



(12) 发明专利申请

(10) 申请公布号 CN 114360540 A
(43) 申请公布日 2022. 04. 15

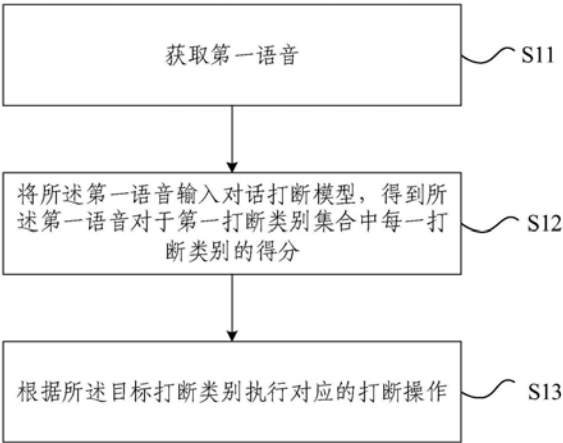
(21) 申请号 202210011325.4
(22) 申请日 2022.01.06
(71) 申请人 京东科技信息技术有限公司
地址 100176 北京市大兴区北京经济技术
开发区科创十一街18号院2号楼6层
601
(72) 发明人 杨久东 陈蒙 王培英 冯明超
(74) 专利代理机构 北京路浩知识产权代理有限
公司 11002
代理人 王宇杨

(51) Int.Cl.
G10L 15/22 (2006.01)
G10L 15/06 (2013.01)
G06K 9/62 (2022.01)
G06F 16/35 (2019.01)

权利要求书2页 说明书8页 附图3页

(54) 发明名称
一种对话管理方法、装置、电子设备及存储
介质

(57) 摘要
本发明涉及一种对话管理方法、装置、电子
设备及存储介质,该方法包括:获取第一语音;将
所述第一语音输入对话打断模型,得到所述第一
语音对于第一打断类别集合中每一打断类别的
得分;所述对话打断模型为利用第一训练数据集
预先训练基于深度学习的分类模型得到的;根据
所述每一打断类别的得分确定所述第一语音的
目标打断类别;根据所述目标打断类别执行对应
的打断操作。本发明预设多种打断类别及每一打
断类别相应的打断操作,通过对话打断模型智能
对输入的第一语音进行评分,确定出第一语音的
目标打断类别,进而执行相应的打断操作,对输
入的语音进行了更精确的判断,执行的打断操作
也更为精细、合理,适应了语音输入以及用户需
求的多样性。



1. 一种对话管理方法,其特征在于,包括:

获取第一语音;

将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分;所述对话打断模型为利用第一训练数据集预先训练基于深度学习的分类模型得到的;

根据所述每一打断类别的得分确定所述第一语音的目标打断类别;

根据所述目标打断类别执行对应的打断操作。

2. 根据权利要求1所述的对话管理方法,其特征在于,所述第一打断类别集合至少包括下列类别之一:环境噪声、背景人声、问候语、语气词、补充上一轮、抢答当前轮、抢答未播报轮、预设话题范围之外的话题、机器语音。

3. 根据权利要求2所述的对话管理方法,其特征在于,所述根据所述目标打断类别执行对应的打断操作,包括:

在所述目标打断类别为环境噪声的情况下,恢复语音播报;

在所述目标打断类别为背景人声的情况下,恢复语音播报;

在所述目标打断类别为问候语的情况下,恢复语音播报;

在所述目标打断类别为语气词的情况下,恢复语音播报;

在所述目标打断类别为补充上一轮的情况下,跳转至上一轮,重新识别用户意图;

在所述目标打断类别为抢答当前轮的情况下,停止当前轮语音播报,识别用户意图;

在所述目标打断类别为抢答未播报轮的情况下,重播当前轮语音,后续不再播报已抢答轮语音;

在所述目标打断类别为预设话题范围之外的话题的情况下,针对所述预设话题范围之外的话题进行应答,应答完毕后重播当前轮语音;

在所述目标打断类别为机器语音的情况下,结束语音播报。

4. 根据权利要求2或3所述的对话管理方法,其特征在于,在所述将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分之前,还包括:

确定所述第一语音的强度大于所述第一阈值,则打断当前轮语音播报,并将所述第一语音输入所述对话打断模型。

5. 根据权利要求4所述的对话管理方法,其特征在于,在所述目标类别为环境噪声或背景人声的情况下,则增大所述第一阈值。

6. 根据权利要求1所述的对话管理方法,其特征在于,所述获取第一语音之前,还包括:

获取所述第一训练数据集;所述第一训练数据集包括样本语音以及每一样本语音对应的标签数据;

将所述第一训练数据集拆分为训练集、验证集、测试集;

根据所述训练集对分类模型进行训练,优化模型的参数;

根据所述验证集对训练优化后的分类模型进行验证,优化模型的超参数;

根据所述测试集对验证优化后的分类模型进行测试评估,得到验证优化后的分类模型的泛化误差;

确定所述泛化误差小于预设阈值,则取所述验证优化后的分类模型为所述对话打断模

型。

7. 根据权利要求1所述的对话管理方法,其特征在于,所述对话打断模型采用wide&deep模型。

8. 一种对话管理装置,其特征在于,包括:

采集模块,用于获取第一语音;

评分模块,用于将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分;所述对话打断模型为利用第一训练数据集预先训练基于深度学习的分类模型得到的;

判断模块,用于根据所述每一打断类别的得分确定所述第一语音的目标打断类别;

执行模块,用于根据所述目标打断类别执行对应的打断操作。

9. 一种电子设备,包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,其特征在于,所述处理器执行所述程序时实现如权利要求1至7任一项所述对话管理方法的步骤。

10. 一种非暂态计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现如权利要求1至7任一项所述对话管理方法的步骤。

一种对话管理方法、装置、电子设备及存储介质

技术领域

[0001] 本发明涉及人工智能技术领域,尤其涉及一种对话管理方法、装置、电子设备及存储介质。

背景技术

[0002] 口语对话系统广泛应用于智能人机交互场景,打断是口语对话系统中一个常见的问题,系统需要在检测到用户语音输入后,判断是否停止播放语音。现有技术中,常见的打断后处理策略仅仅是判断打断的真假,对于假打断就恢复播报,对于真打断就走正常流程,但是这样是远远不够且并不合理,无法应对语音输入的多样性,以及用户需求的多样性。

[0003] 现有口语对话系统中,对语音输入的判断不够精细,无法执行合理的打断策略,导致客户体验较差的问题。针对此缺陷,亟需一种更为精细、合理的对话管理方案。

发明内容

[0004] 本发明提供一种对话管理方法、装置、电子设备及存储介质,用以解决现有技术中对语音输入的判断不够精细,无法执行合理的打断策略,导致客户体验较差的问题。

[0005] 本发明提供一种对话管理方法,包括:

[0006] 获取第一语音;

[0007] 将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分;所述对话打断模型为利用第一训练数据集预先训练基于深度学习的分类模型得到的;

[0008] 根据所述每一打断类别的得分确定所述第一语音的目标打断类别;

[0009] 根据所述目标打断类别执行对应的打断操作。

[0010] 根据本发明提供的一种对话管理方法,所述第一打断类别集合至少包括下列类别之一:环境噪声、背景人声、问候语、语气词、补充上一轮、抢答当前轮、抢答未播报轮、预设话题范围之外的话题、机器语音。

[0011] 根据本发明提供的一种对话管理方法,所述根据所述目标打断类别执行对应的打断操作,包括:

[0012] 在所述目标打断类别为环境噪声的情况下,恢复语音播报;

[0013] 在所述目标打断类别为背景人声的情况下,恢复语音播报;

[0014] 在所述目标打断类别为问候语的情况下,恢复语音播报;

[0015] 在所述目标打断类别为语气词的情况下,恢复语音播报;

[0016] 在所述目标打断类别为补充上一轮的情况下,跳转至上一轮,重新识别用户意图;

[0017] 在所述目标打断类别为抢答当前轮的情况下,停止当前轮语音播报,识别用户意图;

[0018] 在所述目标打断类别为抢答未播报轮的情况下,重播当前轮语音,后续不再播报已抢答轮语音;

[0019] 在所述目标打断类别为预设话题范围之外的话题的情况下,针对所述预设话题范围之外的话题进行应答,应答完毕后重播当前轮语音;

[0020] 在所述目标打断类别为机器语音的情况下,结束语音播报。

[0021] 根据本发明提供的一种对话管理方法,在所述将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分之前,还包括:

[0022] 确定所述第一语音的强度大于所述第一阈值,则打断当前轮语音播报,并将所述第一语音输入所述对话打断模型。根据本发明提供的一种对话管理方法,在所述目标类别为环境噪声或背景人声的情况下,则增大所述第一阈值。

[0023] 根据本发明提供的一种对话管理方法,所述获取第一语音之前,还包括:

[0024] 获取所述第一训练数据集;所述第一训练数据集包括样本语音以及每一样本语音对应的标签数据;

[0025] 将所述第一训练数据集拆分为训练集、验证集、测试集;

[0026] 根据所述训练集对分类模型进行训练,优化模型的参数;

[0027] 根据所述验证集对训练优化后的分类模型进行验证,优化模型的超参数;

[0028] 根据所述测试集对验证优化后的分类模型进行测试评估,得到验证优化后的分类模型的泛化误差;

[0029] 确定所述泛化误差小于预设阈值,则取所述验证优化后的分类模型为所述对话打断模型。

[0030] 根据本发明提供的一种对话管理方法,所述对话打断模型采用wide&deep模型。

[0031] 本发明还提供一种对话管理装置,包括:

[0032] 采集模块,用于获取第一语音;

[0033] 评分模块,用于将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分;所述对话打断模型为利用第一训练数据集预先训练基于深度学习的分类模型得到的;

[0034] 判断模块,用于根据所述每一打断类别的得分确定所述第一语音的目标打断类别;

[0035] 执行模块,用于根据所述目标打断类别执行对应的打断操作。

[0036] 本发明还提供一种电子设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现如上述任一种所述对话管理方法的步骤。

[0037] 本发明还提供一种非暂态计算机可读存储介质,其上存储有计算机程序,该计算机程序被处理器执行时实现如上述任一种所述对话管理方法的步骤。

[0038] 本发明提供的一种对话管理方法、装置、电子设备及存储介质,预设多种打断类别及每一打断类别相应的打断操作,通过对话打断模型智能对输入的第一语音进行评分,确定出第一语音的目标打断类别,进而执行相应的打断操作,对输入的语音进行了更精确的判断,执行的打断操作也更为精细、合理,适应了语音输入的多样性以及用户需求的多样性。

附图说明

[0039] 为了更清楚地说明本发明或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,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。

[0040] 图1是本发明提供的一种对话管理方法的流程示意图;

[0041] 图2是本发明提供的一种对话管理方法中对话打断模型的预训练流程示意图;

[0042] 图3是本发明提供的一种对话管理装置的结构示意图;

[0043] 图4是本发明提供的电子设备的结构示意图。

具体实施方式

[0044] 为使本发明的目的、技术方案和优点更加清楚,下面将结合本发明中的附图,对本发明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发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0045] 打断是指口语对话系统中系统播报语音时,用户开始了讲话,系统语音和用户语音有重叠的情况。发明人在实际工作中发现,现有口语对话系统中,打断之后仅仅只判断打断的真假。而现实的应用场景中,用户的打断往往是多样的,比如回答以前的问题,或者抢先回答了当前问题,或者提前回答了后续还未提出的问题,又或者回答了话题无关的问题等,这里涉及复杂的流程间跳转的问题。对于用户的打断类型,仅仅进行二分类的真假打断判断是不够的。对此,本发明提出一种更为精细、合理的对话管理方案,以对具体的打断类型进行更细化的分类,对于每个细分的类别进行不同的应答策略。

[0046] 下面结合图1-图4描述本发明提供的一种对话管理方法、装置、电子设备及存储介质。

[0047] 图1为本发明提供的一种对话管理方法的流程示意图,如图1所示,该方法包括:

[0048] S11、获取第一语音;

[0049] 具体地,口语对话系统在语音播报的过程中还需要采集语音,通过采集到的语音识别用户意图。当然采集到的语音中可能包括噪声,如环境噪声,还会包括当前播报的语音,因此可以根据实际需求对采集到的语音进行预处理,进行降噪处理或根据播报记录从采集到的语音中滤除口语对话系统本身播报的语音信息得到第一语音信息;还可以直接将采集到的语音作为第一语音,由后续的对话打断模型完成降噪、滤除本身播报的语音信息的过程。声音采集设备可采用如录音机、麦克风等,具体的声音采集设备类型、采样频率本发明不作限制。

[0050] S12、将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分;所述对话打断模型为利用第一训练数据集预先训练基于深度学习的分类模型得到的;

[0051] 具体地,本发明中利用第一训练数据集预先训练好了对话打断模型,对话打断模型为基于深度学习的分类模型,如逻辑回归、决策树、随机森林、梯度提升树、多层感知机、线性SVM、朴素贝叶斯模型等。投入使用前需要通过训练数据对分类模型进行训练,提升分

类模型泛化能力。训练完成的对话打断模型能够对输入的语音进行多个预设打断类别的评分。可以理解的是,第一打断类别集合中的预设打断类别至少为两个,以满足基本业务需求,具体的打断类别及类别数量可以根据业务需求以及模型训练成本进行相应设置。

[0052] 将第一语音输入预训练好的对话打断模型,以得到第一语音对于第一打断类别集合中每一打断类别的得分,用于进行后续对话管理过程。

[0053] S13、根据所述每一打断类别的得分确定所述第一语音的目标打断类别;

[0054] 具体地,在得到第一语音对于每一打断类别的得分后即可确定第一语音的目标打断类别,例如优选最高得分对应的打断类别为目标打断类别;

[0055] S14、根据所述目标打断类别执行对应的打断操作。

[0056] 具体地,对于第一打断类别集合中的每一打断类别,预先设置了待执行的打断操作。在确定第一语音的目标打断类别后,即可根据该目标类别执行对应的打断操作。

[0057] 本实施例中预设第一打断类别集合中的多种打断类别及每一打断类别相应的打断操作,通过对话打断模型智能对输入的第一语音进行评分,确定出第一语音的目标打断类别,进而执行相应的打断操作,对输入的语音进行了更精确的判断,执行的打断操作也更为精细、合理,适应了语音输入的多样性以及用户需求的多样性。

[0058] 基于上述实施例,在一个实施例中,所述第一打断类别集合至少包括下列类别之一:环境噪声、背景人声、问候语、语气词、补充上一轮、抢答当前轮、抢答未播报轮、预设话题范围之外的话题、机器语音。

[0059] 具体地,“环境噪声”是指周边生产、施工、交通和社会活动产生的干扰噪声,如鸣笛声、警铃声、空调噪声、音乐播放声等;“背景人声”是指当前对话用户所处的环境中他人发出的干扰噪声,背景人声的强度一般比用户声小且通常由多人发出;“问候语”常见于会话刚开始时用户对机器人打招呼,比如“喂”、“你好”、“hello”等;“语气词”是表示语气的虚词,常用在句尾或句中停顿处表示种种语气,例如“对”、“嗯”、“是”等;“补充上一轮”是指用户在当前轮语音播报中对上一轮语音播报中的问题进行补充回答或提出疑问等。“抢答当前轮”是指系统在播报当前轮问题但未播报完时,用户抢先回答了当前轮问题;“抢答未播报轮”是指用户并未回答当前轮问题,而是回答了后续未播报轮中的问题;“预设话题范围之外的话题”表明用户的回答并未被系统的预设话题范围覆盖,属于话题范围之外的话题,可能是闲聊或新问题等等;“机器语音”表明该语音是由机器程序(如语音助手、留言等)响应播放的语音,并不是由人发出的语音。

[0060] 本实施例中通过多种具体的打断类别对输入的语音进行精确分类判断,降低了真假打断的误判率,同时为后续执行打断操作提供了依据。

[0061] 基于上述任一实施例,在一个实施例中,所述根据所述目标打断类别执行对应的打断操作,包括:

[0062] 在所述目标打断类别为环境噪声的情况下,恢复语音播报;

[0063] 在所述目标打断类别为背景人声的情况下,恢复语音播报;

[0064] 在所述目标打断类别为问候语的情况下,恢复语音播报;

[0065] 在所述目标打断类别为语气词的情况下,恢复语音播报;

[0066] 在所述目标打断类别为补充上一轮的情况下,跳转至上一轮,重新识别用户意图;

[0067] 在所述目标打断类别为抢答当前轮的情况下,停止当前轮语音播报,识别用户意

图;

[0068] 在所述目标打断类别为抢答未播报轮的情况下,重播当前轮语音,后续不再播报已抢答轮语音;

[0069] 在所述目标打断类别为预设话题范围之外的话题的情况下,针对所述预设话题范围之外的话题进行应答,应答完毕后重播当前轮语音;

[0070] 在所述目标打断类别为机器语音的情况下,结束语音播报。

[0071] 具体地,“环境噪声”是指周边生产、施工、交通和社会活动产生的干扰噪声,如鸣笛声、警铃声、空调噪声、音乐播放声等,环境噪音对应属于假打断,因此系统需要恢复语音播报;“背景人声”是指当前对话用户所处的环境中他人发出的干扰噪声,背景人声对应属于假打断,因此系统需要恢复语音播报;“问候语”常见于会话刚开始时用户对机器人打招呼,比如“喂”、“你好”、“hello”等,问候语也属于假打断,用户只是在打招呼并不是真的想要说话,所以这时系统需要恢复播报;“语气词”是表示语气的虚词,常用在句尾或句中停顿处表示种种语气,例如“对”、“嗯”、“是”等,用户只是习惯性表达或停顿并不是真的想要说话,所以这时系统需要恢复播报;“补充上一轮”是指用户在当前轮语音播报中对上一轮语音播报中的问题进行补充回答或提出疑问等,因此系统需要跳回上一轮,并结合补充回答重新理解上一轮的用户语义,并把当前轮状态改为未执行,在上一轮完成后再次执行当前轮语音播报。“抢答当前轮”是指系统在播报当前轮问题但未播报完时,用户抢先回答了当前轮问题,此时系统需要结束当前轮的播报,并把当前轮状态改为已执行;“抢答未播报轮”是指用户并未回答当前轮问题,而是回答了后续未播报轮中的问题,此时需要把用户抢先回答的轮次状态置为已执行,而由于当前轮并未被回答,所以需要当前轮从头开始播报,后续不再播报已抢答轮的语音;“预设话题范围之外的话题”表明用户的回答并未被系统的问题列表覆盖,属于话题之外的问题,可能是闲聊或新问题等等,这时可以使用知识库进行回答,之后重复播报当前轮;“机器语音”表明该语音是由机器程序(如语音助手、留言等)响应播放的语音,并不是由人发出的语音,所以对于机器语音的打断,系统应该结束通话,后续再联系用户。

[0072] 本实施例中针对每一具体的打断类别执行了具体的打断操作,执行的打断操作更为精细、合理,进一步适应了语音输入的多样性以及用户需求的多样性。

[0073] 基于上述任一实施例,在一个实施例中,在所述将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分之前,还包括:

[0074] 确定所述第一语音的强度大于所述第一阈值,则打断当前轮语音播报,并将所述第一语音输入所述对话打断模型。

[0075] 具体地,在将语音输入对话打断模型中进行评分之前还需要进行阈值判断,判断采集到的第一语音的强度是否大于第一阈值,在确定第一语音的强度大于第一阈值时,则打断当前轮语音播报,并将所述第一语音输入对话打断模型,以便确定目标打断类别进而执行相应打断操作;在确定第一语音的强度小于或等于第一阈值时,则不打断当前轮语音播报,相应地也无需将第一语音输入对话打断模型。

[0076] 本实施例中,通过第一阈值对第一语音的强度进行判断,确定第一语音的强度大于第一阈值,则打断当前轮语音播报,并将所述第一语音输入对话打断模型,有效防止了低强度噪声对口语对话系统语音播报的频繁打断。

[0077] 基于上述任一实施例,在一个实施例中,在所述目标类别为环境噪声或背景人声的情况下,则增大所述第一阈值。

[0078] 具体地,由于口语对话系统对强度小于或等于第一阈值的输入语音不中断当前轮语音播报,因此,在判断输入的第一语音为环境噪声或背景人声的情况下,则增大第一阈值,可以有效减少噪声语音对口语对话系统语音播报的频繁打断。

[0079] 图2是本发明提供的一种对话管理方法中对话打断模型的预训练流程示意图,如图2所示,基于上述任一实施例,在一个实施例中,所述获取第一语音之前,还包括:

[0080] S21、获取所述第一训练数据集;所述第一训练数据集包括样本语音以及每一样本语音对应的标签数据;

[0081] 具体地,第一训练数据集中包括多条经过数据标注的样本语音数据,每一条样本语音数据对应有标注的打断类别评分信息。例如,一段风扇转动声音被标注为“环境噪声评分:1,其余打断类别评分:0”,又例如“你好”的发音语音被标注为“问候语评分:1,其余打断类别评分:0”。第一训练集中的语音数据覆盖了第一打断类别集合中的所有打断类别,用于训练分类模型,使得分类模型具备对输入的语音进行打断类别判断能力。

[0082] S22、将所述第一训练数据集拆分为训练集、验证集、测试集;

[0083] 具体地,将第一训练数据集进行上述拆分是为了防止模型过拟合,降低泛化误差。对第一训练数据集拆分的具体比例,可以根据第一训练数据的总数量按需求设置。

[0084] S23、根据所述训练集对分类模型进行训练,优化模型的参数;

[0085] 具体地,根据所述训练集对基于深度学习的分类模型进行训练,优化模型的参数,使得模型具备学习能力。

[0086] S24、根据所述验证集对训练优化后的分类模型进行验证,优化模型的超参数;

[0087] 具体地,根据验证集对训练优化后的分类模型进行验证,优化模型的超参数,验证集的作用是为了防止模型在训练集上过拟合,即防止模型过多学到训练集上特有的特征。其中,超参数可以包含正则化参数、神经网络的层数、每一隐含层中神经元的个数等。

[0088] S25、根据所述测试集对验证优化后的分类模型进行测试评估,得到验证优化后的分类模型的泛化误差;

[0089] 具体地,测试集用于测试分类模型的性能,对验证优化后的分类模型进行测试评估,得到验证优化后的分类模型的泛化误差。

[0090] S26、确定所述泛化误差小于预设阈值,则取所述验证优化后的分类模型为所述对话打断模型。

[0091] 具体地,确定所述泛化误差小于预设阈值,则取所述验证优化后的分类模型为对话打断模型,即选择泛化能力更优的分类模型作为对话打断模型。

[0092] 本实施例中将第一训练数据集进行拆分,并利用拆分得到的训练集、验证集、测试集分类模型,筛选出了泛化能力更优的分类模型作为对话打断模型,提升了对话打断模型的泛化能力。

[0093] 基于上述任一实施例,在一个实施例中,所述对话打断模型采用wide&deep模型。

[0094] 具体地,现有技术通常采用卷积神经网络模型(CNN),长短期记忆网络(LSTM)模型对打断语音进行二分类判断。本发明采用wide&deep模型,wide&deep模型是一种融合浅层(wide)模型和深层(deep)模型进行联合训练的框架,综合利用浅层模型的记忆能力和深层

模型的泛化能力,实现单模型对分类系统准确性和扩展性的兼顾。其中wide部分是线性模型,其具有很强的记忆能力,deep部分采用DNN的结构,具有很强的泛化能力,因此wide&deep模型兼具记忆能力与泛化能力。wide&deep模型在训练过程中整体进行训练。

[0095] 本实施例中对话打断模型采用wide&deep模型综合利用了浅层模型的记忆能力和深层模型的泛化能力,语音判断能力更强。

[0096] 下面对本发明提供的一种对话管理装置进行描述,下文描述的对话管理装置与上文描述的对话管理方法可相互对应参照。

[0097] 图3为本发明提供的一种对话管理装置的结构示意图,如图3所示,该装置包括:采集模块31、评分模块32、判断模块33、执行模块34。

[0098] 其中,采集模块31,用于获取第一语音;评分模块32,用于将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分;所述对话打断模型为利用第一训练数据集预先训练基于深度学习的分类模型得到的;判断模块33,用于根据所述每一打断类别的得分确定所述第一语音的目标打断类别;执行模块34,用于根据所述目标打断类别执行对应的打断操作。

[0099] 本实施例中预设第一打断类别集合中的多种打断类别及每一打断类别相应的打断操作,通过对话打断模型智能对输入的第一语音进行评分,确定出第一语音的目标打断类别,进而执行相应的打断操作,对输入的语音进行了更精确的判断,执行的打断操作也更为精细、合理,适应了语音输入的多样性以及用户需求的多样性。

[0100] 基于上述任一实施例,在一个实施例中,该装置还包括:确定模块。

[0101] 其中,确定模块,用于确定所述第一语音的强度大于所述第一阈值,则打断当前轮语音播报,并将所述第一语音输入所述对话打断模型。

[0102] 本实施例中通过第一阈值对第一语音的强度进行判断,确定第一语音的强度大于第一阈值,则打断当前轮语音播报,并将所述第一语音输入对话打断模型,有效防止了低强度噪声对口语对话系统语音播报的频繁打断。

[0103] 基于上述任一实施例,在一个实施例中,该装置还包括:更新模块。

[0104] 其中,更新模块,用于在所述目标类别为环境噪声或背景人声的情况下,则增大所述第一阈值。

[0105] 本实施例中通过更新模块,在判断输入的第一语音为环境噪声或背景人声的情况下,则增大第一阈值,可以有效减少噪声语音对口语对话系统语音播报的频繁打断。

[0106] 本发明还提供一种电子设备,图4示例了一种电子设备的结构示意图,如图4所示,该电子设备可以包括:处理器(processor) 410、通信接口(Communications Interface) 420、存储器(memory) 430和通信总线440,其中,处理器410,通信接口420,存储器430通过通信总线440完成相互间的通信。处理器410可以调用存储器430中的逻辑指令,以执行一种对话管理方法,该方法包括:获取第一语音;将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分;所述对话打断模型为利用第一训练数据集预先训练基于深度学习的分类模型得到的;根据所述每一打断类别的得分确定所述第一语音的目标打断类别;根据所述目标打断类别执行对应的打断操作。

[0107] 此外,上述的存储器430中的逻辑指令可以通过软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本

发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM, Read-Only Memory)、随机存取存储器(RAM, Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0108] 另一方面,本发明还提供一种计算机程序产品,所述计算机程序产品包括存储在非暂态计算机可读存储介质上的计算机程序,所述计算机程序包括程序指令,当所述程序指令被计算机执行时,计算机能够执行上述各提供的一种对话管理方法,该方法包括:获取第一语音;将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分;所述对话打断模型为利用第一训练数据集预先训练基于深度学习的分类模型得到的;根据所述每一打断类别的得分确定所述第一语音的目标打断类别;根据所述目标打断类别执行对应的打断操作。

[0109] 又一方面,本发明还提供一种非暂态计算机可读存储介质,其上存储有计算机程序,该计算机程序被处理器执行时实现以执行上述各提供的一种对话管理方法,该方法包括:获取第一语音;将所述第一语音输入对话打断模型,得到所述第一语音对于第一打断类别集合中每一打断类别的得分;所述对话打断模型为利用第一训练数据集预先训练基于深度学习的分类模型得到的;根据所述每一打断类别的得分确定所述第一语音的目标打断类别;根据所述目标打断类别执行对应的打断操作。

[0110] 以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性的劳动的情况下,即可以理解并实施。

[0111] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到各实施方式可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件。基于这样的理解,上述技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品可以存储在计算机可读存储介质中,如ROM/RAM、磁碟、光盘等,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行各个实施例或者实施例的某些部分所述的方法。

[0112] 最后应说明的是:以上实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

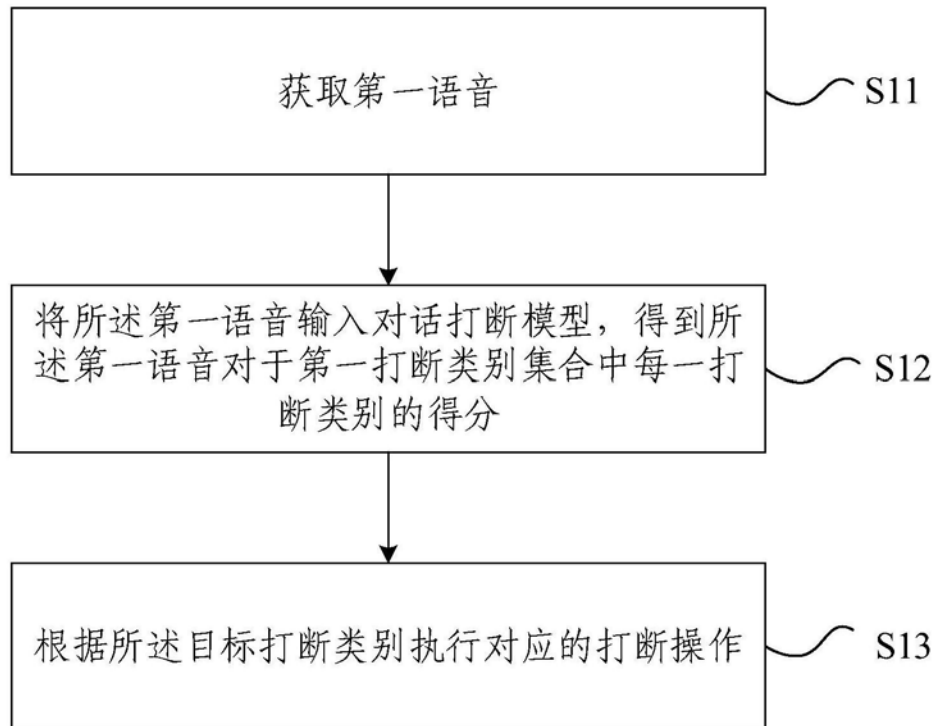


图1

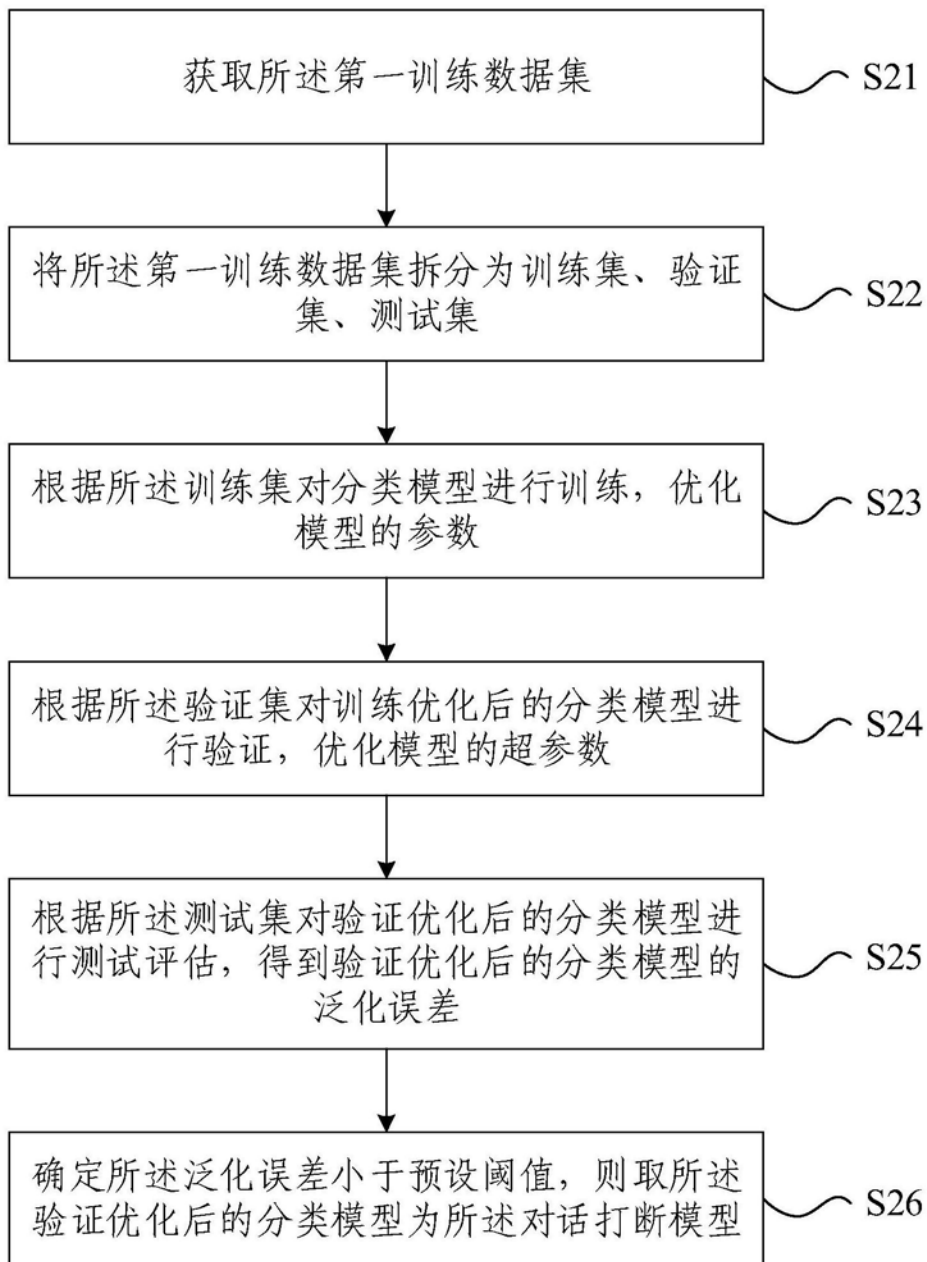


图2

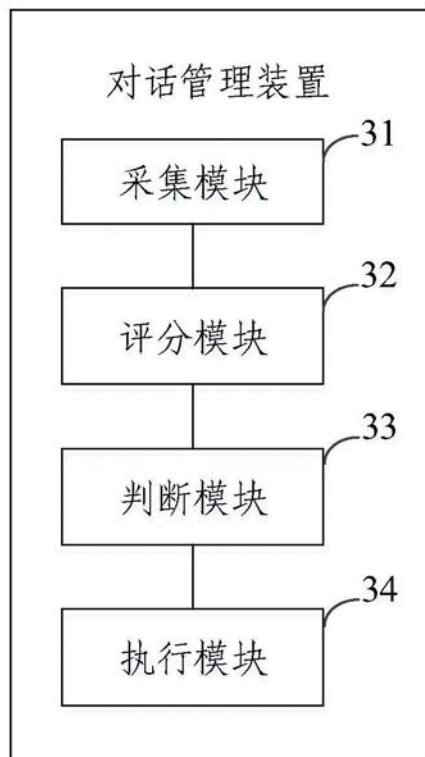


图3

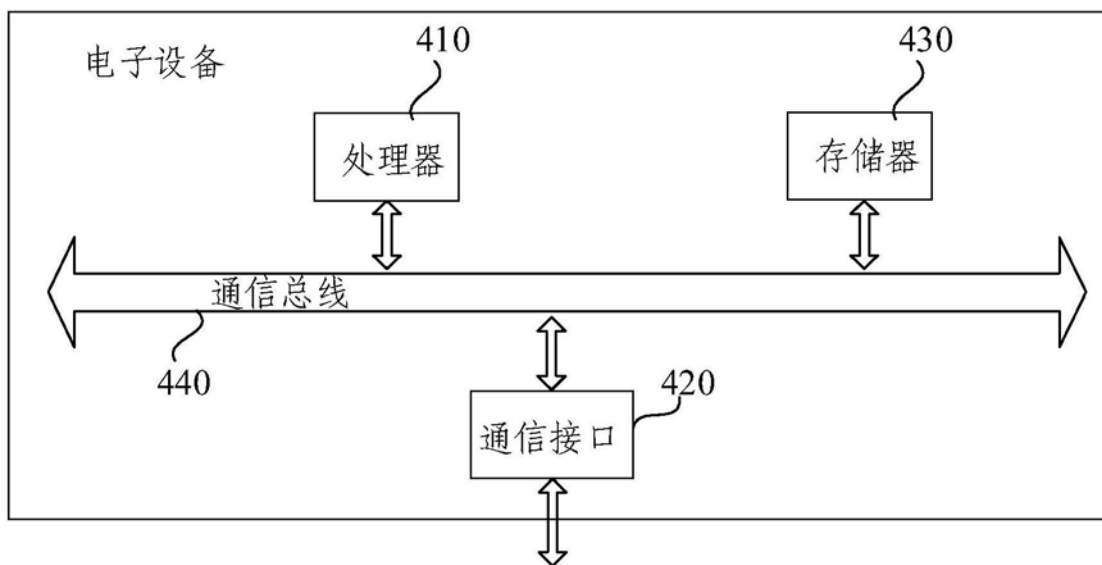


图4