

(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
2022 年 9 月 15 日 (15.09.2022)



(10) 国际公布号
WO 2022/188773 A1

(51) 国际专利分类号:
G06F 16/35 (2019.01)

(21) 国际申请号: PCT/CN2022/079721

(22) 国际申请日: 2022 年 3 月 8 日 (08.03.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202110268045.7 2021年3月12日 (12.03.2021) CN

(71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

(72) 发明人: 张先礼 (ZHANG, Xianli); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 管冲 (GUAN, Chong); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 陈曦 (CHEN, Xi); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

(74) 代理人: 北京派特恩知识产权代理有限公司 (CHINA PAT INTELLECTUAL PROPERTY OFFICE); 中国北京市海淀区海淀南路21号中关村知识产权大厦B座2层, Beijing 100080 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,

(54) **Title:** TEXT CLASSIFICATION METHOD AND APPARATUS, DEVICE, COMPUTER-READABLE STORAGE MEDIUM, AND COMPUTER PROGRAM PRODUCT

(54) 发明名称: 文本分类方法、装置、设备、计算机可读存储介质及计算机程序产品

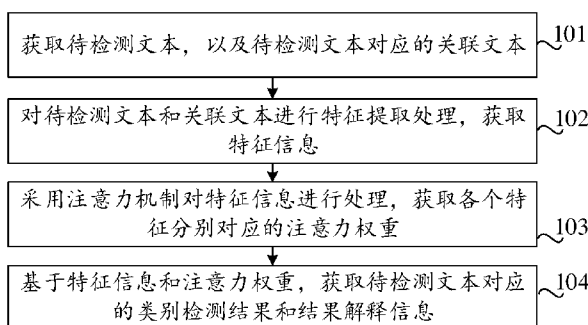


图 1

- 101 Obtain a text to be detected and an associated text corresponding to said text
- 102 Perform feature extraction on the text to be detected and the associated text to obtain feature information
- 103 Process the feature information by using an attention mechanism, to obtain attention weights respectively corresponding to the features
- 104 Obtain, on the basis of the feature information and the attention weights, a category detection result corresponding to the text to be detected and result interpretation information

(57) **Abstract:** Provided are a text classification method and apparatus, a device, and a storage medium, relating to the technical field of artificial intelligence. The text classification method comprises: obtaining a text to be detected and an associated text corresponding to said text (101); performing feature extraction on the text to be detected and the associated text to obtain feature information (102); processing the feature information by using an attention mechanism, to obtain attention weights respectively corresponding to the

GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JM, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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:

一 包括国际检索报告(条约第21条(3))。

features (103); and obtaining, on the basis of the feature information and the attention weights, a category detection result corresponding to the text to be detected and result interpretation information (104). The category detection result comprises a category distribution probability corresponding to the text to be detected, and the result interpretation information is used for describing the degree of influence of each feature on the category detection result.

(57) 摘要: 提供一种文本分类方法、装置、设备及存储介质, 属于人工智能技术领域。文本分类方法包括: 获取待检测文本, 以及待检测文本对应的关联文本(101); 对待检测文本和关联文本进行特征提取处理, 获取特征信息(102); 采用注意力机制对特征信息进行处理, 获取各个特征分别对应的注意力权重(103); 基于特征信息和注意力权重, 获取待检测文本对应的类别检测结果和结果解释信息(104); 其中, 类别检测结果包括待检测文本对应的类别分布概率, 结果解释信息用于描述各个特征针对类别检测结果的影响程度。

文本分类方法、装置、设备、计算机可读存储介质及计算机程序产品

相关申请的交叉引用

5 本申请基于申请号为 202110268045.7、申请日为 2021 年 03 月 12 日的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的全部内容在此引入本申请作为参考。

技术领域

本申请涉及人工智能技术领域，涉及但不限于一种文本分类方法、装置、设备、计算机可读存储介质及计算机程序产品。

10 背景技术

目前，随着人工智能的发展，通过计算机设备结合机器学习模型进行文本分类的运用也越来越多。

在相关技术中，技术人员可以通过大量训练样本进行模型训练，进而采用训练完成后的深度学习模型对各个文本进行分类。例如，在医疗领域的疾病预测场景中，医护人员在对目标对象进行疾病检测时，可以将目标对象的病症描述文本输入至深度学习模型中，由深度学习模型根据病症描述文本进行疾病预测，输出目标对象对应的病理信息数据，以供医护人员作为参考。再例如，在教育领域的书籍分类场景中，工作人员在对书籍所属科目进行分类时，可以将书籍的总结描述文本（前言、后记、摘要等）
15 输入至深度学习模型中，由深度学习模型根据总结描述文本进行书籍分类，输出书籍对应的科目信息数据作为科目分类结果。当然，基于深度学习模型的文本分类还可应用在其他多个场景中，此处不一一举例。

然而，在上述相关技术中，深度学习模型输出的文本分类结果的可信度不够高。

25 发明内容

本申请实施例提供一种文本分类方法、装置、设备、计算机可读存储

介质及计算机程序产品，能够提高类别检测结果的可信度。

本申请实施例提供一种文本分类方法，所述方法包括：

获取待检测文本，以及所述待检测文本对应的关联文本；

对所述待检测文本和所述关联文本进行特征提取处理，得到特征信息；

5 其中，所述特征信息中包括至少一个特征；

采用注意力机制对所述特征信息进行处理，得到各个所述特征分别对应的注意力权重，所述注意力权重用于指示所述特征针对所述待检测文本和所述关联文本的重要程度；

10 基于所述特征信息和所述注意力权重，获取所述待检测文本对应的类别检测结果和结果解释信息；其中，所述类别检测结果包括所述待检测文本对应的类别分布概率，所述结果解释信息用于描述各个所述特征针对所述类别检测结果的影响程度。

本申请实施例提供一种文本分类模型的训练方法，所述方法包括：

获取多个样本文本，以及各个所述样本文本对应的类别信息；

15 从各个所述样本文本中分别提取样本特征信息；

采用所述样本特征信息和所述类别信息对所述文本分类模型进行训练；

其中，所述文本分类模型中引入注意力机制，所述注意力机制用于确定所述样本特征信息中的各个样本特征针对所述样本文本中的重要程度，所述重要程度用于确定各个所述样本特征对模型输出结果的影响程度。

20 本申请实施例提供一种文本分类装置，所述装置包括：

文本获取模块，配置为获取待检测文本，以及所述待检测文本对应的关联文本；

特征提取模块，配置为对所述待检测文本和所述关联文本进行特征提取处理，得到特征信息；其中，所述特征信息中包括至少一个特征；

25 权重获取模块，配置为采用注意力机制对所述特征信息进行处理，得到各个所述特征分别对应的注意力权重，所述注意力权重用于指示所述特征针对所述待检测文本和所述关联文本的重要程度；

30 结果获取模块，配置为基于所述特征信息和所述注意力权重，获取所述待检测文本对应的类别检测结果和结果解释信息；其中，所述类别检测结果包括所述待检测文本对应的类别分布概率，所述结果解释信息用于描

述各个所述特征针对所述类别检测结果的影响程度。

本申请实施例提供一种文本分类模型的训练装置，所述装置包括：

样本获取模块，配置为获取多个样本文本，以及各个所述样本文本对应的类别信息；

- 5 特征获取模块，配置为从各个所述样本文本中分别提取样本特征信息；
 模型训练模块，配置为采用所述样本特征信息和所述类别信息对所述文本分类模型进行训练；

- 其中，所述文本分类模型中引入注意力机制，所述注意力机制用于确定所述样本特征信息中的各个样本特征针对所述样本文本中的重要程度，
10 所述重要程度用于确定各个所述样本特征对模型输出结果的影响程度。

- 本申请实施例提供一种计算机设备，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条指令、至少一段程序、代码集或指令集，所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现上述文本分类方法，或实现上述文本分类模型的
15 训练方法。

 本申请实施例提供一种计算机可读存储介质，所述可读存储介质中存储有至少一条指令、至少一段程序、代码集或指令集，所述至少一条指令、所述至少一段程序、所述代码集或指令集由处理器加载并执行以实现上述文本分类方法，或实现上述文本分类模型的训练方法。

- 20 本申请实施例提供一种计算机程序产品或计算机程序，该计算机程序产品或计算机程序包括计算机指令，该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令，处理器执行该计算机指令，使得该计算机设备执行上述文本分类方法，或实现上述文本分类模型的训练方法。

- 25 本申请实施例提供的技术方案可以带来如下有益效果：

- 通过注意力机制对特征信息进行处理，确定用于指示各个特征针对文本的重要程度的注意力权重，进而依据特征信息和注意力权重获取类别检测结果和结果解释信息，一方面，类别检测结果中包括类别分布概率，使得类别检测结果具有多样性，避免仅仅输出一个类别造成的输出单一；另
30 一方面，结果解释信息用于描述各个特征针对类别检测结果的影响程度，

为类别检测结果提供可解释性，提高了类别检测结果的可信度，为基于结果解释信息确定类别检测结果提供了直观的依据。

附图说明

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，显而易见地，下面描述中的附图仅仅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是本申请实施例提供的一种文本分类方法的流程图；

图2是本申请实施例提供的文本分类模型的训练方法的流程图；

图3是本申请实施例提供的另一种文本分类方法的流程的示意图；

图4是本申请实施例提供的一种文本分类模型的检测流程的示意图；

图5是本申请实施例提供的一种用户界面的示意图；

图6是本申请实施例提供的一个文本分类装置的框图；

图7是本申请实施例提供的另一个文本分类装置的框图；

图8是本申请实施例提供的文本分类模型的训练装置的框图；

图9是本申请实施例提供的计算机设备的结构框图。

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进行描述。

人工智能(Artificial Intelligence, AI)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。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。

5 自然语言处理(Nature Language Processing, NLP)是计算机科学领域与人工智能领域中的一个重要方向。它研究能实现人与计算机之间用自然语言进行有效通信的各种理论和方法。自然语言处理是一门融语言学、计算机科学、数学于一体的科学。因此,这一领域的研究将涉及自然语言,即人们日常使用的语言,所以它与语言学的研究有着密切的联系。自然语言处理技术通常包括文本处理、语义理解、机器翻译、机器人问答、知识图谱等技术。

机器学习(Machine Learning, ML)是一门多领域交叉学科,涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为,以获取新的知识或技能,重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心,是使
15 计算机具有智能的根本途径,其应用遍及人工智能的各个领域。机器学习和深度学习通常包括人工神经网络、置信网络、强化学习、迁移学习、归纳学习、示教学习等技术。

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、智能穿戴设备、虚拟助理、智能音箱、智能
20 营销、无人驾驶、自动驾驶、无人机、机器人、智能医疗、智能客服等,相信随着技术的发展,人工智能技术将在更多的领域得到应用,并发挥越来越重要的价值。

本申请实施例提供的方案涉及人工智能的机器学习等技术,采用从多个样本文本中提取的样本特征信息,以及样本文本对应的类别信息,对文本分类模型进行训练。在一些实施例中,单个样本文本中包括第一文本和
25 第二文本,且第二文本为第一文本的关联文本。示例性地,若第一文本为用户的当前病症文本,则第二文本为用户的历史病历文本;若第一文本为用户的当前提交文章,则第二文本为用户的历史提交文本;若第一文本为用户的当前浏览网页,则第二文本为用户在最近浏览的一个或多个历史浏览网页。在获取样本特征信息时,对第一文本和第二文本进行特征提取以
30

获取第一文本对应的样本特征信息。之后，在文本分类模型训练完成后，在获取待检测文本的情况下，基于待检测文本，以及待检测文本对应的关联文本，提取出待检测文本对应的特征信息，并采用文本分类模型对该特征信息进行处理，得到待检测文本对应的类别检测结果和结果解释信息。

- 5 其中，该类别检测结果包括待检测文本对应的类别分布概率，结果描述信息用于描述特征信息中的各个特征针对类别检测结果的影响程度。除此之外，在本申请实施例中，上述文本分类模型中引入空间注意力机制和稀疏访问级注意力机制。其中，空间注意力机制用于确定单个特征的重要程度，稀疏访问级注意力机制用于确定特征所处特征集合的重要程度，特征集合
- 10 中包括来自相同文本的特征。空间注意力机制的本质就是定位目标并进行一些变换或者获取权重，通过空间注意力机制能够确定出特征信息各个特征的空间注意力权重，从而从特征信息中确定出需要被关注的特征。基于空间注意力机制确定特征的空间注意力权重时，首先获取特征信息中各个特征的嵌入向量，得到嵌入矩阵，然后通过编码器将嵌入矩阵映射到隐藏
- 15 状态，得到特征的隐藏状态矩阵，最后通过空间注意力模型的模型参数以及特征的隐藏状态矩阵确定出各个特征的空间注意力权重。基于稀疏访问级注意力机制确定特征所处特征集合的重要程度，也即确定特征所处特征集合的稀疏访问级注意力权重，在实现时，可以基于特征的隐藏状态矩阵以及稀疏访问级注意力模型的模型参数确定各个特征的输出矩阵，然后再
- 20 利用激活函数 $\text{softmax}()$ 、稀疏概率激活函数 $\text{sparemax}()$ 以及各个特征的输出矩阵，确定特征所处特征集合的重要程度。

需要说明的一点是，本申请实施例提供的文本分类模型，可以广泛应用于各个领域。例如可以至少应用于如下领域：

- (1) 在医疗领域的疾病预测场景中，医护人员依据从目标对象中观察
- 25 获取的病症，确定当前病症描述文本，并将目标对象的标识信息和病症描述文本输入至计算机设备中。例如，计算机设备在获取病症描述文本之后，基于目标对象的标识信息获取该目标对象的历史病历文本，并对当前病症描述文本和历史病历文本进行特征提取处理，得到目标对象的特征信息，该特征信息中包括至少一个特征（如病症名称、病症用药等）。之后，采用
- 30 文本分类模型，对目标对象的特征信息进行处理，得到文本分类模型所输

出的目标对象对应的疾病分布概率和结果解释信息。其中，疾病分布概率用于指示目标对象可能具有的疾病，以及各个疾病的出现概率值；结果解释信息用于描述疾病分布概率的获取依据，示例性地，结果解释信息中包括各个特征在疾病分布概率确定过程中的影响程度。例如，医护人员根据目标对象对应的疾病分布概率和结果解释信息能够快速确定用户对应的疾病，并根据该疾病采取相应的措施，降低医护人员的人工负担。

(2) 在教育领域的书籍分类场景中，工作人员将待分类书籍的总结描述文本（前言、后记、摘要等）作为待检测文本，并将待分类书籍的标识信息（待分类书籍的作者）和待检测文本输入至计算机设备中。计算机设备基于待检测数据的标识信息，获取关联书籍（与待分类书籍具有相同作者的书籍），并对待分类书籍的总结性描述文本和关联书籍的总结性描述文本进行特征提取，确定待检测书籍的特征信息，该特征信息中包括至少一个特征（关键词）。之后，采用文本分类模型对待分类书籍的特征信息进行处理，得到待分类书籍的类型分布概率和结果解释信息。其中，类型分布概率用于指示待分类书籍可能属于的科目，以及属于各个科目的概率值；结果解释信息用于描述类型分布概率的获取依据，示例性地，结果解释信息中包括各个特征在类型分布概率确定过程中的影响程度。在实现时，根据待分类书籍的类型分布概率和结果解释信息能够快速确定该书籍所属的科目，并将该书籍放置在合适的位置，便于后续学生的查找与学习，书籍的快速归类从侧面保证学生所查找的信息的完整性。

(3) 在人工智能领域的用户意图确定场景中，智能机器人在检测到待服务用户之后，与待服务用户进行交谈，获取用户所说的话生成待检测文本，并依据待服务用户的用户身份信息确定待服务用户是否为旧用户。例如，机器人可以启动身份识别系统确定用户身份以获取用户身份信息，也可以由待服务用户主动提供用户身份。若待服务用户为旧用户，则基于用户身份信息获取待服务用户的历史服务信息生成关联文本，并对待检测文本和关联文本进行特征提取，得到待服务用户的特征信息，该特征信息中包括至少一个特征（关键词）。之后，采用文本分类模型对待服务用户的特征信息进行处理，得到待服务用户的意图分布概率和结果解释信息。其中，意图分布概率用于指示待服务用户可能的意图，以及该意图对应的概率值；

结果解释信息用于描述意图分布概率的获取依据，示例性地，结果解释信息中包括各个特征在意图分布概率确定过程中的影响程度。之后智能机器人根据意图分布概率确定待服务用户的用户需求，并将该用户引导至合适的位置。

5 当然，本申请实施例中的文本分类模型还可以运用于其它各个领域，在此不一一举例。

10 为了便于说明，在下述方法实施例中，仅以各步骤的执行主体为计算机设备为例进行介绍说明，计算机设备可以是任何具备计算和存储能力的电子设备。例如，该计算机设备可以是服务器，该服务器可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN(Content Delivery Network, 内容分发网络)、以及大数据和人工智能平台等基础云计算服务的云服务器。再例如，该计算机设备也可以是终端，该终端可以是智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表、智能车载设备等，但并不局限于此。终端以及服务器可以通过有线或无线通信方式进行直接或间接地连接，本申请实施例对此不做限制。需要说明的一点是，在本申请实施例中，各步骤的执行主体可以是同一计算机设备，也可以由多个不同的计算机设备交互配合执行，此处不作限定。还需要说明的一点是，在本申请实施例中，下述文本分类方法的执行主体与下述文本分类模型的训练方法的执行主体可以是相同的计算机设备，也可以是不同的计算机设备，本申请实施例对此不作限定。

25 请参考图 1，其示出了本申请实施例提供的一种文本分类方法的流程图。该方法可以包括以下几个步骤：

步骤 101，获取待检测文本，以及待检测文本对应的关联文本。

30 待检测文本即为未知类别的文本。在本申请实施例中，计算机设备在对待检测文本进行分类之前，获取待检测文本，并获取该待检测文本对应的关联文本。其中，关联文本是指与待检测文本具有关联关系的文本。在一些实施例中，该关联关系包括但不限于以下至少一项：具有相同的作者、

针对相同的对象、具有相同的操作者、具有相同的属性信息等，本申请实施例对此不作限定。其中，上述属性信息为文本浏览时长、文本浏览时刻、文本浏览次数等。

示例性地，以上述关联关系为针对相同的对象为例，计算机设备在获取待检测文本之后，确定该待检测文本所描述的目标对象（人、动物、物体、天气等），进而基于该目标对象，获取针对该目标对象的其它文本作为待检测文本的关联文本。

在一种可能的实施方式中，由用户主动向计算机设备提供上述待检测文本。在一些实施例中，用户根据实际情况确定待检测信息，并向计算机设备或计算机设备的关联设备输入该待检测信息对应的文本内容，计算机设备获取该文本内容作为待检测文本。其中，文本内容的输入方式可以是文字输入、语音输入、图像输入或手势输入，等等，本申请实施例对此不作限定。

在另一种可能的实施方式中，由计算机设备主动获取上述待检测文本。在一些实施例中，计算机设备可以按照一定的时间间隔从网络环境或特定设备中获取上述待检测文本。其中，上述时间间隔可以为 1 秒(second, s)、1 分钟(minute, min)、1 小时(hour, h)、1 天或 1 月等，本申请实施例对此不作限定。

当然，计算机设备可以根据实际情况对待检测文本的获取方式进行灵活设置和调整。示例性地，若待检测文本为待检测病症，则计算机设备可以实时获取该待检测文本，并基于该待检测文本确定用户具有的疾病，保证病症检测的及时性；若待检测文本为待分类文章，则计算机设备按照一定的时间间隔获取待检测文本，并基于该待检测文本确定文章分类，进而依据文章分类将文章存储至合适的位置，在对待检测文本进行分类的同时，使得计算机设备按照一定的时间周期进行运行，避免某个时刻计算机设备负荷过大。

其中，上述待检测文本可以为直接输入文本，也可以为间接输入文本。

在一种可能的实施方式中，上述待检测文本为用户直接输入的文本，计算机设备在获取用户输入内容之后，直接将该输入内容作为待检测文本。在这种情况下，该用户输入内容为文本内容。在一些实施例中，上述内容

的输入方式可以是文字输入、语音输入、图像输入或手势输入，等等，本申请实施例对此不作限定。

在另一种可能的实施方式中，上述待检测文本为用户间接输入的文本，计算机设备在获取用户输入内容之后，对该输入内容进行提取，并将提取出的文本作为待检测文本。在这种情况下，用户输入内容可以为语音内容、图像内容、文字内容等，本申请实施例对此不作限定。示例性地，为了简化用户操作，用户可以直接将输入图像，进而计算机设备根据该图像提取出待检测文本，此时，待检测文本可以为图像中的文字信息。

步骤 102，对待检测文本和关联文本进行特征提取处理，得到特征信息。

在本申请实施例中，计算机设备在获取上述待检测文本和关联文本之后，对该待检测文本和关联文本进行特征提取处理，得出特征信息。其中，特征信息中包括从检测文本和关联文本中提取出的至少一个关键词。一个关键词可以表征一个特征。示例性地，若上述待检测文本为病症描述信息，则特征信息中包括的关键词有病症名称、病症用药等，在实现时可以将关键词病症名称、病症用药等确定为待检测文本的不同特征。

在一些实施例中，计算机设备在提取特征信息时，过滤去除文本（待检测文本和关联文本）中的噪声信息，如标点符号、表情符号等，然后对去除噪声的文本进行分词处理，并提取出实体作为关键词，进而生成上述特征信息。

步骤 103，采用注意力机制对特征信息进行处理，得到各个特征分别对应的注意力权重。

在本申请实施例中，计算机设备在获取上述特征信息之后，采用注意力机制对该特征信息进行处理，得到各个特征分别对应的注意力权重。其中，该注意力权重用于指示特征针对上述待检测文本和上述关联文本的重要程度。在一些实施例中，计算机设备可以依据特征在上述待检测文本和上述关联文本中至少之一的出现频率，确定该特征针对待检测文本和关联文本的重要程度。当然，计算机设备也可以依据各个特征之间的关联关系，确定特征针对待检测文本和关联文本的重要程度，示例性地，与某个特征具有相关关系的特征越多，则该特征的重要程度越高，反之，与某个特征具有相关关系的特征越少，则该特征的重要程度越低。也即，与一个特征

相关联的特征数与该特征的重要程度是正相关的。

在本申请实施例中，上述注意力机制包括空间注意力机制和稀疏访问级注意力机制。其中，空间注意力机制用于衡量单个特征的重要程度，稀疏访问级注意力机制用于衡量特征集合的重要程度，该特征集合中包括来自相同文本的特征。

在一些实施例中，上述步骤 103 包括以下几个步骤：1、采用空间注意力机制，对特征信息中的各个特征进行处理，得到各个特征分别对应的空间注意力权重；2、基于各个特征的来源，将来自相同文本的特征进行组合划分，得到多组特征集合；其中，不同特征集合中包括来自不同文本的特征；3、采用稀疏访问级注意力机制，对多组特征集合进行处理，得到各个特征集合分别对应的稀疏访问级注意力权重。

在本申请实施例中，上述注意力权重包括空间注意力权重和稀疏访问级注意力权重。

在一些实施例中，计算机设备在获取空间注意力权重时，采用空间注意力机制对特征信息中的各个特征进行处理，得到各个特征分别对应的空间注意力权重。其中，空间注意力权重用于指示特征信息中单个特征的重要程度。

在一些实施例中，计算机设备在获取稀疏访问级注意力权重时，基于各个特征的来源，将来自相同文本的特征进行组合划分，得到多组特征集合。其中，不同特征集合中包括来自不同文件的特征，相同特征集合中包括来自相同文件的特征。计算机设备采用稀疏访问级注意力机制，对多组特征集合进行处理，得到各个特征集合分别对应的稀疏访问级注意力权重。其中，稀疏访问级注意力权重用于指示特征集合的重要程度。

当然，在实际运用中，计算机设备可以在获取特征信息时，确定多组特征集合。示例性地，计算机设备在获取各个特征之后，采用特征向量来表示特征集合，一个特征向量对应一个特征集合，并将多个特征集合组成特征矩阵，并以该特征矩阵来表示特征信息。在这种情况下，计算机设备可以直接采用稀疏访问级注意力机制对特征信息进行处理，不需要对各个特征进行组合划分。

步骤 104，基于特征信息和注意力权重，获取待检测文本对应的类别检

测结果和结果解释信息。

在本申请实施例中，计算机设备在获取上述特征信息和上述注意力权重之后，基于特征信息和注意力权重，获取待检测文本对应的类别检测结果和结果解释信息。其中，上述类别检测结果包括待检测文本对应的类别分布概率，上述结果解释信息用于描述各个特征针对类别检测结果的影响程度。

其中，上述注意力机制包括空间注意力机制和稀疏访问级注意力机制，该空间注意力机制用于指示单个特征的重要程度，该稀疏访问级注意力机制用于指示特征集合的重要程度，特征集合中包括来自相同文本的特征。

在一些实施例中，上述步骤 104 中的“基于所述特征信息和所述注意力权重，获取所述待检测文本对应的结果解释信息”可以通过以下几个步骤实现：获取特征对应的空间注意力权重和稀疏访问级注意力权重；根据空间注意力权重和稀疏访问级注意力权重，确定特征对应的最终注意力权重；基于最终注意力权重和特征对应的嵌入向量，确定特征对应的贡献度系数。将各个所述特征分别对应的贡献度系数确定为所述待检测文本对应的结果解释信息。

在本申请实施例中，计算机设备在获取上述结果解释信息时，对特征信息中的各个特征进行遍历，获取各个特征对应的贡献度系数，进而生成结果解释信息。其中，该结果解释信息中包括各个特征分别对应的贡献度系数。

以单个特征为例，计算机设备在获取特征对应的贡献度系数时，获取特征对应的空间注意力权重和稀疏访问级注意力权重。其中，该空间注意力权重用于指示单个特征的重要程度，稀疏访问级注意力权重用于指示特征所处特征集合的重要程度。然后根据空间注意力权重和稀疏访问级注意力权重，确定特征对应的最终注意力权重，进而基于该最终注意力权重和特征对应的嵌入向量，确定特征对应的贡献度系数。

在一些实施例中，计算机设备在获取类别检测结果和结果解释信息之后，可以向用户展示该类别检测结果和结果解释信息，用户能够基于类别检测结果确定待检测文本可能的分类，以及待检测文本的各个分类分别对应的概率值，而且，用户能够基于结果解释信息确定类别检测结果的获取

依据。

综上所述，本申请实施例提供的技术方案中，通过注意力机制对特征信息进行处理，得到用于指示各个特征针对文本的重要程度的注意力权重，进而依据特种信息和注意力权重获取类别检测结果和结果解释信息，一方面，类别检测结果中包括类别分布概率，使得类别检测结果具有多样性，避免仅仅输出一个类别造成的输出单一；另一方面，结果解释信息用于描述各个特征针对类别检测结果的影响程度，为类别检测结果提供可解释性，提高了类别检测结果的可信度，用户能够基于结果解释信息确定类别检测结果的获取依据，使得用户能够更加直观地理解类别检测结果。

另外，在对待检测文本进行分类时，结合关联文本，保证类别检测结果的准确性；而且，对于单个特征，采用空间注意力机制衡量单个特征的重要程度，采用稀疏访问级注意力机制衡量特征集合的重要程度，且特征集合中包括来自相同文本的特征，后续采用空间注意力权重和稀疏访问级注意力权重确定特征的贡献度系数，保证针对特征的贡献度评估的准确性，使得针对类别检测结果的结果解释信息更加可靠。

在本申请实施例中，上述方法也可以通过文本分类模型来实现，即上述类别检测结果和上述结果解释信息是文本分类模型通过上述特征信息获取的。其中，文本分类模型采用不同的模型参数对所述特征信息进行多次处理。在一些实施例中，上述类别检测结果中包括 n 组类别检测子结果，上述结果解释信息中包括 n 组类别检测子结果对应的 n 个结果解释子信息， n 为正整数。需要说明的一点是，上述类别检测子结果与上述结果解释子信息具有一一对应的关系，也就是说，文本分类模型在采用某一组模型参数对特征信息进行处理后，可以得到本次处理的类别检测子结果和结果解释子信息。

在本申请实施例中，计算机设备在获取上述类别检测结果和结果解释信息之后，对类别检测结果中的类别检测子结果进行求平均处理，并对结果解释信息中的结果解释子信息进行求平均处理。

在一些实施例中，针对类别检测结果的处理步骤如下：从 n 组类别检测子结果中分别获取目标类别对应的概率值；对目标类别的概率值进行求平均处理，得到目标类别对应的最终概率值；生成待检测文本对应的最终

类别检测结果，最终类别检测结果中包括各个类别分别对应的最终概率值。

类别检测子结果中包括待检测文本对应的类别分布概率。其中，不同的类别检测结果为文本分类模型采用不同模型参数对待检测文本进行处理后得到的，也就是说，不同的类别检测结果中包含的类别分布概率存在差异。在本申请实施例中，计算机设备在获取上述 n 组类别检测子结果之后，对该 n 组类别检测子结果中的类别概率分布进行求平均处理。

在一些实施例中，计算机设备在获取上述 n 组类别检测子结果之后，从 n 组类别检测子结果中分别获取目标类别对应的概率值。其中，该目标类别可以是针对待检测文本的任意检测结果中的任意类别，计算机设备对目标类别的概率值进行求平均处理，得到目标类别对应的最终概率值，进而生成待检测文本对应的最终类别检测结果。其中，最终类别检测结果中包括各个类别分别对应的最终概率值。

在一些实施例中，针对结果解释信息的处理步骤如下：从 n 个结果解释子信息中分别获取目标特征对应的贡献度系数；对目标特征对应的贡献度系数进行求平均处理，得到目标特征对应的最终贡献度系数；生成待检测文本对应的最终结果解释信息，最终结果解释信息中包括各个所述特征分别对应的最终贡献度系数。

结果解释子信息中包括各个特征分别对应的贡献度系数。其中，不同的结果解释子信息为文本分类模型采用不同模型参数对待检测文本进行处理后获取的，也就是说，不同的结果解释子信息中包含的各个特征分别对应的贡献度系数存在差异。在本申请实施例中，计算机设备在获取上述 n 组结果解释子信息之后，对该 n 个结果解释子信息中的贡献度系数进行求平均处理。

在一些实施例中，计算机设备在获取上述 n 个结果解释子信息之后，从 n 个结果解释子信息中分别获取目标特征对应的贡献度系数。其中，该目标特征可以是任意特征，然后计算机设备对目标特征的贡献度系数进行求平均处理，得到目标特征对应的最终贡献度系数，进而生成待检测文本对应的最终结果解释信息。其中，最终结果解释信息中包括各个特征分别对应的最终贡献度系数。

在本申请实施例中，计算机设备获取上述最终类别检测结果之后，可以依

据该最终类别检测结果确定文本分类模型针对待检测文本的置信度，实现步骤如下：基于各个类别分别对应的最终概率值，获取最终类别检测结果的度量参数；其中，度量参数用于度量最终类别检测结果的混乱程度；根据度量参数，确定文本分类模型针对待检测文本的置信度；响应于置信度满足条件，确定最终类别检测结果为准确检测结果；响应于置信度不满足条件，生成提示信息，提示信息用于提醒对待检测文本进行人工检测。

上述度量参数可以为最终类别检测结果中类别分布概率的方差或熵，该度量参数与混乱程度呈正相关关系，即度量参数越大，最终类别检测结果越混乱。在本申请实施例中，计算机设备在获取上述最终类别检测结果之后，基于该最终类别检测结果中各个类别分别对应的最终概率值，获取最终检测结果的度量参数，并根据该度量参数，确定文本分类模型针对待检测文本的置信度。其中，该置信度与文本分类模型针对待检测文本的可靠程度呈正比，即置信度越高，文本分类模型针对待检测文本的分类越可靠。

上述条件是指针对置信度的判断条件。示例性地，该条件为置信度大于阈值。在本申请实施例中，计算机设备在获取上述置信度之后，对该置信度进行检测。若置信度满足条件，则确定文本分类模型针对待检测文本的可靠程度满足要求，确定上述最终类别检测结果为准确检测结果；若置信度不满足条件，则确定文本分类模型针对待检测文本的可靠程度不满足要求，进而生成提示信息，提醒工作人员对待检测文本进行人工检测。当然，在实际运用中，工作人员可以对上述条件进行调整，如在依据度量参数确定不确定度的情况下，将条件设置为不确定度小于目标值。在实现时，可以按照公式（1-1）利用所得到的类别分布概率的均值的熵来量化预测的不确定度：

$$E p_{i_{\epsilon}} = - \sum_{i=1}^2 p_{i \log} (p_i) \quad (1-1);$$

其中， p_i 是类别分布概率的均值。

如此，可以量化认知不确定度，与相关技术中的其它模型相比，可以确定本申请实施例所提出文本分类模型的置信度，即可以确定文本分类模型对当前预测结果是否自信，从而提高文本分类模型的可信度。

在一些实施例中，上述不同的模型参数是根据多次随机采样获取的。示例性地，在对待检测模型进行处理时，在每次对待检测文本进行处理之前，采用蒙特卡洛采样获取一组模型参数，进而基于该模型参数对待检测文本进行处理。当然，在实际运用中，可以根据实际情况对多组不同的模型参数的获取方式进行灵活设备与调整，本申请实施例对此不作限定。

下面，对本申请实施例提供的中文本分类模型的训练方法进行介绍。

请参考图 2，其示出了本申请实施例提供的文本分类模型的训练方法的流程图。该方法可以包括以下几个步骤：

步骤 201，获取多个样本文本，以及各个样本文本对应的类别信息。

样本文本是指已知类别的文本。在本申请实施例中，计算机设备在对文本分类模型进行训练之前，获取多个样本文本，并获取各个样本文本对应的类别信息。其中，单个样本文本中包括第一文本和第二文本，且第二文本为第一文本的关联文本。示例性地，若第一文本为用户的当前病症文本，则第二文本为用户的历史病历文本；若第一文本为用户的当前提交文章，则第二文本为用户的历史提交文本；若第一文本为用户的当前浏览网页，则第二文本为用户在最近浏览的有限个历史浏览网页。

步骤 202，从各个样本文本中分别提取样本特征信息。

在本申请实施例中，计算机设备在获取上述样本文本之后，从各个样本文本中分别提取样本特征信息。其中，该样本特征信息中包括至少一个样本特征，样本特征是指依据文本记录内容提取出的关键词。在一些实施例中，计算机设备在提取样本特征信息时，过滤去除样本文本中的噪声信息，如标点符号、表情符号等，并对去除噪声的样本文本进行分词处理，并提取出实体词作为关键词，进而生成上述样本特征信息。

步骤 203，采用样本特征信息和类别信息对文本分类模型进行训练。

在本申请实施例中，计算机设备在获取上述样本特征信息之后，采用样本特征信息和类别信息对文本分类模型进行训练。在一些实施例中，计算机设备依据样本文本与样本特征信息之间的对应关系，将类别信息作为样本特征信息的标签信息，生成训练样本，并采用训练样本对文本分类模型进行训练。

在一些实施例中，计算机设备在对文本分类模型进行训练的过程中，采用多组模型参数，对样本特征信息进行多次处理，获取多组模型输出结果，并基于多组模型输出结果，以及类别信息，确定文本分类模型的损失函数值。其中，损失函数用于指示文本分类模型的输出结果的准确性。之后，计算机设备基于损失函数值对文本分类模型的参数进行调整，直至损失函数收敛。

示例性地，文本分类模型的损失函数为可以用公式 (1-2) 表示：

$$L_{esm} = \log \frac{1}{T_{mc}} \sum_{i=1}^N \sum_c^C \sum_{t=1}^{T_{mc}} \exp \left(\hat{y}_{i,c,t} - \log \sum_{j=1}^l \hat{y}_{i,c,t,j} \right) + \lambda \sum_{k=1}^L (\|W_k\|_2^2 + \|b_k\|_2^2) \quad (1-2);$$

其中， T_{mc} 是指文本分类模型针对样本特征信息的处理次数，且每次采用不同的模型参数对样本特征信息进行处理， N 是指样本特征信息的数量， C 是指样本特征信息， c 是指样本特征信息中的特征， $\hat{y}_{i,c,t}$ 是指第 i 个样本特征信息对应的类别信息， $\hat{y}_{i,c,t,j}$ 是指文本分类模型的第 j 个输出结果， l 是指文本分类模型输出的类别数量， λ 是指文本分类模型的平衡参数， L 是指文本分类模型的网络层数， W_k 、 b_k 是指第 k 层网络对应的模型参数。

在本申请实施中，上述文本分类模型中引入注意力机制，该注意力机制用于确定样本特征信息中的各个样本特征针对样本文本中的重要程度，该重要程度用于确定各个样本特征对模型输出结果的影响程度。在一些实施例中，上述注意力机制包括空间注意力机制和稀疏访问级注意力机制。在实际运用中，空间注意力机制用于衡量单个特征的重要程度，稀疏访问级注意力机制用于衡量特征集合的重要程度，特征集合中包括来自相同文本的特征。

在本申请实施例中，还会获取验证文本，在利用样本文本训练网络之后，会利用验证文本验证训练之后模型的评估指标，按照此流程迭代。观察在每轮验证集上的结果，如果验证集上的结果连续五轮没有提升，保留之前验证集上最好的结果的模型参数。

综上所述，本申请实施例提供的技术方案中，通过样本特征信息和类别信息训练得到文本分类模型，且文本分类模型中引入注意力机制，在模型运用过程中，采用注意力机制对特征信息进行处理，确定用于指示各个特征针对文本的重要程度的注意力权重，使得模型在输出结果时，能够同

时输出结果的获取依据，使得用户能够更加直观地理解模型输出的结果。

下面，结合参考图 3，以医疗领域为例，对本申请实施例提供的文本分类方法进行介绍。在医疗领域中，上述文本分类模型可以用于疾病预测。

5 步骤 301，获取目标对象的当前病症描述文本，以及目标对象的历史病历文本。

目标对象是任意进行疾病预测的用户。当前病症描述文本用于指示目标对象当前出现的病症。在一些实施例中，该当前病症描述文本为医护人员根据目标对象的实际情况所确定，并输入至计算机设备的文本信息。

10 在本申请实施例中，在对目标对象进行疾病预测之前，获取目标对象的当前病症描述文本，以及目标对象的历史病历文本。在一些实施例中，计算机设备在确定对目标对象进行疾病预测之后，基于目标对象的标识信息，从病历存储库中获取目标对象的历史病历文本。其中，上述标识信息用于指示唯一的对象，示例性地，标识信息为身份证号码、医院分配的
15 的独特编号、手机号码等。

当然，在实际运用中，上述历史病历文本也可以是由医护人员根据纸质文件总结获取，并输入至计算机设备的。

步骤 302，对所述当前病历描述文本，以及历史病历描述文本进行特征提取处理，获取病历特征信息。

20 在本申请实施例中，计算机设备在获取上述当前病历描述文本，以及历史病历描述文本之后，对当前病历描述文本，以及历史病历描述文本进行特征提取处理，获取病历特征信息。在实现时，可以利用自然语言处理技术分析出当前病历描述文本，以及历史病历描述文本的病历特征信息，其中，该病历特征信息可以为结构化的离散特征，并且该病历特征信息中
25 包括病历特征，该病历特征是指依据当前病历描述文本或历史病历描述文本的记录内容提取出的关键词，如病症名称、病症用药等。

在本申请实施例中，计算机设备在对各个病历文本进行特征提取时，将从相同病历文本中提取的病历特征作为一个病历特征集合，进而由多个病历特征集合组合获取病历特征信息。示例性地，如图 4 所示，以矩阵的形式对病历特征信息 41 进行表示，每一列中的病历特征为一个病历特征集
30

合，依据各个病历文本的生成时间，对病历特征集合进行排序，最早生成的病历文本对应的病历特征集合排在第一列，最晚生成的病历文本对应的病历特征集合排在最后一列。

步骤 303，利用文本分类模型对病历特征信息进行预测处理，获取疾病
5 预测结果和结果解释信息。

文本分类模型是通过训练得到的深度学习模型。在本申请实施例中，
计算机设备在获取上述病历特征信息之后，利用文本分类模型对该病历特
征信息进行预测处理，进而获取文本分类模型输出的疾病预测结果和结果
解释信息。其中，疾病预测结果中包括针对目标对象的疾病预测概率分布，
10 结果解释信息用于描述各个病症特征针对疾病预测结果的影响程度。

在一些实施例中，文本分类模型中引入空间注意力机制和稀疏访问级
注意力机制。示例性地，如图 4 所示，将病历特征信息 41 输入至文本分类
模型之后，通过全连接层 42 获取每个病历文本对应的病症特征集合的嵌入
向量，再由特征编码器 43 捕获输入病症特征复杂的关联关系，将目标对象
15 的病症特征信息的嵌入矩阵映射到隐藏状态，进而依据空间注意力机制 44
生成用于衡量单个病症特征的重要程度的空间注意力权重，依据稀疏访问
级注意力机制 45 生成用于衡量病症特征集合的重要程度的稀疏访问级注
意力权重，之后，将病症特征信息的嵌入矩阵、空间注意力权重和稀疏访问
级注意力权重输入至患者表示模块 46 获取目标对象的表示信息，进而通过
20 线性层 47 将表示信息映射到逻辑值，最后依据分类器 48 将逻辑值转化为
概率值，进而生成疾病预测结果。另外，在文本分类模型对病症特征信息
41 进行处理的过程中，采用病症特征的空间注意力权重、稀疏访问级注
意力机制和嵌入向量，确定病症特征对应的贡献度系数。

示例性地，假设病症特征信息的特征矩阵为 X ，则依据全连接层获取
25 的嵌入向量 E 如公式 (1-3) 所示：

$$E = W_e X \quad (1-3);$$

其中， W_e 是可学习的参数矩阵；

病症特征信息的嵌入矩阵映射到隐藏状态后 H 如公式 (1-4) 所示：

$$H = f(E) \quad (1-4);$$

其中， $f()$ 是任意合适的基于神经网络的特征编码器，如循环神经网络

络、卷积神经网络等；

空间注意力权重 β 如公式(1-5)所示：

$$\beta = \tanh(W_{\beta}H + b_{\beta}) \quad (1-5);$$

其中， W_{β} 、 b_{β} 为模型参数；

5 利用稀疏访问级注意力机制确定稀疏访问级注意力权重 α 可以通过公式(1-6)得出：

$$\alpha = (\text{softmax}(\delta) + \text{sparsemax}(\delta)) / 2 \quad (1-6);$$

其中， $\delta = W_{\delta}H + b_{\delta}$ ， W_{δ} 、 b_{δ} 为模型参数；

患者表示模块获取的病人表示 E_r 如公式(1-7)所示：

$$E_r = \alpha(\beta \odot E)^T \quad (1-7);$$

E_r 通过线性层映射得到的逻辑值 $\tilde{y}$ 如公式(1-8)所示：

$$\tilde{y} = W_c E_r + b_c \quad (1-8);$$

其中， W_c 、 b_c 为模型参数；

逻辑值 $\tilde{y}$ 通过分类器转换获取的概率值 y^* 如公式(1-9)所示：

$$y^* = \text{softmax}(\tilde{y}) \quad (1-9);$$

15 针对第 m 个病症特征在第 t 次检测获取的贡献度系数 $CM[t, m]$ 如公式(1-10)所示：

$$CM[t, m] = W_c^T (\alpha[t] \beta[:, t] \odot W_e[:, m]) \quad (1-10);$$

20 其中， $\alpha[t]$ 是指针对病症特征信息的第 t 次检测获取的第 m 个病症特征的稀疏访问级注意力权重， $\beta[:, t]$ 是指针对病症特征信息的第 t 次检测获取的第 m 个病症特征的空间注意力权重。 $W_e[:, m]$ 是指第 m 个病症特征的嵌入向量。

需要说明的一点是，上述部分内容是对文本分类模型单次检测的介绍，在实际运用中，文本分类模型采用不同的模型参数对同一病症特征信息进行多次检测后，对多次检测获取的结果进行求平均处理，得到最终疾病预测结果和最终结果解释信息。

步骤304，在用户界面中显示疾病预测结果和结果解释信息。

在本申请实施例中，计算机设备在通过文本分类模型获取疾病预测结果和结果解释信息之后，在用户界面中显示疾病预测结果和结果解释信息。

30 其中，结果解释信息可以与疾病预测结果显示在同一界面中，也可以显示

在不同界面中，本申请实施例对此不作限定。

示例性地，如图 5 所示，在用户界面 50 中包括目标对象的当前病症描述文本 51，以及通过文本分类模型所获取的疾病预测结果 52。在一些实施例中，在用户点击控件 53 之后，在用户界面中显示结果解释信息。

5

下述为本申请装置实施例，可以用于执行本申请方法实施例。对于本申请装置实施例中未披露的细节，请参照本申请方法实施例。

请参考图 6，其示出了本申请实施例提供的文本分类装置的框图。该装置具有实现上述文本分类方法的功能，所述功能可以由硬件实现，也可以由硬件执行相应的软件实现。该装置可以是计算机设备，也可以设置于计算机设备中。该装置 600 可以包括：

文本获取模块 610，配置为获取待检测文本，以及所述待检测文本对应的关联文本；特征提取模块 620，配置为对所述待检测文本和所述关联文本进行特征提取处理，得到特征信息；其中，所述特征信息中包括至少一个特征；权重获取模块 630，配置为采用注意力机制对所述特征信息进行处理，得到各个所述特征分别对应的注意力权重，所述注意力权重用于指示所述特征针对所述待检测文本和所述关联文本的重要程度；结果获取模块 640，配置为基于所述特征信息和所述注意力权重，获取所述待检测文本对应的类别检测结果和结果解释信息；其中，所述类别检测结果包括所述待检测文本对应的类别分布概率，所述结果解释信息用于描述各个所述特征针对所述类别检测结果的影响程度。

在一些实施例中，所述结果获取模块 640，配置为获取所述特征对应的空间注意力权重和稀疏访问级注意力权重；其中，所述空间注意力权重用于指示单个所述特征的重要程度，所述稀疏访问级注意力权重用于指示所述特征所处特征集合的重要程度，所述特征集合中包括来自相同文本的特征；根据所述空间注意力权重和所述稀疏访问级注意力权重，确定所述特征对应的最终注意力权重；基于所述最终注意力权重和所述特征对应的嵌入向量，确定所述特征对应的贡献度系数；将各个所述特征分别对应的贡献度系数确定为所述待检测文本对应的结果解释信息。

在一些实施例中，所述权重获取模块 630，配置为采用空间注意力机制，

对所述特征信息中的各个所述特征进行处理，得到各个所述特征分别对应的空间注意力权重；基于各个所述特征的来源，将来自相同文本的所述特征进行组合划分，得到多组特征集合；其中，不同特征集合中包括来自不同文本的特征；采用稀疏访问级注意力机制，对所述多组特征集合进行处理，得到各个所述特征集合分别对应的稀疏访问级注意力权重。

在一些实施例中，所述类别检测结果是利用文本分类模型对所述特征信息进行预测处理得到的；其中，所述文本分类模型采用不同的模型参数对所述特征信息进行多次处理，所述类别检测结果中包括 n 组类别检测子结果，所述结果解释信息中包括所述 n 组类别检测子结果对应的 n 个结果解释子信息，所述类别检测子结果与所述结果解释子信息具有一一对应的关系， n 为正整数。

在一些实施例中，如图 7 所示，所述装置 600，还包括：结果处理模块 650，配置为从所述 n 组类别检测子结果中分别获取目标类别对应的概率值；对所述目标类别的概率值进行求平均处理，得到所述目标类别对应的最终概率值；生成所述待检测文本对应的最终类别检测结果，所述最终类别检测结果中包括各个类别分别对应的最终概率值。

在一些实施例中，如图 7 所示，所述装置 600，还包括：置信度检测模块 660，配置为基于各个所述类别分别对应的最终概率值，获取所述最终类别检测结果的度量参数；其中，所述度量参数用于度量所述最终类别检测结果的混乱程度；根据所述度量参数，确定所述文本分类模型针对所述待检测文本的置信度；响应于所述置信度满足条件，确定所述最终类别检测结果为准确检测结果；响应于所述置信度不满足条件，生成提示信息，所述提示信息用于提醒对所述待检测文本进行人工检测。

在一些实施例中，所述不同的模型参数是根据多次随机采样获取的。

综上所述，本申请实施例提供的技术方案中，通过注意力机制对特征信息进行处理，确定用于指示各个特征针对文本的重要程度的注意力权重，进而依据特种信息和注意力权重获取类别检测结果和结果解释信息，一方面，类别检测结果中包括类别分布概率，使得类别检测结果具有多样性，避免仅仅输出一个类别造成的输出单一；另一方面，结果解释信息用于描述各个特征针对类别检测结果的影响程度，为类别检测结果提供可解释性，

提高了类别检测结果的可信度，用户能够基于结果解释信息确定类别检测结果的获取依据，使得用户能够更加直观地理解类别检测结果。

请参考图 8，其示出了本申请实施例提供的文本分类模型的训练装置的框图。该装置具有实现上述文本分类模型的训练方法的功能，所述功能可以由硬件实现，也可以由硬件执行相应的软件实现。该装置可以是计算机设备，也可以设置于计算机设备中。该装置 800 可以包括：

样本获取模块 810，配置为获取多个样本文本，以及各个所述样本文本对应的类别信息；特征获取模块 820，配置为从各个所述样本文本中分别提取样本特征信息；模型训练模块 830，配置为采用所述样本特征信息和所述类别信息对所述文本分类模型进行训练；其中，所述文本分类模型中引入注意力机制，所述注意力机制用于确定所述样本特征信息中的各个样本特征针对所述样本文本中的重要程度，所述重要程度用于确定各个所述样本特征对模型输出结果的影响程度。

在一些实施例中，所述模型训练模块 830，配置为采用多组模型参数，对所述样本特征信息进行多次处理，得到多组模型输出结果；基于所述多组模型输出结果，以及所述类别信息，确定所述文本分类模型的损失函数值；其中，损失函数用于指示所述文本分类模型的输出结果的准确性；基于所述损失函数值对所述文本分类模型的参数进行调整，直至所述损失函数收敛。

在一些实施例中，所述注意力机制包括空间注意力机制和稀疏访问级注意力机制；其中，所述空间注意力机制用于衡量单个特征的重要程度，所述稀疏访问级注意力机制用于衡量特征集合的重要程度，所述特征集合中包括来自相同文本的特征。

综上所述，本申请实施例提供的技术方案中，通过样本特征信息和类别信息训练得到文本分类模型，且文本分类模型中引入注意力机制，在模型运用过程中，采用注意力机制对特征信息进行处理，确定用于指示各个特征针对文本的重要程度的注意力权重，使得模型在输出结果时，能够同时输出结果的获取依据，使得用户能够更加直观地理解模型输出的结果。

需要说明的是，上述实施例提供的装置，在实现其功能时，仅以上述

各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将设备的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。另外，上述实施例提供的装置与方法实施例属于同一构思，其实现过程可参考对应的方法实施例。

5

请参考图 9，其示出了本申请实施例提供的计算机设备的结构框图。该计算机设备可用于实现上述文本分类方法或文本分类模型的训练方法的功能。如图 9 所示，该计算机设备 900 包括：中央处理单元（Central Processing Unit, CPU）901、包括随机存取存储器（Random Access Memory, RAM）902 和只读存储器（Read Only Memory, ROM）903 的系统存储器 904，以及连接系统存储器 904 和中央处理单元 901 的系统总线 905。计算机设备 900 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, I/O 系统）906，和用于存储操作系统 913、应用程序 914 和其他程序模块 915 的大容量存储设备 907。

10

基本输入/输出系统 906 包括显示器 908 和输入设备 909。其中显示器 908 和输入设备 909 都通过连接到系统总线 905 的输入输出控制器 910 连接到中央处理单元 901。

15

大容量存储设备 907 通过连接到系统总线 905 的大容量存储控制器（未示出）连接到中央处理单元 901。大容量存储设备 907 及其相关联的计算机可读介质为计算机设备 900 提供非易失性存储。

20

不失一般性，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、EPROM（Erasable Programmable Read Only Memory，可擦除可编程只读存储器）、EEPROM（Electrically Erasable Programmable Read Only Memory，电可擦可编程只读存储器）、闪存或其他固态存储其技术，CD-ROM、DVD（Digital Video Disc，高密度数字视频光盘）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计算机存储介质不局限于上述几种。上述的系统存储器 904 和大容量存储设备 907 可以统称为存储器。

25

在一些实施例中，计算机设备 900 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 900 可以通过连接在系统总线 905 上的网络接口单元 911 连接到网络 912，或者说，也可以使用网络接口单元 911 来连接到其他类型的网络或远程计算机系统（未示出）。

5 所述存储器还包括计算机程序，该计算机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，以实现上述文本分类方法，或实现上述文本分类模型的训练方法。

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中存储有至少一条指令、至少一段程序、代码集或指令集，所述至少一条指令、所述至少一段程序、所述代码集或所述指令集在被处理器执行时以实现上述文本分类方法，或实现上述文本分类模型的训练方法。

在一些实施例中，该计算机可读存储介质可以包括：ROM（Read Only Memory，只读存储器）、RAM（Random Access Memory，随机存取记忆体）、SSD（Solid State Drives，固态硬盘）或光盘等。其中，随机存取记忆体可以包括 ReRAM（Resistance Random Access Memory，电阻式随机存取记忆体）和 DRAM（Dynamic Random Access Memory，动态随机存取存储器）。

本申请实施例提供一种计算机程序产品或计算机程序，该计算机程序产品或计算机程序包括计算机指令，该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令，处理器执行该计算机指令，使得该计算机设备执行上述文本分类方法，或执行上述文本分类模型的训练方法。

应当理解的是，在本文中提及的“多个”是指两个或两个以上。“和/或”，描述关联对象的关联关系，表示可以存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况。字符“/”一般表示前后关联对象是一种“或”的关系。另外，本文中描述的步骤编号，仅示例性示出了步骤间的一种可能的执行先后顺序，在一些其它实施例中，上述步骤也可以不按照编号顺序来执行，如两个不同编号的步骤同时执行，或者两个不同编号的步骤按照与图示相反的顺序执行，本申请实施例对此不作限定。

30 以上所述仅为本申请的示例性实施例，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、等同替换、改进等，均应包含在本申请的保护范围之内。

权利要求书

1.一种文本分类方法，所述方法包括：

获取待检测文本，以及所述待检测文本对应的关联文本；

对所述待检测文本和所述关联文本进行特征提取处理，得到特征信息；

5 其中，所述特征信息中包括至少一个特征；

采用注意力机制对所述特征信息进行处理，得到各个所述特征对应的注意力权重，所述注意力权重用于指示所述特征针对所述待检测文本和所述关联文本的重要程度；

10 基于所述特征信息和所述注意力权重，获取所述待检测文本对应的类别检测结果和结果解释信息；其中，所述类别检测结果包括所述待检测文本对应的类别分布概率，所述结果解释信息用于描述各个所述特征针对所述类别检测结果的影响程度。

2.根据权利要求1所述的方法，其中，基于所述特征信息和所述注意力权重，获取所述待检测文本对应的结果解释信息，包括：

15 确定所述特征对应的空间注意力权重和稀疏访问级注意力权重；其中，所述空间注意力权重用于指示单个所述特征的重要程度，所述稀疏访问级注意力权重用于指示所述特征所处特征集合的重要程度，所述特征集合中包括来自相同文本的特征；

20 根据所述空间注意力权重和所述稀疏访问级注意力权重，确定所述特征对应的最终注意力权重；

基于所述最终注意力权重和所述特征对应的嵌入向量，确定所述特征对应的贡献度系数；

将各个所述特征分别对应的贡献度系数确定为所述待检测文本对应的结果解释信息。

25 3.根据权利要求1所述的方法，其中，所述采用注意力机制对所述特征信息进行处理，得到各个所述特征分别对应的注意力权重，包括：

采用空间注意力机制，对所述特征信息中的各个所述特征进行处理，得到各个所述特征分别对应的空间注意力权重；

基于各个所述特征的来源，将来自相同文本的所述特征进行组合划分，

得到多组特征集合；其中，不同特征集合中包括来自不同文本的特征；

采用稀疏访问级注意力机制，对所述多组特征集合进行处理，得到各个所述特征集合分别对应的稀疏访问级注意力权重。

4.根据权利要求 1 至 3 任一项所述的方法，其中，所述类别检测结果是
5 利用文本分类模型对所述特征信息进行预测处理得到的；

其中，所述文本分类模型采用不同的模型参数对所述特征信息进行多次处理，得到所述类别检测结果，所述类别检测结果中包括 n 组类别检测子结果，所述结果解释信息中包括所述 n 组类别检测子结果对应的 n 个结果解释子信息，所述类别检测子结果与所述结果解释子信息具有一一对应的
10 的关系， n 为正整数。

5.根据权利要求 4 所述的方法，其中，所述基于所述特征信息和所述注意力权重，获取所述待检测文本对应的类别检测结果之后，还包括：

从所述 n 组类别检测子结果中分别获取目标类别对应的概率值，所述目标类别为针对待检测文本的任意检测结果中的任意类别；

15 对所述目标类别的概率值进行求平均处理，得到所述目标类别对应的最终概率值；

生成所述待检测文本对应的最终类别检测结果，所述最终类别检测结果中包括各个类别分别对应的最终概率值。

6.根据权利要求 5 所述的方法，其中，所述生成所述待检测文本对应的
20 最终类别检测结果之后，还包括：

基于各个所述类别分别对应的最终概率值，获取所述最终类别检测结果的度量参数；其中，所述度量参数用于度量所述最终类别检测结果的混乱程度；

根据所述度量参数，确定所述文本分类模型针对所述待检测文本的置
25 信度；

响应于所述置信度满足条件，确定所述最终类别检测结果为准确检测结果；

响应于所述置信度不满足条件，生成提示信息，所述提示信息用于提醒对所述待检测文本进行人工检测。

30 7.根据权利要求 4 所述的方法，其中，所述基于所述特征信息和所述注

注意力权重，确定所述待检测文本对应的结果解释信息之后，还包括：

从所述 n 个结果解释子信息中分别获取目标特征对应的贡献度系数，
所述目标特征为所述特征信息中的任意特征；

对所述目标特征对应的贡献度系数进行求平均处理，得到所述目标特
5 征对应的最终贡献度系数；

生成所述待检测文本对应的最终结果解释信息，所述最终结果解释信息
中包括各个所述特征分别对应的最终贡献度系数。

8.根据权利要求 4 所述的方法，其中，所述不同的模型参数是根据多次
随机采样获取的。

10 9.一种文本分类模型的训练方法，所述方法包括：

获取多个样本文本，以及各个所述样本文本对应的类别信息；

从各个所述样本文本中分别提取样本特征信息；

采用所述样本特征信息和所述类别信息对所述文本分类模型进行训练；

其中，所述文本分类模型中引入注意力机制，所述注意力机制用于确
15 定所述样本特征信息中的各个样本特征针对所述样本文本中的重要程度，
所述重要程度用于确定各个所述样本特征对模型输出结果的影响程度。

10.根据权利要求 9 所述的方法，其中，所述采用所述样本特征信息和
所述类别信息对所述文本分类模型进行训练，包括：

采用多组模型参数，对所述样本特征信息进行多次处理，得到多组模
20 型输出结果；

基于所述多组模型输出结果，以及所述类别信息，确定所述文本分类
模型的损失函数值；其中，损失函数用于指示所述文本分类模型的输出结
果的准确性；

基于所述损失函数值对所述文本分类模型的参数进行调整，直至所述
25 损失函数收敛。

11.根据权利要求 9 所述的方法，其中，所述注意力机制包括空间注意
力机制和稀疏访问级注意力机制；

其中，所述空间注意力机制用于衡量单个特征的重要程度，所述稀疏
访问级注意力机制用于衡量特征集合的重要程度，所述特征集合中包括来
30 自相同文本的特征。

12.一种文本分类装置，所述装置包括：

文本获取模块，配置为获取待检测文本，以及所述待检测文本对应的关联文本；

5 特征提取模块，配置为对所述待检测文本和所述关联文本进行特征提取处理，得到特征信息；其中，所述特征信息中包括至少一个特征；

权重获取模块，配置为采用注意力机制对所述特征信息进行处理，得到各个所述特征分别对应的注意力权重，所述注意力权重用于指示所述特征针对所述待检测文本和所述关联文本的重要程度；

10 结果获取模块，配置为基于所述特征信息和所述注意力权重，获取所述待检测文本对应的类别检测结果和结果解释信息；其中，所述类别检测结果包括所述待检测文本对应的类别分布概率，所述结果解释信息用于描述各个所述特征针对所述类别检测结果的影响程度。

13.一种文本分类模型的训练装置，所述装置包括：

15 样本获取模块，配置为获取多个样本文本，以及各个所述样本文本对应的类别信息；

特征获取模块，配置为从各个所述样本文本中分别提取样本特征信息；

模型训练模块，配置为采用所述样本特征信息和所述类别信息对所述文本分类模型进行训练；

20 其中，所述文本分类模型中引入注意力机制，所述注意力机制用于确定所述样本特征信息中的各个样本特征针对所述样本文本中的重要程度，所述重要程度用于确定各个所述样本特征对模型输出结果的影响程度。

25 14.一种计算机设备，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条指令、至少一段程序、代码集或指令集，所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现如权利要求 1 至 8 任一项所述的文本分类方法，或实现如权利要求 9 至 11 任一项所述的文本分类模型的训练方法。

30 15.一种计算机可读存储介质，所述存储介质中存储有至少一条指令、至少一段程序、代码集或指令集，所述至少一条指令、所述至少一段程序、所述代码集或指令集由处理器加载并执行以实现如权利要求 1 至 8 任一项所述的文本分类方法，或实现如权利要求 9 至 11 任一项所述的文本分类模

型的训练方法。

16.一种计算机程序产品或计算机程序，所述计算机程序产品或计算机程序包括计算机指令，所述计算机指令存储在计算机可读存储介质中；

5 当计算机设备的处理器从所述计算机可读存储介质读取所述计算机指令，并执行所述计算机指令时，实现权利要求 1 至 8 任一项或者权利要求 9 至 11 任一项所述的方法。

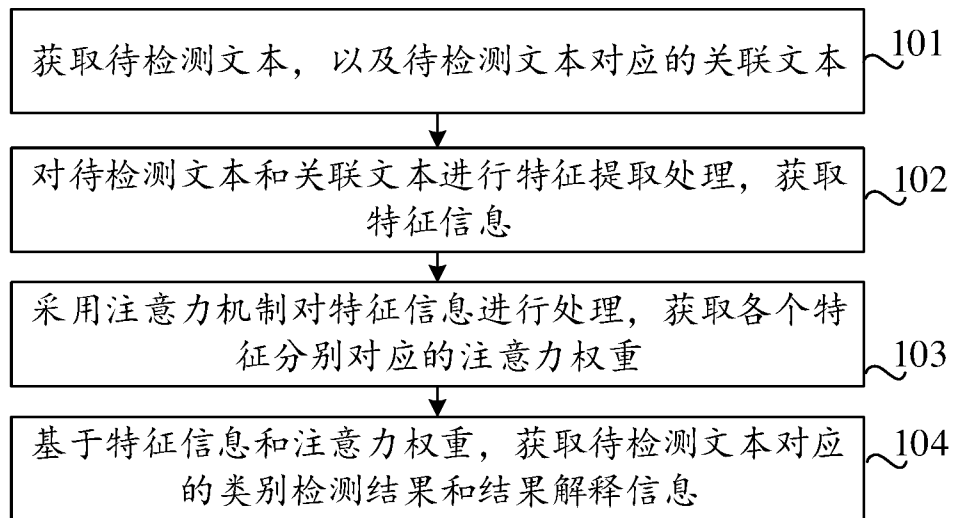


图 1

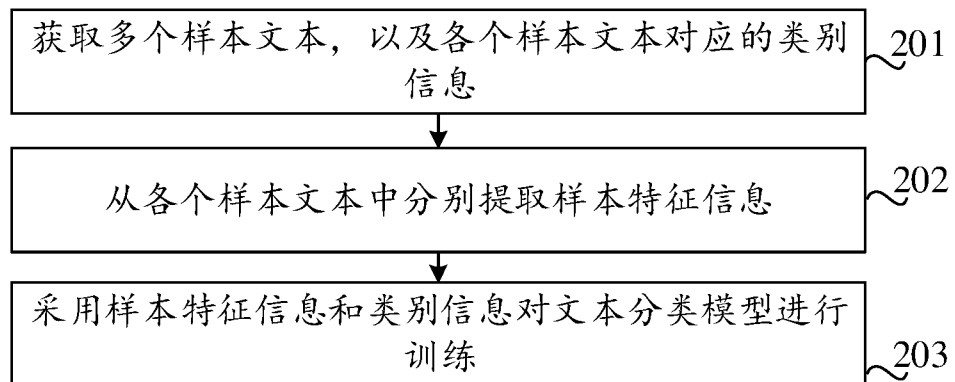


图 2

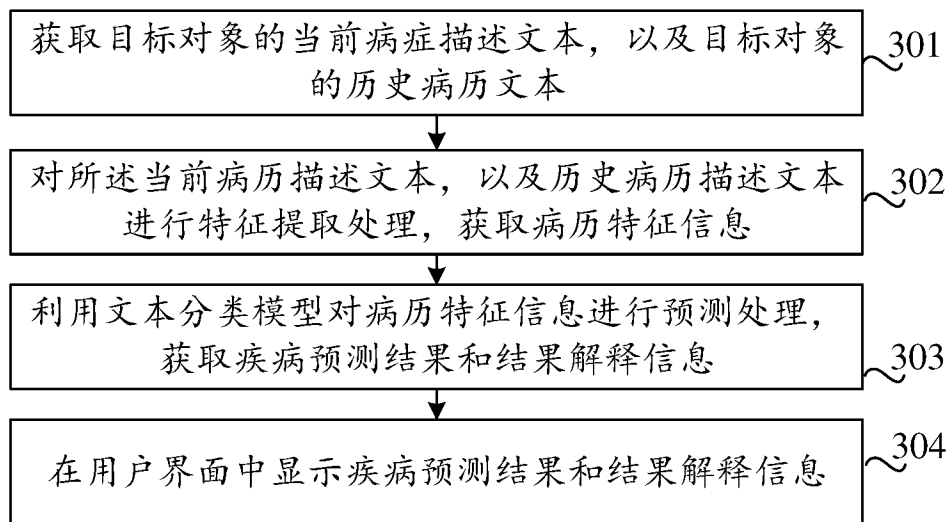


图 3

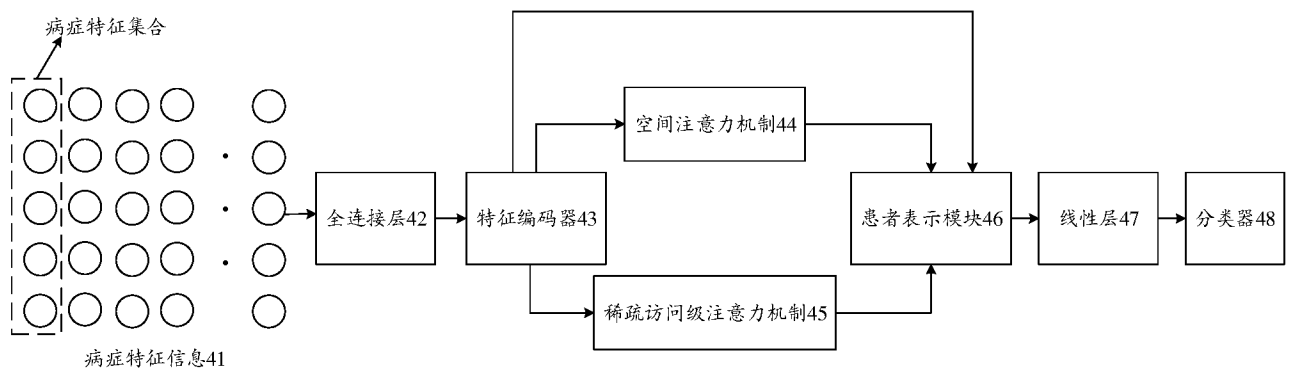


图 4

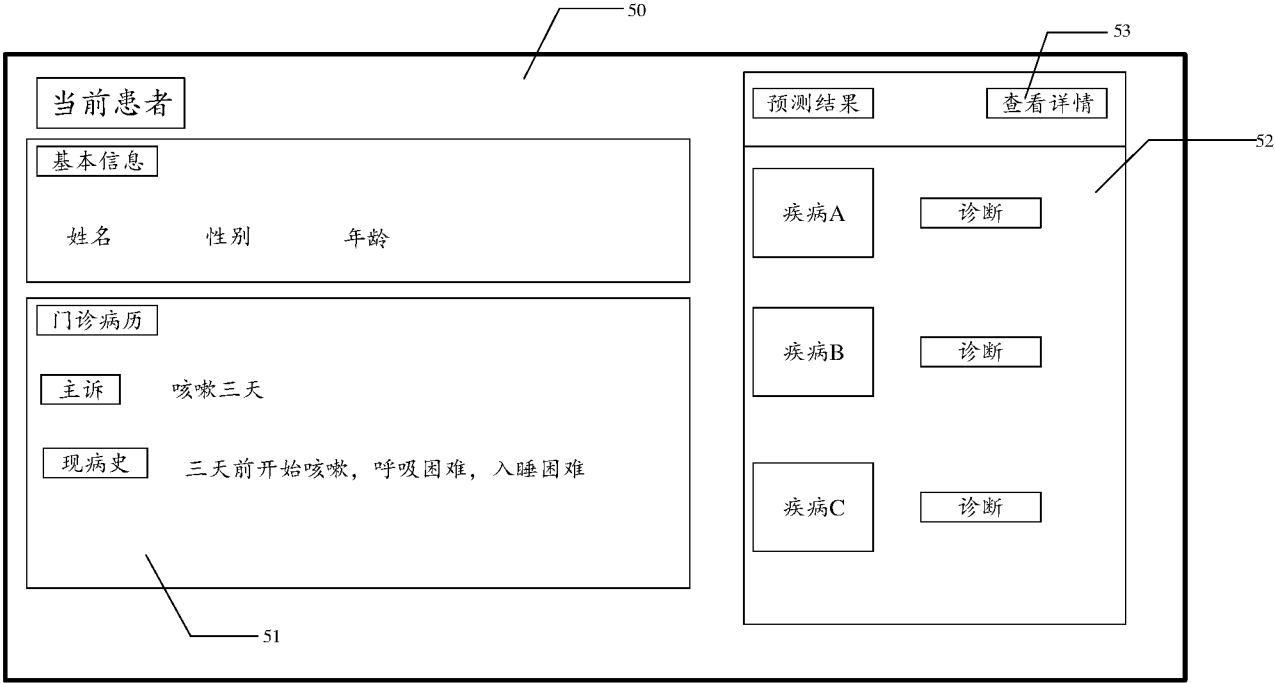


图 5

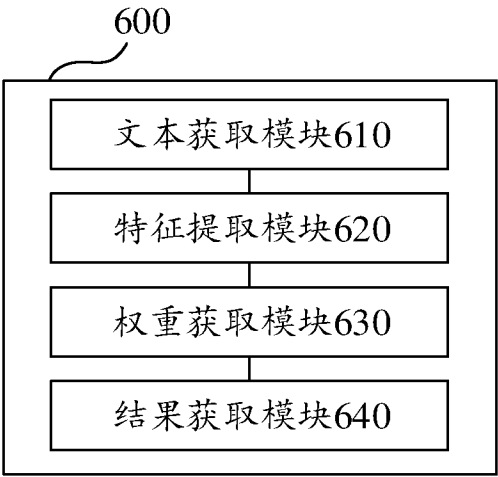


图 6

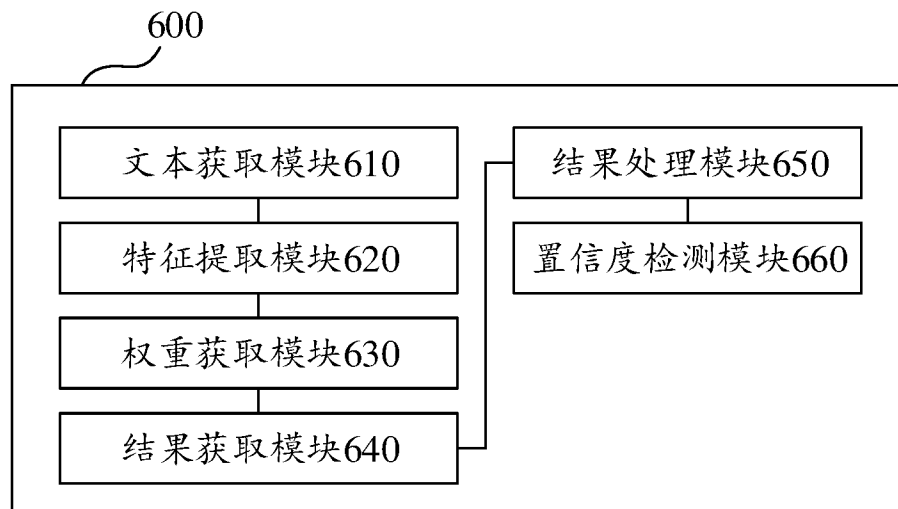


图 7

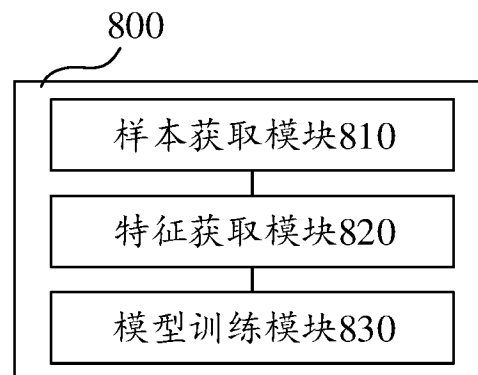


图 8

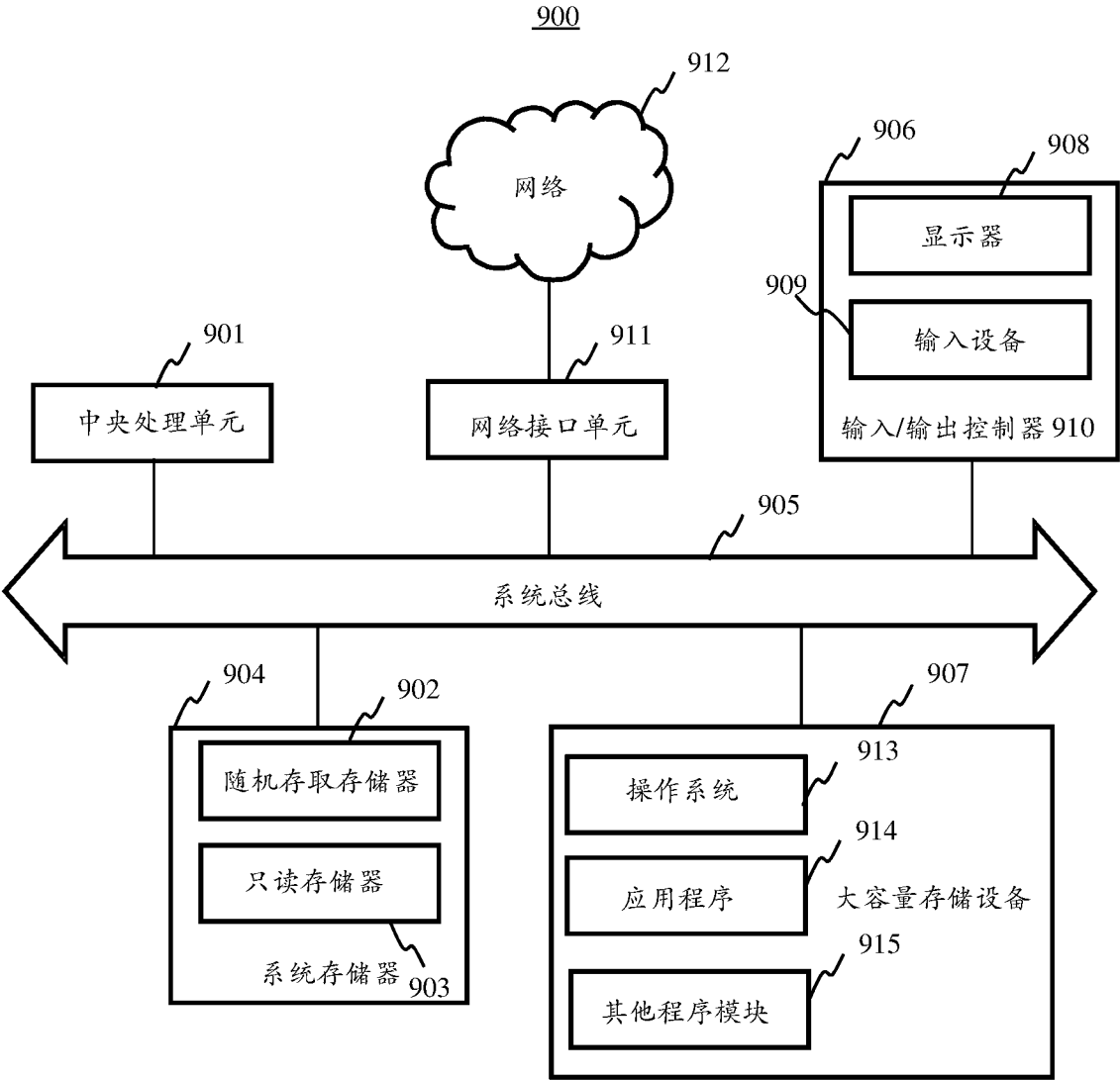


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/079721

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/35(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)

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; WPABSC; ENTXT; CJFD; ENTXTC; VEN; CNKI; IEEE: 文本, 语句, 分类, 归类, 结果, 解释, 训练, 原因, 注意力, 贡献, 权重, 关联, 历史, 输出, 样本, text, statement, classification, sort, result, interpretation, train, reason, attention, contribution, weight, relevance, history, output, sample

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 113722474 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 30 November 2021 (2021-11-30) description, paragraphs 5-198, and figures 1-4	1-16
Y	CN 111858919 A (CHINA MOBILE (SUZHOU) SOFTWARE TECHNOLOGY CO., LTD.; CHINA MOBILE COMMUNICATIONS GROUP CO., LTD.) 30 October 2020 (2020-10-30) description, paragraphs 4-134	1-16
Y	CN 112347252 A (JILIN UNIVERSITY) 09 February 2021 (2021-02-09) description, paragraphs 5-87	1-16
A	CN 110874411 A (FUZHOU UNIVERSITY) 10 March 2020 (2020-03-10) entire document	1-16
A	CN 112269874 A (BEIJING WUZI UNIVERSITY) 26 January 2021 (2021-01-26) entire document	1-16
A	US 2020210526 A1 (NETAPP INC.) 02 July 2020 (2020-07-02) entire document	1-16



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

26 April 2022

Date of mailing of the international search report

23 May 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
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/CN2022/079721

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	113722474	A	30 November 2021	None	
CN	111858919	A	30 October 2020	None	
CN	112347252	A	09 February 2021	None	
CN	110874411	A	10 March 2020	None	
CN	112269874	A	26 January 2021	None	
US	2020210526	A1	02 July 2020	US 10824815 B2	03 November 2020

国际检索报告

国际申请号

PCT/CN2022/079721

A. 主题的分类

G06F 16/35 (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)

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; CNTEXT; WPABSC; ENTXT; CJFD; ENTXTC; VEN; CNKI; IEEE: 文本, 语句, 分类, 归类, 结果, 解释, 训练, 原因, 注意力, 贡献, 权重, 关联, 历史, 输出, 样本, text, statement, classification, sort, result, interpretation, train, reason, attention, contribution, weight, relevance, history, output, sample

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 113722474 A (腾讯科技深圳有限公司) 2021年11月30日 (2021 - 11 - 30) 说明书第5-198段, 附图1-4	1-16
Y	CN 111858919 A (中移苏州软件技术有限公司 中国移动通信集团有限公司) 2020年10月30日 (2020 - 10 - 30) 说明书第4-134段	1-16
Y	CN 112347252 A (吉林大学) 2021年2月9日 (2021 - 02 - 09) 说明书第5-87段	1-16
A	CN 110874411 A (福州大学) 2020年3月10日 (2020 - 03 - 10) 全文	1-16
A	CN 112269874 A (北京物资学院) 2021年1月26日 (2021 - 01 - 26) 全文	1-16
A	US 2020210526 A1 (NETAPP INC) 2020年7月2日 (2020 - 07 - 02) 全文	1-16

☐ 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

2022年4月26日

国际检索报告邮寄日期

2022年5月23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

授权官员

韩洁

电话号码 (86-27)-59183851

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/079721

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	113722474	A	2021年11月30日	无	
CN	111858919	A	2020年10月30日	无	
CN	112347252	A	2021年2月9日	无	
CN	110874411	A	2020年3月10日	无	
CN	112269874	A	2021年1月26日	无	
US	2020210526	A1	2020年7月2日	US 10824815 B2	2020年11月3日