

跨域推荐模型的训练方法、装置、设备、介质及产品

本申请要求于 2023 年 01 月 13 日提交的，申请名称为“跨域推荐模型的训练方法、装置、设备、介质及产品”的、中国专利申请号为“202310084391.9”的优先权，该中国专利申请的全部内容通过引用结合在本申请中。

5 技术领域

本申请涉及人工智能领域，特别涉及一种跨域推荐模型的训练方法、装置、设备、介质及产品。

背景技术

随着互联网技术的快速发展，用户可以在网络上选择自己喜欢的内容进行浏览，服务器也可以根据用户的日常喜好为用户推荐具有相似度的内容。

10 相关技术中，通过使用直接为用户推荐目标内容的推荐方式，召回与用户的喜好内容的相似度较高的目标内容，向用户进行推荐。或，通过使用根据用户的喜好内容的内容特征推荐目标内容的推荐方式，召回内容特征与用户的喜好内容的内容特征相似度较高的目标内容，向用户进行推荐。但是上述推荐方式均需要引入用户节点，在一定程度上会受到用户数量规模的影响，当用户数量规模较小时，推荐模型能够使用的数据就少，从而可能出现推荐的目标内容与用户的喜好内容相关性不足的问题。

15 发明内容

本申请提供了一种跨域推荐模型的训练方法、装置、设备、介质及产品，所述技术方案如下：

根据本申请的一方面，提供了一种跨域推荐模型的训练方法，所述方法包括：构建异质网络，所述异质网络包括样本源域内容节点和样本目标域内容节点之间的节点二部图、所述样本源域内容节点和样本源域语义标签之间的第一标签二部图、所述样本目标域内容节点和样本目标域语义标签之间的第二标签二部图；基于所述节点二部图中存在连接边的所述样本源域内容节点和所述样本目标域内容节点、所述样本源域内容节点在所述第一标签二部图中对应的样本源域语义标签，以及所述样本目标域内容节点在所述第二标签二部图中对应的样本目标域语义标签，生成训练样本；基于所述训练样本对所述跨域推荐模型进行训练。

25 根据本申请的另一方面，提供了一种跨域推荐方法，所述方法包括：获取用户账号的历史行为；基于所述用户账号的所述历史行为，确定所述用户账号存在历史互动的源域内容；基于源域内容向量和目标域内容向量之间的相似性，确定与所述源域内容对应的目标域内容；向所述用户账号推荐所述目标域内容；其中，所述源域内容向量是所述源域内容的特征向量，所述目标域内容向量是所述目标域内容的特征向量，所述源域内容向量是基于所述源域内容和所述源域内容在第一标签二部图中对应的源域语义标签构建的，所述目标域内容向量是基于所述目标域内容和所述目标域内容在第二标签二部图中对应的目标域语义标签构建的，所述第一标签二部图是基于所述源域内容和所述源域语义标签构建的，所述第二标签二部图是基于所述目标域内容和所述目标域语义标签构建的。

30 根据本申请的另一方面，提供了一种跨域推荐模型的训练装置，所述装置包括：构建模块，用于构建异质网络，所述异质网络包括样本源域内容节点和样本目标域内容节点之间的节点二部图、所述样本源域内容节点和样本源域语义标签之间的第一标签二部图、所述样本目标域内容节点和样本目标域语义标签之间的第二标签二部图；生成模块，用于基于所述节点二部图中存在连接边的所述样本源域内容节点和所述样本目标域内容节点、所述样本源域内容节点在所述第一标签二部图中对应的样本源域语义标签，以及所述样本目标域内容节点在所述第二标签二部图中对应的样本目标域语义标签，生成训练样本；训练模块，用于基于所述训练样本对所述跨域推荐模型进行训练。

40 根据本申请的另一方面，提供了一种跨域推荐装置，所述装置包括：获取模块，用于获取用户账号的历史行为；确定模块，用于基于所述用户账号的所述历史行为，确定所述用户账号存在历史互动的源域内容；所述确定模块，还用于基于源域内容向量和目标域内容向量之间的相似性，确定与所述源域内容对应的目标域内容；推荐模块，用于向所述用户账号推荐所述目标域内容；其中，所述源域内容向量是所述源域内容的特征向量，所述目标域内容向量是所述目标域内容的特征向量，所述源域内容向量是基于所述源域内容和所述源域内容在第一标签二部图中对应的源域语义标签构建的，所述目标域内容向量是基于所述目标域内容和所述目标域内容在第二标签二部图中对应的目标域语义标签构建的，所述第一标签二部图是基于所述源域内容和所述源域语义标签构建的，所述第二标签二部图是基于所述目标域

内容和所述目标域语义标签构建的。

根据本申请的另一方 面，提供了一种计算机设备，该计算机设备包括：处理器和存储器，所述存储器中存储有至少一段程序，所述至少一段程序由所述处理器加载并执行，以实现如上方面所述的跨域推荐模型的训练方法和跨域推荐方法。

5 根据本申请的另一方 面，提供了一种计算机可读存储介质，该计算机可读存储介质中存储有至少一段程序，所述至少一段程序由处理器加载并执行以实现如上方面所述的跨域推荐模型的训练方法和跨域推荐方法。

根据本申请的另一方 面，提供了一种计算机程序产品，该计算机程序产品包括至少一段程序，所述至少一段程序存储在计算机可读存储介质中；计算机设备的处理器从所述计算机可读存储介质中读取所述至少一段程序，所述处理器执行所述至少一段程序，使得所述计算机设备执行如上方面所述的跨域推荐模型的训练方法和跨域推荐方法。

本申请提供的技术方案带来的有益效果至少包括：通过构建包括样本源域内容节点和样本目标域内容节点之间的节点二部图、样本源域内容节点和样本源域语义标签之间的标签二部图、样本目标域内容节点和样本目标域语义标签之间的标签二部图的异质网络，将源域内容节点和目标域内容节点进行关联；同时利用源域语义标签和目标域语义标签增强了源域内容节点和目标域内容节点的特征表示。通过基于样本源域内容节点、样本目标域内容节点、源域语义标签和目标域语义标签生成的训练样本训练跨域推荐模型，可以使跨域推荐模型充分学习到将要进行推荐的目标内容的嵌入表示，克服由于用户数据稀疏造成的影响，从而更好的使用该跨域推荐模型实现向用户推荐相关性较高的目标内容。

附图说明

20 为了更清楚地说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍。

- 图 1 是本申请的一个示例性实施例提供的计算机系统的结构图；
- 图 2 是本申请的一个示例性实施例提供的跨域推荐模型的训练方法的流程图；
- 图 3 是本申请的一个示例性实施例提供的跨域推荐模型的结构图；
- 25 图 4 是本申请的一个示例性实施例提供的跨域推荐模型的训练方法的流程图；
- 图 5 是本申请的一个示例性实施例提供的跨域推荐模型的结构图；
- 图 6 是本申请的一个示例性实施例提供的跨域推荐模型的训练方法的流程图；
- 图 7 是本申请的一个示例性实施例提供的跨域推荐模型的结构图；
- 图 8 是本申请的一个示例性实施例提供的跨域推荐模型的结构图；
- 30 图 9 是本申请的一个示例性实施例提供的跨域推荐模型的结构图；
- 图 10 是本申请的一个示例性实施例提供的跨域推荐模型的结构图；
- 图 11 是本申请的一个示例性实施例提供的跨域推荐模型的训练方法的流程图；
- 图 12 是本申请的一个示例性实施例提供的跨域推荐模型的结构图；
- 图 13 是本申请的一个示例性实施例提供的跨域推荐模型的结构图；
- 35 图 14 是本申请的一个示例性实施例提供的跨域推荐模型的结构图；
- 图 15 是本申请的一个示例性实施例提供的异质网络的结构图；
- 图 16 是本申请的一个示例性实施例提供的跨域推荐方法的流程图；
- 图 17 是本申请的一个示例性实施例提供的跨域推荐方法的流程图；
- 图 18 是本申请的一个示例性实施例提供的跨域推荐方法的流程图；
- 40 图 19 是本申请的一个示例性实施例提供的跨域推荐方法的流程图；
- 图 20 是本申请的一个示例性实施例提供的跨域推荐方法的流程图；
- 图 21 是本申请的一个示例性实施例提供的跨域推荐方法的流程图；
- 图 22 是本申请的一个示例性实施例提供的跨域推荐方法的流程图；
- 图 23 是本申请的一个示例性实施例提供的跨域推荐方法的流程图；
- 45 图 24 是本申请的一个示例性实施例提供的跨域推荐模型的训练装置的框图；
- 图 25 是本申请的一个示例性实施例提供的跨域推荐装置的框图；

图 26 是本申请的一个示例性实施例提供的计算机设备的框图。

具体实施方式

为使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请实施方式作进一步地详细描述。

5 首先，对本申请实施例中涉及的名词进行简单介绍：

跨域推荐：是指一种包括源域和目标域的推荐方法。通过使用跨域推荐方法，能够实现从源域中的内容推荐得到目标域中的内容。

语义标签：是指与内容相关的离散的文本标签，如与内容主题相关的标签、与内容中的一个或多个关键词相关的标签、与内容概念相关的标签。

10 推荐系统：是指一种信息过滤系统，根据用户的历史行为，比如观看视频的行为、发表评论的行为、浏览网页的行为等，推荐符合用户喜好的内容；或者，根据用户喜好内容的内容特征，比如文章的标题、视频内容的分类、音乐的作者等，推荐具有相似特征的内容。

15 同现网络：也被称为共现网络，是指基于至少两个节点构建的该至少两个节点在同一场景中同时出现的网络。当两个节点在同一场景中同时出现，就在两节点中连接一条边，边上的权重代表两节点共同出现的次数，边上的权重越大，则表明两个节点间联系越强。

嵌入表示：是指基于神经网络模型生成的对输入对象的分布式表示。主要作用是将原有对象的高维稀疏向量转化为低维、稠密向量，使得这些低维、稠密向量能够表达相应对象的一个或多个特征。同时，不同向量之间的距离能反应对象之间的相似性。

20 图 1 示出了本申请一个示例性实施例提供的计算机系统的结构框图。该计算机系统 100 包括：终端 120 和服务器 140。

25 终端 120 安装和运行有支持源域内容和目标域内容的平台。该源域内容是指在终端 120 上能够进行浏览、观看、阅读或聆听的多媒体内容，比如在终端 120 上能够通过视频播放平台观看的视频，或在终端 120 上能够通过新闻平台阅读的新闻，或在终端 120 上能够通过音乐平台聆听的音乐等等。该目标域内容是指在终端 120 上能够进行浏览、观看、阅读或聆听的多媒体内容，比如在终端 120 上能够通过视频播放平台观看的视频，或在终端 120 上能够通过新闻平台阅读的新闻，或在终端 120 上能够通过音乐平台聆听的音乐等等。该平台上登录有用户账号，用户账号通过该平台对该平台展示的源域内容或目标域内容进行浏览和互动，如用户账号可以点击源域内容或目标域内容进行观看或阅读，可以对正在观看或浏览的源域内容或目标域内容进行点赞、评论、分享、举报和不看等操作。获取上述用户账号的信息是在得到许可并符合法律的相关规定的前提下进行的。

30 终端 120 通过无线网络或有线网络与服务器 140 相连。

服务器 140 包括一台服务器、多台服务器、云计算平台和虚拟化中心中的至少一种。示例性的，服务器 140 包括处理器 144 和存储器 142，存储器 142 中包括用于执行不同步骤的至少一个模块，本申请实施例中以包括接收模块 1421、控制模块 1422 和发送模块 1423 为例进行说明，接收模块 1421 用于接收来自终端 120 的请求，如对源域内容或目标域内容进行点赞的请求；控制模块 1422 用于控制源域内容和目标域内容的播放和展示；发送模块 1423 用于向终端 120 发送响应，如向终端 120 发送对源域内容或目标域内容进行点赞是否成功的反馈。本领域技术人员可以知晓，上述存储器 142 中的模块的数量可以更多或更少，本申请实施例中对该数量不加以限定。

40 可选地，服务器 140 承担主要计算工作，终端 120 承担次要计算工作；或者，服务器 140 承担次要计算工作，终端 120 承担主要计算工作；或者，服务器 140 和终端 120 通过协同的方式承担计算工作。

可选地，上述支持源域内容和目标域内容的平台是不同操作系统平台（安卓或 IOS）上相同的平台。可选的，终端 120 的设备类型相同或不同，该设备类型包括：智能手机、智能手表、车载终端、可穿戴设备、智能电视、平板电脑、电子书阅读器、MP3 播放器、MP4 播放器、膝上型便携计算机和台式计算机中的至少一种。

45 本领域技术人员可以知晓，上述终端 120 的数量可以更多或更少。比如上述终端 120 可以仅为一个，或者上述终端为几十个或几百个，或者更多数量。本申请实施例对终端的数量和设备类型不加以限定。

随着人工智能技术研究和进步，人工智能技术在多个领域展开研究和应用，例如常见的智能家居、

无人驾驶、机器人、智能医疗、智能客服等。人工智能(Artificial Intelligence, AI)技术是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说,人工智能是计算机科学的一个综合技术,它企图了解智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法,使机器具有感知、推理与决策的功能。人工智能技术是一门综合学科,涉及领域广泛,既有硬件层面的技术也有软件层面的技术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。相关技术中的人工智能软件技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

本申请实施例提供的方案涉及人工智能的机器学习(Machine Learning, ML)技术,具体通过如下实施例进行说明:

图2是本申请一个示例性实施例提供的跨域推荐模型的训练方法的示意图。该方法由计算机设备执行,该计算机设备可以是图1所示的终端120或服务器140。该方法包括:

步骤220:构建异质网络,包括样本源域内容节点和样本目标域内容节点之间的节点二部图、样本源域内容节点和样本源域语义标签之间的第一标签二部图、样本目标域内容节点和样本目标域语义标签之间的第二标签二部图;

构建包括样本源域内容节点、样本目标域内容节点、样本源域语义标签和样本目标域语义标签的异质网络。其中,样本源域内容节点和样本目标域内容节点之间通过连接边进行连接能够构成节点二部图,样本源域内容节点和样本源域语义标签之间通过连接边能够构成标签二部图,样本目标域内容节点和样本目标域语义标签之间通过连接边能够构成标签二部图。

该样本源域内容节点是根据样本源域内容获得的。样本源域内容是指在如图1所示的终端120上能够进行浏览、观看、阅读、聆听中的至少一种操作的多媒体内容,如样本源域内容可以是视频、新闻、音乐等。该样本目标域内容节点是根据样本目标域内容获得的。样本目标域内容是指在如图1所示的终端120上能够进行浏览、观看、阅读、聆听中的至少一种操作的多媒体内容,如样本目标域内容可以是视频、新闻、音乐等。该样本源域内容和样本目标域内容是终端120上的用户账号在一个时间段内先后进行操作的内容,如用户账号在一个时间段内先观看了长视频后观看了短视频,则样本源域内容是长视频,样本目标域内容是短视频。

在一些实施例中,样本源域内容和样本目标域内容可以是同一个域中的内容,如样本源域内容和样本目标域内容都是短视频内容,或样本源域内容和样本目标域内容都是长视频内容,样本源域内容和样本目标域内容都是新闻内容,样本源域内容和样本目标域内容都是商品内容。

在一些实施例中,样本源域内容和样本目标域内容可以是不同域中的内容,如样本源域内容是短视频内容而样本目标域内容是长视频内容,或样本源域内容是长视频内容而样本目标域内容是短视频内容,或样本源域内容是视频内容而样本目标域内容是新闻内容,或样本源域内容是视频内容而样本目标域内容是商品内容。

步骤240:基于节点二部图中存在连接边的样本源域内容节点和样本目标域内容节点、样本源域内容节点在第一标签二部图中对应的样本源域语义标签,以及样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签,生成训练样本;

基于节点二部图中存在连接边的样本源域内容节点和样本目标域内容节点、样本源域内容节点在第一标签二部图中对应的样本源域语义标签,以及样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签,生成训练样本。将样本源域内容节点以及样本源域内容节点在第一标签二部图中对应的样本源域语义标签确定为样本源域数据;将样本目标域内容节点以及样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签确定为样本目标域数据;将样本源域数据和样本目标域数据确定为训练样本。该训练样本至少包括从样本源域内容节点推荐得到样本目标域内容节点的正例样本,以及从样本目标域内容节点推荐得到样本源域内容节点的负例样本。

在一些实施例中,样本源域内容节点和样本目标域内容节点可以是一一对应的关系,一个样本源域内容节点对应一个样本目标域内容节点。

在一些实施例中，样本源域内容节点和样本目标域内容节点可以是一对多的关系，一个样本源域内容节点对应至少两个样本目标域内容节点。

在一些实施例中，样本源域内容节点和样本目标域内容节点可以是多对一的关系，至少两个样本源域内容节点对应同一个样本目标域内容节点。

5 **步骤 260：**基于训练样本对跨域推荐模型进行训练。

 将训练样本输入跨域推荐模型，对该跨域推荐模型进行训练。该训练样本是多组样本数据的集合，其中任意一组样本数据包括一个样本源域内容节点、一个样本目标域内容节点以及样本源域内容节点在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签。通过将样本源域内容节点、样本目标域内容节点以及样本源域内容节点在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签输入跨域推荐模型，对该跨域模型进行训练。该跨域推荐模型是用于进行从源域内容推荐得到目标域内容的跨域推荐的模型。

 综上所述，本实施例提供的方法，通过构建包括样本源域内容节点和样本目标域内容节点之间的节点二部图、样本源域内容节点和样本源域语义标签之间的第一标签二部图、样本目标域内容节点和样本目标域语义标签之间的第二标签二部图的异质网络，将源域内容节点和目标域内容节点进行关联；同时利用源域语义标签和目标域语义标签增强了源域内容节点和目标域内容节点的特征表示。通过基于样本源域内容节点、样本目标域内容节点、源域语义标签和目标域语义标签生成的训练样本训练跨域推荐模型，可以使跨域推荐模型充分学习到将要进行推荐的目标内容的嵌入表示，克服由于用户数据稀疏造成的影响，从而更好的使用该跨域推荐模型实现向用户推荐相关性较高的目标内容。

20 在一些实施例中，上述跨域推荐模型包括源域语义塔、目标域语义塔和匹配层。如图 3 所示，跨域推荐模型 300，包括源域语义塔 320、目标域语义塔 340 和匹配层 360。其中，源域语义塔 320 的输出端和匹配层 360 的输入端连接，目标域语义塔 340 的输出端和匹配层 360 的输入端连接。

 图 4 是本申请一个示例性实施例提供的跨域推荐模型的训练方法的示意图。上述步骤 260 还包括如下至少一个步骤：

25 **步骤 261：**将样本源域内容节点和样本源域内容节点对应的样本源域语义标签输入源域语义塔，得到样本源域内容向量；

 跨域推荐模型中的源域语义塔是用于生成源域内容的嵌入表征的网络模型。将样本源域内容节点和样本源域内容节点对应的样本源域语义标签输入源域语义塔，能够得到样本源域内容对应的嵌入表征，即样本源域内容向量。该样本源域内容节点是包括样本源域内容的内容特征的节点，例如样本源域内容是一段长视频，该长视频的内容特征包括视频属性。样本源域内容节点对应的样本源域语义标签是与样本源域内容的内容特征对应的标签中的至少一个，例如与视频属性对应的标签包括标题标签、主题标签、作者标签、演员标签等等，则样本源域语义标签可以是其中至少一种。

30 **步骤 262：**将样本目标域内容节点和样本目标域内容节点对应的样本目标域语义标签输入目标域语义塔，得到样本目标域内容向量；

35 跨域推荐模型中的目标域语义塔是用于生成目标域内容的嵌入表征的网络模型。将样本目标域内容节点和样本目标域内容节点对应的样本目标域语义标签输入目标域语义塔，能够得到样本目标域内容对应的嵌入表征，即样本目标域内容向量。该样本目标域内容节点是包括样本目标域内容的内容特征的节点，例如样本目标域内容是一段长视频，该长视频的内容特征包括视频属性。样本源域内容节点对应的样本目标域语义标签是与样本目标域内容的内容特征对应的标签中的至少一个，例如与视频属性对应的标签包括标题标签、主题标签、作者标签、演员标签等等，则样本目标域语义标签可以是其中至少一种。

40 **步骤 263：**将样本源域内容向量和样本目标域内容向量输入匹配层，得到预测相似度；

 跨域推荐模型中的匹配层用于计算源域语义塔输出的源域内容向量和目标域语义塔输出的目标域内容向量的相关性。将基于源域语义塔获得的样本源域内容向量和基于目标域语义塔获得的样本目标域内容向量输入匹配层，通过 cosine 余弦函数计算这两个向量的余弦相似度，将该余弦相似度作为样本源域内容和样本目标域内容的预测相似度。

45 **步骤 264：**计算预测相似度和真实样本相似度之间的误差损失；

将训练样本中的正例样本（从源域内容到目标域内容）标注为 1，负例样本（从目标域内容到源域内容）标注为 0，则得到二分类训练样本。二分类训练样本的概率和为 1，即对于同样的源域内容和目标域内容，一个用户只能先源域内容后目标域内容或者先目标域内容后源域内容。

5 使用二分类交叉熵作为损失函数计算预测相似度和真实样本相似度之间的误差损失。损失函数的公式为：

$$\text{Loss} = -(y \cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y}))$$

其中，Loss 表示误差损失，y 表示真实样本相似度， $\hat{y}$ 表示预测相似度。

步骤 265：基于误差损失对跨域推荐模型进行训练。

10 基于样本源域内容节点和样本目标域内容节点的连接边上的权重，确定样本源域数据和样本目标域数据之间的真实样本相似度，该权重是基于样本源域内容节点和样本目标域内容节点在一个时间段内同时出现的次数确定的。基于二分类交叉熵损失函数确定跨域推荐模型预测相似度和真实样本相似度之间的误差损失，将该误差反馈给跨域推荐模型，对跨域推荐模型进行训练。

15 综上所述，本实施例提供的方法，通过将样本源域内容节点和样本源域内容节点对应的样本源域语义标签输入源域语义塔；以及，将样本目标域内容节点和样本目标域内容节点对应的样本目标域语义标签输入目标域语义塔，得到对应的内容向量，可以得到更丰富、更细致的内容表示。将对应的内容向量输入匹配层进行匹配，得到预测的两个内容向量的预测相似度，可以更好地衡量两个内容向量之间的相似程度。通过损失函数计算预测相似度和真实样本相似度之间的误差损失，对跨域推荐模型进行训练，使跨域推荐模型计算的预测相似度无限接近真实样本相似度，可以提高推荐系统的准确性和鲁棒性。

20 在一些实施例中，上述源域语义塔包括级联的源域节点特征提取网络、源域标签特征提取网络、源域拼接层和源域表示层。如图 5 所示，源域语义塔 320 包括级联的源域节点特征提取网络 321、源域标签特征提取网络 322、源域拼接层 323 和源域表示层 324。其中，源域节点特征提取网络 321 的输入端和源域内容节点连接，源域节点特征提取网络 321 的输出端和源域拼接层 323 连接，源域标签特征提取网络 322 的输出端和源域拼接层 323 的输入端连接，源域拼接层 323 的输出端和源域表示层 324 的输入端连接。

图 6 是本申请一个示例性实施例提供的跨域推荐模型的训练方法的示意图。上述步骤 261 还包括如下至少一个步骤：

25 步骤 2611：将样本源域内容节点输入源域节点特征提取网络，得到第一样本内容向量；

30 在一些实施例中，上述源域节点特征提取网络包括级联的源域节点输入层、源域节点嵌入层和源域节点表示层。如图 7 所示，源域节点特征提取网络 321 包括级联的源域节点输入层 3211、源域节点嵌入层 3212 和源域节点表示层 3213。其中，源域节点输入层 3211 的输入端和源域内容节点连接，源域节点输入层 3211 的输出端和源域节点嵌入层 3212 的输入端连接，源域节点嵌入层 3212 的输出端和源域节点表示层 3213 的输入端连接，源域节点表示层 3213 的输出端和源域拼接层 323 的输入端连接。

通过源域节点输入层将样本源域内容节点的第一样本内容特征输入。将源域节点输入层中的第一样本内容特征输出，通过源域节点嵌入层对第一样本内容特征进行嵌入表示，得到第一样本内容的嵌入表征向量。将源域节点嵌入层中的第一样本内容的嵌入表征向量输出，通过源域节点表示层对第一样本内容的嵌入表征向量进行学习，得到第一样本内容向量。

35 步骤 2612：将样本源域语义标签输入源域标签特征提取网络，得到第一样本语义标签向量；

40 在一些实施例中，上述源域标签特征提取网络包括级联的源域语义标签输入层、源域嵌入表征编码器和源域标签表示层。如图 8 所示，源域标签特征提取网络 322 包括级联的源域语义标签输入层 3221、源域嵌入表征编码器 3222 和源域标签表示层 3223。其中，源域语义标签输入层 3221 的输入端和源域语义标签连接，源域语义标签输入层 3221 的输出端和源域嵌入表征编码器 3222 的输入端连接，源域嵌入表征编码器 3222 的输出端和源域标签表示层 3223 的输入端连接，源域标签表示层 3223 的输出端和源域拼接层 323 的输入端连接。

45 通过源域语义标签输入层将样本源域语义标签的第一样本语义特征输入。将源域语义标签输入层中的第一样本语义特征输出，通过源域嵌入表征编码器对第一样本语义特征进行嵌入表示，得到第一样本语义的嵌入表征向量。将源域嵌入表征编码器中的第一样本语义的嵌入表征向量输出，通过源域标签表示层对第一样本语义的嵌入表征向量进行学习，得到第一样本语义标签向量。

步骤 2613: 将第一样本内容向量和第一样本语义标签向量输入源域拼接层, 得到样本源域拼接向量;

如图 9 所示, 上述源域语义塔 320 通过级联的源域节点输入层 3211、源域节点嵌入层 3212 和源域节点表示层 3213, 输出第一样本内容向量; 通过级联的源域节点输入层 3211、源域节点嵌入层 3212 和源域节点表示层 3213, 输出第一样本语义标签向量。

- 5 将第一样本内容向量和第一样本语义标签向量输入源域拼接层 323, 将第一样本内容向量和第一语义标签向量进行向量合并, 得到样本源域拼接向量。

步骤 2614: 将样本源域拼接向量输入源域表示层, 得到样本源域内容向量。

- 如图 9 所示, 将源域拼接层 323 输出的样本源域拼接向量输入源域表示层 324, 得到样本源域内容向量。该样本源域内容向量包括样本源域内容的完整特征向量, 该完整特征向量是基于样本源域内容节点和与样本源域内容节点对应的样本源域语义标签得到的。

- 10 在一些实施例, 上述目标域语义塔包括级联的目标域节点特征提取网络、目标域标签特征提取网络、目标域拼接层和目标域表示层。如图 10 所示, 目标域语义塔 340 包括级联的目标域节点特征提取网络 341、目标域标签特征提取网络 342、目标域拼接层 343 和目标域表示层 344。其中, 目标域节点特征提取网络 341 的输入端和目标域内容节点连接, 目标域节点特征提取网络 341 的输出端和目标域拼接层 343 连接, 目标域标签特征提取网络 342 的输出端和目标域拼接层 343 的输入端连接, 目标域拼接层 343 的输出端和目标域表示层 344 的输入端连接。

图 11 是本申请一个示例性实施例提供的跨域推荐模型的训练方法的示意图。上述步骤 262 还包括如下至少一个步骤:

步骤 2621: 将样本目标域内容节点输入目标域节点特征提取网络, 得到第二样本内容向量;

- 20 在一些实施例, 上述目标域节点特征提取网络包括级联的目标域节点输入层、目标域节点嵌入层和目标域节点表示层。如图 12 所示, 目标域节点特征提取网络 341 包括级联的目标域节点输入层 3411、目标域节点嵌入层 3412 和目标域节点表示层 3413。其中, 目标域节点输入层 3411 的输入端和目标域内容节点连接, 目标域节点输入层 3411 的输出端和目标域节点嵌入层 3412 的输入端连接, 目标域节点嵌入层 3412 的输出端和目标域节点表示层 3413 的输入端连接, 目标域节点表示层 3413 的输出端和目标域拼接层 343 的输入端连接。

通过目标域节点输入层将样本目标域内容节点的第二样本内容特征输入。将目标域节点输入层中的第二样本内容特征输出, 通过目标域节点嵌入层对第二样本内容特征进行嵌入表示, 得到第二样本内容的嵌入表征向量。将目标域节点嵌入层中的第二样本内容的嵌入表征向量输出, 通过目标域节点表示层对第二样本内容的嵌入表征向量进行学习, 得到第二样本内容向量。

- 30 **步骤 2622:** 将样本目标域语义标签输入目标域标签特征提取网络, 得到第二样本语义标签向量;

- 35 在一些实施例, 上述目标域标签特征提取网络包括级联的目标域语义标签输入层、目标域嵌入表征编码器和目标域标签表示层。如图 13 所示, 目标域标签特征提取网络 342 包括级联的目标域语义标签输入层 3421、目标域嵌入表征编码器 3422 和目标域标签表示层 3423。其中, 目标域语义标签输入层 3421 的输入端和目标域语义标签连接, 目标域语义标签输入层 3421 的输出端和目标域嵌入表征编码器 3422 的输入端连接, 目标域嵌入表征编码器 3422 的输出端和目标域标签表示层 3423 的输入端连接, 目标域标签表示层 3423 的输出端和目标域拼接层 343 的输入端连接。

- 40 通过目标域语义标签输入层将样本目标域语义标签的第二样本语义特征输入。将目标域语义标签输入层中的第二样本语义特征输出, 通过目标域嵌入表征编码器对第二样本语义特征进行嵌入表示, 得到第二样本语义的嵌入表征向量。将目标域嵌入表征编码器中的第二样本语义的嵌入表征向量输出, 通过目标域标签表示层对第二样本语义的嵌入表征向量进行学习, 得到第二样本语义标签向量。

步骤 2623: 将第二样本内容向量和第二样本语义标签向量输入目标域拼接层, 得到样本目标域拼接向量;

- 45 如图 14 所示, 上述目标域语义塔 340 通过级联的目标域节点输入层 3411、目标域节点嵌入层 3412 和目标域节点表示层 3413, 输出第二样本内容向量; 通过级联的目标域节点输入层 3411、目标域节点嵌入层 3412 和目标域节点表示层 3413, 输出第二样本语义标签向量。

将第二样本内容向量和第二样本语义标签向量输入目标域拼接层 343, 将第二样本内容向量和第二语

义标签向量进行向量合并，得到样本目标域拼接向量。

步骤 2624：将样本目标域拼接向量输入目标域表示层，得到样本目标域内容向量。

如图 14 所示，将目标域拼接层 343 输出的样本目标域拼接向量输入目标域表示层 344，得到样本目标域内容向量。该样本目标域内容向量包括样本目标域内容的完整特征向量，该完整特征向量是基于样本目标域内容节点和与样本目标域内容节点对应的样本目标域语义标签得到的。

综上所述，本实施例提供的方法，通过源域语义塔和目标域语义塔，实现了将第一样本内容向量和对应的第一样本语义标签向量进行拼接，得到包括样本源域内容的完整特征的样本源域内容向量；以及将第二样本内容向量和对应的第二样本语义标签向量进行拼接，得到包括样本目标域内容的完整特征的样本目标域内容向量。通过将内容节点和对应的语义标签分别输入到源域语义塔和目标域语义塔中，可以得到更全面、更准确的内容表示和语义标签表示。通过拼接层将内容向量和语义标签向量进行拼接，可以得到更完整的特征表示。

图 15 是本申请一个示例性实施例提供的异质网络 400 的结构图。

异质网络 400 中包括样本源域内容节点、样本目标域内容节点、与样本源域内容节点关联的样本源域语义标签节点和与样本目标域内容节点关联的样本目标域语义标签节点。样本源域内容节点和样本目标域内容节点之间通过连接边构成节点二部图，样本源域内容节点和样本源域语义标签之间通过连接边构成第一标签二部图，样本目标域内容节点和样本目标域语义标签之间通过连接边构成第二标签二部图。

在一些实施例中，异质网络 400 还包括基于具有同现关系的样本源域内容节点构建的源域同现网络 420 以及基于具有同现关系的样本目标域内容节点构建的目标域同现网络 440。

源域同现网络 420：

基于多个用户账号在源域的历史行为，构建源域同现网络。基于多个用户账号在源域的历史行为，确定多个用户账号存在历史互动的样本源域内容。例如通过用户账号点击观看短视频的历史行为，确定用户账号存在历史互动的样本源域内容为观看的短视频。基于样本源域内容在同一个用户账号的第二时间段内同时出现的次数，当该样本源域内容同时出现的次数大于阈值时，将样本源域内容对应的样本源域内容节点通过连接边进行连接，得到源域同现网络。该一个时间段可以是系统预设的，也可以是根据具体情况随机调整的，如可以是 1 分钟、5 分钟、10 分钟。该阈值可以是系统预设的，也可以是根据具体情况随机调整的，如可以是 3 次、5 次、10 次。假设一个时间段为 5 分钟，同时出现的次数的阈值为 3 次，用户账号在源域的历史行为是观看短视频，则基于一个用户账号在 5 分钟内观看短视频的行为，确定多个短视频同时出现的次数，将同时出现的次数超过 3 次的短视频进行连接，得到源域同现网络。

在一些实施例中，源域同现网络中的样本源域内容节点是一一对应的关系，例如样本源域内容节点 2 对应样本源域内容节点 1。

在一些实施例中，源域同现网络中的样本源域内容节点是一对多的关系，例如样本源域内容节点 4 对应样本源域内容节点 3 和样本源域内容节点 5。

目标域同现网络 440：

基于多个用户账号在目标域的历史行为，构建目标域同现网络。基于多个用户账号在目标域的历史行为，确定多个用户账号存在历史互动的样本目标域内容。例如通过用户账号点击观看短视频的历史行为，确定用户账号存在历史互动的样本目标域内容为观看的短视频。基于样本目标域内容在同一个用户账号的第三时间段内同时出现的次数，当该样本目标域内容同时出现的次数大于阈值时，将样本目标域内容对应的样本目标域内容节点通过连接边进行连接，得到目标域同现网络。该一个时间段可以是系统预设的，也可以是根据具体情况随机调整的，如可以是 1 分钟、5 分钟、10 分钟。该阈值可以是系统预设的，也可以是根据具体情况随机调整的，如可以是 3 次、5 次、10 次。假设一个时间段为 5 分钟，同时出现的次数的阈值为 3 次，用户账号在目标域的历史行为是观看短视频，则基于一个用户账号在 5 分钟内观看短视频的行为，确定多个短视频同时出现的次数，将同时出现的次数超过 3 次的短视频进行连接，得到目标域同现网络。

在一些实施例中，目标域同现网络中的样本目标域内容节点是一一对应的关系，例如样本目标域内容节点 1 对应样本目标域内容节点 2。

在一些实施例中，目标域同现网络中的样本目标域内容节点是一对多的关系，例如样本目标域内容

节点5对应样本目标域内容节点3和样本目标域内容节点6。

样本源域内容节点和样本源域语义标签之间的第一标签二部图：

基于源域中的语义标签系统，构建样本源域内容节点和样本源域语义标签之间的第一标签二部图。该语义标签系统用于对用户或内容进行语义标签建模，得到用户或内容对应的语义标签。例如，对于一个短视频，通过语义标签系统得到与短视频内容对应的主题标签、作者标签、关键词标签、人物标签等等。将具有对应关系的样本源域语义标签与样本源域内容节点通过连接边进行连接，得到样本源域内容节点和样本源域语义标签之间的第一标签二部图。

在一些实施例中，样本源域语义标签节点和样本源域内容节点是一一对应的关系，例如样本源域语义标签节点1对应样本源域内容节点1，样本源域语义标签节点3对应样本源域内容节点3。

在一些实施例中，样本源域语义标签节点和样本源域内容节点是一对多的关系，例如样本源域语义标签节点5对应样本源域内容节点4和样本源域内容节点5。

在一些实施例中，样本源域语义标签节点和样本源域内容节点是多对一的关系，例如样本源域语义标签节点2和样本源域语义标签节点4对应样本源域内容节点2，样本源域语义标签节点5和样本源域语义标签节点6对应样本源域内容节点5。

样本目标域内容节点和样本目标域语义标签之间的第二标签二部图：

基于目标域中的语义标签系统，构建样本目标域内容节点和样本目标域语义标签之间的第二标签二部图。该语义标签系统用于对用户或内容进行语义标签建模，得到用户或内容对应的语义标签。例如，对于一个短视频，通过语义标签系统得到与短视频内容对应的主题标签、作者标签、关键词标签、人物标签等等。将具有对应关系的样本目标域语义标签与样本目标域内容节点通过连接边进行连接，得到样本目标域内容节点和样本目标域语义标签之间的第二标签二部图。

在一些实施例中，样本目标域语义标签节点和样本目标域内容节点是一一对应的关系，例如样本目标域语义标签节点2对应样本目标域内容节点2。

在一些实施例中，样本目标域语义标签节点和样本目标域内容节点是一对多的关系，例如样本目标域语义标签节点1对应样本目标域内容节点1和样本目标域内容节点3，样本目标域语义标签节点5对应样本目标域内容节点5和样本目标域内容节点6。

在一些实施例中，样本目标域语义标签节点和样本目标域内容节点是多对一的关系，例如样本目标域语义标签节点3和样本目标域语义标签节点6对应样本目标域内容节点4，样本目标域语义标签节点4和样本目标域语义标签节点5对应样本目标域内容节点5。

样本源域内容节点和样本目标域内容节点之间的节点二部图：

基于多个用户账号在源域和目标域的历史行为，构建样本源域内容节点和样本目标域内容节点之间的节点二部图。通过多个用户账号在源域的历史行为，确定多个用户账号存在历史互动的样本源域内容。例如通过用户账号点击观看短视频的历史行为，确定用户账号存在历史互动的样本源域内容为观看的短视频。基于多个用户账号在目标域的历史行为，确定多个用户账号存在历史互动的样本目标域内容。例如通过用户账号点击观看长视频的历史行为，确定用户账号存在历史互动的样本目标域内容为观看的长视频。基于样本源域内容和样本目标域内容在同一个用户账号下的第一时间段内同时出现的次数，将样本源域内容对应的样本源域内容节点和样本目标域内容对应的样本目标域内容节点通过连接边进行连接，得到样本源域内容节点和样本目标域内容节点之间的节点二部图。该一个时间段可以是系统预设的，也可以是根据具体情况随机调整的，如可以是1分钟、5分钟、10分钟。假设一个时间段为5分钟，用户账号在源域的历史行为是观看短视频，用户账号在目标域的历史行为是阅读新闻，则基于一个用户账号在5分钟内既观看短视频又阅读新闻的行为，构建短视频节点和新闻节点之间的节点二部图。

在一些实施例中，样本源域内容节点和样本目标域内容节点是一一对应的关系，例如样本源域内容节点5对应样本目标域内容节点6。

在一些实施例中，样本源域内容节点和样本目标域内容节点是一对多的关系，例如样本源域内容节点1对应样本目标域内容节点2和样本目标域内容节点4，样本源域内容节点2对应样本目标域内容节点1和样本目标域内容节点3，样本源域内容节点4对应样本目标域内容节点4和样本目标域内容节点5。

在一些实施例中，样本源域内容节点和样本目标域内容节点是多对一的关系，例如样本源域内容节点

点1、样本目标域内容节点3和样本目标域内容节点4对应样本目标域内容节点4。

节点二部图中包括存在连接边的第一样本目标域内容节点和第一样本目标域内容节点，如将样本源域内容节点1作为第一样本源域内容节点，将样本目标域内容节点2作为第一样本目标域内容节点，则一组训练样本为样本源域内容节点1和样本目标域内容节点2，以及样本源域内容节点1在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点2在第二标签二部图中对应的样本目标域语义标签。

在一些实施例中，在源域同现网络中存在与第一样本源域内容节点具有同现关系的第二样本源域内容节点的情况下，使用第二样本源域内容节点和第一样本目标域内容节点，第二样本源域内容节点在第一标签二部图中对应的样本源域语义标签和第一样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签，生成训练样本。例如，在源域同现网络420中存在与样本源域内容节点1具有同现关系的样本源域内容节点2和样本源域内容节点3，则一组训练样本为样本源域内容节点2和样本目标域内容节点2，以及样本源域内容节点2在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点2在第二标签二部图中对应的样本目标域语义标签；一组训练样本为样本源域内容节点3和样本目标域内容节点2，以及样本源域内容节点3在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点2在第二标签二部图中对应的样本目标域语义标签。

在一些实施例中，在目标域同现网络中存在与第一样本目标域内容节点具有同现关系的第二样本目标域内容节点的情况下，使用第二样本目标域内容节点和第一样本源域内容节点，以及第一样本源域内容节点在第一标签二部图中对应的样本源域语义标签和第二样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签，生成训练样本。例如，在目标域同现网络440中存在与样本目标域内容节点2具有同现关系的样本目标域内容节点1和样本目标域内容节点3，则一组训练样本为样本源域内容节点1和样本目标域内容节点1，以及样本源域内容节点1在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点1在第二标签二部图中对应的样本目标域语义标签；一组训练样本为样本源域内容节点1和样本目标域内容节点3，以及样本源域内容节点1在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点3在第二标签二部图中对应的样本目标域语义标签。

在一些实施例中，在源域同现网络中存在与第一样本源域内容节点具有同现关系的第二样本源域内容节点且目标域同现网络中存在与第一样本目标域内容节点具有同现关系的第二样本目标域内容节点的情况下，使用第二样本源域内容节点和第二样本目标域内容节点，以及第二样本源域内容节点在第一标签二部图中对应的样本源域语义标签和第二样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签。例如，源域同现网络420中存在与样本源域内容节点1具有同现关系的样本源域内容节点2和样本源域内容节点3，且在目标域同现网络440中存在与样本目标域内容节点2具有同现关系的样本目标域内容节点1和样本目标域内容节点3，则一组训练样本为样本源域内容节点2和样本目标域内容节点1，以及样本源域内容节点2在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点1在第二标签二部图中对应的样本目标域语义标签；一组训练样本为样本源域内容节点2和样本目标域内容节点3，以及样本源域内容节点2在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点3在第二标签二部图中对应的样本目标域语义标签；一组训练样本为样本源域内容节点3和样本目标域内容节点1，以及样本源域内容节点3在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点1在第二标签二部图中对应的样本目标域语义标签；一组训练样本为样本源域内容节点3和样本目标域内容节点3，以及样本源域内容节点3在第一标签二部图中对应的样本源域语义标签和样本目标域内容节点3在第二标签二部图中对应的样本目标域语义标签。

综上所述，本实施例提供的方法，通过构建源域同现网络和目标域同现网络，可以挖掘出源域和目标域中的语义关系和内容关联，通过构建样本源域内容节点和样本源域语义标签之间的第一标签二部图与样本目标域内容节点和样本目标域语义标签之间的第二标签二部图，可以将样本源域内容节点和样本源域语义标签、以及样本目标域内容节点和样本目标域语义标签之间的关系模型化，通过沟通样本源域内容节点和样本目标域内容节点之间的节点二部图，可以将样本源域内容节点和样本目标域内容节点之间的关联考虑进来，进而获得更多的训练样本数据，使跨域推荐模型能够学习到更多数据，从而能够利用该跨域推荐模型更好地进行内容推荐。

图16是本申请一个示例性实施例提供的跨域推荐方法的示意图。该方法由计算机设备执行，该计算

机设备可以是图 1 所示的终端 120 或服务器 140。该方法包括：

步骤 520：获取用户账号的历史行为；

服务器 140 通过获取与其通过无线网络或有线网络相连的终端 120 上的用户账号的历史行为。该历史行为包括用户账号点击、触摸、滑动、评论、长按、指纹识别、面部识别中的至少一种。用户通过用户账号实施的针对终端 120 的全部操作，均认为是用户账号的历史行为。

步骤 540：确定用户账号存在历史互动的源域内容；

基于用户账号的历史行为，确定用户账号存在历史互动的源域内容。例如通过用户账号点击观看短视频的历史行为，确定用户账号存在历史互动的样本源域内容为观看的短视频。

步骤 560：基于源域内容向量和目标域内容向量之间的相似性，确定与源域内容对应的目标域内容；

通过源域内容向量和目标域内容向量之间的相似性，确定与源域内容对应的目标域内容。该相似性是由上述训练得到的跨域推荐模型计算得到的。该源域内容向量是源域内容的特征向量，该目标域内容向量是目标域内容的特征向量，源域内容向量是基于源域内容和源域内容在第一标签二部图中对应的源域语义标签构建的，目标域内容向量是基于目标域内容和目标域内容在第二标签二部图中对应的目标域语义标签构建的。其中，第一标签二部图是基于源域内容和源域语义标签构建的，第二标签二部图是基于目标域内容和目标域语义标签构建的。

在一些实施例，源域内容和目标域内容可以是同一个域中的内容，如源域内容和目标域内容都是短视频内容，或源域内容和目标域内容都是长视频内容。

在一些实施例，源域内容和目标域内容可以是不同域中的内容，如源域内容是短视频内容而目标域内容是长视频内容，或源域内容是长视频内容而目标域内容是短视频内容。

步骤 580：向用户账号推荐目标域内容。

服务器 140 在确定了与源域内容对应的目标域内容后，服务器 140 将该目标域内容发送给终端 120，由终端 120 显示该目标域内容，实现基于用户账号在终端 120 上的历史行为对应的源域内容向用户账号推荐目标域内容。

综上所述，本实施例提供的方法，通过获取用户账号的历史行为，确定用户账号存在历史互动的源域内容。基于源域内容对应的源域内容向量和目标域内容对应的目标域内容向量之间的相似性，确定与源域内容对应的目标域内容，并将该目标域内容推荐给用户账号。通过利用用户账号的历史行为数据，可以获得用户账号的偏好和兴趣，从而更加准确地推荐目标域内容，该目标域内容是基于源域内容获得的，能够避免因用户数据稀疏造成无法推荐或推荐的相似性较低的问题。

图 17 是本申请一个示例性实施例提供的跨域推荐方法的示意图。上述步骤 560 还包括如下至少一个步骤：

步骤 561：获取源域内容向量；

源域内容向量是基于第一内容向量和第一语义标签向量构建的。其中，第一内容向量是源域内容对应的内容向量，第一语义标签向量是源域内容在第一标签二部图中对应的源域语义标签对应的语义标签向量。基于用户账号的历史行为对应的源域内容，从服务器中获取与该源域内容对应的源域内容向量。

步骤 562：获取多个目标域内容向量；

目标域内容向量是基于第二内容向量和第二语义标签向量构建的。其中，第二内容向量是目标域内容对应的内容向量，第二语义标签向量是目标域内容在第二标签二部图中对应的目标域语义标签对应的语义标签向量。从服务器中获取与目标域内容对应的目标域内容向量。

步骤 563：计算源域内容向量和每个目标域内容向量之间的相似度；

针对用户账号的历史行为对应的源域内容的源域内容向量，计算与多个目标域内容向量的多个相似度。例如，目标域内容向量包括目标域内容向量 1、目标域内容向量 2 和目标域内容向量 3，则计算源域内容向量和目标域内容向量 1 的相似度 1，计算源域内容向量和目标域内容向量 2 的相似度 2，计算源域内容向量和目标域内容向量 3 的相似度 3。

步骤 564：召回相似度超过阈值或排名在前 n 位的目标内容向量对应的目标域内容，作为源域内容对应的目标域内容。

通过利用向量检索工具，召回与源域内容向量的相似度超过阈值或排名在前 n 位的目标内容向量对

应的目标域内容,将该目标内容向量确定为源域内容对应的目标域内容。其中,n的取值为正整数,可以是系统预设的,也可以是根据具体情况随机调整的,如n为1,或n为2,或n为3。例如,n为2,在目标域内容向量包括目标域内容向量1、目标域内容向量2和目标域内容向量3的情况下,包括源域内容向量和目标域内容向量1的相似度1,源域内容向量和目标域内容向量2的相似度2,源域内容向量和目标域内容向量3的相似度3。若相似度1大于相似度3大于相似度2,则召回与源域内容对应的目标域内容1和目标域内容3。

综上所述,本实施例提供的方法,通过获取源域内容向量和多个目标域内容向量,并计算源域内容向量和每个目标域内容向量之间的相似度,召回相似度超过阈值或排名在前n位的目标内容向量对应的目标域内容作为源域内容对应的目标域内容,可以更准确地推荐与源域内容相关的目标域内容。该目标域内容是基于源域内容获得的,能够避免因用户数据稀疏造成无法推荐或推荐的相似性较低的问题。

在一些实施例中,服务器140中运行有跨域推荐模型,该跨域推荐模型包括源域语义塔和目标域语义塔。如图3所示,跨域推荐模型300,包括源域语义塔320和目标域语义塔340。上述源域语义塔包括级联的源域节点特征提取网络、源域标签特征提取网络、源域拼接层和源域表示层。如图5所示,源域语义塔320包括级联的源域节点特征提取网络321、源域标签特征提取网络322、源域拼接层323和源域表示层324。其中,源域节点特征提取网络321的输入端和源域内容节点连接,源域节点特征提取网络321的输出端和源域拼接层323连接,源域标签特征提取网络322的输出端和源域拼接层323的输入端连接,源域拼接层323的输出端和源域表示层324的输入端连接。

图18是本申请一个示例性实施例提供的跨域推荐方法的示意图。上述步骤561还包括如下至少一个步骤:

步骤5611:通过源域节点特征提取网络对源域内容进行特征提取,得到第一内容向量;

在一些实施例中,上述源域节点特征提取网络包括级联的源域节点输入层、源域节点嵌入层和源域节点表示层。如图7所示,源域节点特征提取网络321包括级联的源域节点输入层3211、源域节点嵌入层3212和源域节点表示层3213。其中,源域节点输入层3211的输入端和源域内容节点连接,源域节点输入层3211的输出端和源域节点嵌入层3212的输入端连接,源域节点嵌入层3212的输出端和源域节点表示层3213的输入端连接,源域节点表示层3213的输出端和源域拼接层323的输入端连接。

通过源域节点输入层将源域内容节点的第一内容特征输入。将源域节点输入层中的第一内容特征输出,通过源域节点嵌入层对第一内容特征进行嵌入表示,得到第一内容的嵌入表征向量。将源域节点嵌入层中的第一内容的嵌入表征向量输出,通过源域节点表示层对第一内容的嵌入表征向量进行学习,得到第一内容向量。

步骤5612:通过源域标签特征提取网络对源域语义标签进行特征提取,得到第一语义标签向量;

在一些实施例中,上述源域标签特征提取网络包括级联的源域语义标签输入层、源域嵌入表征编码器和源域标签表示层。如图8所示,源域标签特征提取网络322包括级联的源域语义标签输入层3221、源域嵌入表征编码器3222和源域标签表示层3223。其中,源域语义标签输入层3221的输入端和源域语义标签连接,源域语义标签输入层3221的输出端和源域嵌入表征编码器3222的输入端连接,源域嵌入表征编码器3222的输出端和源域标签表示层3223的输入端连接,源域标签表示层3223的输出端和源域拼接层323的输入端连接。

通过源域语义标签输入层将源域语义标签的第一语义特征输入。将源域语义标签输入层中的第一语义特征输出,通过源域嵌入表征编码器对第一语义特征进行嵌入表示,得到第一语义的嵌入表征向量。将源域嵌入表征编码器中的第一语义的嵌入表征向量输出,通过源域标签表示层对第一语义的嵌入表征向量进行学习,得到第一语义标签向量。

步骤5613:通过源域拼接层对第一内容向量和第一语义标签向量进行拼接,得到源域拼接向量;

如图9所示,上述源域语义塔320通过级联的源域节点输入层3211、源域节点嵌入层3212和源域节点表示层3213,输出第一内容向量;通过级联的源域语义标签输入层3221、源域嵌入表征编码器3222和源域标签表示层3223,输出第一语义标签向量。

将第一内容向量和第一语义标签向量输入源域拼接层323,将第一内容向量和第一语义标签向量进行向量合并,得到源域拼接向量。

步骤 5614: 通过源域表示层对源域拼接向量进行特征提取, 得到源域内容向量。

如图 9 所示, 将源域拼接层 323 输出的源域拼接向量输入源域表示层 324, 得到源域内容向量。该源域内容向量是包括源域内容节点和与源域内容节点对应的源域语义标签的该源域内容的完整特征向量。

在一些实施例中, 上述目标域语义塔包括级联的目标域节点特征提取网络、目标域标签特征提取网络、目标域拼接层和目标域表示层。如图 10 所示, 目标域语义塔 340 包括级联的目标域节点特征提取网络 341、目标域标签特征提取网络 342、目标域拼接层 343 和目标域表示层 344。其中, 目标域节点特征提取网络 341 的输入端和目标域内容节点连接, 目标域节点特征提取网络 341 的输出端和目标域拼接层 343 连接, 目标域标签特征提取网络 342 的输出端和目标域拼接层 343 的输入端连接, 目标域拼接层 343 的输出端和目标域表示层 344 的输入端连接。

图 19 是本申请一个示例性实施例提供的跨域推荐方法的示意图。上述步骤 562 还包括如下至少一个步骤:

步骤 5621: 通过目标域节点特征提取网络对目标域内容进行特征提取, 得到第二内容向量;

在一些实施例中, 上述目标域节点特征提取网络包括级联的目标域节点输入层、目标域节点嵌入层和目标域节点表示层。如图 12 所示, 目标域节点特征提取网络 341 包括级联的目标域节点输入层 3411、目标域节点嵌入层 3412 和目标域节点表示层 3413。其中, 目标域节点输入层 3411 的输入端和目标域内容节点连接, 目标域节点输入层 3411 的输出端和目标域节点嵌入层 3412 的输入端连接, 目标域节点嵌入层 3412 的输出端和目标域节点表示层 3413 的输入端连接, 目标域节点表示层 3413 的输出端和目标域拼接层 343 的输入端连接。

通过目标域节点输入层将目标域内容节点的第二内容特征输入。将目标域节点输入层中的第二内容特征输出, 通过目标域节点嵌入层对第二内容特征进行嵌入表示, 得到第二内容的嵌入表征向量。将目标域节点嵌入层中的第二内容的嵌入表征向量输出, 通过目标域节点表示层对第二内容的嵌入表征向量进行学习, 得到第二内容向量。

步骤 5622: 通过目标域标签特征提取网络对目标域语义标签进行特征提取, 得到第二语义标签向量;

在一些实施例中, 上述目标域标签特征提取网络包括级联的目标域语义标签输入层、目标域嵌入表征编码器和目标域标签表示层。如图 13 所示, 目标域标签特征提取网络 342 包括级联的目标域语义标签输入层 3421、目标域嵌入表征编码器 3422 和目标域标签表示层 3423。其中, 目标域语义标签输入层 3421 的输入端和目标域语义标签连接, 目标域语义标签输入层 3421 的输出端和目标域嵌入表征编码器 3422 的输入端连接, 目标域嵌入表征编码器 3422 的输出端和目标域标签表示层 3423 的输入端连接, 目标域标签表示层 3423 的输出端和目标域拼接层 343 的输入端连接。

通过目标域语义标签输入层将目标域语义标签的第二语义特征输入。将目标域语义标签输入层中的第二语义特征输出, 通过目标域嵌入表征编码器对第二语义特征进行嵌入表示, 得到第二语义的嵌入表征向量。将目标域嵌入表征编码器中的第二语义的嵌入表征向量输出, 通过目标域标签表示层对第二语义的嵌入表征向量进行学习, 得到第二语义标签向量。

步骤 5623: 通过目标域拼接层对第二内容向量和第二语义标签向量进行拼接, 得到目标域拼接向量;

如图 14 所示, 上述目标域语义塔 340 通过级联的目标域节点输入层 3411、目标域节点嵌入层 3412 和目标域节点表示层 3413, 输出第二内容向量; 通过级联的目标域语义标签输入层 3421、目标域嵌入表征编码器 3422 和目标域标签表示层 3423, 输出第二语义标签向量。

将第二内容向量和第二语义标签向量输入目标域拼接层 343, 将第二内容向量和第二语义标签向量进行向量合并, 得到目标域拼接向量。

步骤 5624: 通过目标域表示层对目标域拼接向量进行特征提取, 得到目标域内容向量。

如图 14 所示, 将目标域拼接层 343 输出的目标域拼接向量输入目标域表示层 344, 得到目标域内容向量。该目标域内容向量是包括目标域内容节点和与目标域内容节点对应的目标域语义标签的该目标域内容的完整特征向量。

综上所述, 本实施例提供的方法, 通过源域语义塔和目标域语义塔, 实现了将第一内容向量和对应的第一语义标签向量进行拼接, 得到包括源域内容的完整特征的源域内容向量; 以及将第二内容向量和对应的第二语义标签向量进行拼接, 得到包括目标域内容的完整特征的目标域内容向量。可以获得更完

整的特征表示，从而提高了推荐系统的准确性和鲁棒性。

在一些实施例中，上述源域节点特征提取网络包括级联的源域节点输入层、源域节点嵌入层和源域节点表示层。如图 7 所示，源域节点特征提取网络 321 包括级联的源域节点输入层 3211、源域节点嵌入层 3212 和源域节点表示层 3213。其中，源域节点输入层 3211 的输入端和源域内容节点连接，源域节点输入层 3211 的输出端和源域节点嵌入层 3212 的输入端连接，源域节点嵌入层 3212 的输出端和源域节点表示层 3213 的输入端连接，源域节点表示层 3213 的输出端和源域拼接层 323 的输入端连接。

图 20 是本申请一个示例性实施例提供的跨域推荐方法的示意图。上述步骤 5611 还包括如下至少一个步骤：

步骤 56111：通过源域节点输入层将源域内容的内容特征输入；

10 将源域内容的内容特征作为源域节点输入层的输入，该源域内容的内容特征包括与源域内容对应的特征，或对应源域内容的内容属性。

示例性的，假设源域内容为某电影，则该源域内容的内容特征包括某电影所属的领域、某电影的主题、某电影的时长中的至少一种，如该电影属于动作片，或该电影属于文艺片。假设源域内容为某新闻，则该源域内容的内容特征包括某新闻的关联对象、某新闻所属的领域、某新闻出现的频次中的至少一种，
15 如该新闻属于财经类，或该新闻属于生活类。

步骤 56112：通过源域节点嵌入层对源域内容的内容特征进行嵌入表示，得到源域内容的嵌入表征向量；

20 将源域内容的内容特征经源域节点输入层输入到源域节点嵌入层，该源域节点嵌入层用于对输入对象进行嵌入表示，即基于神经网络模型生成对输入对象的分布式表示。主要作用是将原有对象的高维稀疏向量转化为低维、稠密向量，使得这些低维、稠密向量能够表达相应对象的一个或多个特征。同时，不同向量之间的距离能反应对象之间的相似性。通过对源域内容的内容特征进行嵌入表示，得到源域内容的嵌入表征向量。

步骤 56113：通过源域节点表示层对源域内容的嵌入表征向量进行学习，得到第一内容向量。

25 将源域内容的嵌入表征向量输入源域节点表示层，该源域节点表示层用于对输入对象进行学习，一般使用多层叠加的全连接层实现。通过对源域内容的嵌入表征向量进行学习，得到第一内容向量。

图 21 是本申请一个示例性实施例提供的跨域推荐方法的示意图。上述步骤 5612 还包括如下至少一个步骤：

步骤 56121：通过源域语义标签输入层将源域语义标签的语义特征输入；

30 将源域语义标签的语义特征作为源域语义标签输入层的输入，该源域语义标签的语义特征包括与源域内容对应的语义标签特征。

示例性的，假设源域内容为某电影，则源域语义标签包括某电影的名称、某电影的角色名、某电影的制作人中的至少一种。假设源域内容为某新闻，则该源域语义标签包括某新闻的标题、某新闻的通讯作者、某新闻中的关键词中的至少一种。

35 **步骤 56122：**通过源域嵌入表征编码器对源域语义标签的语义特征进行嵌入表示，得到源域语义标签的嵌入表征向量；

40 将源域语义标签的语义特征经源域语义标签输入层输入到源域嵌入表征编码器，该源域嵌入表征编码器用于对输入对象进行嵌入表示，即基于神经网络模型生成对输入对象的分布式表示。主要作用是将原有对象的高维稀疏向量转化为低维、稠密向量，使得这些低维、稠密向量能够表达相应对象的一个或多个特征。同时，不同向量之间的距离能反应对象之间的相似性。通过对源域语义标签的语义特征进行嵌入表示，得到源域语义标签的嵌入表征向量。

步骤 56123：通过源域标签表示层对源域语义标签的嵌入表征向量进行学习，得到第一语义标签向量。

将源域语义标签的嵌入表征向量输入源域标签表示层，该源域标签表示层用于对输入对象进行学习，一般使用多层叠加的全连接层实现。通过对源域语义标签的嵌入表征向量进行学习，得到第一语义标签向量。

45 图 22 是本申请一个示例性实施例提供的跨域推荐方法的示意图。上述步骤 5621 还包括如下至少一个步骤：

步骤 56211: 通过目标域节点输入层将目标域内容的内容特征输入:

将目标域内容的内容特征作为目标域节点输入层的输入, 该目标域内容的内容特征包括与目标域内容对应的特征, 或对应目标域内容的内容属性。

示例性的, 假设目标域内容为某电影, 则该目标域内容的内容特征包括某电影所属的领域、某电影的主题、某电影的时长中的至少一种, 如该电影属于动作片, 或该电影属于文艺片。假设目标域内容为某新闻, 则该目标域内容的内容特征包括某新闻的关联对象、某新闻所属的领域、某新闻出现的频次中的至少一种, 如该新闻属于财经类, 或该新闻属于生活类。

步骤 56212: 通过目标域节点嵌入层对目标域内容的内容特征进行嵌入表示, 得到目标域内容的嵌入表征向量;

将目标域内容的内容特征经目标域节点输入层输入到目标域节点嵌入层, 该目标域节点嵌入层用于对输入对象进行嵌入表示, 即基于神经网络模型生成对输入对象的分布式表示。主要作用是将原有对象的高维稀疏向量转化为低维、稠密向量, 使得这些低维、稠密向量能够表达相应对象的一个或多个特征。同时, 不同向量之间的距离能反应对象之间的相似性。通过对目标域内容的内容特征进行嵌入表示, 得到目标域内容的嵌入表征向量。

步骤 56213: 通过目标域节点表示层对目标域内容的嵌入表征向量进行学习, 得到第二内容向量。

将目标域内容的嵌入表征向量输入目标域节点表示层, 该目标域节点表示层用于对输入对象进行学习, 一般使用多层叠加的全连接层实现。通过对目标域内容的嵌入表征向量进行学习, 得到第二内容向量。

图 23 是本申请一个示例性实施例提供的跨域推荐方法的示意图。上述步骤 5622 还包括如下至少一个步骤:

步骤 56221: 通过目标域语义标签输入层将目标域语义标签的语义特征输入;

将目标域语义标签的语义特征作为目标域语义标签输入层的输入, 该目标域语义标签的语义特征包括与目标域内容对应的语义标签特征。

示例性的, 假设目标域内容为某电影, 则目标域语义标签包括某电影的名称、某电影的角色名、某电影的制作者中的至少一种。假设目标域内容为某新闻, 则该目标域语义标签包括某新闻的标题、某新闻的通讯作者、某新闻中的关键词中的至少一种。

步骤 56222: 通过目标域嵌入表征编码器对目标域语义标签的语义特征进行嵌入表示, 得到目标域语义标签的嵌入表征向量;

将目标域语义标签的语义特征经目标域语义标签输入层输入到目标域嵌入表征编码器, 该目标域嵌入表征编码器用于对输入对象进行嵌入表示, 即基于神经网络模型生成对输入对象的分布式表示。主要作用是将原有对象的高维稀疏向量转化为低维、稠密向量, 使得这些低维、稠密向量能够表达相应对象的一个或多个特征。同时, 不同向量之间的距离能反应对象之间的相似性。通过对目标域语义标签的语义特征进行嵌入表示, 得到目标域语义标签的嵌入表征向量。

步骤 56223: 通过目标域标签表示层对目标域语义标签的嵌入表征向量进行学习, 得到第二语义标签向量。

将目标域语义标签的嵌入表征向量输入目标域标签表示层, 该目标域标签表示层用于对输入对象进行学习, 一般使用多层叠加的全连接层实现。通过对目标域语义标签的嵌入表征向量进行学习, 得到第二语义标签向量。

综上所述, 本实施例提供的方法, 通过节点输入层、节点嵌入层、节点表示层、语义标签输入层、嵌入表征编码器和、标签表示层, 对输入的内容和与内容对应的语义标签进行特征提取, 可以得到更全面、更准确的内容向量和语义标签向量, 使模型充分学习到源域内容的内容特征和目标域内容的内容特征, 从而在推荐过程中更好地考虑内容的语义信息和内容之间的关联, 实现更好的推荐。

本申请实施例提出了一种新的跨域推荐的图神经网络模型, 除了引入图中节点的拓扑关系, 还引入了节点关联的语义标签以增强节点的表征能力。同时集中训练源域与目标域中节点的嵌入表征能力, 克服源域或目标域中由于数据稀疏造成的影响。本申请实施例可以广泛适用根据用户账号的历史行为进行的推荐服务, 适用相同域之间的推荐, 如从长视频到长视频的推荐、从新闻到新闻的推荐; 还适用不同

域之间的推荐，如从长视频到短视频的推荐、从短视频到长视频的推荐、从视频到商品的推荐，从视频到音乐的推荐等。

本申请的主要流程如下：根据用户账号的历史行为构建出一个从源域内容到目标域内容的异质网络。该异质网络包括由源域内容构建的同现网络、由目标域内容构建的同现网络、源域内容节点与源域内容
5 对应的语义标签之间的标签二部图、目标域内容节点与目标域内容对应的语义标签之间的标签二部图、源域内容节点与目标域内容节点之间的节点二部图。

源域同现网络以源域视频同现网络为例，将每个用户在源域中所观看的视频按照观看时间排序。同时预先设定一个视频同现的时间窗口，如 10 分钟、20 分钟等均可，同时假设同一用户在同一时间窗口内观看的所有视频的内容具有一定的相似性。根据该时间窗口对所有用户的视频浏览序列构建同现网络。
10 对同一用户、同一时间窗口内的视频，如果没有边则连接一条边，如果已经有边则不断累加边上的权重。然后，将所有用户的同现边及边的同现次数进行累加，就得到整个源域视频同现网络。最后，为了过滤掉那些相关性不高的视频间的边，再设定一个同现权重的阈值，如 3 次，这样就得到了源域视频的同现网络了。对于目标域同现网络的构建，采用与构建源域同现网络相同的方法即可获得，只是目标域上如果用户账号的历史行为稀疏，在同现阈值的设定上可以调整，无需与源域上的阈值一致。

15 在源域同现网络和目标域同现网络构建完成后，即可通过常见的嵌入表示算法得到同现网络中节点的嵌入表征。

标签二部图只需要将源域与目标域中的内容节点与其在推荐系统中已有的语义标签关联即可。通过使用常见的词向量方法即可得到语义标签的嵌入表征。相关方法在很多自然语言处理技术（Natural Language Processing, NLP）的文献中已经提及，本文在此就不再赘述。

20 节点二部图通过选择同时浏览源域内容和目标域内容的用户，将每个用户浏览的内容按照时间进行排序，得到一个内容序列。该内容序列中既包含源域内容，也包含目标域内容。如果内容序列中邻接的两个内容一个属于源域，另一个属于目标域，则在这两个跨域节点间连接一条边，同时累加边上的同现权重。该二部图将后续跨域推荐模型训练的正例样本使用，而跨域推荐模型的负例样本则通过随机负采样得到。

25 在构建出异质网络后，本申请将构建从源域到目标域的图神经网络模型，即本申请实施例所述的跨域推荐模型，用于学习源域内容与目标域内容的相关性，从而最终得到源域内容节点与目标域内容节点的嵌入表征。

该跨域推荐模型的主要结构包括源域语义塔与目标域语义塔，其中源域语义塔与目标域语义塔的结构相同，因此以源域语义塔为例进行介绍。

30 经过节点输入层和语义标签输入层，分别用于输入节点以及节点关联的语义标签的合集。其中，节点输入层与语义标签输入层输入的均是经过标识化后节点以及语义标签，以便相关特征可以在节点嵌入层与嵌入表征编码器中取得嵌入向量。通过节点嵌入层把节点映射为低维、稠密的嵌入表示，从而提供给后续的深度学习网络。通过嵌入表征编码器把与内容节点连接的语义标签的文本特征向量映射为低维、稠密的嵌入表示，从而提供给后续的深度学习网络。节点嵌入层和嵌入表征编码器可以是各类详解循环神经网络(Recurrent Neural Network, RNN)，也可以是多层的叠加与组合。节点表示层和语义标签表示层
35 分别用于学习前面节点嵌入层与嵌入表征编码器输入的节点与语义标签集合的嵌入表征向量，一般使用多层叠加的全连接层实现。拼接层的功能是将节点的语义表示与该节点的图嵌入表征相拼接，从而得到该节点的完整嵌入表征。

40 当跨域推荐模型训练完毕后，即可分别用源域语义塔和目标域语义塔抽取源域内容与目标域内容的嵌入表征。再通过向量检索工具，召回与每个内容向量邻近的目标向量，从而实现跨域推荐。

本申请实施例能够实现的有益效果至少包括：

本申请是较为轻量化的跨域推荐模型，除了引入内容节点在各自域的同现图中的嵌入表征，还引入了节点关联的语义标签以增强了对应节点的特征能力，提高跨域推荐模型的准确性。同时使用单独的语义塔以训练数据稀疏的目标域节点的嵌入表征能力，可以极大克服目标域的数据稀疏问题造成的影响。
45 本申请不要求源域与目标域的节点的语义标签在同一特征空间。可以广泛应用于源域与目标域是不同域的跨域推荐中。

图 24 示出了本申请一个示例性实施例提供的跨域推荐模型的训练装置的框图。该装置包括：

构建模块 2410，用于构建异质网络，异质网络包括样本源域内容节点和样本目标域内容节点之间的节点二部图、样本源域内容节点和样本源域语义标签之间的第一标签二部图和样本目标域内容节点和样本目标域语义标签之间的第二标签二部图。

- 5 生成模块 2420，用于基于节点二部图中存在连接边的样本源域内容节点和样本目标域内容节点、样本源域内容节点在第一标签二部图中对应的样本源域语义标签，以及样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签，生成训练样本。

训练模块 2430，用于基于训练样本对跨域推荐模型进行训练。

跨域推荐模型包括源域语义塔、目标域语义塔和匹配层。

- 10 训练模块 2430，还用于针对任一训练样本，将样本源域内容节点和样本源域内容节点对应的样本源域语义标签输入源域语义塔，得到样本源域内容向量。

训练模块 2430，还用于将样本目标域内容节点和样本目标域内容节点对应的样本目标域语义标签输入目标域语义塔，得到样本目标域内容向量。

训练模块 2430，还用于将样本源域内容向量和样本目标域内容向量输入匹配层，得到预测相似度。

- 15 训练模块 2430，还用于计算预测相似度和真实样本相似度之间的误差损失。

训练模块 2430，还用于基于误差损失对跨域推荐模型进行训练。

源域语义塔包括级联的源域节点特征提取网络、源域标签特征提取网络、源域拼接层和源域表示层，目标域语义塔包括级联的目标域节点特征提取网络、目标域标签特征提取网络、目标域拼接层和目标域表示层。

- 20 训练模块 2430，还用于将样本源域内容节点输入源域节点特征提取网络，得到第一样本内容向量。

训练模块 2430，还用于将样本源域语义标签输入源域标签特征提取网络，得到第一样本语义标签向量。

训练模块 2430，还用于将第一样本内容向量和第一样本语义标签向量输入源域拼接层，得到样本源域拼接向量。

- 25 训练模块 2430，还用于将样本源域拼接向量输入源域表示层，得到样本源域内容向量。

训练模块 2430，还用于将样本目标域内容节点输入目标域节点特征提取网络，得到第二样本内容向量；将样本目标域语义标签输入目标域标签特征提取网络，得到第二样本语义标签向量。

训练模块 2430，还用于将第二样本内容向量和第二样本语义标签向量输入目标域拼接层，得到样本目标域拼接向量。

- 30 训练模块 2430，还用于将样本目标域拼接向量输入目标域表示层，得到样本目标域内容向量。

生成模块 2420，还用于将样本源域内容节点以及在第一标签二部图中对应的样本源域语义标签确定为样本源域数据。

生成模块 2420，还用于将样本目标域内容节点以及在第二标签二部图中对应的目标域语义标签确定为样本目标域数据。

- 35 生成模块 2420，还用于基于连接边上的权重，确定样本源域数据和样本目标域数据之间的真实样本相似度，权重是基于样本源域内容节点和样本目标域内容节点在一个时间段内同时出现的次数确定的。

生成模块 2420，还用于将样本源域数据和样本目标域数据确定为训练样本。

异质网络还包括基于具有同现关系的样本源域内容节点构建的源域同现网络，基于具有同现关系的样本目标域内容节点构建的目标域同现网络。

- 40 节点二部图中包括存在连接边的第一样本源域内容节点和第一样本目标域内容节点。

生成模块 2420，还用于在源域同现网络中存在与第一样本源域内容节点具有同现关系的第二样本源域内容节点的情况下，使用第二样本源域内容节点和第一样本目标域内容节点，以及第二样本源域内容节点在第一标签二部图中对应的样本源域语义标签和第一样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签，生成训练样本。

- 45 生成模块 2420，还用于在目标域同现网络中存在与第一样本目标域内容节点具有同现关系的第二样本目标域内容节点的情况下，使用第一样本源域内容节点和第二样本目标域内容节点，以及第一样本源

域内容节点在第一标签二部图中对应的样本源域语义标签和第二样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签，生成训练样本。

生成模块 2420，还用于在源域同现网络中存在与第一样本源域内容节点具有同现关系的第二样本源域内容节点且目标域同现网络中存在与第一样本目标域内容节点具有同现关系的第二样本目标域内容节点的情况下，使用第二样本源域内容节点和第二样本目标域内容节点，以及第二样本源域内容节点在第一标签二部图中对应的样本源域语义标签和第二样本目标域内容节点在第二标签二部图中对应的样本目标域语义标签，生成训练样本。

构建模块 2410，用于基于源域中的语义标签系统，构建样本源域内容节点和样本源域语义标签之间的第一标签二部图。

构建模块 2410，还用于基于目标域中的语义标签系统，构建样本目标域内容节点和样本目标域语义标签之间的第二标签二部图。

构建模块 2410，还用于基于多个用户账号在源域和目标域的历史行为，构建样本源域内容节点和样本目标域内容节点之间的节点二部图。

构建模块 2410，还用于基于源域中的语义标签系统，得到样本源域语义标签。

构建模块 2410，还用于将具有对应关系的样本源域语义标签与样本源域内容节点通过连接边进行连接，得到样本源域内容节点和样本源域语义标签之间的第一标签二部图。

构建模块 2410，还用于基于目标域中的语义标签系统，得到样本目标域语义标签。

构建模块 2410，还用于将具有对应关系的样本目标域语义标签与样本目标域内容节点通过连接边进行连接，得到样本目标域内容节点和样本目标域语义标签之间的第二标签二部图。

构建模块 2410，还用于基于多个用户账号在源域的历史行为，确定多个用户账号存在历史互动的样本源域内容。

构建模块 2410，还用于基于多个用户账号在目标域的历史行为，确定多个用户账号存在历史互动的样本目标域内容。

构建模块 2410，还用于基于样本源域内容和样本目标域内容在同一个用户账号下的第一时间段内同时出现的次数，将样本源域内容对应的样本源域内容节点和样本目标域内容对应的样本目标域内容节点通过连接边进行连接，得到样本源域内容节点和样本目标域内容节点之间的节点二部图。

构建模块 2410，还用于基于多个用户账号在源域的历史行为，构建源域同现网络。

构建模块 2410，还用于基于多个用户账号在目标域的历史行为，构建目标域同现网络。

构建模块 2410，还用于基于多个用户账号在源域的历史行为，确定多个用户账号存在历史互动的样本源域内容。

构建模块 2410，还用于基于样本源域内容在同一个用户账号的第二时间段内同时出现的次数，将样本源域内容对应的样本源域内容节点通过连接边进行连接，得到源域同现网络。

构建模块 2410，还用于基于多个用户账号在目标域的历史行为，确定多个用户账号存在历史互动的样本目标域内容。

构建模块 2410，还用于基于样本目标域内容在同一个用户账号的第三时间段内同时出现的次数，将样本目标域内容对应的样本目标域内容节点通过连接边进行连接，得到目标域同现网络。

图 25 示出了本申请一个示例性实施例提供的跨域推荐装置的框图。该装置包括：

获取模块 2510，用于获取用户账号的历史行为。

确定模块 2520，用于基于用户账号的历史行为，确定用户账号存在历史互动的源域内容。

确定模块 2520，还用于基于源域内容向量和目标域内容向量之间的相似性，确定与源域内容对应的目标域内容。

推荐模块 2530，用于向用户账号推荐目标域内容。

源域内容向量是源域内容的特征向量，目标域内容向量是目标域内容的特征向量，源域内容向量是基于源域内容和源域内容在第一标签二部图中对应的源域语义标签构建的，目标域内容向量是基于目标域内容和目标域内容在第二标签二部图中对应的目标域语义标签构建的，第一标签二部图是基于源域内容和源域语义标签构建的，第二标签二部图是基于目标域内容和目标域语义标签构建的。

确定模块 2520, 还用于获取源域内容向量, 源域内容向量是基于第一内容向量和第一语义标签向量构建的, 第一内容向量是源域内容对应的内容向量, 第一语义标签向量是源域内容在第一标签二部图中对应的源域语义标签对应的语义标签向量。

5 确定模块 2520, 还用于获取多个目标域内容向量, 目标域内容向量是基于第二内容向量和第二语义标签向量构建的, 第二内容向量是目标域内容对应的内容向量, 第二语义标签向量是目标域内容在第二标签二部图中对应的目标域语义标签对应的语义标签向量。

确定模块 2520, 还用于计算源域内容向量和每个目标域内容向量之间的相似度。

确定模块 2520, 还用于召回相似度超过阈值或排名在前 n 位的目标内容向量对应的目标域内容, 作为源域内容对应的目标域内容。其中, n 的取值为正整数。

10 服务器运行有跨域推荐模型, 跨域推荐模型包括源域语义塔和目标域语义塔; 源域语义塔包括源域节点特征提取网络、源域标签特征提取网络、源域拼接层和源域表示层, 目标域语义塔包括目标域节点特征提取网络、目标域标签特征提取网络、目标域拼接层和目标域表示层。

确定模块 2520, 还用于通过源域节点特征提取网络对源域内容进行特征提取, 得到第一内容向量。

15 确定模块 2520, 还用于通过源域标签特征提取网络对源域语义标签进行特征提取, 得到第一语义标签向量。

确定模块 2520, 还用于通过源域拼接层对第一内容向量和第一语义标签向量进行拼接, 得到源域拼接向量。

确定模块 2520, 还用于通过源域表示层对源域拼接向量进行特征提取, 得到源域内容向量。

20 确定模块 2520, 还用于通过目标域节点特征提取网络对目标域内容进行特征提取, 得到第二内容向量; 通过目标域标签特征提取网络对目标域语义标签进行特征提取, 得到第二语义标签向量。

确定模块 2520, 还用于通过目标域拼接层对第二内容向量和第二语义标签向量进行拼接, 得到目标域拼接向量。

确定模块 2520, 还用于通过目标域表示层对目标域拼接向量进行特征提取, 得到目标域内容向量。

25 源域节点特征提取网络包括级联的源域节点输入层、源域节点嵌入层和源域节点表示层, 源域标签特征提取网络包括级联的源域语义标签输入层、源域嵌入表征编码器和源域标签表示层。

确定模块 2520, 还用于通过源域节点输入层将源域内容的内容特征输入。

确定模块 2520, 还用于通过源域节点嵌入层对源域内容的内容特征进行嵌入表示, 得到源域内容的嵌入表征向量。

30 确定模块 2520, 还用于通过源域节点表示层对源域内容的嵌入表征向量进行学习, 得到第一内容向量。

确定模块 2520, 还用于通过源域语义标签输入层将源域语义标签的语义特征输入。

确定模块 2520, 还用于通过源域嵌入表征编码器对源域语义标签的语义特征进行嵌入表示, 得到源域语义标签的嵌入表征向量。

35 确定模块 2520, 还用于通过源域标签表示层对源域语义标签的嵌入表征向量进行学习, 得到第一语义标签向量。

目标域节点特征提取网络包括级联的目标域节点输入层、目标域节点嵌入层和目标域节点表示层, 目标域标签特征提取网络包括级联的目标域语义标签输入层、目标域嵌入表征编码器和目标域标签表示层。

确定模块 2520, 还用于通过目标域节点输入层将目标域内容的内容特征输入。

40 确定模块 2520, 还用于通过目标域节点嵌入层对目标域内容的内容特征进行嵌入表示, 得到目标域内容的嵌入表征向量。

确定模块 2520, 还用于通过目标域节点表示层对目标域内容的嵌入表征向量进行学习, 得到第二内容向量。

确定模块 2520, 还用于通过目标域语义标签输入层将目标域语义标签的语义特征输入。

45 确定模块 2520, 还用于通过目标域嵌入表征编码器对目标域语义标签的语义特征进行嵌入表示, 得到目标域语义标签的嵌入表征向量。

确定模块 2520, 还用于通过目标域标签表示层对目标域语义标签的嵌入表征向量进行学习, 得到第二语义标签向量。

图 26 示出了本申请一个示例性实施例提供的计算机设备的结构示意图。该计算机设备可以是终端, 也可以是服务器。示例性的, 计算机设备 2600 包括中央处理单元 (Central Processing Unit, CPU) 2601、包括随机存取存储器 (Random Access Memory, RAM) 2602 和只读存储器 (Read-Only Memory, ROM) 2603 的系统存储器 2604, 以及连接系统存储器 2604 和中央处理单元 2601 的系统总线 2605。所述计算机设备 2600 还包括帮助计算机内的各个器件之间传输信息的基本输入/输出系统 (Input/Output 系统) 2606, 和用于存储操作系统 2613、客户端 2614 和其他程序模块 2615 的大容量存储设备 2607。

在一些实施例中, 所述基本输入/输出系统 2606 包括有用于显示信息的显示器 2608 和用于用户输入信息的诸如鼠标、键盘之类的输入设备 2609。其中所述显示器 2608 和输入设备 2609 都通过连接到系统总线 2605 的输入/输出控制器 2610 连接到中央处理单元 2601。所述基本输入/输出系统 2606 还可以包括输入/输出控制器 2610 以用于接收和处理来自键盘、鼠标、或电子触控笔等多个其他设备的输入。类似地, 输入/输出控制器 2610 还提供输出到显示屏、打印机或其他类型的输出设备。

所述大容量存储设备 2607 通过连接到系统总线 2605 的大容量存储控制器 (未示出) 连接到中央处理单元 2601。所述大容量存储设备 2607 及其相关联的计算机可读介质为计算机设备 2600 提供非易失性存储。也就是说, 所述大容量存储设备 2607 可以包括诸如硬盘或者只读光盘 (Compact Disc Read-Only Memory, CD-ROM) 驱动器之类的计算机可读介质 (未示出)。

所述计算机可读介质可以包括计算机存储介质和通信介质。计算机存储介质包括以用于存储诸如计算机可读指令、数据结构、程序模块或其他数据等信息的任何方法或技术实现的易失性和非易失性、可移动和不可移动介质。计算机存储介质包括 RAM、ROM、可擦除可编程只读存储器 (Erasable Programmable Read-Only Memory, EPROM)、电可擦除可编程只读存储器 (Electrically Erasable Programmable Read-Only Memory, EEPROM)、闪存或其他固态存储其技术, CD-ROM、数字通用光盘 (Digital Versatile Disc, DVD) 或其他光学存储、磁带盒、磁带、磁盘存储或其他磁性存储设备。当然, 本领域技术人员可知所述计算机存储介质不局限于上述几种。上述的系统存储器 2604 和大容量存储设备 2607 可以统称为存储器。

根据本申请的各种实施例, 所述计算机设备 2600 还可以通过诸如因特网等网络连接到网络上的远程计算机运行。也即计算机设备 2600 可以通过连接在所述系统总线 2605 上的网络接口单元 2616 连接到网络 2617, 或者说, 也可以使用网络接口单元 2616 来连接到其他类型的网络或远程计算机系统 (未示出)。

本申请一示例性实施例还提供了一种计算机可读存储介质, 该计算机可读存储介质中存储有至少一段程序, 至少一段程序由处理器加载并执行以实现上述各个方法实施例提供的跨域推荐模型的训练方法和跨域推荐方法。

本申请一示例性实施例还提供了一种计算机程序产品, 计算机程序产品包括至少一段程序, 至少一段程序存储在可读存储介质中; 通信设备的处理器从可读存储介质中读取指令, 处理器执行指令, 使得通信设备执行以实现上述各个方法实施例提供的跨域推荐模型的训练方法和跨域推荐方法。

应当理解的是, 在本文中提及的“多个”是指两个或两个以上。本领域技术人员在考虑说明书及实践这里公开的发明后, 将容易想到本申请的其它实施方案。本申请旨在涵盖本申请的任何变型、用途或者适应性变化, 这些变型、用途或者适应性变化遵循本申请的一般性原理并包括本申请未公开的本技术领域的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的, 本申请的真正范围和精神由下面的权利要求指出。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成, 也可以通过程序来指令相关的硬件完成, 所述的程序可以存储于一种计算机可读存储介质中, 上述提到的存储介质可以是只读存储器, 磁盘或光盘等。

以上所述仅为本申请的可选实施例, 并不用以限制本申请, 凡在本申请的精神和原则之内, 所作的任何修改、等同切换、改进等, 均应包含在本申请的保护范围之内。

权利要求书

1.一种跨域推荐模型的训练方法，其特征在于，所述方法包括：

构建异质网络，所述异质网络包括样本源域内容节点和样本目标域内容节点之间的节点二部图、所述样本源域内容节点和样本源域语义标签之间的第一标签二部图、所述样本目标域内容节点和样本目标域语义标签之间的第二标签二部图；

基于所述节点二部图中存在连接边的所述样本源域内容节点和所述样本目标域内容节点、所述样本源域内容节点在所述第一标签二部图中对应的样本源域语义标签，以及所述样本目标域内容节点在所述第二标签二部图中对应的样本目标域语义标签，生成训练样本；

基于所述训练样本对所述跨域推荐模型进行训练。

2.根据权利要求1所述的方法，其特征在于，所述跨域推荐模型包括源域语义塔、目标域语义塔和匹配层；

所述基于所述训练样本对所述跨域推荐模型进行训练，包括：

针对任一所述训练样本，将所述样本源域内容节点和所述样本源域内容节点对应的所述样本源域语义标签输入所述源域语义塔，得到样本源域内容向量；以及，将所述样本目标域内容节点和所述样本目标域内容节点对应的所述样本目标域语义标签输入所述目标域语义塔，得到样本目标域内容向量；

将所述样本源域内容向量和所述样本目标域内容向量输入所述匹配层，得到预测相似度；

计算所述预测相似度和真实样本相似度之间的误差损失；

基于所述误差损失对所述跨域推荐模型进行训练。

3.根据权利要求2所述的方法，其特征在于，所述源域语义塔包括级联的源域节点特征提取网络、源域标签特征提取网络、源域拼接层和源域表示层，所述目标域语义塔包括级联的目标域节点特征提取网络、目标域标签特征提取网络、目标域拼接层和目标域表示层；

所述将所述样本源域内容节点和所述样本源域内容节点对应的源域语义标签输入所述源域语义塔，得到样本源域内容向量，包括：

将所述样本源域内容节点输入所述源域节点特征提取网络，得到第一样本内容向量；将所述样本源域语义标签输入所述源域标签特征提取网络，得到第一样本语义标签向量；

将所述第一样本内容向量和所述第一样本语义标签向量输入所述源域拼接层，得到样本源域拼接向量；

将所述样本源域拼接向量输入所述源域表示层，得到所述样本源域内容向量；

所述将所述样本目标域内容节点和所述样本目标域内容节点对应的目标域语义标签输入所述目标域语义塔，得到样本目标域内容向量，包括：

将所述样本目标域内容节点输入所述目标域节点特征提取网络，得到第二样本内容向量；将所述样本目标域语义标签输入所述目标域标签特征提取网络，得到第二样本语义标签向量；

将所述第二样本内容向量和所述第二样本语义标签向量输入所述目标域拼接层，得到样本目标域拼接向量；

将所述样本目标域拼接向量输入所述目标域表示层，得到所述样本目标域内容向量。

4.根据权利要求1至3任一所述的方法，其特征在于，所述基于所述节点二部图中存在连接边的所述样本源域内容节点和所述样本目标域内容节点，所述样本源域内容节点在所述第一标签二部图中对应的样本源域语义标签，以及所述样本目标域内容节点在所述第二标签二部图中对应的样本目标域语义标签，生成训练样本，包括：

将所述样本源域内容节点以及在所述第一标签二部图中对应的所述样本源域语义标签确定为样本源域数据；将所述样本目标域内容节点以及在所述第二标签二部图中对应的所述样本目标域语义标签确定为样本目标域数据；

基于所述连接边上的权重，确定所述样本源域数据和所述样本目标域数据之间的真实样本相似度，所述权重是基于所述样本源域内容节点和所述样本目标域内容节点在一个时间段内同时出现的次数确定的；

将所述样本源域数据、所述样本目标域数据和所述真实样本相似度确定为所述训练样本。

5.根据权利要求1至3任一所述的方法,其特征在于,所述异质网络还包括:基于具有同现关系的所述样本源域内容节点构建的源域同现网络,基于具有同现关系的所述样本目标域内容节点构建的目标域同现网络;

5 所述节点二部图中包括存在连接边的第一样本源域内容节点和第一样本目标域内容节点,所述方法还包括:

在所述源域同现网络中存在与所述第一样本源域内容节点具有同现关系的第二样本源域内容节点的情况下,使用所述第二样本源域内容节点和所述第一样本目标域内容节点,以及所述第二样本源域内容节点在所述第一标签二部图中对应的所述样本源域语义标签和所述第一样本目标域内容节点在所述第二标签二部图中对应的所述样本目标域语义标签,生成所述训练样本;

10 在所述目标域同现网络中存在与所述第一样本目标域内容节点具有同现关系的第二样本目标域内容节点的情况下,使用所述第一样本源域内容节点和所述第二样本目标域内容节点,以及所述第一样本源域内容节点在所述第一标签二部图中对应的所述样本源域语义标签和所述第二样本目标域内容节点在所述第二标签二部图中对应的所述样本目标域语义标签,生成所述训练样本;

15 在所述源域同现网络中存在与所述第一样本源域内容节点具有同现关系的第二样本源域内容节点且所述目标域同现网络中存在与所述第一样本目标域内容节点具有同现关系的第二样本目标域内容节点的情况下,使用所述第二样本源域内容节点和所述第二样本目标域内容节点,以及所述第二样本源域内容节点在所述第一标签二部图中对应的所述样本源域语义标签和所述第二样本目标域内容节点在所述第二标签二部图中对应的所述样本目标域语义标签,生成所述训练样本。

6.根据权利要求1至3任一所述的方法,其特征在于,所述方法还包括:

20 基于源域中的语义标签系统,构建所述样本源域内容节点和所述样本源域语义标签之间的所述第一标签二部图;

基于目标域中的语义标签系统,构建所述样本目标域内容节点和所述样本目标域语义标签之间的所述第二标签二部图;

25 基于多个用户账号在所述源域和所述目标域的历史行为,构建所述样本源域内容节点和所述样本目标域内容节点之间的节点二部图。

7.根据权利要求6所述的方法,其特征在于,所述基于源域中的语义标签系统,构建所述样本源域内容节点和所述样本源域语义标签之间的所述第一标签二部图,包括:

基于所述源域中的所述语义标签系统,得到所述样本源域语义标签;

30 将具有对应关系的所述样本源域语义标签与所述样本源域内容节点通过连接边进行连接,得到所述样本源域内容节点和样本源域语义标签之间的所述第一标签二部图。

8.根据权利要求6所述的方法,其特征在于,所述基于目标域中的语义标签系统,构建所述样本目标域内容节点和所述样本目标域语义标签之间的所述第二标签二部图,包括:

基于所述目标域中的所述语义标签系统,得到所述样本目标域语义标签;

35 将具有对应关系的所述样本目标域语义标签与所述样本目标域内容节点通过连接边进行连接,得到所述样本目标域内容节点和样本目标域语义标签之间的所述第二标签二部图。

9.根据权利要求6所述的方法,其特征在于,所述基于多个用户账号在所述源域和所述目标域的历史行为,构建所述样本源域内容节点和样本目标域内容节点之间的节点二部图,包括:

基于所述多个用户账号在所述源域的所述历史行为,确定所述多个用户账号存在历史互动的样本源域内容;

40 基于所述多个用户账号在所述目标域的所述历史行为,确定所述多个用户账号存在历史互动的样本目标域内容;

基于所述样本源域内容和所述样本目标域内容在同一个所述用户账号下的第一时间段内同时出现的次数,将所述样本源域内容对应的所述样本源域内容节点和所述样本目标域内容对应的所述样本目标域内容节点通过连接边进行连接,得到所述样本源域内容节点和样本目标域内容节点之间的节点二部图。

45 10.根据权利要求6所述的方法,其特征在于,所述方法还包括:

基于多个用户账号在源域的历史行为,构建源域同现网络;

基于多个用户账号在目标域的历史行为，构建目标域同现网络。

11.一种跨域推荐方法，其特征在于，所述方法包括：

获取用户账号的历史行为；

基于所述用户账号的所述历史行为，确定所述用户账号存在历史互动的源域内容；

5 基于源域内容向量和目标域内容向量之间的相似性，确定与所述源域内容对应的目标域内容；

向所述用户账号推荐所述目标域内容；

其中，所述源域内容向量是所述源域内容的特征向量，所述目标域内容向量是所述目标域内容的特征向量，所述源域内容向量是基于所述源域内容和所述源域内容在第一标签二部图中对应的源域语义标签构建的，所述目标域内容向量是基于所述目标域内容和所述目标域内容在第二标签二部图中对应的目标域语义标签构建的，所述第一标签二部图是基于所述源域内容和所述源域语义标签构建的，所述第二标签二部图是基于所述目标域内容和所述目标域语义标签构建的。

12.根据权利要求 11 所述的方法，其特征在于，所述基于源域内容向量和目标域内容向量之间的相似性，确定与所述源域内容对应的目标域内容，包括：

获取所述源域内容向量，所述源域内容向量是基于第一内容向量和第一语义标签向量构建的，所述第一内容向量是所述源域内容对应的内容向量，所述第一语义标签向量是所述源域内容在所述第一标签二部图中对应的所述源域语义标签对应的语义标签向量；

获取多个所述目标域内容向量，所述目标域内容向量是基于第二内容向量和第二语义标签向量构建的，所述第二内容向量是所述目标域内容对应的内容向量，所述第二语义标签向量是所述目标域内容在所述第二标签二部图中对应的所述目标域语义标签对应的语义标签向量；

20 计算所述源域内容向量和每个所述目标域内容向量之间的相似度；

召回所述相似度超过阈值或排名在前 n 位的目标内容向量对应的目标域内容，作为所述源域内容对应的所述目标域内容；

其中，n 的取值为正整数。

13.根据权利要求 12 所述的方法，其特征在于，服务器运行有跨域推荐模型，所述跨域推荐模型包括源域语义塔和目标域语义塔；所述源域语义塔包括源域节点特征提取网络、源域标签特征提取网络、源域拼接层和源域表示层，所述目标域语义塔包括目标域节点特征提取网络、目标域标签特征提取网络、目标域拼接层和目标域表示层；

所述获取所述源域内容向量，包括：

通过所述源域节点特征提取网络对所述源域内容进行特征提取，得到所述第一内容向量；通过所述源域标签特征提取网络对所述源域语义标签进行特征提取，得到所述第一语义标签向量；

通过所述源域拼接层对所述第一内容向量和所述第一语义标签向量进行拼接，得到源域拼接向量；

通过所述源域表示层对所述源域拼接向量进行特征提取，得到所述源域内容向量；

所述获取所述目标域内容向量，包括：

通过所述目标域节点特征提取网络对所述目标域内容进行特征提取，得到所述第二内容向量；通过所述目标域标签特征提取网络对所述目标域语义标签进行特征提取，得到所述第二语义标签向量；

通过所述目标域拼接层对所述第二内容向量和所述第二语义标签向量进行拼接，得到目标域拼接向量；

通过所述目标域表示层对所述目标域拼接向量进行特征提取，得到所述目标域内容向量。

14.根据权利要求 13 所述的方法，其特征在于，所述源域节点特征提取网络包括级联的源域节点输入层、源域节点嵌入层和源域节点表示层，所述源域标签特征提取网络包括级联的源域语义标签输入层、源域嵌入表征编码器和源域标签表示层；

所述通过所述源域节点特征提取网络对所述源域内容进行特征提取，得到所述第一内容向量，包括：

通过所述源域节点输入层将所述源域内容的内容特征输入；

通过所述源域节点嵌入层对所述源域内容的内容特征进行嵌入表示，得到所述源域内容的嵌入表征向量；

通过所述源域节点表示层对所述源域内容的嵌入表征向量进行学习，得到所述第一内容向量；

- 所述通过所述源域标签特征提取网络对所述源域语义标签进行特征提取，得到所述第一语义标签向量，包括：
- 通过所述源域语义标签输入层将所述源域语义标签的语义特征输入；
- 通过所述源域嵌入表征编码器对所述源域语义标签的语义特征进行嵌入表示，得到所述源域语义标签的嵌入表征向量；
- 通过所述源域标签表示层对所述源域语义标签的嵌入表征向量进行学习，得到所述第一语义标签向量。
- 15.根据权利要求 13 所述的方法，其特征在于，所述目标域节点特征提取网络包括级联的目标域节点输入层、目标域节点嵌入层和目标域节点表示层，所述目标域标签特征提取网络包括级联的目标域语义标签输入层、目标域嵌入表征编码器和目标域标签表示层；
- 所述通过所述目标域节点特征提取网络对所述目标域内容进行特征提取，得到所述第二内容向量，包括：
- 通过所述目标域节点输入层将所述目标域内容的内容特征输入；
- 通过所述目标域节点嵌入层对所述目标域内容的内容特征进行嵌入表示，得到所述目标域内容的嵌入表征向量；
- 通过所述目标域节点表示层对所述目标域内容的嵌入表征向量进行学习，得到所述第二内容向量；
- 所述通过所述目标域标签特征提取网络对所述目标域语义标签进行特征提取，得到所述第二语义标签向量，包括：
- 通过所述目标域语义标签输入层将所述目标域语义标签的语义特征输入；
- 通过所述目标域嵌入表征编码器对所述目标域语义标签的语义特征进行嵌入表示，得到所述目标域语义标签的嵌入表征向量；
- 通过所述目标域标签表示层对所述目标域语义标签的嵌入表征向量进行学习，得到所述第二语义标签向量。
- 16.一种跨域推荐模型的训练装置，其特征在于，所述装置包括：
- 构建模块，用于构建异质网络，所述异质网络包括样本源域内容节点和样本目标域内容节点之间的节点二部图、所述样本源域内容节点和样本源域语义标签之间的第一标签二部图、所述样本目标域内容节点和样本目标域语义标签之间的第二标签二部图；
- 生成模块，用于基于所述节点二部图中存在连接边的所述样本源域内容节点和所述样本目标域内容节点、所述样本源域内容节点在所述第一标签二部图中对应的样本源域语义标签，以及所述样本目标域内容节点在所述第二标签二部图中对应的样本目标域语义标签，生成训练样本；
- 训练模块，用于基于所述训练样本对所述跨域推荐模型进行训练。
- 17.一种跨域推荐装置，其特征在于，所述装置包括：
- 获取模块，用于获取用户账号的历史行为；
- 确定模块，用于基于所述用户账号的所述历史行为，确定所述用户账号存在历史互动的源域内容；
- 所述确定模块，还用于基于源域内容向量和目标域内容向量之间的相似性，确定与所述源域内容对应的目标域内容；
- 推荐模块，用于向所述用户账号推荐所述目标域内容；
- 其中，所述源域内容向量是所述源域内容的特征向量，所述目标域内容向量是所述目标域内容的特征向量，所述源域内容向量是基于所述源域内容和所述源域内容在第一标签二部图中对应的源域语义标签构建的，所述目标域内容向量是基于所述目标域内容和所述目标域内容在第二标签二部图中对应的目标域语义标签构建的，所述第一标签二部图是基于所述源域内容和所述源域语义标签构建的，所述第二标签二部图是基于所述目标域内容和所述目标域语义标签构建的。
- 18.一种计算机设备，其特征在于，所述计算机设备包括：处理器和存储器，所述存储器中存储有至少一段程序，所述至少一段程序由所述处理器加载并执行，以实现如权利要求 1 至 10 任一项所述的跨域推荐模型的训练方法和权利要求 11 至 15 任一项所述的跨域推荐方法。
- 19.一种计算机可读存储介质，其特征在于，所述可读存储介质中存储有至少一段程序，所述至少一

段程序由处理器加载并执行以实现如权利要求 1 至 10 任一项所述的跨域推荐模型的训练方法和权利要求 11 至 15 任一项所述的跨域推荐方法。

20. 一种计算机程序产品，其特征在于，所述计算机程序产品包括至少一段程序，所述至少一段程序存储在计算机可读存储介质中；通信设备的处理器从所述计算机可读存储介质中读取所述至少一段程序，
- 5 所述处理器执行所述至少一段程序，使得所述通信设备执行如权利要求 1 至 10 任一项所述的跨域推荐模型的训练方法和权利要求 11 至 15 任一项所述的跨域推荐方法。