(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 114153975 A
(43)申请公布日 2022.03.08

(21)申请号 202111494286.X

(22)申请日 2021.12.08

(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 王泽勋 陈蒙 乐雨泉 赵宇明 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 王治东

(51)Int.Cl.
G06F 16/35(2019.01)
G06K 9/62(2022.01)
G10L 15/22(2006.01)
G06N 3/04(2006.01)

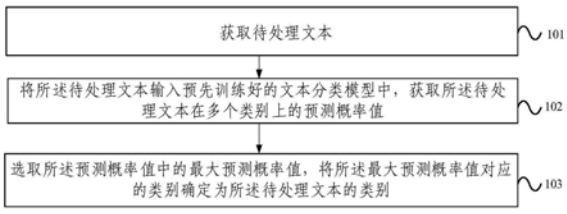
权利要求书3页 说明书11页 附图2页

(54)发明名称

文本分类方法、装置、电子设备及存储介质

(57)摘要

本公开提供一种文本分类方法、装置、电子设备及存储介质，所述方法：获取待处理文本；将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值；选取所述预测概率值中的最大预测概率值，将所述最大预测概率值对应的类别确定为所述待处理文本的类别；其中，所述文本分类模型是根据三元组样本数据进行训练得到的，所述文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的。本公开提供的文本分类方法能够提高NLP系统的鲁棒性，提高语音意图的准确性，提升用户的体验。



CN 114153975 A

1.一种文本分类方法,其特征在于,包括:

获取待处理文本;其中,所述待处理文本为待识别语音数据通过语音识别系统转换得到的文本;

将所述待处理文本输入预先训练好的文本分类模型中,获取所述待处理文本在多个类别上的预测概率值;

选取所述预测概率值中的最大预测概率值,将所述最大预测概率值对应的类别确定为所述待处理文本的类别;

其中,所述文本分类模型是根据三元组样本数据进行训练得到的,其中,所述三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及所述样本语音数据的语音识别结果数据;

所述文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的,其中,所述相对损失值是根据第一预测概率分布值与第二预测概率分布值确定的;所述第一预测概率分布值是基于所述样本语音数据的转换数据得到的,所述第二预测概率分布值是基于所述样本语音数据的语音识别结果数据得到的。

2.根据权利要求1所述的文本分类方法,其特征在于,在所述获取待处理文本之前,方法还包括:

在文本分类模型中的参数为第一参数值的情况下,将所述样本语音数据的转换数据输入所述文本分类模型,得到第一预测概率分布值;其中,所述第一参数值为文本分类模型在训练过程中所使用参数值中的任意一个;

根据所述第一预测概率分布值以及所述样本语音数据的类别标签数据,计算交叉熵损失值;

在文本分类模型中的参数为第一参数值的情况下,将所述样本语音数据的语音识别结果数据输入所述文本分类模型,得到第二预测概率分布值;

根据所述第一预测概率分布值与所述第二预测概率分布值计算相对损失值;

根据所述交叉熵损失值与所述相对损失值,在文本分类模型的参数为第一参数值的情况下计算所述文本分类模型的损失函数值;

根据所述损失函数值,从所述文本分类模型在训练过程中所使用的多个参数值中确定所述文本分类模型的最终值。

3.根据权利要求2所述的文本分类方法,其特征在于,所述根据所述第一预测概率分布值与所述第二预测概率分布值计算相对损失值,包括:

根据所述第一预测概率分布值与所述第二预测概率分布值计算KL散度值;

将所述KL散度值作为所述相对损失值。

4.根据权利要求2所述的文本分类方法,其特征在于,所述根据所述交叉熵损失值与所述相对损失值,在文本分类模型的参数为第一参数值的情况下计算文本分类模型的损失函数值,具体包括:

$Loss=\beta \cdot CE(p(Trans))+(1-\beta)KL(p(Trans),p(ASR))$

其中,Loss表示所述文本分类模型的损失函数值;β表示所述文本分类模型的超参数;CE(p(Trans))表示所述交叉熵损失值,p(Trans)表示所述第一预测概率分布值;KL(p(Trans),p(ASR))表示所述相对损失值,p(ASR)表示所述第二预测概率分布值。

5.根据权利要求2所述的文本分类方法，其特征在于，在所述在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的转换数据输入所述文本分类模型之前，方法还包括：

获取样本语音数据；

播放所述样本语音数据，监听并记录所述样本语音数据，得到所述样本语音数据的转换数据以及类别标签数据；

将所述样本语音数据输入语音识别系统中，得到所述样本语音数据的语音识别结果数据。

6.根据权利要求1所述的文本分类方法，其特征在于，所述文本分类模型包括嵌入层、卷积层、池化层和全连接层；

所述将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值，包括：

将所述待处理文本输入到所述嵌入层中进行词嵌入处理，获取所述待处理文本的表征矩阵；

将所述表征矩阵输入所述卷积层中进行卷积处理，获取卷积结果；

将所述卷积结果输入所述池化层中进行最大池化处理，获取高维特征向量；

将所述高维特征向量输入所述全连接层中，计算所述高维特征向量在多个类别上的预测概率值。

7.一种文本分类装置，其特征在于，包括：

第一获取模块，用于获取待处理文本；其中，所述待处理文本为待识别语音数据通过语音识别系统转换得到的文本；

第二获取模块，用于将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值；

选取与确定模块，用于选取所述预测概率值中的最大预测概率值，将所述最大预测概率值对应的类别确定为所述待处理文本的类别；

其中，所述文本分类模型是根据三元组样本数据进行训练得到的，其中，所述三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及所述样本语音数据的语音识别结果数据；

所述文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的，其中，所述相对损失值是根据第一预测概率分布值与第二预测概率分布值确定的；所述第一预测概率分布值是基于所述样本语音数据的转换数据得到的，所述第二预测概率分布值是基于所述样本语音数据的语音识别结果数据得到的。

8.根据权利要求7所述的文本分类装置，其特征在于，所述装置还包括训练模块，所述训练模块用于：

在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的转换数据输入所述文本分类模型，得到第一预测概率分布值；其中，所述第一参数值为文本分类模型在训练过程中所使用参数值中的任意一个；

根据所述第一预测概率分布值以及所述样本语音数据的类别标签数据，计算交叉熵损失值；

在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的语音识别结果数据输入所述文本分类模型，得到第二预测概率分布值；

根据所述第一预测概率分布值与所述第二预测概率分布值计算相对损失值；

根据所述交叉熵损失值与所述相对损失值，在文本分类模型的参数为第一参数值的情况下计算所述文本分类模型的损失函数值；

根据所述损失函数值，从所述文本分类模型在训练过程中所使用的多个参数值中确定所述文本分类模型的最终值。

9.一种电子设备，其特征在于，包括：

处理器、存储器和总线，其中，

所述处理器和所述存储器通过所述总线完成相互间的通信；

所述存储器存储有可被所述处理器执行的程序指令，所述处理器调用所述程序指令能够执行如权利要求1至6中任一项所述文本分类方法的步骤。

10.一种非暂态计算机可读存储介质，其特征在于，所述非暂态计算机可读存储介质存储计算机指令，所述计算机指令使所述计算机执行如权利要求1至6中任一项所述文本分类方法的步骤。

文本分类方法、装置、电子设备及存储介质

技术领域

[0001] 本公开涉及信息处理技术领域，尤其涉及一种文本分类方法、装置、电子设备及存储介质

背景技术

[0002] 随着语音识别技术的不断发展，如今语音智能系统已经深入应用到工业的各个方面，使得各个产品能够自然地和用户进行交互。

[0003] 现有技术中，语音智能系统中识别用户意图主要采取管道结构，即用户—ASR(Automatic Speech Recognition，自动语音识别技术)系统—NLP(Natural Language Processing，自然语言处理)系统，其中，ASR系统将用户的音频转化为文本，然后NLP系统对转换出来的文本进行类别分类，通过确定文本的类别标签的方式得到用户的文本意图。由于NLP系统的输入是ASR系统的输出，若ASR系统将用户的语音转换错误，NLP系统的识别意图会与原始用户的真实意图相差甚远，使得NLP系统的鲁棒性不高，导致用户的语音意图判断结果不准确。

发明内容

[0004] 本公开提供一种文本分类方法、装置、电子设备及存储介质，用以解决现有技术中由于NLP系统的鲁棒性不高导致语音意图判断准确性低的技术问题，实现提高系统的鲁棒性和语音意图准确性的目的。

[0005] 第一方面，本公开提供一种文本分类方法，包括：

[0006] 获取待处理文本；其中，所述待处理文本为待识别语音数据通过语音识别系统转换得到的文本；

[0007] 将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值；

[0008] 选取所述预测概率值中的最大预测概率值，将所述最大预测概率值对应的类别确定为所述待处理文本的类别；

[0009] 其中，所述文本分类模型是根据三元组样本数据进行训练得到的，其中，所述三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及所述样本语音数据的语音识别结果数据；

[0010] 所述文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的，其中，所述相对损失值是根据第一预测概率分布值与第二预测概率分布值确定的；所述第一预测概率分布值是基于所述样本语音数据的转换数据得到的，所述第二预测概率分布值是基于所述样本语音数据的语音识别结果数据得到的。

[0011] 进一步地，根据本公开提供的一种文本分类方法，在所述获取待处理文本之前，方法还包括：

[0012] 在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的转换数

据输入所述文本分类模型，得到第一预测概率分布值；其中，所述第一参数值为文本分类模型在训练过程中所使用参数值中的任意一个；

[0013] 根据所述第一预测概率分布值以及所述样本语音数据的类别标签数据，计算交叉熵损失值；

[0014] 在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的语音识别结果数据输入所述文本分类模型，得到第二预测概率分布值；

[0015] 根据所述第一预测概率分布值与所述第二预测概率分布值计算相对损失值；

[0016] 根据所述交叉熵损失值与所述相对损失值，在文本分类模型的参数为第一参数值的情况下计算所述文本分类模型的损失函数值；

[0017] 根据所述损失函数值，从所述文本分类模型在训练过程中所使用的多个参数值中确定所述文本分类模型的最终值。

[0018] 进一步地，根据本公开提供的一种文本分类方法，所述根据所述第一预测概率分布值与所述第二预测概率分布值计算相对损失值，包括：

[0019] 根据所述第一预测概率分布值与所述第二预测概率分布值计算KL散度值；

[0020] 将所述KL散度值作为所述相对损失值。

[0021] 进一步地，根据本公开提供的一种文本分类方法，所述根据所述交叉熵损失值与所述相对损失值，在文本分类模型的参数为第一参数值的情况下计算文本分类模型的损失函数值，具体包括：

[0022] $Loss = \beta \cdot CE(p(Trans)) + (1-\beta) KL(p(Trans), p(ASR))$

[0023] 其中，Loss表示所述文本分类模型的损失函数值；β表示所述文本分类模型的超参数；CE(p(Trans))表示所述交叉熵损失值，p(Trans)表示所述第一预测概率分布值；KL(p(Trans),p(ASR))表示所述相对损失值，p(ASR)表示所述第二预测概率分布值。

[0024] 进一步地，根据本公开提供的一种文本分类方法，在所述在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的转换数据输入所述文本分类模型之前，方法还包括：

[0025] 获取样本语音数据；

[0026] 播放所述样本语音数据，监听并记录所述样本语音数据，得到所述样本语音数据的转换数据以及类别标签数据；

[0027] 将所述样本语音数据输入语音识别系统中，得到所述样本语音数据的语音识别结果数据。

[0028] 进一步地，根据本公开提供的一种文本分类方法，所述文本分类模型包括嵌入层、卷积层、池化层和全连接层；

[0029] 所述将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值，包括：

[0030] 将所述待处理文本输入到所述嵌入层中进行词嵌入处理，获取所述待处理文本的表征矩阵；

[0031] 将所述表征矩阵输入所述卷积层中进行卷积处理，获取卷积结果；

[0032] 将所述卷积结果输入所述池化层中进行最大池化处理，获取高维特征向量；

[0033] 将所述高维特征向量输入所述全连接层中，计算所述高维特征向量在多个类别上

的预测概率值。

[0034] 第二方面，本公开还提供一种文本分类装置，包括：

[0035] 第一获取模块，用于获取待处理文本；其中，所述待处理文本为待识别语音数据通过语音识别系统转换得到的文本；

[0036] 第二获取模块，用于将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值；

[0037] 选取与确定模块，用于选取所述预测概率值中的最大预测概率值，将所述最大预测概率值对应的类别确定为所述待处理文本的类别；

[0038] 其中，所述文本分类模型是根据三元组样本数据进行训练得到的，其中，所述三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及所述样本语音数据的语音识别结果数据；

[0039] 所述文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的，其中，所述相对损失值是根据第一预测概率分布值与第二预测概率分布值确定的；所述第一预测概率分布值是基于所述样本语音数据的转换数据得到的，所述第二预测概率分布值是基于所述样本语音数据的语音识别结果数据得到的。

[0040] 进一步地，根据本公开提供的文本分类装置，所述装置还包括训练模块，所述训练模块用于：

[0041] 在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的转换数据输入所述文本分类模型，得到第一预测概率分布值；其中，所述第一参数值为文本分类模型在训练过程中所使用参数值中的任意一个；

[0042] 根据所述第一预测概率分布值以及所述样本语音数据的类别标签数据，计算交叉熵损失值；

[0043] 在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的语音识别结果数据输入所述文本分类模型，得到第二预测概率分布值；

[0044] 根据所述第一预测概率分布值与所述第二预测概率分布值计算相对损失值；

[0045] 根据所述交叉熵损失值与所述相对损失值，在文本分类模型的参数为第一参数值的情况下计算所述文本分类模型的损失函数值；

[0046] 根据所述损失函数值，从所述文本分类模型在训练过程中所使用的多个参数值中确定所述文本分类模型的最终值。

[0047] 第三方面，本公开还提供一种电子设备，包括：

[0048] 处理器、存储器和总线，其中，

[0049] 所述处理器和所述存储器通过所述总线完成相互间的通信；

[0050] 所述存储器存储有可被所述处理器执行的程序指令，所述处理器调用所述程序指令能够执行如上任一项中所述文本分类方法的步骤。

[0051] 第四方面，本公开还提供一种非暂态计算机可读存储介质，所述非暂态计算机可读存储介质存储计算机指令，所述计算机指令使所述计算机执行如上所述文本分类方法的步骤。

[0052] 本公开提供的一种文本分类方法、装置、电子设备及存储介质，所述方法通过将获取的待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的

预测概率值;选取所述预测概率值中的最大预测概率值,将所述最大预测概率值对应的类别确定为所述待处理文本的类别,其中,文本分类模型是根据三元组样本数据训练得到的,且该模型的损失函数值是根据交叉熵损失值以及相对损失值确定的。本公开通过将样本语音识别的识别结果作为扰动样本对文本分类模型进行训练,能够提高NLP系统的鲁棒性,提高语音意图判断的准确性,提升用户的体验。

附图说明

[0053] 为了更清楚地说明本公开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公开的一些实施例,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。

[0054] 图1是本公开提供的一种文本分类方法的流程示意图;

[0055] 图2是本公开提供的文本分类模型的训练过程的流程示意图;

[0056] 图3是本公开提供的一种获取样本数据的流程示意图;

[0057] 图4是本公开提供的一种文本分类装置的结构示意图;

[0058] 图5是本公开提供的电子设备的结构示意图。

具体实施方式

[0059] 为使本公开的目的、技术方案和优点更加清楚,下面将结合本公开中的附图,对本公开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公开一部分实施例,而不是全部的实施例。基于本公开中的实施例,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,都属于本公开保护的范围。

[0060] 图1为本公开提供的文本分类方法,如图1所示,本公开提供的文本分类方法,包括以下步骤:

[0061] 步骤101:获取待处理文本;其中,所述待处理文本为待识别语音数据通过语音识别系统转换得到的文本;

[0062] 步骤102:将所述待处理文本输入预先训练好的文本分类模型中,获取所述待处理文本在多个类别上的预测概率值;

[0063] 步骤103:选取所述预测概率值中的最大预测概率值,将所述最大预测概率值对应的类别确定为所述待处理文本的类别;

[0064] 其中,所述文本分类模型是根据三元组样本数据进行训练得到的,其中,所述三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及所述样本语音数据的语音识别结果数据;

[0065] 所述文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的,其中,所述相对损失值是根据第一预测概率分布值与第二预测概率分布值确定的;所述第一预测概率分布值是基于所述样本语音数据的转换数据得到的,所述第二预测概率分布值是基于所述样本语音数据的语音识别结果数据得到的。

[0066] 具体地,损失函数(Loss Function)为将随机事件或其有关随机变量的取值映射为非负实数以表示该随机事件的“风险”或“损失”的函数。在应用中,损失函数通常作为学

习准则与优化问题相联系,即通过最小化损失函数求解和评估模型。

[0067] 交叉熵(Cross Entropy)是信息论中的一个重要概念,主要用于度量两个概率分布间的差异性信息,语言模型的性能通常用交叉熵来衡量。

[0068] 在步骤101中,获取待处理文本是指语音场景下,待识别语音数据通过语音识别系统转换得到的文本,需要说明的是,待处理文本用于判断语音识别结果的真实意图。

[0069] 在步骤102中,将待处理文本输入预先训练好的文本分类模型中,得到待处理文本在多个类别上的预测概率值,如待处理文本为“我想查看今年童装秋季最新款”,输入文本分类模型中,得到在男装类别上的预测概率值为0.2,在女装类别上的预测概率值为0.1,在童装类别上的预测概率值为0.7,其中,所述文本分类模型是根据三元组样本数据进行训练得到的,三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及样本语音数据的语音识别结果数据;文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的。需要说明的是,若用户在目标明确的场景下,多个类别可以设置为浏览、选购、下单、完成等类别,若用户在随意浏览的场景下,也可以设置为男装、女装、童装等,可以根据实际需要进行设定,在此不作具体限定。

[0070] 在步骤103中,根据步骤102中得到的多个类别上的预测概率值选取其中最大预测概率值,将最大预测概率值对应的类别确定为待处理文本的类别,如上述得到的预测概率值中,童装类别的预测概率值最大,将童装确定为待处理文本的类别,也就是说,用户的真实意图是想浏览童装。

[0071] 本公开实施例中,通过预先训练好的文本分类模型对待处理文本的处理实现用户真实意图的判断,其中,文本分类模型根据三元组样本数据训练得到的,三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及样本语音数据的语音识别结果数据,将样本语音数据的转换数据作为原始样本,ASR系统得到的语音识别结果数据作为扰动样本,而且该模型的损失函数值是根据交叉熵损失值和相对损失值确定的。本公开提供的方法能够提高NLP系统的鲁棒性,提高语音意图判断的准确性,提升用户体验。

[0072] 在本公开的一个实施例中,如图2所示,在所述获取待处理文本之前,方法还包括:

[0073] 在文本分类模型中的参数为第一参数值的情况下,将所述样本语音数据的转换数据输入所述文本分类模型,得到第一预测概率分布值;其中,所述第一参数值为文本分类模型在训练过程中所使用参数值中的任意一个;

[0074] 根据所述第一预测概率分布值以及所述样本语音数据的类别标签数据,计算交叉熵损失值;

[0075] 在文本分类模型中的参数为第一参数值的情况下,将所述样本语音数据的语音识别结果数据输入所述文本分类模型,得到第二预测概率分布值;

[0076] 根据所述第一预测概率分布值与所述第二预测概率分布值计算相对损失值;

[0077] 根据所述交叉熵损失值与所述相对损失值,在文本分类模型的参数为第一参数值的情况下计算所述文本分类模型的损失函数值;

[0078] 根据所述损失函数值,从所述文本分类模型在训练过程中所使用的多个参数值中确定所述文本分类模型的最终值。

[0079] 具体地,相对损失值是指用来分析两个预测概率分布值之间差异得到的数值。

[0080] 在本公开实施例中,需要确定出文本分类模型的最终参数值,得到最终的文本分

类模型，完成文本分类模型的训练。其中文本分类模型的训练分两步进行：第一步，在第一参数值为θ的情况下，将样本语音数据的转换数据输入文本分类模型中，得到样本语音数据的转换数据的第一预测概率分布值p(Trans)，根据样本语音数据的转换数据的第一预测概率分布值和类别标签数据计算得到交叉熵损失值$label_{loss}$；第二步，在同样的第一参数值为θ的情况下，将ASR系统转换得到的语音识别结果数据输入文本分类模型中，得到第二预测概率分布值p(ASR)，根据第一预测概率分布值p(Trans)和第二预测概率分布值p(ASR)计算得到相对损失值，最后根据得到的交叉熵损失值和相对损失值输入文本分类模型的损失函数计算公式中，得到模型的损失函数值。调整第一参数值θ的大小，重复上述操作，当得到的损失函数值满足预设要求时，则将损失函数值对应的参数值确定为文本分类模型的最终值，如选用最小损失函数值对应的参数值确定为文本分类模型的参数值，得到最终的文本分类模型。

[0081] 需要说明的是，文本分类模型可以选用TextCNN模型，也可以选用其他的模型作为文本分类模型，可以根据实际需要进行选用，在此不作具体限定。

[0082] 本公开实施例中，通过在获取待处理文本之前完成文本分类模型的训练，具体根据样本语音数据的转换数据、样本语句的类别标签数据以及样本语音数据的语音识别结果数据对模型进行训练，得到文本分类模型。本公开实施例提供的方法提高了文本分类模型的鲁棒性，提高用户语音意图判断的准确性，提升用户体验。

[0083] 在本公开的另一个实施例中，所述根据所述第一预测概率分布值与所述第二预测概率分布值计算相对损失值，包括：

[0084] 根据所述第一预测概率分布值与所述第二预测概率分布值计算KL散度值；

[0085] 将所述KL散度值作为所述相对损失值。

[0086] 具体地，KL(全称Kullback-Leibler)散度，又称为相对熵，是两个概率分布间差异的非对称性度量。

[0087] 在本公开实施例中，根据第一预测概率分布值和第二预测概率分布值计算KL散度，将KL散度作为相对损失值。其中，$KL(p,q)=\sum_{i=1}^{n}p(x_i)\,log\,\frac{p(x_i)}{q(x_i)}$，可以将第一预测概率分布值p(Trans)和第二预测概率分布值p(ASR)输入该公式中，得到相对损失值。

[0088] 本公开实施例中，将KL散度值作为相对损失值，能够量化出第一预测概率分布值和第二预测概率分布值之间的差异，拉近两个预测概率分布，使其具有相同的预测结果，保证文本分类模型的鲁棒性。

[0089] 在本公开的一个实施例中，所述根据所述交叉熵损失值与所述相对损失值，在文本分类模型的参数为第一参数值的情况下计算文本分类模型的损失函数值，具体包括：

[0090] Loss＝β•CE(p(Trans))+(1-β)KL(p(Trans),p(ASR))

[0091] 其中，Loss表示所述文本分类模型的损失函数值；β表示所述文本分类模型的超参数；CE(p(Trans))表示所述交叉熵损失值，p(Trans)表示所述第一预测概率分布值；KL(p(Trans),p(ASR))表示所述相对损失值，p(ASR)表示所述第二预测概率分布值。

[0092] 具体地，超参数是在开始学习过程之前设置的参数，而不是通过训练得到的参数数据。

[0093] 在本公开实施例中，文本分类模型的损失函数具体可以通过Loss＝β•$label_{loss}$+

(1-β)kl_{loss}推导得到上述公式，其中，交叉熵损失函数值可以通过$CE\left(p(y;X,\theta)\right)=-\sum_{l=1}^{L}\hat{p}(y_l|X)log(p(y_l|X))$计算得到，其中，L表示类别标签的总数，y表示预测类别标签，X表示输入样本语音数据的转换数据的向量，θ表示模型的第一参数值。$\hat{p}$表示样本语音数据的转换数据的真实分布值，p表示样本语音数据的转换数据的预测分布值。

[0094] 本公开实施例中，通过上述具体的表征公式计算得到该文本分类模型的损失函数值，根据损失函数值调整文本分类模型的参数值，确定出文本分类模型的最优参数值。

[0095] 在本公开的另一个实施例中，在所述在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的转换数据输入所述文本分类模型之前，方法还包括：

[0096] 获取样本语音数据；

[0097] 播放所述样本语音数据，监听并记录所述样本语音数据，得到所述样本语音数据的转换数据以及类别标签数据；

[0098] 将所述样本语音数据输入语音识别系统中，得到所述样本语音数据的语音识别结果数据。

[0099] 具体地，样本语音数据是指语音场景下的样本数据。

[0100] 在本公开实施例中，将获取语音场景下的样本语音数据进行播放，人工进行监听并记录监听结果，将通过人工标注、监听转写得到的数据设定为转换数据。整体的训练样本数据的构建如图3所示，首先获取一批样本语音数据，人工监听，并编写听到的样本语音数据，得到人工转写数据(Trans)，同时，人工在转写数据的时候也需要标注听到的样本语音数据对应的类别标签(Label)，这里的标签是转写文本对应的标签；然后，利用ASR系统对样本语音数据进行识别，得到语音识别结果数据；最后，将人工转写数据、语音识别结果数据和类别标签数据组成对抗三元组数据，即(Trans，ASR，Label)。另外，ASR系统对一段样本语音数据的识别可能会得到n个结果，在此情况下，可以看作是同一个Trans和Label对应多个语音识别结果数据，这样得到多个三元组数据。需要说明的是，本公开实施例中，样本语音数据的人工转写数据设为天然的原始样本，ASR系统的语音识别结果数据设为扰动样本。具体样本的设定可以根据实际需要进行设定，在此不作具体限定。

[0101] 本公开实施例中通过将获取样本语音数据人工监听并记录得到人工转写数据作为原始样本，ASR系统对样本语音数据的语音识别结果数据作为扰动样本，对文本分类模型进行训练，确定该模型的最终参数值，提高模型的鲁棒性和语音意图判断的准确率。

[0102] 在本公开的另一个实施例中，所述文本分类模型包括嵌入层、卷积层、池化层和全连接层；

[0103] 所述将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值，包括：

[0104] 将所述待处理文本输入到所述嵌入层中进行词嵌入处理，获取所述待处理文本的表征矩阵；

[0105] 将所述表征矩阵输入所述卷积层中进行卷积处理，获取卷积结果；

[0106] 将所述卷积结果输入所述池化层中进行最大池化处理，获取高维特征向量；

[0107] 将所述高维特征向量输入所述全连接层中，计算所述高维特征向量在多个类别上

的预测概率值。

[0108] 具体地，卷积(Convolution)是通过两个函数f和g生成第三个函数的一种数学算子，表征函数f与g经过翻转和平移的重叠部分函数值乘积对重叠长度的积分。最大值池化(max pooling)处理能够更好保留纹理特征。

[0109] 在本公开实施例中，将获取到的待处理文本输入预先训练好的文本分类模型中，首先嵌入层将待处理文本中的字转化为embedding表征，得到待处理文本的表征矩阵，然后卷积层对得到的表征矩阵按照字的方向进行卷积处理，得到卷积结果，池化层对卷积结果进行最大池化处理，获得高维特征向量，最后将高维特征向量接入全连接层中，计算高维特征向量在每个类别上的预测概率值，将多个预测概率值输出，最终将最大预测概率值对应的类别确定为待处理文本的类别。

[0110] 本公开实施例中，通过文本分类模型的嵌入层、卷积层、池化层和全连接层对待处理文本的层级处理，最终得到待处理文本在多个类别上的预测概率值，提高了文本分类模型的精确度和语音意图结果的准确性。

[0111] 在本公开的一个实施例中，文本分类模型训练的整个流程，包括：

[0112] 在语音场景下，样本语音数据的人工转写数据作为天然的原始样本，ASR系统的识别结果数据作为扰动样本。训练样本数据的构建过程为：首先，获取一批样本语音数据，人工监听，并编写听到的样本语音数据，得到人工转写数据(Trans)，同时，人工在转写数据的时候也需要标注听到的样本语音数据对应的类别标签(Label)，这里的标签是转写文本对应的标签；然后，利用ASR系统对样本语音数据进行识别，得到语音识别结果数据(ASR)；最后，将人工转写数据、语音识别结果数据和类别标签数据组成对抗三元组数据，即(Trans，ASR，Label)。另外，ASR系统对一段样本语音数据的识别可能会得到n个结果，在此情况下，可以看作是同一个Trans和Label对应多个ASR，这样得到多个三元组数据。

[0113] 在上述得到训练样本数据之后，进行文本分类模型的训练。训练模型分为两步进行：第一步，在第一参数值为θ的情况下，选取基础模型，如TextCNN，将样本语音数据的转换数据输入基础模型中，得到样本语音数据的转换数据的第一预测概率分布值p(Trans)，根据样本语音数据的转换数据的第一预测概率分布值和类别标签数据计算得到交叉熵损失值$label_{loss}$；第二步，在同样的第一参数值为θ的设置下，再将ASR系统转换得到的语音识别结果数据输入基础模型中，得到第二预测概率分布值p(ASR)，同时，为了拉近第一预测概率分布值p(Trans)和第二预测概率分布值p(ASR)，使其接近相同，根据第一预测概率分布值p(Trans)和第二预测概率分布值p(ASR)计算KL散度作为相对损失值，最后根据得到的交叉熵损失值和相对损失值得到模型的损失函数值：

[0114] $Loss=\beta \cdot label_{loss}+(1-\beta)kl_{loss}$

[0115] $=\beta \cdot CE(p(Trans))+(1-\beta)KL(p(Trans),p(ASR))$

[0116] 其中，β是模型超参数，CE表示交叉熵，KL表示KL散度。

[0117] 然后调整第一参数值θ的大小，重复上述操作，当得到的损失函数值满足预设要求时，则将损失函数值对应的参数值确定为基础模型的最终值，得到最终的文本分类模型。

[0118] 在完成文本分类模型的训练后，将待处理文本输入训练完成的文本分类模型中，首先将待处理文本的各个字转化为embedding表征，得到待处理文本的表征矩阵，将表征矩阵按照字的方向进行卷积操作，进一步对卷积操作的结果进行最大池化处理，得到高维特

征向量，将高维特征向量连接到softmax层，输出高维特征向量在每个类别标签上的预测概率值，将最大预测概率值对应的类别标签确定为最终的类别。需要说明的是，在实际应用场景中，由于不存在Trans数据，且不需要计算损失函数值，因此，直接将ASR系统转换过来的文本输入文本分类模型中，获取在每个类别上的预测概率值，将最大预测概率值对应的类别确定为该文本的类别，便得到文本意图。

[0119] 图4为本公开提供的一种文本分类装置，如图4所示，本公开提供的文本分类装置，包括：

[0120] 第一获取模块401，用于获取待处理文本；其中，所述待处理文本为待识别语音数据通过语音识别系统转换得到的文本；

[0121] 第二获取模块402，用于将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值；

[0122] 选取与确定模块403，用于选取所述预测概率值中的最大预测概率值，将所述最大预测概率值对应的类别确定为所述待处理文本的类别；

[0123] 其中，所述文本分类模型是根据三元组样本数据进行训练得到的其中，所述三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及所述样本语音数据的语音识别结果数据；

[0124] 所述文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的，其中，所述相对损失值是根据第一预测概率分布值与第二预测概率分布值确定的；所述第一预测概率分布值是基于所述样本语音数据的转换数据得到的，所述第二预测概率分布值是基于所述样本语音数据的语音识别结果数据得到的。

[0125] 本公开实施例提供的文本分类装置，能够提高NLP系统的鲁棒性，提高语音意图结果的准确性，提升用户体验。

[0126] 进一步，所述装置还包括训练模块，所述训练模块用于：

[0127] 在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的转换数据输入所述文本分类模型，得到第一预测概率分布值；其中，所述第一参数值为文本分类模型在训练过程中所使用参数值中的任意一个；

[0128] 根据所述第一预测概率分布值以及所述样本语音数据的类别标签数据，计算交叉熵损失值；

[0129] 在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的语音识别结果数据输入所述文本分类模型，得到第二预测概率分布值；

[0130] 根据所述第一预测概率分布值与所述第二预测概率分布值计算相对损失值；

[0131] 根据所述交叉熵损失值与所述相对损失值，在文本分类模型的参数为第一参数值的情况下计算所述文本分类模型的损失函数值；

[0132] 根据所述损失函数值，从所述文本分类模型在训练过程中所使用的多个参数值中确定所述文本分类模型的最终值。

[0133] 进一步，所述训练模块还用于：

[0134] 根据所述第一预测概率分布值与所述第二预测概率分布值计算KL散度值；

[0135] 将所述KL散度值作为所述相对损失值。

[0136] 进一步，所述训练模块还用于：

[0137] 所述根据所述交叉熵损失值与所述相对损失值，在文本分类模型的参数为第一参数值的情况下计算文本分类模型的损失函数值，具体包括：

[0138] $Loss = \beta \cdot CE(p(Trans)) + (1-\beta) KL(p(Trans), p(ASR))$

[0139] 其中，Loss表示所述文本分类模型的损失函数值；β表示所述文本分类模型的超参数；CE(p(Trans))表示所述交叉熵损失值，p(Trans)表示所述第一预测概率分布值；KL(p(Trans),p(ASR))表示所述相对损失值，p(ASR)表示所述第二预测概率分布值。

[0140] 进一步，所述装置在所述在文本分类模型中的参数为第一参数值的情况下，将所述样本语音数据的转换数据输入所述文本分类模型之前，还用于：

[0141] 获取样本语音数据；

[0142] 播放所述样本语音数据，监听并记录所述样本语音数据，得到所述样本语音数据的转换数据以及类别标签数据；

[0143] 将所述样本语音数据输入语音识别系统中，得到所述样本语音数据的语音识别结果数据。

[0144] 进一步，所述装置还用于：

[0145] 所述文本分类模型包括嵌入层、卷积层、池化层和全连接层；

[0146] 所述将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值，包括：

[0147] 将所述待处理文本输入到所述嵌入层中进行词嵌入处理，获取所述待处理文本的表征矩阵；

[0148] 将所述表征矩阵输入所述卷积层中进行卷积处理，获取卷积结果；

[0149] 将所述卷积结果输入所述池化层中进行最大池化处理，获取高维特征向量；

[0150] 将所述高维特征向量输入所述全连接层中，计算所述高维特征向量在多个类别上的预测概率值。

[0151] 由于本公开实施例所述装置与上述实施例所述方法的原理相同，对于更加详细的解释内容在此不再赘述。

[0152] 图5为本公开实施例提供的电子设备实体结构示意图，如图5所示，本公开提供一种电子设备，包括：处理器(processor)501、存储器(memory)502和总线503；

[0153] 其中，处理器501、存储器502通过总线503完成相互间的通信；

[0154] 处理器501用于调用存储器502中的程序指令，以执行上述各方法实施例所提供的方法，例如包括：获取待处理文本；其中，所述待处理文本为待识别语音数据通过语音识别系统转换得到的文本；将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值；选取所述预测概率值中的最大预测概率值，将所述最大预测概率值对应的类别确定为所述待处理文本的类别；其中，所述文本分类模型是根据三元组样本数据进行训练得到的其中，所述三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及所述样本语音数据的语音识别结果数据；所述文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的，其中，所述相对损失值是根据第一预测概率分布值与第二预测概率分布值确定的；所述第一预测概率分布值是基于所述样本语音数据的转换数据得到的，所述第二预测概率分布值是基于所述样本语音数据的语音识别结果数据得到的。

[0155] 本实施例提供一种非暂态计算机可读存储介质，非暂态计算机可读存储介质存储计算机指令，计算机指令使所述计算机执行上述各方法实施例所提供的方法，例如包括：获取待处理文本；其中，所述待处理文本为待识别语音数据通过语音识别系统转换得到的文本；将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值；选取所述预测概率值中的最大预测概率值，将所述最大预测概率值对应的类别确定为所述待处理文本的类别；其中，所述文本分类模型是根据三元组样本数据进行训练得到的其中，所述三元组样本数据包括样本语音数据的转换数据、样本语音数据的类别标签数据以及所述样本语音数据的语音识别结果数据；所述文本分类模型的损失函数值是根据交叉熵损失值以及相对损失值确定的，其中，所述相对损失值是根据第一预测概率分布值与第二预测概率分布值确定的；所述第一预测概率分布值是基于所述样本语音数据的转换数据得到的，所述第二预测概率分布值是基于所述样本语音数据的语音识别结果数据得到的。

[0156] 本领域普通技术人员可以理解：实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成，前述的程序可以存储于一计算机可读取存储介质中，该程序在执行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0157] 最后应说明的是：以上实施例仅用以说明本公开的技术方案，而非对其限制；尽管参照前述实施例对本公开进行了详细的说明，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公开各实施例技术方案的精神和范围。

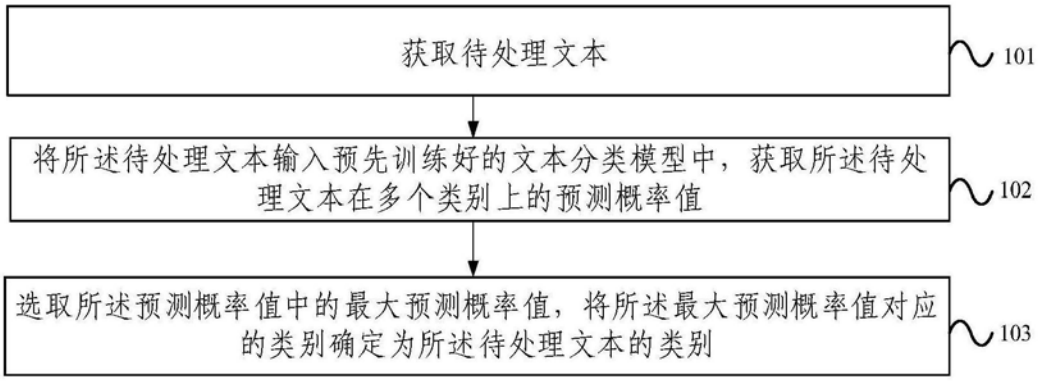

获取待处理文本
101
将所述待处理文本输入预先训练好的文本分类模型中，获取所述待处理文本在多个类别上的预测概率值
102
选取所述预测概率值中的最大预测概率值，将所述最大预测概率值对应的类别确定为所述待处理文本的类别
103

图1

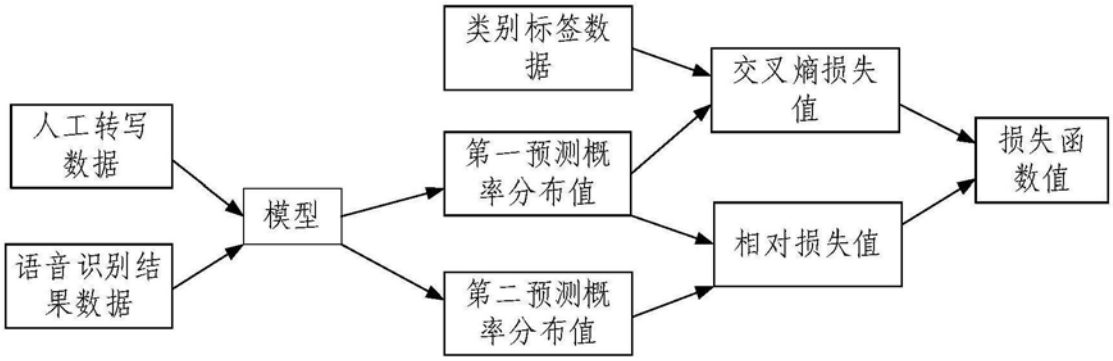

类别标签数据
交叉熵损失值
人工转写数据
第一预测概率分布值
损失函数值
模型
相对损失值
语音识别结果数据
第二预测概率分布值

图2

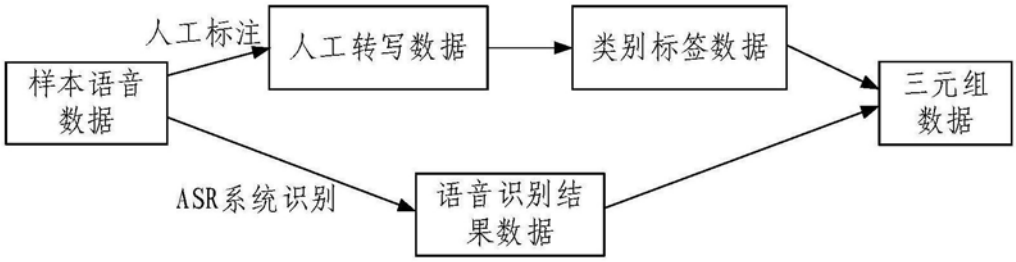

人工标注
人工转写数据
类别标签数据
样本语音数据
三元组数据
ASR系统识别
语音识别结果数据

图3

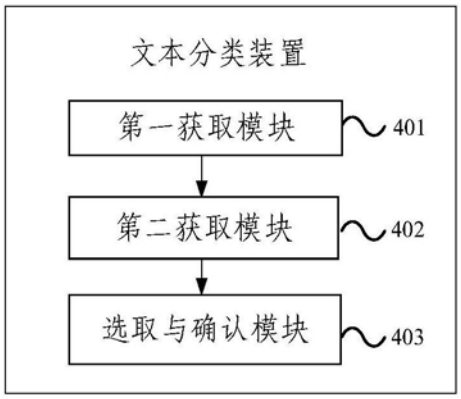

文本分类装置
第一获取模块
401
第二获取模块
402
选取与确认模块
403

图4

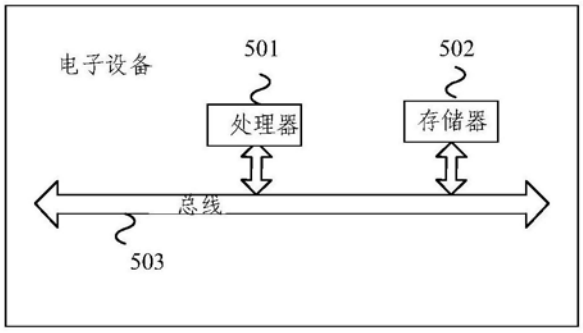

电子设备
501
处理器
502
存储器
总线
503

图5