

主播信息的推送方法、装置、计算机设备和存储介质

本申请要求在2019年12月18日提交中国专利局、申请号为201911311364.0的中国专利申请的优先权，该申请的全部内容通过引用结合在本申请中。

技术领域

本公开涉及直播领域的技术，例如涉及一种主播信息的推送方法、装置、计算机设备和存储介质。

背景技术

随着网络科技的发展，尤其是移动终端的普及，主播用户在直播平台上开设直播间，将身边的事情进行直播，其他观众用户可以进入该直播间观看直播，与主播用户进行互动。

为了提高观众用户的用户体验，直播平台会筛选观众用户可能喜欢的主播用户，将该主播用户的用户信息推送至观众用户，方便观众用户快速进入该主播用户开设的直播间观看直播。

主播用户的筛选方式是由基于机器学习的模型来打分排序决定的，模型输出的分数主要由多维度的特征决定。在直播的场景下，老的主播用户由于具有丰富的特征会得到较多的推送机会，而新的主播用户则因为缺少特征只能得到少量的推送机会。

针对上述主播用户的筛选方式，一方面，会抑制新的主播用户的积极性，影响整个直播生态的健康发展；另一方面，推送的主播用户过于集中在特征丰富的老主播用户中，重复推送主播用户、主播用户特性相近的情况较多，容易让观众用户厌倦，而观众用户想要获取其他主播用户，通常需要使用关键词搜索主播用户并进行筛选，这不仅仅对筛选、推送主播用户所耗费资源造成浪费，也导致观众用户观看直播的操作繁琐。

发明内容

本公开提供一种主播信息的推送方法、装置、计算机设备和存储介质，以解决集中推送特征丰富的老主播用户导致资源浪费、观看直播的操作繁琐的问题。

提供了一种主播信息的推送方法，包括：

召回主播用户；

以直播历史与直播效率作为分类的维度、对所述主播用户划分主播类型；

对所述主播用户计算交互值，其中，所述交互值用于表征观众用户观看所述主播用户直播间的特性；

根据所述主播类型，使用所述交互值对所述主播用户计算综合值；

根据所述综合值将所述主播用户的主播信息推送至观众用户。

还提供了一种主播信息的推送装置，包括：

主播用户召回模块，设置为召回主播用户；

主播用户分类模块，设置为以直播历史与直播效率作为分类的维度、对所述主播用户划分主播类型；

交互值计算模块，设置为对所述主播用户计算交互值，其中，所述交互值用于表征观众用户观看主播用户直播间的特性；

综合值计算模块，设置为根据所述主播类型，使用所述交互值对所述主播用户计算综合值；

主播信息推送模块，设置为根据所述综合值将所述主播用户的主播信息推送至观众用户。

还提供了一种计算机设备，所述计算机设备包括：

一个或多个处理器；

存储器，设置为存储一个或多个程序；

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现上述的主播信息的推送方法。

还提供了一种计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时实现上述的主播信息的推送方法。

附图说明

图1为本发明实施例一提供的一种主播信息的推送方法的流程图；

图2为本发明实施例二提供的一种主播信息的推送方法的流程图；

图3为本发明实施例三提供的一种主播信息的推送方法的流程图；

图4为本发明实施例四提供的一种主播信息的推送方法的流程图；

图5为本发明实施例五提供的一种主播信息的推送装置的结构示意图；

图6为本发明实施例六提供的一种计算机设备的结构示意图。

具体实施方式

下面结合附图和实施例对本公开作说明。

实施例一

图 1 为本发明实施例一提供的一种主播信息的推送方法的流程图，本实施例可适用于训练分类模型的情况，该方法可以由主播信息的推送装置来执行，该主播信息的推送装置可以由软件和/或硬件实现，可配置在计算机设备中，例如，个人电脑、服务器、工作站，等等，该方法包括如下步骤：

S101、确定主播用户。

在本实施例中，可以从直播平台中注册的主播用户中，选择部分或全部主播用户训练分类模型。

主播用户为注册的用户，具有主持直播节目的权限，在计算机设备中，主播用户可以以用户标识进行表示，例如，用户标识号（Identifier，ID）、用户账号，等等。

S102、对主播用户计算效率值。

在本实施例中，可对每个主播用户计算效率值，该效率值用于表征直播效率，直播效率可以用于划分优质的主播用户与普通的主播用户。

在实现中，衡量直播效率的维度可以根据实际情况进行设置，例如，包括观众用户的留驻情况、主播用户的营收情况，等等。

在这些维度上，获取主播用户在指定的时间段进行直播时统计的第一指标，如表征观众用户的留驻情况的人均观看时长、人均关注率，如表征主播用户的营收情况的人均营收值，其中，人均营收值可以指观众用户人均赠送虚拟礼物、消费虚拟代币等的数值。

对第一指标配置第一指标权重，第一指标权重大于或等于 0，为可调参数，可以根据实际的需求进行相应的设置，例如，若倾向高营收的主播用户为优质的主播用户，则可以对人均营收值设置较大的第一指标权重。通过根据实际的需求配置第一指标权重，分类模型会适应性地动态调整，更可能把高营收的主播用户划分为优质的主播用户。

在本实施例中，可使用权重和的方式计算效率值，即计算第一指标与第一指标权重的乘积，作为第一单项效率，计算所有第一单项效率的和值，作为表征直播效率的效率值。

以人均观看时长 `avg_stay_time`，人均关注率 `avg_follow`，人均营收值 `avg_gift` 作为第一指标的示例，可以如下公式计算效率值 `effecitve_score`：

$$\text{effectitve_score} = x \times \text{avg_stay_time} + y \times \text{avg_follow} + z \times \text{avg_gift}$$

其中，x 为人均观看时长 avg_stay_time 的第一指标权重，y 为人均关注率 avg_follow 的第一指标权重，z 为人均营收值 avg_gift 的第一指标权重。

上述计算效率值的方式只是作为示例，在实施本实施例时，可以根据实际情况设置其他计算效率值的方式，例如，在上述计算效率值的方式的基础上，按照第一指标的时间对第一指标权重进行衰减，使得离当前时间越近的第一指标权重的值越大、离当前时间越远的第一指标权重的值越小，或者，第一指标增加表征观众用户活跃度的维度，如人均弹幕数量，等等，本实施例对此不加以限制。

S103、以首次开播时间与效率值作为分类的维度、对主播用户划分主播类型。

主播用户关联首次开播时间，即主播用户第一次主持直播节目的时间，该首次开播时间可用于表征直播历史，直播历史可以用于划分新的主播用户与老的主播用户。

除了首次开播时间之外，还可以采用其他方式表征直播历史，如主播用户的注册时间、主播用户在预设时间段内进行直播的频率，等等，本实施例对此不加以限制。

可对主播用户分析主播直播特征，即可体现在直播间中、主播用户的特征，这些主播直播特征存储在数据库中，在训练分类模型时，可从数据库中读取该主播直播特征。

该主播直播特征包括：

1、主播用户的离线特征

主播用户的离线特征为在非直播时间体现在主播用户主持直播节目时、主播用户的特征，例如，注册时间、平均活跃时间，历史的开播时间，关注人数，营收总值，等等。

2、主播用户的实时特征

主播用户的实时特征为在直播时间体现在主播用户主持直播节目时、主播用户的特征，例如，实时热度、性别、国家、语言、直播类型，等等。

在本实施例中，以首次开播时间与效率值共同作为分类的维度、对主播用户划分主播类型，共可以划分出四个类型，即第一类型、第二类型、第三类型与第四类型。

第一类型与第二类型中主播用户的首次开播时间均晚于第三类型与第四类

型中主播用户的首次开播时间。

第一类型中主播用户的效率值低于第二类型中主播用户的效率值。

第三类型中主播用户的效率值低于第四类型中主播用户的效率值。

第一类型的主播用户可以称之为普通的新主播用户、第二类型的主播用户可以称之为优质的新主播用户、第三类型的主播用户可以称之为普通的老主播用户、第四类型的主播用户可以称之为优质的老主播用户。

在实现中，可将主播用户划分至第一分组、第二分组，其中，第一分组中主播用户的首次开播时间晚于第二分组中主播用户的首次开播时间。

在一种划分方式中，可以预先设置第一阈值、第二阈值，该第一阈值属于 $(0, 1)$ 、该第二阈值属于 $(0, 1)$ 。

将第一分组中主播用户的主播类型划分为第一类型、第二类型，其中，第一类型中主播用户的效率值小于或等于第一阈值、第二类型中主播用户的效率值大于第一阈值，使得第一类型中主播用户的效率值小于第二类型中主播用户的效率值。

将第二分组中主播用户的主播类型划分为第三类型、第四类型，第三类型中主播用户的效率值小于或等于第二阈值、第四类型中主播用户的效率值大于第二阈值，使得第三类型中主播用户的效率值小于第四类型中主播用户的效率值。

在划分第一分组、第二分组时，可统计所有主播用户的首次开播时间的分位数（Quantile），取指定的分位数对应的首次开播时间，作为开播时间阈值，其中，指定的分位数属于 $(0, 1)$ 。

将首次开播时间小于或等于开播时间阈值的主播用户划分至第一分组中。

将首次开播时间大于开播时间阈值的主播用户划分至第二分组中。

S104、根据主播类型，对主播用户设置第一样本权重。

在本实施例中，对于第一类型、第二类型、第三类型与第四类型等主播类型，可以提高第二类型与第四类型中主播用户的第一样本权重，和/或，降低第一类型与第三类型中主播用户的第一样本权重，使得第二类型与第四类型中主播用户的第一样本权重均高于第一类型与第三类型中主播用户的第一样本权重，由于机器学习对第一样本权重高的样本更加敏感，有利于分类模型更加精确地挑选出主播用户中的优质的新主播用户、优质的老主播用户。

在一种设置第一样本权重的示例中，将第二类型中主播用户的效率值，赋值为第二类型中主播用户的第一样本权重。

将第四类型中主播用户的效率值，赋值为第四类型中主播用户的第一样本权重。

将第一类型中主播用户的第一样本权重设置为指定的第一数值（如 1），第一数值小于第一类型中主播用户的效率值。

将第三类型中主播用户的第一样本权重设置为指定的第一数值（如 1），第一数值小于第三类型中主播用户的效率值。

上述设置第一样本权重的方式只是作为示例，在实施本实施例时，可以根据实际情况设置其他设置第一样本权重的方式，本实施例对此不加以限制。

S105、以主播直播特征、第一样本权重与主播类型作为训练样本训练分类模型。

在本实施例中，在确定主播直播特征、第一样本权重与主播类型（即标签 Tag）时，将其作为训练样本训练分类模型，以使得分类模型可用于以直播历史与直播效率作为分类的维度对主播用户进行分类，获得主播类型。

分类模型可以看作是一种函数映射，而分类模型的训练过程是一个函数优化求解的过程。优化求解的目标就是不断更新该分类模型所包含的参数，将已标注的训练样本作为输入的数据，经过分类模型的计算，输出的预测值和标注之间的损失值最小。

由于主播类型的数量一般为两个以上，因此，分类模型一般为多分类模型（multi-class classification），如 softmax 函数、极端梯度提升（eXtreme Gradient Boosting, XGBoost）函数、随机森林（RandomForest）、深度神经网络，等等。

以 softmax 函数为例，softmax 函数将多个神经元的输出，映射到（0,1）的值，而这些值的累和为 1（满足概率的性质），可以看成概率来理解，从而来进行多分类，分类模型表示如下：

$$\text{logit}_i = \sum_{k=1}^n W_{ik} \times X_k + B_i$$

$$\text{score}_i = \frac{e^{\text{logit}_i}}{\sum_{j=1}^n e^{\text{logit}_j}}$$

其中，i 表示第 i 个主播用户，j 表示第 j 个主播用户，k 表示第 k 个主播直播特征，n 表示主播用户的总数量，W 表示第一参数，B 表示第二参数，X 表示主播直播特征，logit 表示第二目标值，score 表示主播用户属于一个主播类型的概率， $\text{score} \in (0, 1)$ 。

第一参数与第二参数属于分类模型中的参数，通过训练分类模型获得，一般取值范围在 $(-\infty, +\infty)$ ，即没有取值范围限定，是因为第一参数的值取决于

主播直播特征自身的分布情况，由于不同的特征提取方法、特征清洗方法等会导致同一个主播直播特征有不同的分布情况，因此第一参数的取值范围也不受限制，一般来说，第一参数的绝对值与主播直播特征的重要性正相关，即第一参数的绝对值越大，代表主播直播特征越重要。

实施例二

图 2 为本发明实施例二提供的一种主播信息的推送方法的流程图，本实施例可适用于训练排序模型的情况，该方法包括如下步骤：

S201、确定主播用户。

S202、获取主播用户关联的主播直播特征、观众用户关联的观众直播特征、该主播用户与该观众用户之间的交互直播特征。

在本实施例中，可对主播用户分析主播直播特征、观众直播特征、与观众用户之间的交互直播特征，主播直播特征、观众直播特征、交互直播特征可存储在数据库中，在训练排序模型时，从数据库提取主播直播特征、观众直播特征、交互直播特征。

观众直播特征为可体现在直播间中、观众用户的特征。

该观众直播特征包括：

1、观众用户的离线特征

观众用户的离线特征为在非直播时间体现在直播间中、观众用户的特征，例如，平均观看时间、历史送礼总数，等等。

2、观众用户的实时特征

观众用户的实时特征为在直播时间体现在直播间中、观众用户的特征，例如，在线时长、性别、国家、语言，等等。

主播用户与观众用户之间的交互直播特征为可体现在主播用户主持直播节目时、该主播用户与观众用户互动的特征，例如，观众用户对主播用户历史观看时长、观众用户历史赠送给主播用户的虚拟礼物总数、观众用户是否关注主播用户，等等。

S203、确定主播用户的交互值。

在本实施例中，可对每个主播用户计算交互值，该交互值用于表征观众用户观看主播用户直播间的特性。

在实现中，可以从主播用户与观众用户交互的维度上，获取主播用户在指

定的时间段进行直播时统计的第二指标，如人均观看时长、人均关注率，人均营收值，等等。

对第二指标配置第二指标权重，第二指标权重大于或等于 0，为可调参数，可以根据实际的需求进行相应的设置，例如，若倾向高营收的主播用户为优质的主播用户，则可以对人均营收值设置较大的第二指标权重。通过根据实际的需求配置第二指标权重，排序模型会适应性地动态调整，更可能把推送高营收的主播用户给观众用户。

在本实施例中，可使用权重和的方式计算交互值，即计算第二指标与第二指标权重的乘积，作为第二单项效率，计算所有第二单项效率的和值，作为主播用户的交互值。

以人均观看时长 avg_stay_time ，人均关注率 avg_follow ，人均营收值 avg_gift 作为第二指标的示例，可以如下公式计算交互值 $interaction_score$ ：

$$interaction_score = a \times avg_stay_time + b \times avg_follow + c \times avg_gift$$

其中， a 为人均观看时长 avg_stay_time 的第二指标权重， b 为人均关注率 avg_follow 的第二指标权重， c 为人均营收值 avg_gift 的第二指标权重。

上述计算交互值的方式只是作为示例，在实施本实施例时，可以根据实际情况设置其他计算交互值的方式，例如，在上述计算交互值的方式的基础上，按照第二指标的时间对第二指标权重进行衰减，使得离当前时间越近的第二指标权重的值越大、离当前时间越远的第二指标权重的值越小，或者，第二指标增加表征观众用户活跃度的维度，如人均弹幕数量，等等，本实施例对此不加限制。

S204、根据交互值，对主播用户设置样本类别。

在实现中，可以将交互值与预设的交互阈值（如 0）进行比较。

若交互值大于交互阈值，则对主播用户设置样本类别为第一类别（如 1）。

若交互值小于或等于交互阈值，则对主播用户设置样本类别为第二类别（如 0）。

S205、按照样本类别对主播用户设置第二样本权重。

在本实施例中，对于第一类别、第二类别的样本类别，可以提高第一类别中主播用户的第二样本权重，和/或，降低第二类别中主播用户的第二样本权重，使得第一类别中主播用户的第二样本权重高于第二类别中主播用户的第二样本权重，由于机器学习对第二样本权重高的样本更加敏感，有利于排序模型更加精确地提高优质的主播用户的排序、降低普通的主播用户的排序。

在一种设置第二样本权重的示例中，将第一类别中主播用户的交互值，赋值为第一类别中主播用户的第二样本权重。

将第二类别中主播用户的第二样本权重设置为指定的第二数值（如 1），第二数值小于第一类别中主播用户的交互值。

上述设置第二样本权重的方式只是作为示例，在实施本实施例时，可以根据实际情况设置其他设置第二样本权重的方式，本实施例对此不加以限制。

S206、以主播直播特征、观众直播特征、交互直播特征、第二样本权重与样本类别作为训练样本，训练排序模型。

在本实施例中，在主播直播特征、观众直播特征、交互直播特征、第二样本权重与样本类别（即标签 Tag）时，将其作为训练样本训练排序模型，以使得排序模型可用于对主播用户计算交互值。

排序模型可以看作是一种函数映射，而排序模型的训练过程是一个函数优化求解的过程。优化求解的目标就是不断更新该排序模型所包含的参数，将已标注的训练样本作为输入的数据，经过排序模型的计算，输出的预测值和标注之间的损失值最小。

由于样本类别的数量一般为两个，因此，排序模型一般为单分类模型（one-class classification），又称二分类模型，单分类模型是一类特殊的分类任务，在该任务中，大多数训练样本为正样本（positive），而其他样本则笼统的划为另一类，如逻辑回归（logistics regression, LR）函数，支持向量机（Support Vector Machine, SVM），卷积神经网络（Convolutional Neural Network, CNN）中的 Lenet、Alexnet、视觉几何组网络（Visual Geometry Group net, VGGnet），孤立森林（Isolation Forest），等等。

以 LR 函数为例，LR 函数通过输入未知类别对象的属性特征序列得到对象所处的类别。由于 LR 函数是一个概率分布函数，因此对于单分类而言，离中心点的距离越远，其属于一类别的可能性就越大，排序模型表示如下：

$$\text{logit}_i = \sum_{i=1}^n Y_i \times Z_i + C_i$$

$$\text{rank_score} = \frac{1}{1 + e^{\text{logit}_i}}$$

其中，i 表示第 i 个主播用户，n 表示主播用户的总数量，Y 表示第三参数，C 表示第四参数，Z 表示主播直播特征、观众直播特征与交互直播特征，logit 表示第四目标值，rank_score 表示主播用户的交互值， $\text{rank_score} \in (0, 1)$ 。

第三参数与第四参数属于排序模型中的参数，通过训练排序模型获得，一

般取值范围在 $(-\infty, +\infty)$ ，即没有取值范围限定，是因为第三参数的值取决于主播直播特征自身的分布情况，由于不同的特征提取方法、特征清洗方法等会导致同一个主播直播特征、观众直播特征、交互直播特征有不同的分布情况，因此第三参数的取值范围也不受限制，一般来说，第三参数的绝对值与主播直播特征、观众直播特征、交互直播特征的重要性正相关，即第三参数的绝对值越大，代表主播直播特征、观众直播特征、交互直播特征越重要。

实施例三

图3为本发明实施例三提供的一种主播信息的推送方法的流程图，本实施例可适用于对主播用户分类计算交互值、并进行推送的情况，该方法包括如下步骤：

S301、召回主播用户。

在直播平台中，注册有大量的主播用户，形成主播池，在待推送主播用户给观众用户时，可从按照不同的业务需求（如召回优质的主播用户、召回符合主播用户个性化需求的主播用户等），使用不同的召回策略从主播池中召回部分主播用户，形成候选池。

在实现中，召回策略包括但不限于：

在线召回（召回在线的主播用户）、订阅召回（召回观众用户订阅的栏目（如一款游戏、餐饮等）的主播用户）、同国家召回（召回与观众用户所属国家相同的主播用户）、同语言召回（召回与观众用户使用的语言相同的主播用户）、协同过滤召回（使用协同过滤算法召回主播用户）、偏好召回（召回与观众用户的偏好相同的主播用户）、相似召回（召回与已召回的主播用户相似的其他主播用户）。

由于使用不同的召回策略可能召回相同的主播用户，因此，在得到召回的主播用户之后，一般会对召回的主播用户进行去重。

S302、以直播历史与直播效率作为分类的维度、对主播用户划分主播类型。

在本实施例中，以直播历史作为分类的维度对主播用户进行分类，可划分新的主播用户与老的主播用户，以直播效率作为分类的维度对主播用户进行分类，可划分优质的主播用户与普通的主播用户。

因此，对主播用户划分的主播类型包括第一类型、第二类型、第三类型与第四类型。

第一类型与第二类型中主播用户的直播历史均少于第三类型与第四类型中

主播用户的直播历史。

第一类型中主播用户的直播效率低于第二类型中主播用户的直播效率。

第三类型中主播用户的直播效率低于第四类型中主播用户的直播效率。

第一类型的主播用户可以称之为普通的新主播用户、第二类型的主播用户可以称之为优质的新主播用户、第三类型的主播用户可以称之为普通的老主播用户、第四类型的主播用户可以称之为优质的老主播用户。

S303、对主播用户计算交互值。

在本实施例中，可对召回的主播用户计算交互值，该交互值用于表征观众用户观看主播用户直播间的特性。

在直播的场景下，主播用户与观众用户之间的信息是特征的主要组成部分，正向的交互往往能带来比较高的推荐。

S304、根据主播类型，使用交互值对主播用户计算综合值。

在本实施例中，在交互值的基础上，针对不同主播类型，可以适应性进行调整，生成综合值，从而按照业务需求对主播用户的排序进行调整。

S305、根据综合值将主播用户的主播信息推送至观众用户。

由于召回的主播用户数量较多，而展示的位置有限，因此，可以参考综合值对从召回的主播用户中选择部分主播用户，将其主播信息（如主播名称、直播间的缩略图，等等）推送给观众用户浏览。

所谓推送给观众用户，是推送给表征观众用户的数据所处的客户端，如浏览器、直播应用、即时通讯工具，等等。

若该观众用户为非注册用户，则可以以设备标识等设备数据表征，如国际移动设备识别码（International Mobile Equipment Identity, IMEI），非注册用户的特征稀疏，在 S301 中可召回非个性化的主播用户，从而推送非个性化的主播用户的主播信息。

若该观众用户为注册用户，则可以以用户 ID、用户账号等用户数据表征，注册用户的特征较为丰富，在 S301 中可召回个性化的主播用户，从而推送个性化的主播用户的主播信息。

在一种选择主播用户的方式中，可以按照综合值对召回的主播用户进行降序排序，选择排序最高的 n （ n 为正整数，属于可调参数，如 20）个主播用户进行推送。

除了选择综合值最高的 n 个主播用户之外，还可以采用其他方式选择主播

用户，例如，选择综合值大于或等于预设的分数阈值的主播用户进行推送，滤除综合值低于预设的分数阈值的主播用户、从保留的主播用户中选择综合值最高的 n 个主播用户进行推送，等等，本实施例对此不加以限制。

在本实施例中，召回主播用户，以直播历史与直播效率作为分类的维度、对主播用户划分主播类型，对主播用户计算交互值，该交互值用于表征观众用户观看主播用户直播间的特性，根据主播类型，使用交互值对主播用户计算综合值，根据综合值将主播用户的主播信息推送至观众用户，通过直播历史可以划分新的主播用户与老的主播用户，通过直播效率可以划分优质的主播用户与普通的主播用户，针对不同主播类型的主播用户，可以适应性地调整交互值，从而对推送机会进行调整，在交互值满足主播用户的质量需求的前提下，可通过调整满足不同的业务需求，使得新的主播用户在较少特征的情况下可以获得较多的推送机会，一方面，提高新的主播用户的积极性，保证整个直播生态的健康发展，另一方面，提高推送的主播用户的多样性，减少重复推送主播用户、主播用户特性相近的情况，提高观众用户感兴趣的主播用户的推送精确度，使得观众用户在推送之后可以直接浏览感兴趣的主播用户主持的直播节目，减少观众用户使用关键词搜索主播用户并进行筛选，从而减少对筛选、推送主播用户所耗费的资源造成的浪费，提高观众用户观看直播的操作简便性。

实施例四

图 4 为本发明实施例四提供的一种主播信息的推送方法的流程图，本实施例以前述实施例为基础，增加使用分类模型对主播用户进行分类、使用排序模型对主播用户计算交互值、对四种主播类型的主播用户调整综合值的处理操作，该方法包括如下步骤：

S401、召回主播用户。

S402、确定分类模型。

在本实施例中，可以预先离线训练分类模型，并存储该分类模型及其参数，分类模型可用于以直播历史与直播效率作为分类的维度对主播用户进行分类。

该分类模型的训练方式与实施例一基本相似，所以描述的比较简单，相关之处参见实施例一的部分说明即可，本实施例在此不加以详述。

在线对主播用户进行分类时，可加载该分类模型及其参数。

S403、获取主播用户关联的主播直播特征。

在本实施例中，可对主播用户分析主播直播特征，即可体现在直播间中、

主播用户的特征，这些主播直播特征存储在数据库中，在使用分类模型对主播用户分类时，可从数据库中可读取该主播直播特征。

该主播直播特征包括：

1、主播用户的离线特征；2、主播用户的实时特征。

S404、将主播直播特征输入至分类模型中进行处理，以输出主播用户属于每个主播类型的概率。

在本实施例中，由于主播类型的数量一般为两个以上，因此，分类模型一般为多分类模型，如 softmax 函数、XGBoost 函数、随机森林、深度神经网络，等等。

以 softmax 函数为例，softmax 函数将多个神经元的输出，映射到 (0,1) 的值，而这些值的累和为 1（满足概率的性质），可以看成概率来理解，从而来进行多分类，分类模型表示如下：

$$\text{logit}_i = \sum_{k=1}^n W_{ik} \times X_k + B_i$$

$$\text{score}_i = \frac{e^{\text{logit}_i}}{\sum_{j=1}^n e^{\text{logit}_j}}$$

其中，i 表示第 i 个主播用户，j 表示第 j 个主播用户，k 表示第 k 个主播直播特征，n 表示主播用户的总数量，W 表示第一参数，B 表示第二参数，X 表示主播直播特征，logit 表示第二目标值，score 表示主播用户属于一个主播类型的概率， $\text{score} \in (0, 1)$ 。

在本示例中，可确定主播直播特征对应的第一参数、第二参数，计算主播直播特征与第一参数的乘积之和，获得第一目标值，在第一目标值的基础上加上第二参数，作为第二目标值。

使用 softmax 函数对第二目标值进行分类，获得主播用户属于每个主播类型的概率。

在使用其他分类模型的情况下，可以使用其他方式计算主播用户属于每个主播类型的概率，本实施例对此不加以限制。

S405、按照概率确定主播用户归属的主播类型。

一实施例中，选择概率值最大的主播类型，作为主播用户归属的主播类型。

例如，在主播类型为第一类型、第二类型、第三类型、第四类型的情况下，把分类模型输出 4 个概率值，用向量表达为 [score_1, score_2, score_3, score_4]，其中，score_1 表示主播用户归属第一类型的概率，score_2 表示主播用户归属第

二类型的概率，score_3 表示主播用户归属第三类型的概率，score_4 表示主播用户归属第四类型的概率，并且，score_1+score_2+score_3+score_4=1，假设将一个主播用户的主播直播特征输入分类模型，输出的概率值为[0.1,0.3,0.15,0.45]，则可以确定该主播用户归属第四类型。

S406、确定排序模型。

在本实施例中，可以预先离线训练排序模型，并存储该排序模型及其参数，排序模型用于对主播用户计算交互值。

该排序模型的训练方式与实施例二基本相似，所以描述的比较简单，相关之处参见实施例二的部分说明即可，本实施例在此不加以详述。

在线对主播用户计算交互值时，可加载该排序模型及其参数。

S407、获取主播用户关联的主播直播特征、观众用户关联的观众直播特征，主播用户与观众用户之间的交互直播特征。

在本实施例中，可对主播用户分析主播直播特征、观众直播特征、与观众用户之间的交互直播特征，主播直播特征、观众直播特征、交互直播特征可存储在数据库中，在使用排序模型对主播用户计算交互值时，从数据库提取主播直播特征、观众直播特征、交互直播特征。

观众直播特征为可体现在直播间中、观众用户的特征。

该观众直播特征包括：

- 1、观众用户的离线特征
- 2、观众用户的实时特征

主播用户与观众用户之间的交互直播特征为可体现在主播用户主持直播节目时、该主播用户与观众用户互动的特征。

S408、将主播直播特征、观众直播特征与交互直播特征输入至排序模型中进行处理，以输出交互值。

在本实施例中，排序模型一般为单分类模型（one-class classification），又称二分类模型，如 LR 函数，SVM，CNN 中的 Lenet、Alexnet、VGGnet，孤立森林，等等。

以 LR 函数为例，LR 函数通过输入未知类别对象的属性特征序列得到对象所处的类别。由于 LR 函数是一个概率分布函数，因此对于单分类而言，离中心点的距离越远，其属于一类的可能性就越大，排序模型表示如下：

$$\text{logit}_i = \sum_{i=1}^n Y_i \times Z_i + C_i$$

$$\text{rank_score} = \frac{1}{1 + e^{\text{logit}_i}}$$

其中, i 表示第 i 个主播用户, n 表示主播用户的总数量, Y 表示第三参数, C 表示第四参数, Z 表示主播直播特征、观众直播特征与交互直播特征, logit 表示第四目标值, rank_score 表示主播用户的交互值, $\text{rank_score} \in (0, 1)$ 。

因此, 在本示例中, 可确定主播直播特征、观众直播特征与交互直播特征分别对应的第三参数、第四参数, 分别计算主播直播特征、观众直播特征、交互直播特征与第三参数的乘积之和, 获得第三目标值, 在第三目标值的基础上加上第四参数, 作为第四目标值。

使用逻辑回归函数对第四目标值进行分类, 获得交互值。

在使用其他排序模型的情况下, 可以使用其他方式计算主播用户的交互值, 本实施例对此不加以限制。

S409、若主播类型为第二类型, 则对主播用户计算调整值。

S410、在交互值的基础上加上调整值, 作为主播用户的综合值。

若一个主播用户的主播类型为第二类型, 即为优质的新主播用户, 新的主播用户由于缺乏与观众用户之间的交互, 导致其交互值一般比老的主播用户要低, 从而在传统的流量分配过程中得到较少流量, 从而抑制其活跃性。

在本实施例中, 对于优质的新主播用户, 适当地对其进行调整 (boost), 提高其排序, 但同时, 新的主播用户直播效率普遍低于老的主播用户, 过多的 boost 也会带来负向效果, 因此, 会对 boost 有所限制。

对于优质的新主播用户的期望是, 通过初期的 boost 后, 优质的新主播用户在成为老的主播用户后凭借其优秀的能力, 获得较高的交互值, 从而成为优质的老主播用户, 不需要 boost 便能得到更多的流量, 提高整个直播生态的主播质量。

在实现中, 可在每次 boost 时, 对主播用户增加调整值, 该调整值与计算调整值的次数负相关, 即与 boost 的次数负相关, 计算调整值 (boost) 的次数越少, 调整值越大, 反之, 计算调整值 (boost) 的次数越多, 调整值越小, 直至为 0。

在一个计算调整值的方式中, 一方面, 可确定基础值, 该基础值小于处于指定排位的主播用户的交互值, 从而控制对优质的新主播用户 boost 时, 对优质的老主播用户的冲击, 控制可能带来的负向影响。

另一方面, 可确定调整系数, 调整系数与计算调整值的次数负相关, 即与 boost 的次数负相关, 计算调整值 (boost) 的次数越少, 调整系数越大, 反之,

计算调整值 (boost) 的次数越多, 调整系数越小, 直至为 0。

计算基础分数与调整系数之间的乘积, 作为调整值。

可通过如下公式实现上述计算调整值的方式:

$$\text{final_score} = \text{rank_score} + \text{boost_score}$$

$$\text{boost_score} = \text{base_score} \times \left(1 - \frac{\text{process}}{\text{goal}}\right)$$

$$\text{base_score} = \max(\text{rank_score}_{\text{mid}} - \text{rank_score}, 0)$$

其中, final_score 为主播用户的综合值, rank_score 为该主播用户的交互值, boost_score 为该主播用户的调整值, base_score 为该主播用户的基础值, $\left(1 - \frac{\text{process}}{\text{goal}}\right)$ 为该主播用户的调整系数, process 为计算调整值 (boost) 的次数, goal 为计算调整值 (boost) 的上限值, rank_score_{mid} 为所有主播用户的交互值中的中间值。

S411、若主播类型为第四类型, 则确定主播用户的效率值。

S412、对效率值调整比重。

S413、计算交互值与调整比重之后的效率值之间的乘积, 作为主播用户的综合值。

若一个主播用户的主播类型为第四类型, 即为优质的老主播用户, 优秀的效率值使得其在流量分配过程中对于其他主播用户更应该得到流量, 因此, 对优质的老主播用户也可以进行 boost。

若主播用户的效率值在训练分类模型时已经计算, 则可以直接提取, 若主播用户的效率值在训练分类模型时未计算, 则可以按照参照实施例一的方式进行计算, 本实施例在此不加以详述。

在本实施例中, 通过调整比重可对普通的老主播用户与优质的老主播用户之间进行衡量, 通过调整比重可以动态控制优质的老主播用户的综合值。

可通过如下公式实现上述计算调整值的方式:

$$\text{final_score} = (1 + (n \times \text{effective_score})) \times \text{rank_score}$$

其中, final_score 为主播用户的综合值, effective_score 为该主播用户的效率值, rank_score 为该主播用户的交互值, $n \in (0, 1)$, 用于衡量普通的老主播用户与优质的老主播用户之间的比重。

S414、若主播类型为第一类型或第三类型, 则将交互值设置为主播用户的

综合值。

若一个主播用户的主播类型为第一类型或第三类型，即为普通的新主播用户或普通的老主播用户，对于这两个主播类型的主播用户，可以不进行 boost，按照其交互值作为其综合值。

上述计算综合值的方式只是作为示例，在实施本实施例时，可以根据实际情况设置其他设置计算综合值的方式，本实施例对此不加以限制。

S415、根据综合值将主播用户的主播信息推送至观众用户。

实施例五

图 5 为本发明实施例五提供的一种主播信息的推送装置的结构示意图，该装置可以包括如下模块：

主播用户召回模块 501，设置为召回主播用户；主播用户分类模块 502，设置为以直播历史与直播效率作为分类的维度、对所述主播用户划分主播类型；交互值计算模块 503，设置为对所述主播用户计算交互值，所述交互值用于表征观众用户观看所述主播用户直播间的特性；综合值计算模块 504，设置为根据所述主播类型，使用所述交互值对所述主播用户计算综合值；主播信息推送模块 505，设置为根据所述综合值将所述主播用户的主播信息推送至观众用户。

实施例六

图 6 为本发明实施例六提供的一种计算机设备的结构示意图。如图 6 所示，该计算机设备包括处理器 600、存储器 601、通信模块 602、输入装置 603 和输出装置 604；计算机设备中处理器 600 的数量可以是一个或多个，图 6 中以一个处理器 600 为例；计算机设备中的处理器 600、存储器 601、通信模块 602、输入装置 603 和输出装置 604 可以通过总线或其他方式连接，图 6 中以通过总线连接为例。

存储器 601 作为一种计算机可读存储介质，可用于存储软件程序、计算机可执行程序以及模块，如本实施例中的主播信息的推送方法对应的模块（例如，如图 5 所示的主播信息的推送装置中的主播用户召回模块 501、主播用户分类模块 502、交互值计算模块 503、综合值计算模块 504 与主播信息推送模块 505）。处理器 600 通过运行存储在存储器 601 中的软件程序、指令以及模块，从而执行计算机设备的多种功能应用以及数据处理，即实现上述的主播信息的推送方法。

处理器 600 通过运行存储在存储器 601 中的软件程序、指令以及模块，从而执行设备的多种功能应用以及数据处理，即实现上述的主播信息的推送方法。

实施例七

本发明实施例七还提供一种计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时实现一种主播信息的推送方法，该方法包括：

召回主播用户；以直播历史与直播效率作为分类的维度、对所述主播用户划分主播类型；对所述主播用户计算交互值，所述交互值用于表征观众用户观看所述主播用户直播间的特性；根据所述主播类型，使用所述交互值对所述主播用户计算综合值；根据所述综合值将所述主播用户的主播信息推送至观众用户。