经过较复杂的处理，结合部分业务逻辑或者规则、模型得到的特征，例如人工打分，模型打分等特征。一般前者用于简单的模型，后者用于复杂的模型。
2. 稳定特征与动态特征。稳定特征是变化频率（更新频率）较少的特征，例如评价平均分、团购单价格等，在较长的时间段内都不会发生变化。动态特征是更新变化比较频繁的特征，有些甚至是实时计算得到的特征，例如距离特征、2 小时销量等特征。针对两类特征的不同可以针对性地设计特征存储和更新方式。

3. 二值特征、连续特征、枚举特征。二值特征主要是 0/1 特征，即特征只取两种值：0 或者 1；连续值特征是取值为有理数的特征，特征取值个数不定，例如距离特征；枚举值特征主要是特征有固定个数个可能值。

### 3.4 特征处理与分析

对特征进行分类后，就需要进行处理和分析。特征有多种处理方法，本章节介绍其中一种，即特征选择方法。特征选择的目的是从众多特征中挑选出少许有用特征，也即寻找最优特征子集。该方法能够剔除冗余的以及不相关的数据，从而实现目标。特征选择的过程一般如下图所示，主要分为产生过程，评估过程，停止条件和验证过程。



特征选择开销大、特征构造成本高，随着美团外卖业务量的增大，推荐业务的业务需求量也呈现跃迁式的增长，特征工程的巨大投入对于效果的提升已经不能满足业务需求，因而需要引入其他深度学习方法提升特征选择的效率。下文简要介绍一些基于深度学习的特征选择优化方案（仅介绍方案的实现逻辑，避免具体的实现方法以及优化方法介绍）：

#### 1. 组合特征

因为不同特征之间的组合是非常有效的，并有很好的可解释性，比如将” 商户是否在用户常驻地”、“用户是否在常驻地”以及” 商户与用户当前距离” 进行组合，再将数据进行离散化，通过组合特征，就可以很好的抓住离散特征中的内在联系，为线性模型增加更多的非线性表述。

#### 2. 归一化

归一化是依照特征矩阵的行处理数据，其目的在于样本向量在点乘运算或其他核函数计算相似性时，拥有统一的标准，也就是说都转化为“单位向量”。美团外卖主要实践的归一化方法包括 Min-max 函数法以及 CDF 函数

法。

$$x' = \frac{x - \min}{\max - \min}$$

$$x' = \int_{-\infty}^x f(x) dx$$

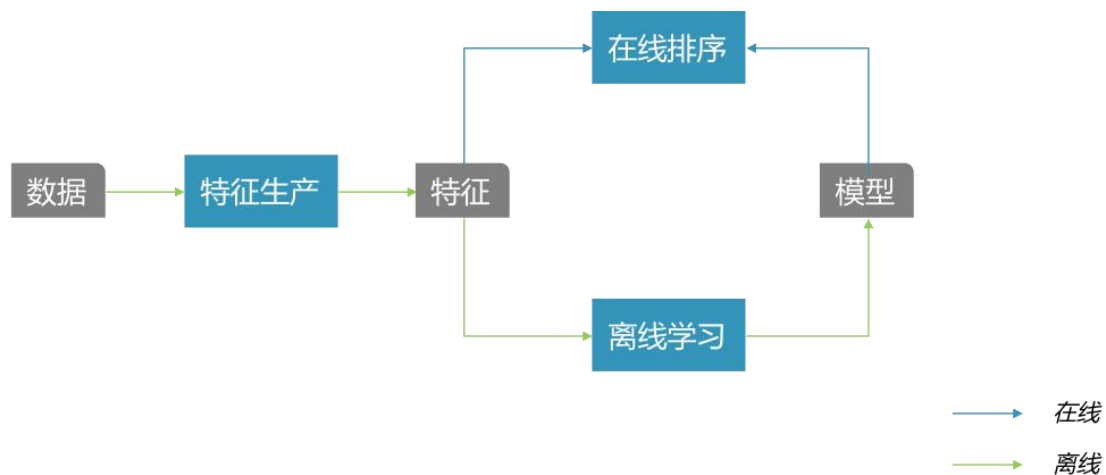
## 四 具体业务执行

接下来我们在本章节来描述在获取到经过标准化的、特征化的、经过清洗的数据后，美团如何对数据进行处理以实现对用户的外卖商品的推荐。

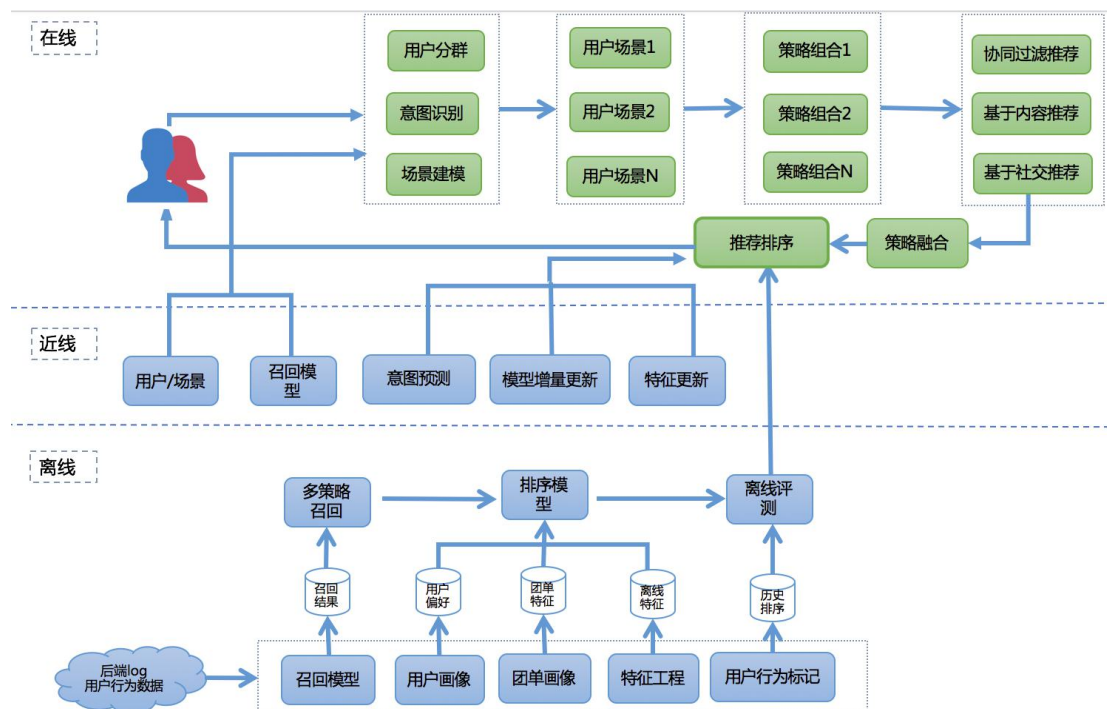
### 4.1 推荐内容排序

在已经构建起用户画像并获取到特征性的用户的喜恶倾向，以及特征化的外卖产品数据的基础上，美团相关系统通过推荐系统以及推荐算法，获取到可能符合客户喜好的产品。目前，市面上大多数的推荐系统都是基于排序实现的，也即对于推荐对象，通过算法将推荐对象所映射的，经过预处理的数据进行基于深度学习、强化学习的加工，获取特征数据，再对特征数据排序以实现推荐效果。排序同时也被应用于搜索结果的展示中，用以完善搜索业务。

美团的具体业务执行模型如下图所示，将数据用于特征生产后通过机器学习后产生模型用以结合特征进行排序。



推荐流程如下：



其中，特征工程以及用户行为数据的转换已经在前文进行了一定程度的描述。从整体框架的角度看，当用户每次请求时，系统就会将当前请求的数据写入到日志当中，利用各种数据处理工具对原始日志进行清洗，格式化，落地到不同类型的存储系统中。在训练时，利用特征工程，从处理过后的数据集中选出训练、测试样本集，并借此进行线下模型的训练和预估。采用多种机器学习算法，并通过线下 AUC、NDCG、Precision 等指标来评估他们的表现。线下模型经过训练和评估后，如果在测试集有比较明显的提高，会将其上线进行线上 AB 测试。同时，也有多种维度的报表对模型进行数据上的支持。

对于不同召回策略所产生的候选集，如果只是根据算法的历史效果决定算法产生的 Item 的位置显得有些简单粗暴，同时，在每个算法的内部，不同 Item 的顺序也只是简单的由一个或者几个因素决定，这些排序的方法只能用于第一步的初选过程，最终的排序结果需要借助机器学习的方法，使用相关的排序模型，综合多方面的因素来确定。

## 结语

本文以美团外卖的推荐业务作为互联网企业智能数据应用的案例，以推荐业务的整个执行过程中数据转化路线以及数据处理流程为主要分析对象，调研了美团外卖推荐业务的较为完整的数据处理思路。

根据上文所整理的，美团外卖推荐业务中的数据处理思路可以归结如下：



由于本小组的能力有限，以及部分技术以及工程优化实践可能涉及企业商业机密而缺乏公开资讯，部分内容缺乏具体的讲解，仅止于“有这样的东西存在”，缺乏对于其运作原理以及运作逻辑的充分理解，在此致歉。

## 参考资料

美团外卖特征平台的建设与实践，

<https://tech.meituan.com/2021/03/04/featureplatform-in-mtwaimai.html>,

机器学习中的数据清洗与特征处理综述，

<https://tech.meituan.com/2015/02/10/machinelearning-data-feature-process.html>

外卖商品的标准化建设与应用，

<https://tech.meituan.com/2021/05/27/food-name-standardization.html>

美团外卖美食知识图谱的迭代及应用，

<https://tech.meituan.com/2021/05/27/food-knowledge-graph.html>

个性化推荐技术与实践（美团技术开放日），

<https://gitbook.cn/books/568f50826dbb161d65e9ce37/>

美团外卖产品分析报告：需求、设计、运营，

<http://www.woshipm.com/evaluating/1629802.html>

美团外卖的用户画像怎么设计？用户画像全流程讲解！，

<https://jishuin.proginn.com/p/763bfbd2c89e>

机器学习在用户画像上的应用（SDCC2015），

<https://doc.mbalib.com/view/ff925128e03010ec79ff91456c98cff5.html>

美团机器学习实践-用户画像数据挖掘，

<https://blog.csdn.net/doufangzheng/article/details/90368635>

每天数亿用户行为数据，美团点评怎么实现秒级转化分析？，

<https://tech.meituan.com/2018/03/20/user-funnel-analysis-design-build.html>