(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2020 年 7 月 16 日 (16.07.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/143186 A1

(51) 国际专利分类号:
G06K 9/62 (2006.01) *G06F 16/951* (2019.01)
G06F 16/36 (2019.01)

(21) 国际申请号: PCT/CN2019/093138

(22) 国际申请日: 2019 年 6 月 27 日 (27.06.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910023778.7 2019年1月10日 (10.01.2019) CN

(71) 申请人: 平安科技（深圳）有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(72) 发明人: 金戈 (JIN, Ge); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 徐亮 (XU, Liang); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市精英专利事务所 (SHENZHEN TALENT PATENT SERVICE); 中国广东省深圳市福田区深南中路6009号绿景广场B栋20层B, Guangdong 518000 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU,

(54) **Title:** RECOMMENDATION SYSTEM TRAINING METHOD AND APPARATUS, AND COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 推荐系统训练方法、装置、计算机设备及存储介质

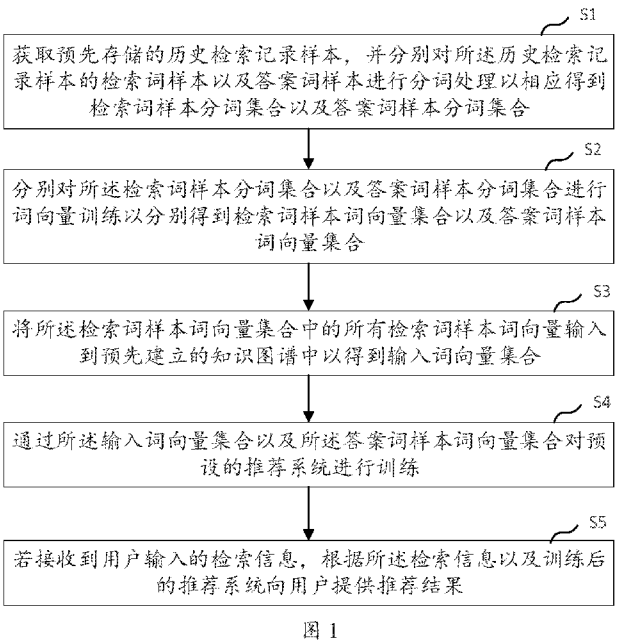

图 1

S1 Obtain pre-stored historical search record samples, and perform word segmentation processing on the search word sample and answer word sample in the historical search record samples respectively to obtain a search word sample segmented word set and an answer word sample word segmented word accordingly

S2 Perform word vector training on the search word sample segmented word set and the answer word sample segmented word set respectively to obtain a search word sample word vector set and an answer word sample word vector set respectively

S3 Input all the search word sample word vectors in the search word sample word vector set into a pre-established knowledge graph to obtain an input word vector set

S4 Train a preset recommendation system by means of the input word vector set and the answer word sample word vector set

S5 If receiving retrieval information input by a user, provide a recommendation result for the user according to the retrieval information and the trained recommendation system

(57) **Abstract:** Disclosed in embodiments of the present application are a recommendation system training method and apparatus, and a computer device and a storage medium. The method relates to an artificial intelligence technology. The method comprises: performing word segmentation processing on a search word sample and an answer word sample to obtain a search word sample segmented word set and an answer word sample segmented word set; performing word vector training on the search word sample segmented word set and the answer word sample segmented word set to obtain a search word sample word vector set and an answer word sample word

WO 2020/143186 A1

[见续页]

CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

vector set; inputting the search word sample word vectors in the search word sample word vector set into a knowledge graph to obtain an input word vector set; training a recommendation system by means of the input word vector set and the answer word sample word vector set; and if receiving retrieval information input by a user, providing a recommendation result for the user according to the trained recommendation system.

(57) 摘要：本申请实施例公开了一种推荐系统训练方法、装置、计算机设备及存储介质。其中，该方法属于人工智能技术，该方法包括：对检索词样本以及答案词样本进行分词处理以得到检索词样本分词集合以及答案词样本分词集合；对检索词样本分词集合以及答案词样本分词集合进行词向量训练以得到检索词样本词向量集合以及答案词样本词向量集合；将检索词样本词向量集合中的检索词样本词向量输入到知识图谱中以得到输入词向量集合；通过输入词向量集合以及答案词样本词向量集合对推荐系统进行训练；若接收到用户输入的检索信息，根据训练后的推荐系统向用户提供推荐结果。

推荐系统训练方法、装置、计算机设备及存储介质

本申请要求于2019年1月10日提交中国专利局、申请号为201910023778.7、申请名称为“推荐系统训练方法、装置、计算机设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及人工智能技术领域，尤其涉及一种推荐系统训练方法、装置、计算机设备及存储介质。

背景技术

随着电子商务的发展，推荐系统的应用越来越广泛。推荐系统能够根据用户输入的检索信息向用户推荐个性化的结果。为了使得推荐的结果更加准确，一般需要大量的数据对推荐系统进行训练。

然而由于用户的数据体量较少，导致推荐系统的训练效果不好，常出现过拟合的现象。进一步地，导致推荐系统的推荐结果不准确。

发明内容

本申请实施例提供了一种推荐系统训练方法、装置、计算机设备及存储介质，旨在提高推荐系统的准确性。

第一方面，本申请实施例提供了一种推荐系统训练方法，其包括：

获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合；

分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合；

将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，其中，所述知识图谱是根据所述历史检索记录样本建立的，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果；

通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练；

若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果。

第二方面，本申请实施例还提供了一种推荐系统训练装置，其包括：

第一分词单元，用于获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合；

第一训练单元，用于分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合；

第一输入单元，用于将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，其中，所述知识图谱是根据所述历史检索记录样本建立的，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果；

第二训练单元，用于通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练；

第一推荐单元，用于若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果。

第三方面，本申请实施例还提供了一种计算机设备，其包括存储器以及与所述存储器相连的处理器；所述存储器用于存储计算机程序；所述处理器用于运行所述存储器中存储的计算机程序，以执行如下步骤：

获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合；

分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合；

将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，其中，所述知识图谱是根据所述历史检

索记录样本建立的，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果；

通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练；

若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果。

第四方面，本申请实施例还提供了一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时使所述处理器执行以下步骤：

获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合；

分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合；

将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，其中，所述知识图谱是根据所述历史检索记录样本建立的，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果；

通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练；

若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果。

附图说明

为了更清楚地说明本申请实施例技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 为本申请实施例提供的一种推荐系统训练方法的流程示意图；

图 2 为本申请实施例提供的一种推荐系统训练方法的子流程示意图；

图 3 为本申请实施例提供的一种推荐系统训练方法的子流程示意图；

图 4 为本申请实施例提供的一种推荐系统训练方法的子流程示意图；

图 5 为本申请实施例提供的一种推荐系统训练方法的子流程示意图；

图 6 为本申请实施例提供的一种推荐系统训练装置的示意性框图；

图 7 为本申请实施例提供的一种推荐系统训练装置的第一分词单元的示意性框图；

图 8 为本申请实施例提供的一种推荐系统训练装置的第一输入单元的示意性框图；

图 9 为本申请实施例提供的一种推荐系统训练装置的第二训练单元的示意性框；

图 10 为本申请实施例提供的一种推荐系统训练装置的第一推荐单元的示意性框图；以及

图 11 为本申请实施例提供的一种计算机设备的示意性框图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请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申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

请参阅图 1，图 1 是本申请实施例提供的推荐系统训练方法的流程示意图。如图 1 所示，该方法包括以下步骤 S1-S5。

S1，获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合。

本申请实施例中，预先收集获取用户的历史检索记录，以得到历史检索记录样本。历史检索记录样本包括检索词样本以及相应的答案词样本。

具体实施中，对所述检索词样本进行分词处理以相应得到检索词样本分词集合。检索词样本分词集合是所述检索词样本的分词结果组成的集合。

相应地，对所述答案词样本进行分词处理以相应得到答案词样本分词集合。答案词样本分词集合是所述答案词样本的分词结果组成的集合。

在一实施例中，如图 2 所示，所述步骤 S1 可包括步骤 S11-S12。

S11，通过预设分词工具对检索词样本以及答案词样本进行分词处理以相应得到初始检索词样本分词集合以及初始答案词样本分词集合。

常用的分词工具为结巴分词工具。本申请实施例中，通过结巴分词工具对检索词样本以及答案词样本进行分词处理以相应得到初始检索词样本分词集合以及初始答案词样本分词集合。

或者在其它实施例中，通过其它分词工具对检索词样本以及答案词样本进行分词处理，本申请对此不做具体限定。

S12，分别将所述初始检索词样本分词集合以及初始答案词样本分词集合中的停止词去除以分别得到检索词样本分词集合以及答案词样本分词集合。

具体实施中，分别将所述初始检索词样本分词集合以及初始答案词样本分词集合中的停止词去除以分别得到检索词样本分词集合以及答案词样本分词集合。

要说明的是，停止词（stop word），常为介词、副词或连词等。停止词没有实际意义，因此需要去除。常用的停止词包括如"在"、"里面"、"也"、"的"、"它"、"为"等。

S2，分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合。

具体实施中，通过词向量工具 word2vec 分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合。

其中，检索词样本词向量集为检索词样本分词集合中各样本的词向量组成的集合；答案词样本词向量集合为答案词样本分词集合中各样本的词向量组成的集合。

需要说明的是，word2vec 是一种常用的自然语言处理工具，其作用就是将自然语言中的字词转为计算机可以理解的词向量。

传统的词向量容易受维数灾难的困扰，且任意两个词之间都是孤立的，不能体现词和词之间的关系，因此本实施例采用 word2vec 来得到词向量，其可通过计算向量之间的距离来体现词与词之间的相似性。

或者，在其他实施例中，可采用其他词向量工具对所述检索词样本分词集

合以及答案词样本分词集合进行词向量训练，本申请对此不作具体限定。

S3，将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合。

具体实施中，预先根据历史检索记录样本建立知识图谱。本方案中，采用自顶向下的方式构建知识图谱。

具体的，借助百科类互联网网站等结构化数据源，从用户的历史检索记录样本中提取本体和模式信息，加入到知识库中得到知识图谱。

需要说明的是，知识图谱的建立方法可参照现有技术资料实现，本申请对此不作具体限定。

在预先建立了知识图谱后，将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合。其中，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果。

在一实施例中，如图 3 所示，所述步骤 S3 可包括步骤 S31-S34。

S31，从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中。

具体实施中，从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中。

S32，将所述目标检索词样本词向量从所述检索词样本词向量集合中移除。

具体实施中，在将目标词向量输入到知识图谱中获得输入词向量后，将所述目标检索词样本词向量从所述检索词样本词向量集合中移除，以避免重复输入。

S33，判断所述检索词样本词向量集合中是否还存在检索词样本词向量。

具体实施中，判断所述检索词样本词向量集合中是否还存在检索词样本词向量。

若所述检索词样本词向量集合中还存在检索词样本词向量，则执行步骤 S31，返回所述从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作

为输入词向量加入到所述输入词向量集合中的步骤。

具体实施中，如果所述检索词样本词向量集合中还存在检索词样本词向量，则返回所述从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中的步骤，如此循环以上步骤，直到将所述检索词样本词向量集合中的所有检索词样本词向量均输入到所述知识图谱中为止。

如果所述检索词样本词向量集合中不存在检索词样本词向量，则执行以下步骤 S4。

S4，通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练。

具体实施中，将所述输入词向量集合中的输入词向量作为输入数据，以及将所述答案词样本词向量集合中的相应的答案词向量作为参考输出数据来对所述推荐系统进行训练。

需要说明的是，在本实施例中，推荐系统是基于深度神经网络（Deep Neural Networks，DNN）建立的推荐系统。或者，在其它实施例中，可采用其它架构的推荐系统，本申请对此不作具体限定。

在一实施例中，如图 4 所示，所述步骤 S4 可包括步骤 S41-S45。

S41，从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中。

具体实施中，从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中。

S42，判断所述推荐系统输出的答案词向量是否与目标答案词向量相同，其中，所述目标答案词向量为所述答案词样本词向量集合中与所述输入词向量相对应的答案词样本词向量。

具体实施中，判断所述推荐系统输出的答案词向量是否与目标答案词向量相同。其中，所述目标答案词向量为所述答案词样本词向量集合中与所述输入词向量相对应的答案词样本词向量。

S43，若所述推荐系统输出的答案词向量与目标答案词向量不相同，将所述推荐系统输出的答案词向量输入到所述知识图谱中，并将所述知识图谱的输出

结果作为新的目标输入词向量输入到所述推荐系统中。

具体实施中，如果所述推荐系统输出的答案词向量与目标答案词向量不相同，将所述推荐系统输出的答案词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为新的目标输入词向量输入到所述推荐系统中，如此循环以上步骤，直到所述推荐系统输出的答案词向量与目标答案词向量相同为止。

S44，若所述推荐系统输出的答案词向量与目标答案词向量相同，将所述目标输入词向量从所述输入词向量集合中移除。

具体实施中，如果所述推荐系统输出的答案词向量与目标答案词向量相同，将所述目标输入词向量从所述输入词向量集合中移除，以避免重复输入。

S45，判断所述输入词向量集合中是否存在输入词向量。

具体实施中，判断所述输入词向量集合中是否存在输入词向量，即判断是否所述输入词向量集合中的所有输入词向量均已输入到推荐系统中。如果输入词向量集合中还存在输入词向量，则继续执行步骤 S41，继续获取输入词向量输入到推荐系统中；如果输入词向量集合中不存在输入词向量，则结束对推荐系统的训练，并执行步骤 S5。

具体实施中，如果所述输入词向量集合中还存在输入词向量，返回所述从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中的步骤，如此循环以上步骤，直到输入词向量集合中不存在输入词向量为止。

S5，若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果。

具体实施中，如果接收到用户输入的检索信息，将所述检索信息转换为词向量后输入到通过上述步骤训练好的推荐系统中，并将推荐系统的输出结果作为提供给用户的推荐结果。

在一实施例中，如图 5 所示，所述步骤 S5 可包括步骤 S51-S53。

S51，分别对所述检索信息进行分词处理以及词向量训练以得到所述检索信息的检索词向量。

具体实施中，分别对所述检索信息进行分词处理后得到检索词，对检索词进行词向量训练以得到检索词向量。

分词处理以及词向量训练的过程在上述步骤 S1-S2 中已经详细描述，在此

不再赘述。

S52，将所述检索词向量输入到所述知识图谱中，并将所述知识图谱的输出结果输入到训练好的推荐系统中。

具体实施中，将所述检索词向量输入到所述知识图谱中，并将所述知识图谱的输出结果输入到训练好的推荐系统中，通过以上步骤可使得推荐系统共享知识图谱中的数据特征，使得推荐系统的推荐结果更加的准确。

S53，将所述推荐系统的输出结果作为推荐结果推荐给用户。

具体实施中，将所述推荐系统的输出结果作为提供给用户的推荐结果。

本申请实施例的技术方案，通过将检索词样本词向量集合中的检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量，并通过输入词向量来对推荐系统进行训练，能够使得推荐系统共享知识图谱中的数据特征，使得推荐系统在训练过程中实现了部分数据增强，提升了推荐系统的精度并一定程度上降低了过拟合的可能性，提高了推荐系统的准确性。

图 6 是本申请实施例提供的一种推荐系统训练装置 60 的示意性框图。如图 6 所示，对应于以上推荐系统训练方法，本申请还提供一种推荐系统训练装置 60。该推荐系统训练装置 60 包括用于执行上述推荐系统训练方法的单元，该装置可以被配置于服务器中。具体地，请参阅图 6，该推荐系统训练装置 60 包括第一分词单元 61、第一训练单元 62、第一输入单元 63、第二训练单元 64 以及第一推荐单元 65。

第一分词单元 61，用于获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合；

第一训练单元 62，用于分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合；

第一输入单元 63，用于将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，其中，所述知识图谱是根据所述历史检索记录样本建立的，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果；

第二训练单元 64，用于通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练；

第一推荐单元 65，用于若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果。

在一实施例中，参见图 7，所述第一分词单元 61 包括第二分词单元 611 以及去除单元 612。

第二分词单元 611，用于通过预设分词工具对检索词样本以及答案词样本进行分词处理以相应得到初始检索词样本分词集合以及初始答案词样本分词集合；

去除单元 612，用于分别将所述初始检索词样本分词集合以及初始答案词样本分词集合中的停止词去除以分别得到检索词样本分词集合以及答案词样本分词集合。

在一实施例中，参见图 8，所述第一输入单元 63 包括第二输入单元 631、第一移除单元 632、第一判断单元 633 以及第一通知单元 634。

第二输入单元 631，用于从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中；

第一移除单元 632，用于将所述目标检索词样本词向量从所述检索词样本词向量集合中移除；

第一判断单元 633，用于判断所述检索词样本词向量集合中是否还存在检索词样本词向量；

第一通知单元 634，用于若所述检索词样本词向量集合中还存在检索词样本词向量，通知第二输入单元返回所述从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中的步骤。

在一实施例中，参见图 9，所述第二训练单元 64 包括第三输入单元 641、第二判断单元 642、第二通知单元 643、第二移除单元 644、第三判断单元 645 以及第三通知单元 646。

第三输入单元 641，用于从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中；

第二判断单元642，用于判断所述推荐系统输出的答案词向量是否与目标答案词向量相同，其中，所述目标答案词向量为所述答案词样本词向量集合中与所述输入词向量相对应的答案词样本词向量；

第二通知单元643，用于若所述推荐系统输出的答案词向量与目标答案词向量不相同，将所述推荐系统输出的答案词向量输入到所述知识图谱中，并通知所述第三输入单元 641 将所述知识图谱的输出结果作为新的目标输入词向量输入到所述推荐系统中；

第二移除单元644，用于若所述推荐系统输出的答案词向量与目标答案词向量相同，将所述目标输入词向量从所述输入词向量集合中移除；

第三判断单元645，用于判断所述输入词向量集合中是否存在输入词向量；

第三通知单元646，用于若所述输入词向量集合中还存在输入词向量，通知所述第三输入单元 641 返回所述从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中的步骤。

在一实施例中，参见图 10，所述第一推荐单元65 包括处理单元651、第四输入单元652 以及第二推荐单元653。

处理单元651，用于分别对所述检索信息进行分词处理以及词向量训练以得到所述检索信息的检索词向量；

第四输入单元652，用于将所述检索词向量输入到所述知识图谱中，并将所述知识图谱的输出结果输入到训练好的推荐系统中；

第二推荐单元653，用于将所述推荐系统的输出结果作为推荐结果推荐给用户。

需要说明的是，所属领域的技术人员可以清楚地了解到，上述推荐系统训练装置 60 和各单元的具体实现过程，可以参考前述方法实施例中的相应描述，为了描述的方便和简洁，在此不再赘述。

上述推荐系统训练装置可以实现为一种计算机程序的形式，该计算机程序可以在如图 11 所示的计算机设备上运行。

请参阅图 11，图 11 是本申请实施例提供的一种计算机设备的示意性框图。该计算机设备 500 可以是终端，也可以是服务器，其中，终端可以是智能手机、平板电脑、笔记本电脑、台式电脑、个人数字助理和穿戴式设备等具有通信功能的电子设备。服务器可以是独立的服务器，也可以是多个服务器组成的服务

器集群。

参阅图 11，该计算机设备 500 包括通过系统总线 501 连接的处理器 502、存储器和网络接口 505，其中，存储器可以包括非易失性存储介质 503 和内存储器 504。

该非易失性存储介质 503 可存储操作系统 5031 和计算机程序 5032。该计算机程序 5032 被执行时，可使得处理器 502 执行一种推荐系统训练方法。

该处理器 502 用于提供计算和控制能力，以支撑整个计算机设备 500 的运行。

该内存储器 504 为非易失性存储介质 503 中的计算机程序 5032 的运行提供环境，该计算机程序 5032 被处理器 502 执行时，可使得处理器 502 执行一种推荐系统训练方法。

该网络接口 505 用于与其它设备进行网络通信。本领域技术人员可以理解，图 11 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备 500 的限定，具体的计算机设备 500 可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

其中，所述处理器 502 用于运行存储在存储器中的计算机程序 5032，以实现本申请实施例的推荐系统训练方法。

应当理解，在本申请实施例中，处理器 502 可以是中央处理单元 (Central Processing Unit，CPU)，该处理器 502 还可以是其他通用处理器、数字信号处理器 (Digital Signal Processor，DSP)、专用集成电路 (Application Specific Integrated Circuit，ASIC)、现成可编程门阵列 (Field-Programmable Gate Array，FPGA) 或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。其中，通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

本领域普通技术人员可以理解的是实现上述实施例的方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成。该计算机程序可存储于一存储介质中，该存储介质为计算机可读存储介质。该计算机程序被该计算机系统中的至少一个处理器执行，以实现上述方法的实施例的流程步骤。

因此，本申请实施例还提供一种存储介质。该存储介质可以为计算机可读存储介质。该存储介质存储有计算机程序，该计算机程序被处理器执行时使处

理器执行以上各实施例中所描述的推荐系统训练方法的步骤。

所述存储介质可以是U盘、移动硬盘、只读存储器（Read-Only Memory，ROM）、磁碟或者光盘等各种可以存储计算机程序的计算机可读存储介质。

本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

以上所述，仅为本申请的具体实施方式，但本申请明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到各种等效的修改或替换，这些修改或替换都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以权利要求的保护范围为准。

权 利 要 求 书

1.一种推荐系统训练方法，包括：

获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合；

分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合；

将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，其中，所述知识图谱是根据所述历史检索记录样本建立的，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果；

通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练；

若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果。

2.根据权利要求 1 所述的方法，其中，所述分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合，包括：

通过预设分词工具对检索词样本以及答案词样本进行分词处理以相应得到初始检索词样本分词集合以及初始答案词样本分词集合；

分别将所述初始检索词样本分词集合以及初始答案词样本分词集合中的停止词去除以分别得到检索词样本分词集合以及答案词样本分词集合。

3.根据权利要求 1 所述的方法，其中，所述将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，包括：

从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中；

将所述目标检索词样本词向量从所述检索词样本词向量集合中移除；

判断所述检索词样本词向量集合中是否还存在检索词样本词向量；

若所述检索词样本词向量集合中还存在检索词样本词向量，返回所述从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中的步骤。

4.根据权利要求 1 所述的方法，其中，所述通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练，包括：

从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中；

判断所述推荐系统输出的答案词向量是否与目标答案词向量相同，其中，所述目标答案词向量为所述答案词样本词向量集合中与所述输入词向量相对应的答案词样本词向量；

若所述推荐系统输出的答案词向量与目标答案词向量不相同，将所述推荐系统输出的答案词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为新的目标输入词向量输入到所述推荐系统中。

5.根据权利要求 4 所述的方法，其中，所述通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练，还包括：

若所述推荐系统输出的答案词向量与目标答案词向量相同，将所述目标输入词向量从所述输入词向量集合中移除；

判断所述输入词向量集合中是否存在输入词向量；

若所述输入词向量集合中还存在输入词向量，返回所述从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中的步骤。

6.根据权利要求 1 所述的方法，其中，所述根据所述检索信息以及训练后的推荐系统向用户提供推荐结果，包括：

分别对所述检索信息进行分词处理以及词向量训练以得到所述检索信息的检索词向量；

将所述检索词向量输入到所述知识图谱中，并将所述知识图谱的输出结果输入到训练好的推荐系统中；

将所述推荐系统的输出结果作为推荐结果推荐给用户。

7、根据权利要求 3 所述的方法，其中，所述将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集

合，还包括：

若所述检索词样本词向量集合中不存在检索词样本词向量，转至所述通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练的步骤。

8.根据权利要求 5 所述的方法，其中，所述通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练，还包括：

若所述输入词向量集合中还存在输入词向量，转至所述若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果的步骤。

9、根据权利要求 1 所述的方法，其中，所述分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合，包括：

通过词向量工具 word2vec 分别对所述检索词样本分词集合以及所述答案词样本分词集合进行词向量训练以分别得到所述检索词样本词向量集合以及所述答案词样本词向量集合。

10.一种推荐系统训练装置，包括：

第一分词单元，用于获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合；

第一训练单元，用于分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合；

第一输入单元，用于将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，其中，所述知识图谱是根据所述历史检索记录样本建立的，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果；

第二训练单元，用于通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练；

第一推荐单元，用于若接收到用户输入的检索信息，根据所述检索信息以

及训练后的推荐系统向用户提供推荐结果。

11.一种计算机设备，包括存储器以及与所述存储器相连的处理器；所述存储器用于存储计算机程序；所述处理器用于运行所述存储器中存储的计算机程序，以执行如下步骤：

获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合；

分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合；

将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，其中，所述知识图谱是根据所述历史检索记录样本建立的，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果；

通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练；

若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果。

12.根据权利要求 11 所述的计算机设备，其中，所述分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合的步骤包括：

通过预设分词工具对检索词样本以及答案词样本进行分词处理以相应得到初始检索词样本分词集合以及初始答案词样本分词集合；

分别将所述初始检索词样本分词集合以及初始答案词样本分词集合中的停止词去除以分别得到检索词样本分词集合以及答案词样本分词集合。

13.根据权利要求 11 所述的计算机设备，其中，所述将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合的步骤包括：

从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中；

将所述目标检索词样本词向量从所述检索词样本词向量集合中移除；

判断所述检索词样本词向量集合中是否还存在检索词样本词向量；

若所述检索词样本词向量集合中还存在检索词样本词向量，返回所述从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中的步骤。

14.根据权利要求 11 所述的计算机设备，其中，所述通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练的步骤包括：

从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中；

判断所述推荐系统输出的答案词向量是否与目标答案词向量相同，其中，所述目标答案词向量为所述答案词样本词向量集合中与所述输入词向量相对应的答案词样本词向量；

若所述推荐系统输出的答案词向量与目标答案词向量不相同，将所述推荐系统输出的答案词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为新的目标输入词向量输入到所述推荐系统中。

15.根据权利要求 14 所述的计算机设备，其中，所述通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练的步骤还包括：

若所述推荐系统输出的答案词向量与目标答案词向量相同，将所述目标输入词向量从所述输入词向量集合中移除；

判断所述输入词向量集合中是否存在输入词向量；

若所述输入词向量集合中还存在输入词向量，返回所述从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中的步骤。

16.根据权利要求 11 所述的计算机设备，其中，所述根据所述检索信息以及训练后的推荐系统向用户提供推荐结果的步骤包括：

分别对所述检索信息进行分词处理以及词向量训练以得到所述检索信息的检索词向量；

将所述检索词向量输入到所述知识图谱中，并将所述知识图谱的输出结果输入到训练好的推荐系统中；

将所述推荐系统的输出结果作为推荐结果推荐给用户。

17、根据权利要求 13 所述的计算机设备，其中，所述将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合的步骤还包括：

若所述检索词样本词向量集合中不存在检索词样本词向量，转至所述通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练的步骤。

18.根据权利要求 15 所述的计算机设备，其中，所述通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练的步骤还包括：

若所述输入词向量集合中还存在输入词向量，转至所述若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果的步骤。

19、根据权利要求 11 所述的计算机设备，其中，所述分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合的步骤包括：

通过词向量工具 word2vec 分别对所述检索词样本分词集合以及所述答案词样本分词集合进行词向量训练以分别得到所述检索词样本词向量集合以及所述答案词样本词向量集合。

20. 一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序被处理器执行时使所述处理器执行以下步骤：

获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合；

分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合；

将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合，其中，所述知识图谱是根据所述历史检索记录样本建立的，所述输入词向量集合中的输入词向量为所述检索词样本词向量集合中的检索词样本词向量输入到所述知识图谱后的输出结果；

通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练；

若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果。

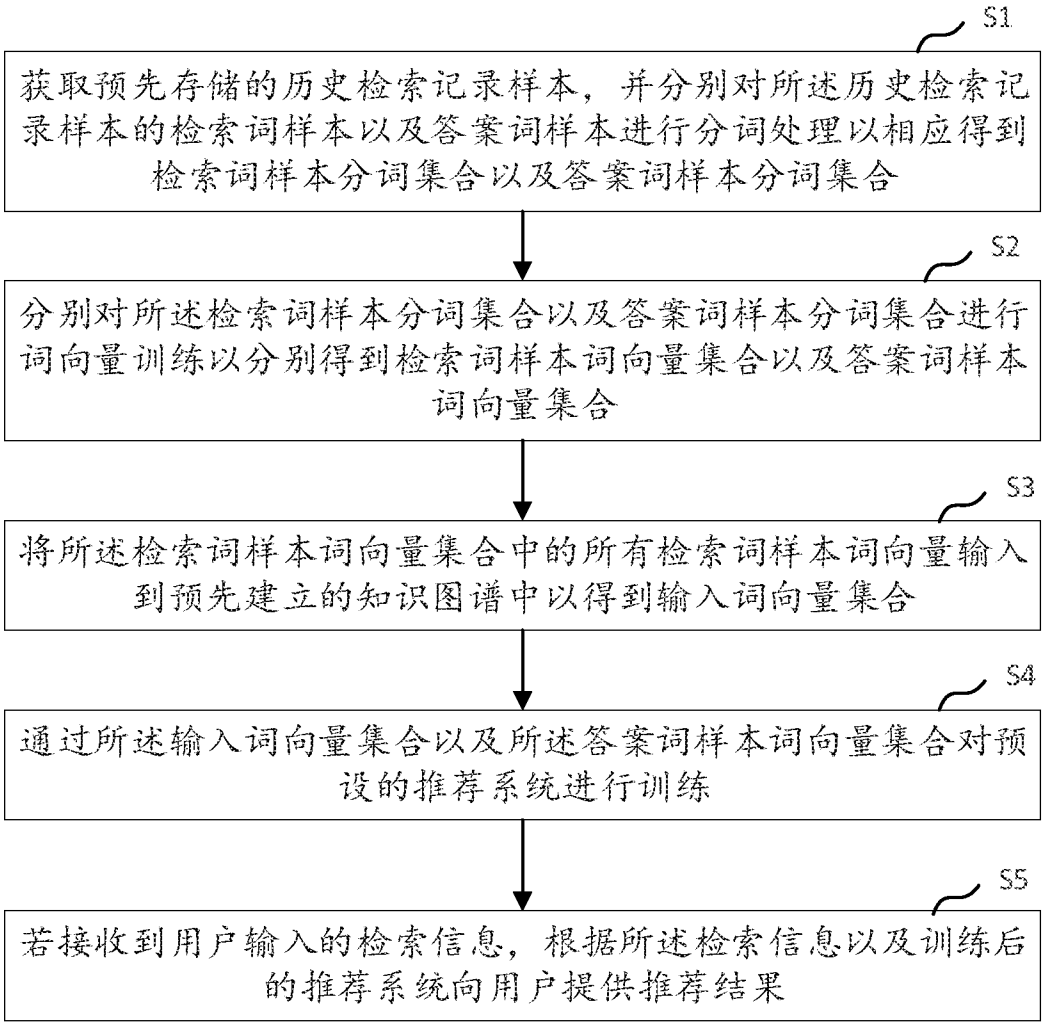
S1
获取预先存储的历史检索记录样本，并分别对所述历史检索记录样本的检索词样本以及答案词样本进行分词处理以相应得到检索词样本分词集合以及答案词样本分词集合
S2
分别对所述检索词样本分词集合以及答案词样本分词集合进行词向量训练以分别得到检索词样本词向量集合以及答案词样本词向量集合
S3
将所述检索词样本词向量集合中的所有检索词样本词向量输入到预先建立的知识图谱中以得到输入词向量集合
S4
通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练
S5
若接收到用户输入的检索信息，根据所述检索信息以及训练后的推荐系统向用户提供推荐结果

图 1

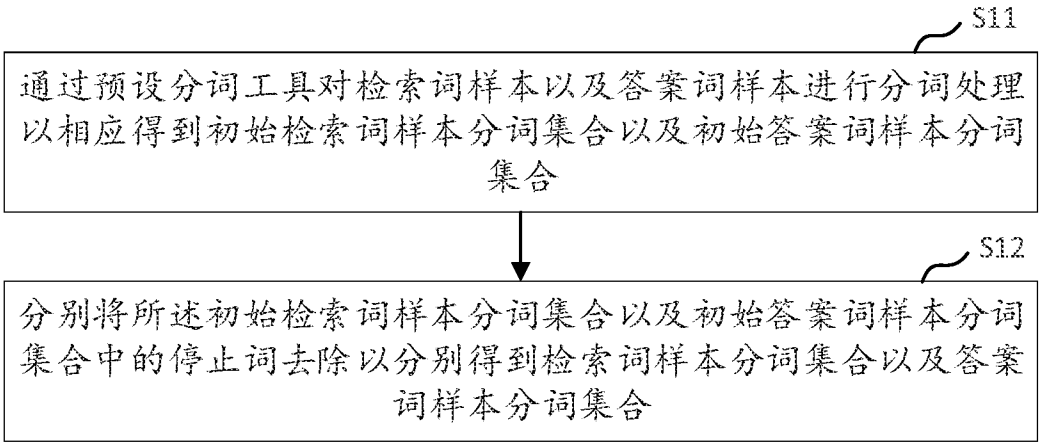
S11
通过预设分词工具对检索词样本以及答案词样本进行分词处理以相应得到初始检索词样本分词集合以及初始答案词样本分词集合
S12
分别将所述初始检索词样本分词集合以及初始答案词样本分词集合中的停止词去除以分别得到检索词样本分词集合以及答案词样本分词集合

图 2

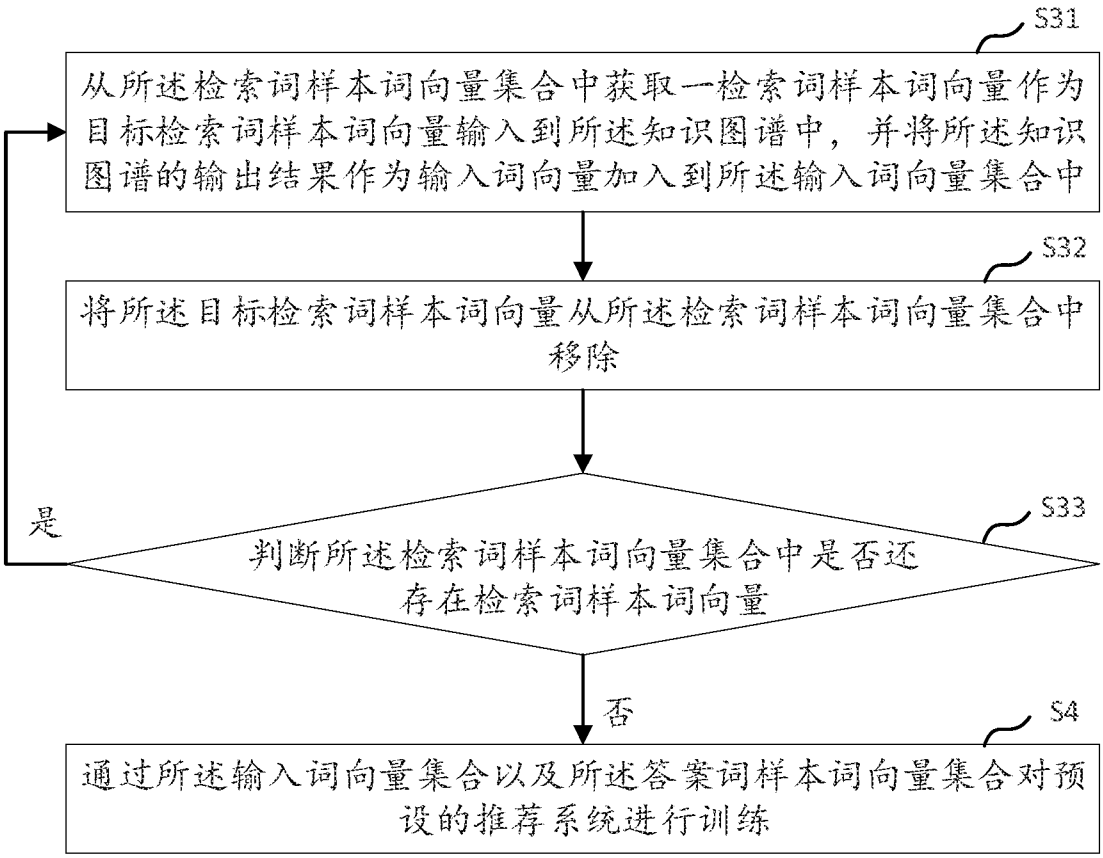
S31
从所述检索词样本词向量集合中获取一检索词样本词向量作为目标检索词样本词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为输入词向量加入到所述输入词向量集合中
S32
将所述目标检索词样本词向量从所述检索词样本词向量集合中移除
S33
判断所述检索词样本词向量集合中是否还存在检索词样本词向量
是
否
S4
通过所述输入词向量集合以及所述答案词样本词向量集合对预设的推荐系统进行训练

图 3

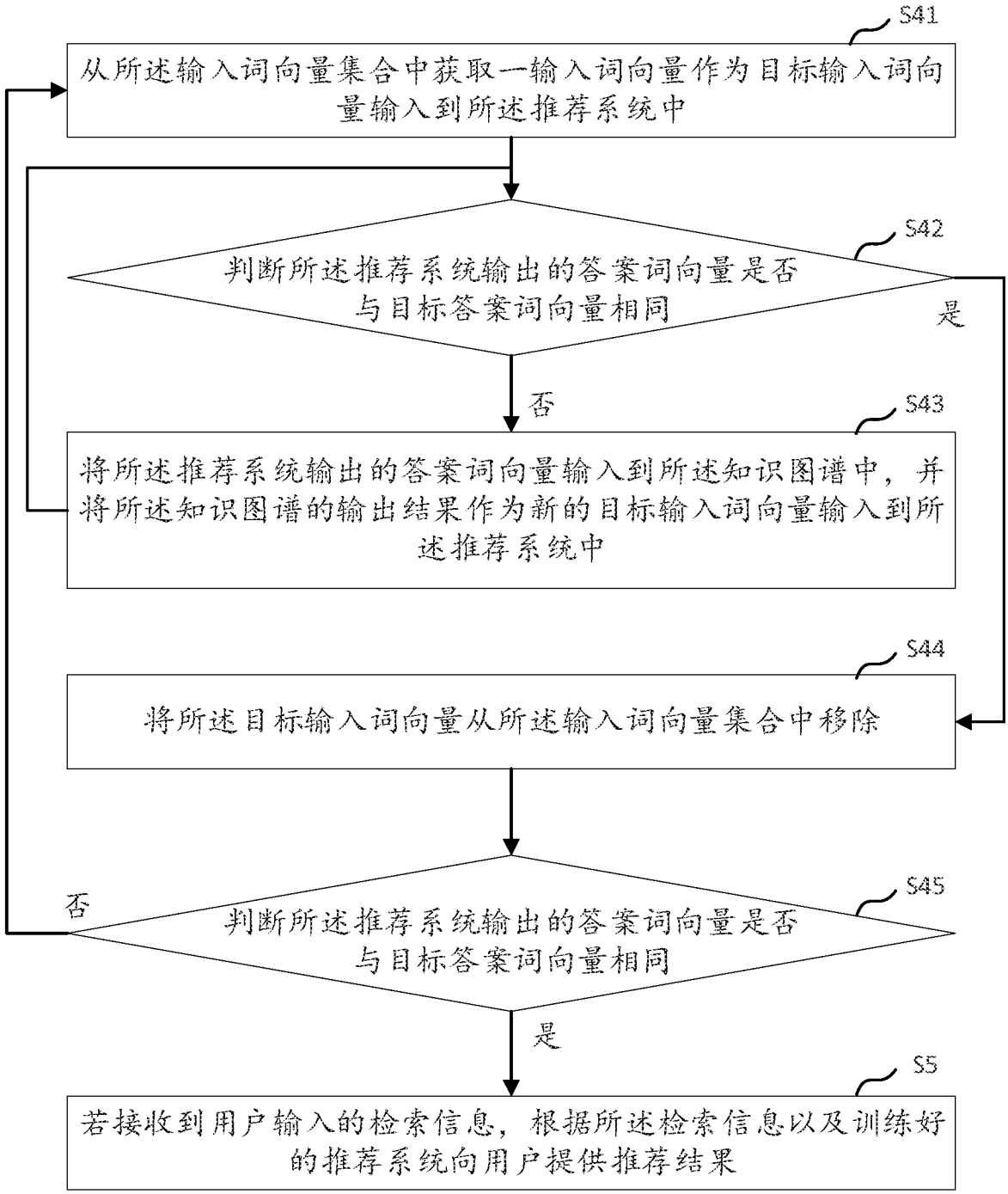
S41
从所述输入词向量集合中获取一输入词向量作为目标输入词向量输入到所述推荐系统中
S42
判断所述推荐系统输出的答案词向量是否与目标答案词向量相同
是
否
S43
将所述推荐系统输出的答案词向量输入到所述知识图谱中，并将所述知识图谱的输出结果作为新的目标输入词向量输入到所述推荐系统中
S44
将所述目标输入词向量从所述输入词向量集合中移除
S45
判断所述推荐系统输出的答案词向量是否与目标答案词向量相同
否
是
S5
若接收到用户输入的检索信息，根据所述检索信息以及训练好的推荐系统向用户提供推荐结果

图 4

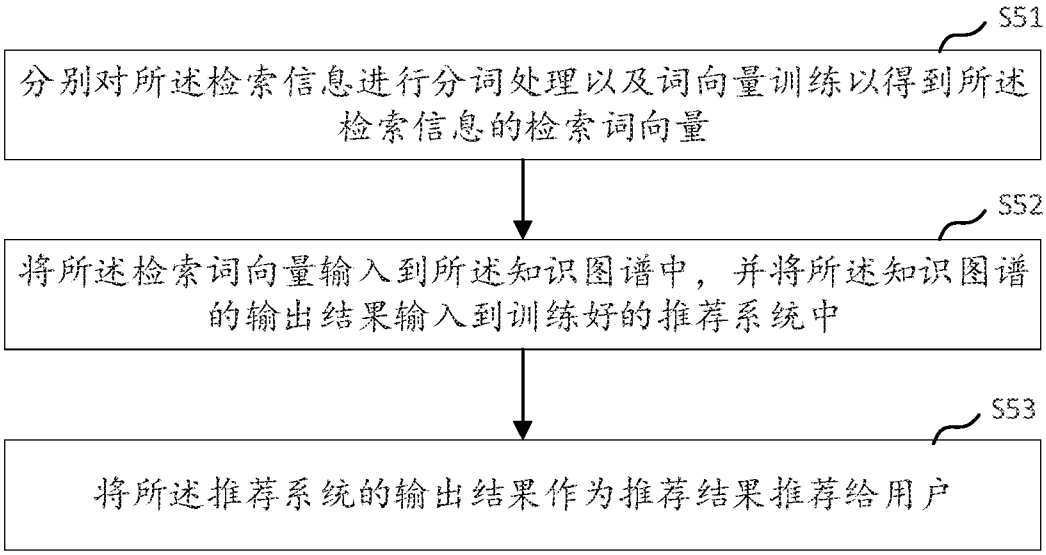
S51
分别对所述检索信息进行分词处理以及词向量训练以得到所述检索信息的检索词向量
S52
将所述检索词向量输入到所述知识图谱中，并将所述知识图谱的输出结果输入到训练好的推荐系统中
S53
将所述推荐系统的输出结果作为推荐结果推荐给用户

图 5

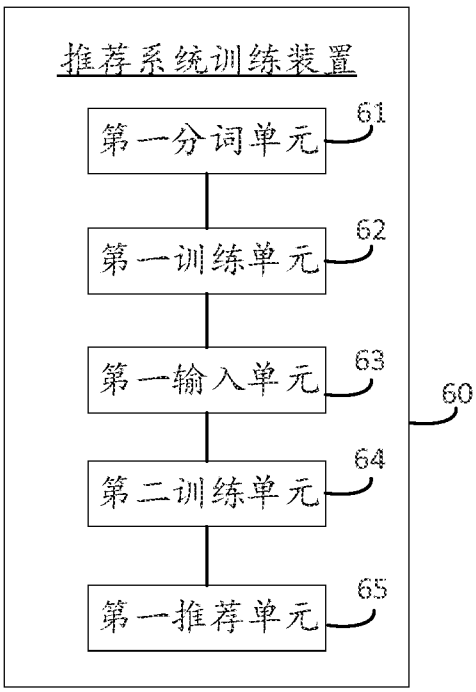
推荐系统训练装置
第一分词单元
61
第一训练单元
62
第一输入单元
63
第二训练单元
64
第一推荐单元
65
60

图 6

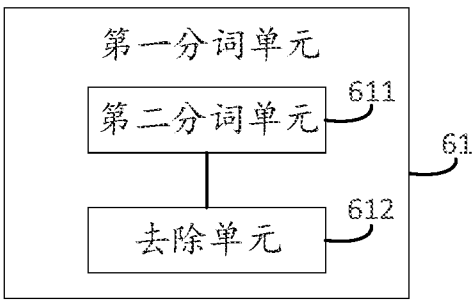
第一分词单元
第二分词单元
611
去除单元
612
61

图 7

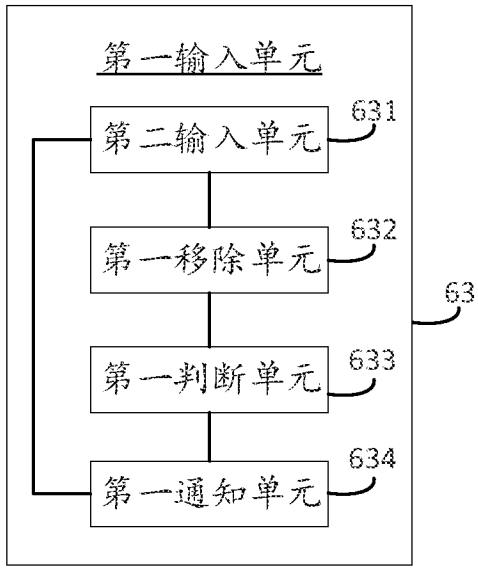
第一输入单元
第二输入单元
631
第一移除单元
632
63
第一判断单元
633
第一通知单元
634

图 8

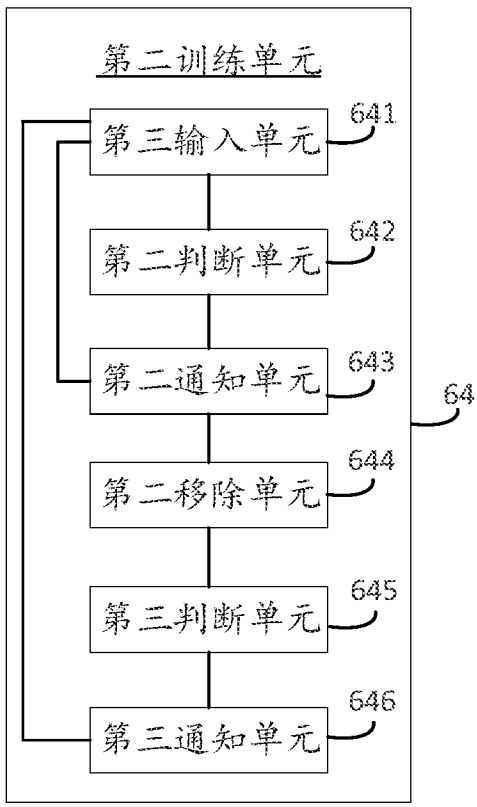
第二训练单元
第三输入单元
641
第二判断单元
642
第二通知单元
643
64
第二移除单元
644
第三判断单元
645
第三通知单元
646

图 9

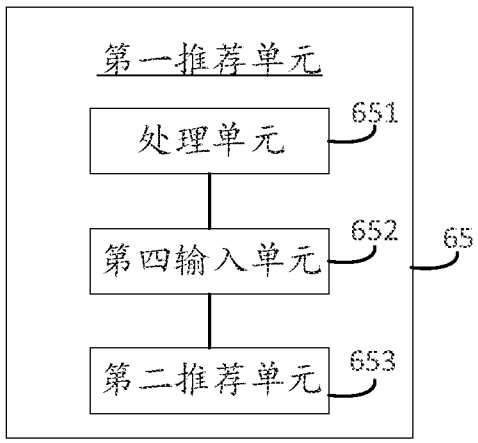
第一推荐单元
处理单元
651
第四输入单元
652
65
第二推荐单元
653

图 10

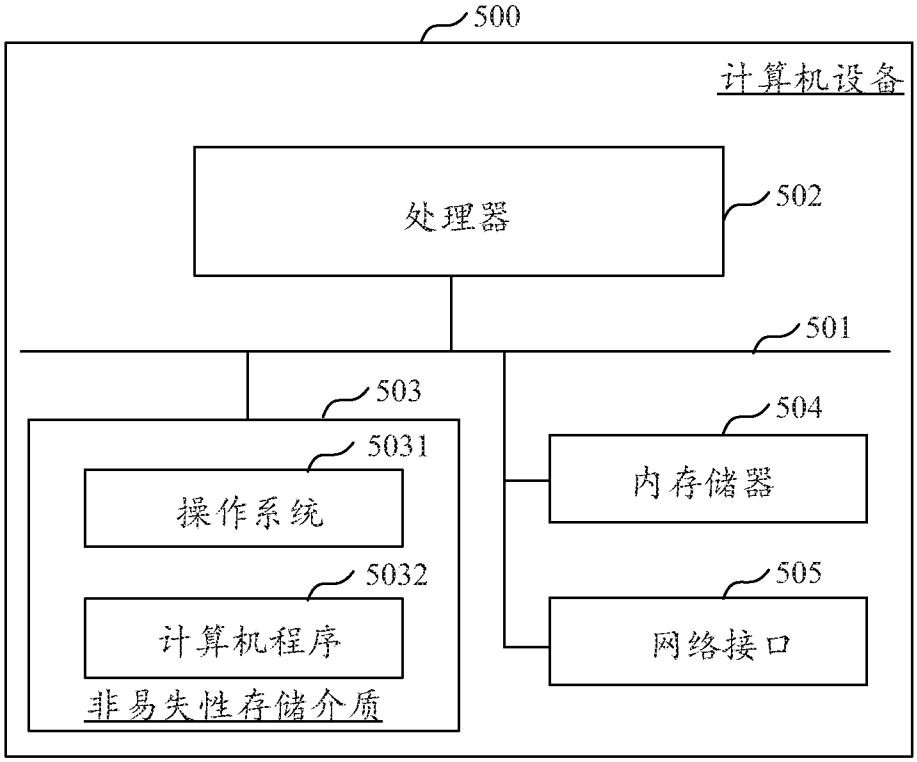
500
计算机设备
处理器
502
501
503
5031
操作系统
5032
计算机程序
非易失性存储介质
504
内存储器
505
网络接口

图 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/093138

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62(2006.01)i; G06F 16/36(2019.01)i; G06F 16/951(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K; G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; VEN; CNKI; USTXT; EPTXT; WOTXT: 历史, 问题, 检索, 结果, 目标, 答案, 分词, 知识图谱, 知识库, 训练, history, question, search, result, target, answer, word, segment, knowledge graph, knowledge map, train

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109858528 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 07 June 2019 (2019-06-07) claims 1-10, and description, paragraphs [0004]-[0186]	1-20
Y	CN 104156359 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 19 November 2014 (2014-11-19) description, paragraphs [0037]-[0134]	1-20
Y	CN 105989040 A (ALIBABA GROUP HOLDING LIMITED) 05 October 2016 (2016-10-05) description, paragraphs [0125]-[0313]	1-20
A	CN 108427707 A (SHENZHEN AXMTEC CO., LTD.) 21 August 2018 (2018-08-21) entire document	1-20
A	US 2007299824 A1 (IBM) 27 December 2007 (2007-12-27) entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
02 September 2019	**20 September 2019**
Name and mailing address of the ISA/CN **China National Intellectual Property Administration** **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China** Facsimile No. **(86-10)62019451**	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/093138

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 109858528 A	07 June 2019	None	
CN 104156359 A	19 November 2014	CN 104156359 B	30 October 2018
CN 105989040 A	05 October 2016	HK 1229484 A0	17 November 2017
CN 108427707 A	21 August 2018	None	
US 2007299824 A1	27 December 2007	US 7533089 B2	12 May 2009

国际检索报告

国际申请号 PCT/CN2019/093138

A. 主题的分类

G06K 9/62(2006.01)i; G06F 16/36(2019.01)i; G06F 16/951(2019.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06K; G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNABS;CNTXT;VEN;CNKI;USTXT;EPTXT;WOTXT:历史，问题，检索，结果，目标，答案，分词，知识图谱，知识库，训练，history，question，search，result，target，answer，word，segment，knowledge graph，knowledge map，train

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 109858528 A (平安科技深圳有限公司) 2019年 6月 7日 (2019-06-07) 权利要求1-10，说明书[0004]-[0186]段	1-20
Y	CN 104156359 A (腾讯科技深圳有限公司) 2014年 11月 19日 (2014-11-19) 说明书[0037]-[0134]段	1-20
Y	CN 105989040 A (阿里巴巴集团控股有限公司) 2016年 10月 5日 (2016-10-05) 说明书[0125]-[0313]段	1-20
A	CN 108427707 A (深圳市阿西莫夫科技有限公司) 2018年 8月 21日 (2018-08-21) 全文	1-20
A	US 2007299824 A1 (IBM) 2007年 12月 27日 (2007-12-27) 全文	1-20

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2019年 9月 2日	2019年 9月 20日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 薛乐梅 电话号码 86-(20)-28950448

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/093138

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	109858528	A	2019年 6月 7日		无		
CN	104156359	A	2014年 11月 19日	CN	104156359	B	2018年 10月 30日
CN	105989040	A	2016年 10月 5日	HK	1229484	A0	2017年 11月 17日
CN	108427707	A	2018年 8月 21日		无		
US	2007299824	A1	2007年 12月 27日	US	7533089	B2	2009年 5月 12日