

一种基于多分类器的推荐方法、装置及电子设备

技术领域

[01] 本说明书涉及软件技术领域，特别涉及一种基于多分类器的推荐方法、装置及电子设备。

5 背景技术

[02] 近年来，随着电子商务、社交网络和移动网络的发展，网络信息数据日益增多，大数据（Big Data）作为一张新兴数据概念而被定义。与此同时，数据的爆炸式增长使得我们进入了大规模数据缝隙的时代。在海量数据环境下，为了快速有效、有针对性地提的利用已有信息，分类器在数据推荐方面得到了快速的发展。

- 10 [03] 传统分类方法大多数是假设数据是线性可分，采用一个分类器对所有的数据进行分类预测。随着大数据时代来临，数据分类要求越来越高，单一的、模糊化的数据分类已无法满足现在的需求，亟需一种新的分类推荐方法，来提数据分类预测的准确性。

发明内容

- 15 [04] 本说明书实施例提供一种基于多分类器的推荐方法、装置及电子设备，用于实现对数据的多元化分类预测，提高数据分类预测的准确性。

[05] 第一方面，本说明书实施例提供一种基于多分类器的推荐方法，包括：

[06] 根据预设的多分类器的主分类器对接收到的特征数据进行预测，获得第一预测数据；

[07] 针对所述特征数据，在预设的多分类器的 k 个子分类器中查找获得目标子分类器，所述目标子分类器的数据分布与所述特征数据之间的相似度最大， $k \geq 2$ ；

- 20 [08] 根据所述目标子分类器对所述特征数据进行预测，获得第二预测数据；

[09] 根据预设算法对所述第一预测数据和所述第二预测数据进行综合处理，获得第三预测数据；

[10] 根据所述第三预测数据对所述特征数据对应的待推荐实体进行推荐。

- 25 [11] 可选地，针对所述特征数据，在预设的多分类器的 k 个子分类器中查找获得目标子分类器，包括：

[12]如果所述 k 个子分类器中每个子分类器均具有数据分布的中心点，查找获取中心点与所述特征数据的数据向量距离最近的子分类器作为所述目标子分类器。

[13]可选地，针对所述特征数据，在预设的多分类器的 k 个子分类器中查找获得目标子分类器，包括：

- 5 [14]如果所述 k 个子分类器中每个子分类器均为具有数据分布参数的高斯混合模型，查找获取所述 k 个子分类器中数据分布参数与所述特征数据的相似度最大的子分类器作为所述目标子分类器。

[15]可选地，在根据所述第三预测数据对所述特征数据对应的待推荐实体进行推荐之后，所述方法还包括：

- 10 [16]检测用户是否选择所述特征数据对应的被推荐实体；

[17]根据用户对被推荐实体的选择与否，更新所述特征数据对应的目标子分类器和主分类器的参数。

[18]可选地，根据用户对被推荐实体的选择与否，更新所述特征数据对应的目标子分类器和主分类器的参数，包括：

- 15 [19]如果用户选择了被推荐实体对所述多分类器进行正反馈，如果用户未选择被推荐实体对所述多分类器进行负反馈；根据所述正反馈的或所述负反馈更新所述特征数据对应的目标子分类器和主分类器的参数。

[20]可选地，在根据所述目标子分类器对所述特征数据进行预测，获得第二预测数据之后，所述方法还包括：

- 20 [21]根据所述特征数据的数据向量更新所述目标子分类器的中心点。

[22]可选地，根据所述特征数据的数据向量更新所述目标子分类器的中心点，包括：

$$P_{t+1} = P_t + \frac{x_t - P_t}{n_i}, n_{i+1} = n_i + 1$$

[23]其中， P_{t+1} 表示所述目标子分类器更新后的中心点， P_t 表示所述目标子分类器更新前的中心点， x_t 表示所述数据向量， n_i 表示更新前被分配到所述目标子分类器的数据数量，

- 25 n_{i+1} 表示所述目标子分类器更新后的数据数量。

[24]第二方面，本说明书实施例提供一种基于多分类器的推荐装置，包括：

[25]第一分类单元，用于根据预设的多分类器的主分类器对接收到的特征数据进行预测，

获得第一预测数据;

[26]查找单元,用于针对所述特征数据,在预设的多分类器的 k 个子分类器中查找获得目标子分类器,所述目标子分类器的数据分布与所述特征数据之间的相似度最大, $k \geq 2$;

5 [27]第二分类单元,用于根据所述目标子分类器对所述特征数据进行预测,获得第二预测数据;

[28]综合处理单元,用于根据预设算法对所述第一预测数据和所述第二预测数据进行综合处理,获得第三预测数据;

[29]推荐单元,用于根据所述第三预测数据对所述特征数据对应的待推荐实体进行推荐。

[30]可选地,所述查找单元用于:

10 [31]如果所述 k 个子分类器中每个子分类器均具有数据分布的中心点,查找获取中心点与所述特征数据的数据向量距离最近的子分类器作为所述目标子分类器。

[32]可选地,所述查找单元用于:

[33]如果所述 k 个子分类器中每个子分类器均为具有数据分布参数的高斯混合模型,查找获取所述 k 个子分类器中数据分布参数与所述特征数据的相似度最大的子分类器作为
15 所述目标子分类器。

[34]可选地,所述装置还包括:

[35]更新单元,用于在根据所述第三预测数据对所述特征数据对应的待推荐实体进行推荐之后,检测用户是否选择所述特征数据对应的被推荐实体;根据用户对被推荐实体的选择与否,更新所述特征数据对应的目标子分类器和主分类器的参数。

20 [36]可选地,所述更新单元用于:

[37]如果用户选择了被推荐实体对所述多分类器进行正反馈,如果用户未选择被推荐实体对所述多分类器进行负反馈;根据所述正反馈的或所述负反馈更新所述特征数据对应的目标子分类器和主分类器的参数。

[38]可选地,所述装置还包括:

25 [39]更新单元,用于在根据所述目标子分类器对所述特征数据进行预测,获得第二预测数据之后,根据所述特征数据的数据向量更新所述目标子分类器的中心点。

[40]可选地,所述更新单元还用于根据如下公式更新目标子分类器的中心点:

$$P_{t+1} = P_t + \frac{x_t - P_t}{n_i}, \quad n_{i+1} = n_i + 1$$

- [41] 其中, P_{t+1} 表示所述目标子分类器更新后的中心点, P_t 表示所述目标子分类器更新前的中心点, x_t 表示所述数据向量, n_i 表示更新前被分配到所述目标子分类器的数据数量, n_{i+1} 表示所述目标子分类器更新后的数据数量。
- 5 [42] 第三方面, 本说明书实施例提供一种计算机可读存储介质, 其上存储有计算机程序, 该程序被处理器执行时实现以下步骤:
- [43] 根据预设的多分类器的主分类器对接收到的特征数据进行预测, 获得第一预测数据;
- [44] 针对所述特征数据, 在预设的多分类器的 k 个子分类器中查找获得目标子分类器, 所述目标子分类器的数据分布与所述特征数据之间的相似度最大, $k \geq 2$;
- 10 [45] 根据所述目标子分类器对所述特征数据进行预测, 获得第二预测数据;
- [46] 根据预设算法对所述第一预测数据和所述第二预测数据进行综合处理, 获得第三预测数据;
- [47] 根据所述第三预测数据对所述特征数据对应的待推荐实体进行推荐。
- [48] 第四方面, 本说明书实施例提供一种电子设备, 包括有存储器, 以及一个或者一个
- 15 以上的程序, 其中一个或者一个以上程序存储于存储器中, 且经配置以由一个或者一个以上处理器执行所述一个或者一个以上程序包含用于进行以下操作的指令:
- [49] 根据预设的多分类器的主分类器对接收到的特征数据进行预测, 获得第一预测数据;
- [50] 针对所述特征数据, 在预设的多分类器的 k 个子分类器中查找获得目标子分类器, 所述目标子分类器的数据分布与所述特征数据之间的相似度最大, $k \geq 2$;
- 20 [51] 根据所述目标子分类器对所述特征数据进行预测, 获得第二预测数据;
- [52] 根据预设算法对所述第一预测数据和所述第二预测数据进行综合处理, 获得第三预测数据;
- [53] 根据所述第三预测数据对所述特征数据对应的待推荐实体进行推荐。
- [54] 本说明书实施例中的上述一个或多个技术方案, 至少具有如下技术效果:
- 25 [55] 本说明书实施例提供一种基于多分类器的推荐方法, 通过采用多分类器的主分类器和目标子分类器来对特征数据进行分类, 实现了对数据的多元化分类预测, 提高了数据分类预测的准确性; 并且, 目标子分类器选择多分类器中数据分布与数据特征相似度最

大的子分类器,获得主分类器对特征数据进行预测的第一预测数据、目标子分类器对特征数据进行分类的第二预测数据,综合第一预测数据和第二预测数据获得第三预测数据,从而获得更为准确的预测数据,解决了现有技术中数据分类预测准确性较低的技术问题;进一步的,根据第三预测数据对特征数据进行推荐处理,从而提高了数据推荐的准确性。

5 附图说明

[56]为了更清楚地说明本说明书实施例中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍,显而易见地,下面描述中的附图是本说明书的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

10 [57]图 1 为本说明书实施例提供一种基于多分类器的推荐方法的流程图;

[58]图 2 为本说明书实施例提供一种基于多分类器的推荐装置的示意图;

[59]图 3 为本说明书实施例提供一种电子设备的示意图。

具体实施方式

15 [60]为使本说明书实施例的目的、技术方案和优点更加清楚,下面将结合本说明书实施例中的附图,对本说明书实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本说明书一部分实施例,而不是全部的实施例。基于本说明书中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本说明书保护的范围。

20 [61]在本说明书实施例提供一种基于多分类器的推荐方法、装置及电子设备,用于实现对数据的多元化分类预测,提高数据分类预测的准确性。

[62]下面结合附图对本说明书实施例技术方案的主要实现原理、具体实施方式及其对应能够达到的有益效果进行详细的阐述。

[63]本说明书实施提供一种预设的多分类器,包含一个主分类器和 k 个子分类器, $k \geq 2$ 。基于预设的多分类器,本申请说明书实施例还提供一种基于多分类器的推荐方法,如图 25 1 所示,所示方法包括:

[64]S101: 根据预设的多分类器的主分类器对接收到的特征数据进行预测,获得第一预测数据;

[65]S102: 针对所述特征数据,在预设的多分类器的 k 个子分类器中获得目标子分类器,所述目标子分类器的数据分布与所述特征数据之间的相似度最大, $k \geq 2$;

[66]S103: 根据所述目标子分类器对所述特征数据进行预测, 获得第二预测数据;

5 [67]S104: 根据预设算法对所述第一预测数据和所述第二预测数据进行综合处理, 获得第三预测数据;

[68]S105: 根据所述第三预测数据对所述特征数据对应的待推荐实体进行推荐。

[69]具体实施过程中, 预设的多分类器可以对待推荐实体的特征进行分类预测, 预测待推荐实体属于哪一类实体。预设的多分类器也可以对用户特征和待推荐实体的特征共同组成的特征进行分类预测, 预测待推荐实体在该用户侧的得分, 按照得分的高低选择是
10 否向该用户推荐该实体。为此, 多分类器接收的特征数据可以是待推荐实体的特征数据, 也可以是待推荐实体特征和用户特征的组合形成的特征数据。

[70]预设的多分类器在接收到特征数据之后执行 S101 和 S102, 其中, S101 和 S102 的执行顺序不分先后, 可以同时执行, 也可以先执行 S101 或 S102。

[71]S101 根据主分类器对接收到的特征数据进行预测获得第一预测数据, 具体可以将特征数据输入主分类器, 由主分类器对特征数据进行打分获得第一得分即第一预测数据。
15

[72]S102 针对接收到的特征数据查找目标子分类器, 从多分类器的 k 个子分类器中查找数据分布与特征数据的相似度最大的子分类器作为目标子分类器。子分类器的数据分布是指该子分类器接收并分类的数据的分布, 或者, 训练形成该子分类器的样本数据的分布。接收到的数据特征与子分类器的数据分布越相似, 该子分类器对该数据特征进行分类的准确性越高。
20

[73]其中, S102 可以通过如下任一方式查找获得目标子分类器:

[74]方式一、如果多分类器的 k 个子分类器中每个子分类器均具有数据分布的中心点, 查找获取中心点与特征数据的数据向量距离最近的子分类器作为目标子分类器。每一个子分类器均具有一个数据分布的中心点, 特征数据的数据向量与哪一个中心点最近, 则
25 表明该特征数据与对应的子分类器的数据分布的相似度最大, 将该子分类器作为目标子分类器。

[75]具体实施过程中, 本说明书实施可以为任一子分类器建立并更新数据分布的中心点。首先, 对 k 个子分类器进行初始化, 将其数据分布的中心点初始化为空。然后, 通过多

分类器接收待分类的特征数据，在多分类器接收的特征数据的数据量小于设定阈值时，随机选取一个子分类器作为目标子分类器，目标子分类器对该特征数据进行分类预测后，根据该特征数据的数据向量更新目标子分类器的中心点：

5 [76]若目标子分类器的中心点为空，将该特征数据的数据向量作为该目标子分类器的中心点；

[77]若目标子分类器的中心点不为空，根据如下公式 1 更新目标子分类器的中心点：

$$P_{t+1} = P_t + \frac{x_t - P_t}{n_i}, n_{i+1} = n_i + 1 \quad \text{公式 1}$$

10 [78]其中， P_{t+1} 表示目标子分类器更新后的中心点， P_t 表示目标子分类器更新前的中心点， x_t 表示特征数据的数据向量， n_i 表示更新前被分配到目标子分类器的数据数量， n_{i+1} 表示目标子分类器更新后的数据数量。

15 [79]在多分类器接收的特征数据的数据量大于等于设定阈值时，计算特征数据的数据向量与每个子分类器的中心点之间的距离，本说明书实施例并不限制数据向量与中心点之间距离的具体算法，可以采用欧式距离、曼哈顿距离、切比雪夫距离等。获取中心点与数据向量之间距离最近的子分类器作为目标子分类器，在目标子分类器对该特征数据的数据向量进行分类预测后，利用公式 1 对目标子分类器进行更新。

[80]方式二、如果多分类器的 k 个子分类器中每个子分类器均为具有数据分布参数的高斯混合模型，查找获取 k 个子分类器中数据分布参数与特征数据的相似度最大的子分类器作为目标子分类器。数据分布参数与特征数据之间的相似度可以采用余弦相似度、皮尔逊相关系数、欧几里德距离等方法计算。

20 [81]在 S102 获得目标子分类器之后，执行 S103 根据获得的目标子分类器对特征数据进行预测获得第二预测数据。具体的，将特征数据输入目标子分类器，由目标子分类器对该特征数据进行打分获得第二得分即第二预测数据。

25 [82]进一步的，在获得第一预测数据和第二预测数据之后，执行 S104 根据预设算法对第一预测数据和第二预测数据进行综合处理，获得第三预测数据。具体的，预设算法可以是将第一预测数据和第二预测数据相加获得第三预测数据，也可以将第一预测数据与第一权重相乘获得第一乘积，将第二预测数据与第二权重相乘获得第二乘积，获得第一乘积和第二乘积之和作为第三预测数据。

[83]S104 之后，执行 S105 根据第三预测数据对特征数据对应的待推荐实体进行推荐。

具体的,对第三预测数据进行排序,将排序靠前的第三预测数据所属的特征数据对应的待推荐实体推荐给用户。或者,判断第三预测数据是否在预设范围内,若在预设范围内,将第三预测数据所属的特征数据对应的待推荐实体推荐给用户。

[84]具体实施过程中,本说明书实施例还基于用户对被推荐实体的选择与否对多分类器进行反馈,以更新多分类器的相关参数。检测用户是否选择特征数据对应的被推荐实体;根据用户对被推荐实体的选择与否,更新特征数据对应的目标子分类器和主分类器的参数。如果用户选择了被推荐实体向多分类器反馈标识1即进行正反馈,如果用户未选择被推荐实体对所述多分类器进行负反馈即进行负反馈;根据所述正反馈的或所述负反馈更新特征数据对应的目标子分类器和主分类器的参数。

10 [85]下面以一具体的更新方法对目标子分类器和主分类器的参数更新进行详细说明,但本说明书实施例并不限于该更新方法。

[86]本说明书实施例为每个分类器(主分类器和每个子分类器)配置一个权值向量和一个关系矩阵。初始化时权值向量为0,关系矩阵初始化时为单位矩阵。假设输入的样本特征维度为d,那么每个分类器的权值向量的维度为d*1,关系矩阵的维度为d*d,为此,多分类器的权值向量、关系矩阵整体可以表示如下:

$$\tilde{\mu} = [\sqrt{\lambda}\mu^T, \mu_1^T, \dots, \mu_j^T, \dots, \mu_K^T]^T \quad \text{公式 2}$$

$$\tilde{\Sigma} = \begin{bmatrix} \sqrt{\lambda}\Sigma & 0 & \dots & 0 \\ 0 & \Sigma_1 & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & \dots & \Sigma_k \end{bmatrix} \quad \text{公式 3}$$

[87]其中, $\tilde{\mu}$ 表示多分类器的权值向量, $\tilde{\Sigma}$ 表示多分类器的关系矩阵; μ 表示主分类器的权值向量; μ_1 至 μ_K 表示第1~k个子分类器的权值向量, μ_j 表示第j个子分类器的权值向量; λ 表示主分类器的权重,一般取值为1; Σ 表示主分类器的关系矩阵, Σ_1 至 Σ_k 表示第1~k个子分类器的关系矩阵。为了便于区分更新前后参数的不同,t时刻多分类器的权值向量可以表示为 $\tilde{\mu}_t$,更新后的权值向量可以表示为 $\tilde{\mu}_{t+1}$,同样的,t时刻关系矩阵表示为 $\tilde{\Sigma}_t$,更新后为 $\tilde{\Sigma}_{t+1}$ 。

25 [88]假设被推荐实体对应的特征数据的数据向量为第i个数据向量 x_t 被分配到了第j个子分类器即目标子分类器,相对于整个多分类器,可以构建新的数据形式:

$$\tilde{x}_t = [x_t^T, 0_1^T, \dots, 0_{j-1}^T, x_t^T, 0_{j+1}^T, \dots, 0_K^T]^T \quad \text{公式 4}$$

[89]通过上述数据形式的构建,使得第i个数据向量 x_t 分别对应主分类器和第j个子分类

器。根据用户对被推荐实体的选择与否，对多分类器进行正反馈或负反馈。例如，用户可以通过点击、长按、重按等任一操作来选择被推荐实体，若用户选择被推荐实体进行正反馈，向多分类器反馈标识 $y_t=+1$ ，若用户未选择被推荐实体进行负反馈，向多分类器反馈标识 $y_t=-1$ 。

- 5 [90]根据正反馈或者负反馈更新主分类器和第 j 个子分类器的权值向量以及主分类器和第 j 个子分类器的权值向量的关系矩阵。具体的，可以先根据反馈的标识 y_t 和第 i 个数据向量，计算对第 i 个数据向量进行预测的损失值，若损失值大于零进行更新，否则不进行更新，提高更新的准确性。

- [91]计算损失值时，可以根据 y_t 、第 i 个数据向量以及主类器和第 j 个子分类器两者的权值向量和关系矩阵，通过如下公式计算获得多分类器对第 i 条数据进行预测的损失值：

$$\ell = \max \left(0, \tilde{\phi} \sqrt{\tilde{\mathbf{x}}_t^T \tilde{\Sigma} \tilde{\mathbf{x}}_t} - y_t \tilde{\mu}^T \tilde{\mathbf{x}}_t \right) \quad \text{公式 5}$$

[92]其中， ℓ 表示损失值， $\tilde{\phi}$ 表示超参数， $\tilde{\mathbf{x}}_t$ 由第 i 个数据向量 x_t 构建而成（详见公式 4）， $\tilde{\mu}$ 、 $\tilde{\Sigma}$ 分别表示多分类器的权重向量、关系矩阵（详见公式 1、2）。 $\tilde{\phi}$ 通常可以为正态分布累加函数的逆函数如 $\tilde{\phi} = \Phi^{-1}(\eta)$ 。

- 15 [93]判断其损失值 ℓ 是否大于零；若损失值 ℓ 大于零，根据 y_t 、第 i 个数据向量，更新主分类器和第 j 个子分类器的权值向量以及主分类器和第 j 个子分类器的关系矩阵。反之，若损失值 ℓ 不大于零，则暂时不对权值向量和关系矩阵进行更新。更新权值向量和关系矩阵时，还可以根据第 i 个数据向量更新多分类器的其它参数，其中， $\tilde{\alpha}_t$ 、 $\tilde{\beta}_t$ 可以为预设参数，也可以根据下述参数进行更新，具体更新公式如下：

$$\tilde{\mu}_{t+1} = \tilde{\mu}_t + \tilde{\alpha}_t y_t \tilde{\Sigma}_t \tilde{\mathbf{x}}_t \quad \text{公式 6}$$

$$\tilde{\Sigma}_{t+1} = \tilde{\Sigma}_t - \tilde{\beta}_t \tilde{\Sigma}_t \tilde{\mathbf{x}}_t \tilde{\mathbf{x}}_t^T \tilde{\Sigma}_t \quad \text{公式 7}$$

$$\tilde{\alpha}_t = \max \left\{ C, \max \left\{ 0, \frac{1}{\tilde{v}_t \tilde{\zeta}} \left(-\tilde{m}_t \tilde{\psi} + \sqrt{\tilde{m}_t^2 \frac{\tilde{\phi}^4}{4} + \tilde{v}_t \tilde{\phi}^2 \tilde{\zeta}} \right) \right\} \right\} \quad \text{公式 8}$$

$$\tilde{\beta}_t = \frac{\tilde{\alpha}_t \tilde{\phi}}{\sqrt{\tilde{u}_t} + \tilde{v}_t \tilde{\alpha}_t \tilde{\phi}} \quad \text{公式 9}$$

$$\tilde{u}_t = \frac{1}{4} \left(-\tilde{\alpha}_t \tilde{v}_t \tilde{\phi} + \sqrt{\tilde{\alpha}_t^2 \tilde{v}_t^2 \tilde{\phi}^2 + 4\tilde{v}_t} \right)^2, \tilde{v}_t = \tilde{\mathbf{x}}_t^T \tilde{\Sigma}_t \tilde{\mathbf{x}}_t \quad \text{公式 10}$$

$$\tilde{\psi} = 1 + \frac{\tilde{\phi}^2}{2} \text{ and } \tilde{\zeta} = 1 + \tilde{\phi}^2 \quad \text{公式 11}$$

$$\tilde{\mathbf{m}}_t = y_t \tilde{\boldsymbol{\mu}}_t^T \tilde{\mathbf{x}}_t \quad \text{公式 12}$$

[94]其中， $\tilde{\boldsymbol{\mu}}_{t+1}$ 表示更新后的多分类器的权值向量， $\tilde{\Sigma}_{t+1}$ 表示更新后的多分类器的关系矩阵。根据公式 1、2 的表示方式，从 $\tilde{\boldsymbol{\mu}}_{t+1}$ 中获得所述主分类器和所述第 j 个子分类器更新后的权值向量 μ 、 μ_j ；从 $\tilde{\Sigma}_{t+1}$ 中获得所述主分类器和所述第 j 个子分类器更新后的关系矩阵 Σ 、 Σ_j 。C 为超参数，一般可以设置为 1。

[95]在上述实施例，通过采用多分类器的主分类器和目标子分类器来对特征数据进行分类，实现了对数据的多元化分类预测，提高了数据分类预测的准确性；并且，目标子分类器选择多分类器中数据分布与数据特征相似度最大的子分类器，获得主分类器对特征数据进行预测的第一预测数据、目标子分类器对特征数据进行分类的第二预测数据，综合第一预测数据和第二预测数据获得第三预测数据，从而获得更为准确的预测数据，解决了现有技术中数据分类预测准确性较低的技术问题；进一步的，根据第三预测数据对特征数据进行推荐处理，从而提高了数据推荐的准确性。

[96]进一步的，上述实施例还基于每一次用户对被推荐实体的选择与否，对多分类器进行反馈，以实现多分类器的在线更新、实时学习，从而可以利用实时数据流，实时预测用户的偏好和兴趣，使得所有接收数据只使用一次，不需要存储任何历史信息，减少了线上的存储压力和计算复杂度。并且，通过中心点和分类器参数的更新，多分类器模型可以在线上实时训练，减少了模型上线和下线的运维压力。

[97]基于上述实施例提供的一种基于多分类器的推荐方法，本实施例还对应提供一种基于多分类器的推荐装置，请参考图 2，所述装置包括：

[98]第一分类单元 21，用于根据预设的多分类器的主分类器对接收到的特征数据进行预测，获得第一预测数据；

[99]查找单元 22，用于针对所述特征数据，在预设的多分类器的 k 个子分类器中查找获得目标子分类器，所述目标子分类器的数据分布与所述特征数据之间的相似度最大， $k \geq 2$ ；

[100] 第二分类单元 23, 用于根据所述目标子分类器对所述特征数据进行预测, 获得第二预测数据;

[101] 综合处理单元 24, 用于根据预设算法对所述第一预测数据和所述第二预测数据进行综合处理, 获得第三预测数据;

5 [102] 推荐单元 25, 用于根据所述第三预测数据对所述特征数据对应的待推荐实体进行推荐。

[103] 作为一种可选的实施方式, 所述查找单元 22 用于:

[104] 如果所述 k 个子分类器中每个子分类器均具有数据分布的中心点, 查找获取中心点与所述特征数据的数据向量距离最近的子分类器作为所述目标子分类器。或者, 如果
10 如果所述 k 个子分类器中每个子分类器均为具有数据分布参数的高斯混合模型, 查找获取所述 k 个子分类器中数据分布参数与所述特征数据的相似度最大的子分类器作为所述目标子分类器。

[105] 作为一种可选的实施方式, 所述装置还包括更新单元 26。更新单元 26 可以用于: 在根据所述第三预测数据对所述特征数据对应的待推荐实体进行推荐之后, 检测用户是
15 否选择所述特征数据对应的被推荐实体; 根据用户对被推荐实体的选择与否, 更新所述特征数据对应的目标子分类器和主分类器的参数。具体的, 如果用户选择了被推荐实体对所述多分类器进行正反馈, 如果用户未选择被推荐实体对所述多分类器进行负反馈; 根据所述正反馈的或所述负反馈更新所述特征数据对应的目标子分类器和主分类器的参数。

20 [106] 更新单元 26 还可以用于: 在根据所述目标子分类器对所述特征数据进行预测, 获得第二预测数据之后, 根据所述特征数据的数据向量更新所述目标子分类器的中心点。具体的, 所述更新单元可以根据如下公式更新目标子分类器的中心点:

$$P_{t+1} = P_t + \frac{x_t - P_t}{n_i}, n_{i+1} = n_i + 1$$

[107] 其中, P_{t+1} 表示所述目标子分类器更新后的中心点, P_t 表示所述目标子分类器更新前的中心点, x_t 表示所述数据向量, n_i 表示更新前被分配到所述目标子分类器的数据数量, n_{i+1} 表示所述目标子分类器更新后的数据数量。
25

[108] 关于上述实施例中的装置, 其中各个单元执行操作的具体方式已经在有关方法的实施例中进行了详细描述, 此处不再详细阐述。

[109] 请参考图 3, 是根据一示例性实施例示出的一种用于实现数据查询方法的电子设备 700 的框图。例如, 电子设备 700 可以是计算机, 数据库控制台, 平板设备, 个人数字助理等。

[110] 参照图 3, 电子设备 700 可以包括以下一个或多个组件: 处理组件 702, 存储器 5 704, 电源组件 706, 多媒体组件 708, 输入/输出 (I/O) 的接口 710, 以及通信组件 712。

[111] 处理组件 702 通常控制电子设备 700 的整体操作, 诸如与显示, 数据通信, 及记录操作相关联的操作。处理元件 702 可以包括一个或多个处理器 720 来执行指令, 以完成上述的方法的全部或部分步骤。此外, 处理组件 702 可以包括一个或多个模块, 便于处理组件 702 和其他组件之间的交互。

10 [112] 存储器 704 被配置为存储各种类型的数据以支持在设备 700 的操作。这些数据的示例包括用于在电子设备 700 上操作的任何应用程序或方法的指令, 联系人数据, 电话簿数据, 消息, 图片, 视频等。存储器 704 可以由任何类型的易失性或非易失性存储设备或者它们的组合实现, 如静态随机存取存储器 (SRAM), 电可擦除可编程只读存储器 (EEPROM), 可擦除可编程只读存储器 (EPROM), 可编程只读存储器 (PROM),
15 只读存储器 (ROM), 磁存储器, 快闪存储器, 磁盘或光盘。

[113] 电源组件 706 为电子设备 700 的各种组件提供电力。电源组件 706 可以包括电源管理系统, 一个或多个电源, 及其他与为电子设备 700 生成、管理和分配电力相关联的组件。

[114] I/O 接口 710 为处理组件 702 和外围接口模块之间提供接口, 上述外围接口模块可以是键盘, 点击轮, 按钮等。这些按钮可包括但不限于: 主页按钮、音量按钮、启动按钮和锁定按钮。
20

[115] 通信组件 712 被配置为便于电子设备 700 和其他设备之间有线或无线方式的通信。电子设备 700 可以接入基于通信标准的无线网络, 如 WiFi, 2G 或 3G, 或它们的组合。在一个示例性实施例中, 通信部件 712 经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中, 所述通信部件 712 还包括近场通信 (NFC) 模块, 以促进短程通信。例如, 在 NFC 模块可基于射频识别 (RFID) 技术, 红外数据协会 (IrDA) 技术, 超宽带 (UWB) 技术, 蓝牙 (BT) 技术和其他技术来实现。
25

[116] 在示例性实施例中, 电子设备 700 可以被一个或多个应用专用集成电路 (ASIC)、

数字信号处理器 (DSP)、数字信号处理设备 (DSPD)、可编程逻辑器件 (PLD)、现场可编程门阵列 (FPGA)、控制器、微控制器、微处理器或其他电子元件实现, 用于执行上述方法。

- [117] 在示例性实施例中, 还提供了一种包括指令的非临时性计算机可读存储介质, 例如包括指令的存储器 704, 上述指令可由电子设备 700 的处理器 720 执行以完成上述方法。例如, 所述非临时性计算机可读存储介质可以是 ROM、随机存取存储器 (RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[118] 一种非临时性计算机可读存储介质, 当所述存储介质中的指令由移动终端的处理器执行时, 使得电子设备能够执行一种基于多分类器的推荐方法, 所述方法包括:

- 10 [119] 根据预设的多分类器的主分类器对接收到的特征数据进行预测, 获得第一预测数据; 针对所述特征数据, 在预设的多分类器的 k 个子分类器中查找获得目标子分类器, 所述目标子分类器的数据分布与所述特征数据之间的相似度最大, $k \geq 2$; 根据所述目标子分类器对所述特征数据进行预测, 获得第二预测数据; 根据预设算法对所述第一预测数据和所述第二预测数据进行综合处理, 获得第三预测数据; 根据所述第三预测数据对
- 15 所述特征数据对应的待推荐实体进行推荐。

[120] 应当理解的是, 本发明并不局限于上面已经描述并在附图中示出的精确结构, 并且可以在不脱离其范围进行各种修改和改变。本发明的范围仅由所附的权利要求来限制

- [121] 以上所述仅为本发明的较佳实施例, 并不用以限制本发明, 凡在本发明的精神和原则之内, 所作的任何修改、等同替换、改进等, 均应包含在本发明的保护范围之内。
- 20