



(12)发明专利申请

(10)申请公布号 CN 111460150 A

(43)申请公布日 2020.07.28

(21)申请号 202010231207.5

(22)申请日 2020.03.27

(71)申请人 北京松果电子有限公司

地址 100085 北京市海淀区清河朱房路临
66号C栋C单元

(72)发明人 唐可欣 齐保元 韩佳乘 孟二利

(74)专利代理机构 北京善任知识产权代理有限公司 11650

代理人 康艳青

(51) Int. Cl.

G06F 16/35(2019.01)

G06N 3/04(2006.01)

G06N 3/08(2006.01)

G06T 7/194(2017.01)

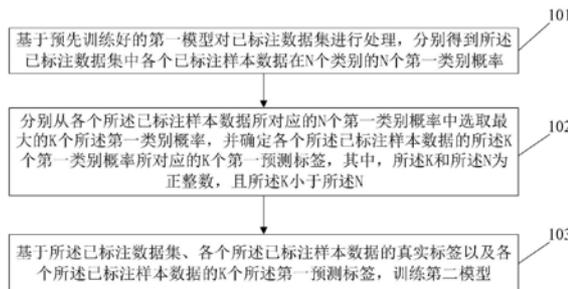
权利要求书3页 说明书15页 附图4页

(54)发明名称

一种分类模型的训练方法、分类方法、装置及存储介质

(57)摘要

本公开是关于一种分类模型的训练方法、分类方法、装置及存储介质,该训练方法包括:基于预先训练好的第一模型对已标注数据集进行处理,分别得到已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率;分别从各个已标注样本数据所对应的N个第一类别概率中选取最大的K个所述第一类别概率,并确定各个所述已标注样本数据的所述K个第一类别概率所对应的K个第一预测标签,其中,所述K和所述N为正整数,且所述K小于所述N;基于所述已标注数据集、各个所述已标注样本数据的真实标签以及各个所述已标注样本数据的K个所述第一预测标签,训练第二模型。相较于基于第一模型输出的所有第一预测标签训练第二模型,能减少存储第一预测标签需要占用的内存空间;当基于第一预测标签计算第二模型的训练损失时,还能提高数据的计算速度。



1. 一种分类模型的训练方法,其特征在于,应用于电子设备,包括:

基于预先训练好的第一模型对已标注数据集进行处理,分别得到所述已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率;

分别从各个所述已标注样本数据所对应的N个第一类别概率中选取最大的K个所述第一类别概率,并确定各个所述已标注样本数据的所述K个第一类别概率所对应的K个第一预测标签,其中,所述K和所述N为正整数,且所述K小于所述N;

基于所述已标注数据集、各个所述已标注样本数据的真实标签以及各个所述已标注样本数据的K个所述第一预测标签,训练第二模型。

2. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

基于所述第一模型对未标注数据集进行处理,分别得到所述未标注数据集中各个未标注样本数据在M个类别的M个第二类别概率;

从各个所述未标注样本数据对应M个第二类别概率中选取最大的H个第二类别概率,并确定各个所述未标注样本数据的所述H个第二类别概率所对应的H个第二预测标签,其中,所述M和所述H为正整数,且所述H小于所述M;

基于所述已标注数据集、所述未标注数据集、各个所述已标注样本数据的真实标签、各个所述已标注样本数据的K个所述第一预测标签、以及各个所述未标注样本数据的H个所述第二预测标签,训练所述第二模型。

3. 根据权利要求2所述的方法,其特征在于,所述基于所述已标注数据集、所述未标注数据集、各个所述已标注样本数据的真实标签、各个所述已标注样本数据的K个所述第一预测标签、以及各个所述未标注样本数据的H个所述第二预测标签,训练所述第二模型,训练所述第二模型,包括:

将所述已标注数据集中的各个所述已标注样本数据输入所述第二模型,得到所述第二模型输出的第三预测标签;

将所述未标注数据集中的各个所述未标注样本数据输入所述第二模型,得到所述第二模型输出的第四预测标签;

利用预设损失函数,基于所述真实标签、各个所述已标注样本数据的K个所述第一预测标签、所述第三预测标签、各个所述未标注样本数据的H个所述第二预测标签、以及所述第四预测标签,确定所述第二模型的训练损失;

基于所述训练损失,调整所述第二模型的模型参数。

4. 根据权利要求3所述的方法,其特征在于,所述基于所述真实标签、各个所述已标注样本数据的K个所述第一预测标签、所述第三预测标签、各个所述未标注样本数据的H个所述第二预测标签、以及所述第四预测标签,确定所述第二模型的训练损失,包括:

基于所述真实标签和所述第三预测标签确定所述第二模型在所述已标注数据集上的第一损失;

基各个所述已标注样本数据的K个所述第一预测标签和所述第三预测标签,确定所述第二模型在所述已标注数据集上的第二损失;

基于各个所述未标注样本数据的H个所述第二预测标签和所述第四预测标签,确定所述第二模型在所述未标注数据集上的第三损失;

基于所述第一损失、所述第二损失和所述第三损失的加权和,确定所述训练损失。

5. 根据权利要求4所述的方法,其特征在于,所述基于所述第一损失、所述第二损失和所述第三损失的加权和,确定所述训练损失,包括:

确定所述第一损失值与第一预设权重的第一乘积;

根据所述第一预设权重确定损失权重,并确定所述第二损失值与所述损失权重的第二乘积;

确定所述第三损失值与第二预设权重的第三乘积,所述第二预设权重小于或者等于所述第一预设权重;

将所述第一乘积、所述第二乘积、以及所述第三乘积相加,得到所述训练损失。

6. 根据权利要求3至5任一项所述的方法,其特征在于,所述方法还包括:

当所述训练损失在设定时长内的数值变化小于设定变化阈值时,停止训练所述第二模型。

7. 一种分类方法,其特征在于,应用于电子设备,包括:

将待分类数据输入采用权利要求1~6任一项分类模型训练方法训练得到的第二模型,输出所述待分类数据在X个类别的X个类别概率;

按照类别概率从大到小,确定X个类别概率中前预设数量个类别概率对应的类别标签;

将所述预设数量个类别标签确定为所述待分类数据的类别标签。

8. 一种分类模型的训练装置,其特征在于,应用于电子设备,包括:

第一确定模块,配置为基于预先训练好的第一模型对已标注数据集进行处理,分别得到所述已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率;

第一选取模块,配置为分别从各个所述已标注样本数据所对应的N个第一类别概率中选取最大的K个所述第一类别概率,并确定各个所述已标注样本数据的所述K个第一类别概率所对应的K个第一预测标签,其中,所述K和所述N为正整数,且所述K小于所述N;

第一训练模块,配置为基于所述已标注数据集、各个所述已标注样本数据的真实标签以及各个所述已标注样本数据的K个所述第一预测标签,训练第二模型。

9. 根据权利要求8所述的装置,其特征在于,所述装置还包括:

第二确定模块,配置为基于所述第一模型对未标注数据集进行处理,分别得到所述未标注数据集中各个未标注样本数据在M个类别的M个第二类别概率;

第二选取模块,配置为从各个所述未标注样本数据对应应有M个第二类别概率中选取最大的H个第二类别概率,并确定各个所述未标注样本数据的所述H个第二类别概率所对应的H个第二预测标签,其中,所述M和所述H为正整数,且所述H小于所述M;

第二训练模块,配置为基于所述已标注数据集、所述未标注数据集、各个所述已标注样本数据的真实标签、各个所述已标注样本数据的K个所述第一预测标签、以及各个所述未标注样本数据的H个所述第二预测标签,训练所述第二模型。

10. 根据权利要求9所述的装置,其特征在于,所述第二训练模块,包括:

第一确定子模块,配置为将所述已标注数据集中的各个所述已标注样本数据输入所述第二模型,得到所述第二模型输出的第三预测标签;

第二确定子模块,配置为将所述未标注数据集中的各个所述未标注样本数据输入所述第二模型,得到所述第二模型输出的第四预测标签;

第三确定子模块,配置为利用预设损失函数,基于所述真实标签、各个所述已标注样本

数据的K个所述第一预测标签、所述第三预测标签、各个所述未标注样本数据的H个所述第二预测标签、以及所述第四预测标签,确定所述第二模型的训练损失;

调整子模块,配置为基于所述训练损失,调整所述第二模型的模型参数。

11. 根据权利要求10所述的装置,其特征在于,所述第三确定子模块,还配置为:

基于所述真实标签和所述第三预测标签确定所述第二模型在所述已标注数据集上的第一损失;

基各个所述已标注样本数据的K个所述第一预测标签和所述第三预测标签,确定所述第二模型在所述已标注数据集上的第二损失;

基于各个所述未标注样本数据的H个所述第二预测标签和所述第四预测标签,确定所述第二模型在所述未标注数据集上的第三损失;

基于所述第一损失、所述第二损失和所述第三损失的加权和,确定所述训练损失。

12. 根据权利要求11所述的装置,其特征在于,所述第三确定子模块,还配置为:

确定所述第一损失值与第一预设权重的第一乘积;

根据所述第一预设权重确定损失权重,并确定所述第二损失值与所述损失权重的第二乘积;

确定所述第三损失值与第二预设权重的第三乘积,所述第二预设权重小于或者等于所述第一预设权重;

将所述第一乘积、所述第二乘积、以及所述第三乘积相加,得到所述训练损失。

13. 根据权利要求10至12任一项所述的装置,其特征在于,所述装置还包括:

停止模块,配置为当所述训练损失在设定时长内的数值变化小于设定变化阈值时,停止训练所述第二模型。

14. 一种分类装置,其特征在于,应用于电子设备,包括:

分类模块,用于将待分类数据输入采用权利要求1~6任一项分类模型训练方法训练得到的第二模型,输出所述待分类数据在X个类别的X个类别概率;

标签确定模块,用于按照类别概率从大到小,确定X个类别概率中前预设数量个类别概率对应的类别标签;

类别确定模块,用于将所述预设数量个类别标签确定为所述待分类数据的类别标签。

15. 一种分类模型的训练装置,其特征在于,包括:

处理器;

配置为存储处理器可执行指令的存储器;

其中,所述处理器配置为:执行时实现上述权利要求1至6中任一种分类模型的训练方法或权利要求7中分类方法中的步骤。

16. 一种非临时性计算机可读存储介质,当所述存储介质中的指令由分类模型的训练装置的处理器执行时,使得所述装置能够执行上述权利要求1至6中任一种分类模型的训练方法或者权利要求7中分类方法中的步骤。

一种分类模型的训练方法、分类方法、装置及存储介质

技术领域

[0001] 本公开涉及数学模型技术领域,尤其涉及一种分类模型的训练方法和装置、分类方法和装置及存储介质。

背景技术

[0002] 文本分类,是指根据任务目标将文档分为N个类别中的一个或多个。目前,随着神经网络语言模型在自然语言处理(Natural Language Processing, NLP)领域的发展,越来越多的研究人员选择对预训练语言模型进行微调的方法来取得高精确度的模型。然而,由于预训练模型复杂的编码结构,模型的微调和实际生产往往伴随着巨额的时间和空间开销。

[0003] 知识蒸馏是一种常见的深度学习模型压缩方法,旨在把一个大模型或者多个模型融合学到的知识迁移到另一个轻量级单模型上。相关技术的知识蒸馏中,对于海量标签文本分类来说,需要保存每个样例的预测标签,而保存每个样例的预测标签需要占用大量的内存空间,并且在实际计算损失函数时,也会因为向量的纬度太高,导致计算过程非常缓慢。

发明内容

[0004] 本公开提供一种分类模型的训练方法、分类方法、装置及存储介质。

[0005] 根据本公开实施例的第一方面,提供一种分类模型的训练方法,应用于电子设备,包括:

[0006] 基于预先训练好的第一模型对已标注数据集进行处理,分别得到所述已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率;

[0007] 分别从各个所述已标注样本数据所对应的N个第一类别概率中选取最大的K个所述第一类别概率,并确定各个所述已标注样本数据的所述K个第一类别概率所对应的K个第一预测标签,其中,所述K和所述N为正整数,且所述K小于所述N;

[0008] 基于所述已标注数据集、各个所述已标注样本数据的真实标签以及各个所述已标注样本数据的K个所述第一预测标签,训练第二模型。

[0009] 可选的,所述方法还包括:

[0010] 基于所述第一模型对未标注数据集进行处理,分别得到所述未标注数据集中各个未标注样本数据在M个类别的M个第二类别概率;

[0011] 从各个所述未标注样本数据对应有M个第二类别概率中选取最大的H个第二类别概率,并确定各个所述未标注样本数据的所述H个第二类别概率所对应的H个第二预测标签,其中,所述M和所述H为正整数,且所述H小于所述M;

[0012] 基于所述已标注数据集、所述未标注数据集、各个所述已标注样本数据的真实标签、各个所述已标注样本数据的K个所述第一预测标签、以及各个所述未标注样本数据的H个所述第二预测标签,训练所述第二模型。

[0013] 可选的,所述基于所述已标注数据集、所述未标注数据集、各个所述已标注样本数据的真实标签、各个所述已标注样本数据的K个所述第一预测标签、以及各个所述未标注样本数据的H个所述第二预测标签,训练所述第二模型,训练所述第二模型,包括:

[0014] 将所述已标注数据集中的各个所述已标注样本数据输入所述第二模型,得到所述第二模型输出的第三预测标签;

[0015] 将所述未标注数据集中的各个所述未标注样本数据输入所述第二模型,得到所述第二模型输出的第四预测标签;

[0016] 利用预设损失函数,基于所述真实标签、各个所述已标注样本数据的K个所述第一预测标签、所述第三预测标签、各个所述未标注样本数据的H个所述第二预测标签、以及所述第四预测标签,确定所述第二模型的训练损失;

[0017] 基于所述训练损失,调整所述第二模型的模型参数。

[0018] 可选的,所述基于所述真实标签、各个所述已标注样本数据的K个所述第一预测标签、所述第三预测标签、各个所述未标注样本数据的H个所述第二预测标签、以及所述第四预测标签,确定所述第二模型的训练损失,包括:

[0019] 基于所述真实标签和所述第三预测标签确定所述第二模型在所述已标注数据集上的第一损失;

[0020] 基于各个所述已标注样本数据的K个所述第一预测标签和所述第三预测标签,确定所述第二模型在所述已标注数据集上的第二损失;

[0021] 基于各个所述未标注样本数据的H个所述第二预测标签和所述第四预测标签,确定所述第二模型在所述未标注数据集上的第三损失;

[0022] 基于所述第一损失、所述第二损失和所述第三损失的加权和,确定所述训练损失。

[0023] 可选的,所述基于所述第一损失、所述第二损失和所述第三损失的加权和,确定所述训练损失,包括:

[0024] 确定所述第一损失值与第一预设权重的第一乘积;

[0025] 根据所述第一预设权重确定损失权重,并确定所述第二损失值与所述损失权重的第二乘积;

[0026] 确定所述第三损失值与第二预设权重的第三乘积,所述第二预设权重小于或者等于所述第一预设权重;

[0027] 将所述第一乘积、所述第二乘积、以及所述第三乘积相加,得到所述训练损失。

[0028] 可选的,所述方法还包括:

[0029] 当所述训练损失在设定时长内的数值变化小于设定变化阈值时,停止训练所述第二模型。

[0030] 根据本公开实施例的第二方面,提供一种分类方法,应用于电子设备,包括:

[0031] 将待分类数据输入采用上述第一方面提供的分类模型训练方法训练得到的第二模型,输出所述待分类数据在X个类别的X个类别概率;

[0032] 按照类别概率从大到小,确定X个类别概率中前预设数量个类别概率对应的类别标签;

[0033] 将所述预设数量个类别标签确定为所述待分类数据的类别标签。

[0034] 根据本公开实施例的第三方面,提供一种分类模型的训练装置,应用于电子设备,

包括：

[0035] 第一确定模块，配置为基于预先训练好的第一模型对已标注数据集进行处理，分别得到所述已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率；

[0036] 第一选取模块，配置为分别从各个所述已标注样本数据所对应的N个第一类别概率中选取最大的K个所述第一类别概率，并确定各个所述已标注样本数据的所述K个第一类别概率所对应的K个第一预测标签，其中，所述K和所述N为正整数，且所述K小于所述N；

[0037] 第一训练模块，配置为基于所述已标注数据集、各个所述已标注样本数据的真实标签以及各个所述已标注样本数据的K个所述第一预测标签，训练第二模型。

[0038] 可选的，所述装置还包括：

[0039] 第二确定模块，配置为基于所述第一模型对未标注数据集进行处理，分别得到所述未标注数据集中各个未标注样本数据在M个类别的M个第二类别概率；

[0040] 第二选取模块，配置为从各个所述未标注样本数据对应有M个第二类别概率中选取最大的H个第二类别概率，并确定各个所述未标注样本数据的所述H个第二类别概率所对应的H个第二预测标签，其中，所述M和所述H为正整数，且所述H小于所述M；

[0041] 第二训练模块，配置为基于所述已标注数据集、所述未标注数据集、各个所述已标注样本数据的真实标签、各个所述已标注样本数据的K个所述第一预测标签、以及各个所述未标注样本数据的H个所述第二预测标签，训练所述第二模型。

[0042] 可选的，所述第二训练模块，包括：

[0043] 第一确定子模块，配置为将所述已标注数据集中的各个所述已标注样本数据输入所述第二模型，得到所述第二模型输出的第三预测标签；

[0044] 第二确定子模块，配置为将所述未标注数据集中的各个所述未标注样本数据输入所述第二模型，得到所述第二模型输出的第四预测标签；

[0045] 第三确定子模块，配置为利用预设损失函数，基于所述真实标签、各个所述已标注样本数据的K个所述第一预测标签、所述第三预测标签、各个所述未标注样本数据的H个所述第二预测标签、以及所述第四预测标签，确定所述第二模型的训练损失；

[0046] 调整子模块，配置为基于所述训练损失，调整所述第二模型的模型参数。

[0047] 可选的，所述第三确定子模块，还配置为：

[0048] 基于所述真实标签和所述第三预测标签确定所述第二模型在所述已标注数据集上的第一损失；

[0049] 基于各个所述已标注样本数据的K个所述第一预测标签和所述第三预测标签，确定所述第二模型在所述已标注数据集上的第二损失；

[0050] 基于各个所述未标注样本数据的H个所述第二预测标签和所述第四预测标签，确定所述第二模型在所述未标注数据集上的第三损失；

[0051] 基于所述第一损失、所述第二损失和所述第三损失的加权和，确定所述训练损失。

[0052] 可选的，所述第三确定子模块，还配置为：

[0053] 确定所述第一损失值与第一预设权重的第一乘积；

[0054] 根据所述第一预设权重确定损失权重，并确定所述第二损失值与所述损失权重的第二乘积；

[0055] 确定所述第三损失值与第二预设权重的第三乘积，所述第二预设权重小于或者等

于所述第一预设权重；

[0056] 将所述第一乘积、所述第二乘积、以及所述第三乘积相加，得到所述训练损失。

[0057] 可选的，其特征在于，所述装置还包括：

[0058] 停止模块，配置为当所述训练损失在设定时长内的数值变化小于设定变化阈值时，停止训练所述第二模型。

[0059] 根据本公开实施例的第四方面，提供一种分类装置，应用于电子设备，包括：

[0060] 分类模块，用于将待分类数据输入采用上述第一方面提供的分类模型训练方法训练得到的第二模型，输出所述待分类数据在X个类别的X个类别概率；

[0061] 标签确定模块，用于按照类别概率从大到小，确定X个类别概率中前预设数量个类别概率对应的类别标签；

[0062] 类别确定模块，用于将所述预设数量个类别标签确定为所述待分类数据的类别标签。

[0063] 根据本公开实施例的第五方面，提供一种分类模型的训练装置，包括：

[0064] 处理器；

[0065] 配置为存储处理器可执行指令的存储器；

[0066] 其中，所述处理器配置为：执行时实现上述第一方面中任一种分类模型的训练方法或者上述第二方面中分类方法中的步骤。

[0067] 根据本公开实施例的第六方面，提供一种非临时性计算机可读存储介质，当所述存储介质中的指令由分类模型的训练装置的处理器执行时，使得所述装置能够执行上述第一方面中任一种分类模型的训练方法或者上述第二方面中分类方法中的步骤。

[0068] 本公开的实施例提供的技术方案可以包括以下有益效果：

[0069] 由上述实施例可知，本公开可以基于第一模型对已标注数据集中的已标注数据进行预测，并输出各个已标注数据的第一类别概率以及各个已标注数据的第一预测标签，然后从第一模型输出的所有第一预测标签中选取概率最大的K个第一类别概率，以及所述K个第一类别概率所对应的K个第一预测标签。

[0070] 由于在基于第一模型对第二模型进行训练的过程中，需要将第一模型输出的第一预测标签保存至设定存储空间，需要基于该第一预测标签训练第二模型时，再从该设定存储空间调用第一预测标签，如果所存储的第一预测标签数量较大，可能会浪费设定存储空间的内存资源。本公开实施例中，通过选取最大的K个第一类别概率所对应的K个第一预测标签训练第二模型，相较于直接基于第一模型输出的所有第一预测标签训练第二模型，第一方面，能够减少存储第一预测标签需要占用的内存空间；第二方面，由于数据量减少，在训练的过程中，如果需要基于第一预测标签计算第二模型的训练损失，能够提高数据的计算速度。

[0071] 应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本公开。

附图说明

[0072] 此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本公开的实施例，并与说明书一起用于解释本公开的原理。

- [0073] 图1是根据一示例性实施例示出的一种分类模型的训练方法的流程示意图。
- [0074] 图2是根据一示例性实施例示出的另一种分类模型的训练方法的流程示意图。
- [0075] 图3是根据一示例性实施例示出的分类模型的训练装置框图。
- [0076] 图4是根据一示例性实施例示出的一种用于分类模型的训练装置的框图。
- [0077] 图5是根据一示例性实施例示出的另一种用于分类模型的训练装置的框图。

具体实施方式

[0078] 这里将详细地对示例性实施例进行说明,其示例表示在附图中。下面的描述涉及附图时,除非另有表示,不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本公开相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本公开的一些方面相一致的装置和方法的例子。

[0079] 本公开实施例中提供了一种分类模型的训练方法,图1是根据一示例性实施例示出的一种分类模型的训练方法的流程示意图,如图1所示,该方法应用于电子设备,主要包括以下步骤:

[0080] 在步骤101中,基于预先训练好的第一模型对已标注数据集进行处理,分别得到所述已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率;

[0081] 在步骤102中,分别从各个所述已标注样本数据所对应的N个第一类别概率中选取最大的K个所述第一类别概率,并确定各个所述已标注样本数据的所述K个第一类别概率所对应的K个第一预测标签,其中,所述K和所述N为正整数,且所述K小于所述N;

[0082] 在步骤103中,基于所述已标注数据集、各个所述已标注样本数据的真实标签以及各个所述已标注样本数据的K个所述第一预测标签,训练第二模型。

[0083] 这里,电子设备包括移动终端和固定终端,其中,移动终端包括:手机、平板电脑、笔记本电脑等;固定终端包括:个人计算机。在其他可选的实施例中,该信息处理方法也可以运行于网络侧设备,其中,网络侧设备包括:服务器、处理中心等。

[0084] 本公开实施例第一模型和第二模型可为实现预定功能的数学模型,包括但不限于以下至少之一:

[0085] 对输入的文本进行分类;

[0086] 对输入图像中目标和背景进行分割的目标分割;

[0087] 对输入图像中目标的分类;

[0088] 基于输入图像的目标跟踪;

[0089] 基于医疗图像的诊断辅助;

[0090] 基于输入语音的语音识别、语音校正等功能。

[0091] 以上仅是对所述第一模型和第二模型所实现预定功能的举例说明,具体实现不局限于上述举例。

[0092] 在其他可选的实施例中,可以基于已标注的训练数据集对预设模型进行训练,得到第一模型,其中,预设模型包括包括预测精度高但数据处理速度慢的预训练模型。例如,Bert模型、知识增强语义表示模型(Enhanced Representation from Knowledge Integration,Ernie模型)、Xlnet模型、神经网络模型、快速文本分类模型以及支持向量机模型等。第二模型包括预测精度低但数据处理速度快的模型,例如,albert模型、tiny模型

等。

[0093] 以第一模型是Bert模型为例,可以基于训练数据集对Bert模型进行训练,得到训练好的目标Bert模型。这时,可以将已标注数据集中的已标注数据输入目标Bert模型,并基于目标Bert模型输出各个已标注数据样本在N各类别的N各第一类别概率,这里,第一类别概率的类型可以包括:非归一化类别概率和归一化类别概率,其中,非归一化概率是未经归一化函数(例如,softmax函数)进行归一化处理的概率值,归一化概率是指经过归一化函数进行归一化处理的概率值。由于非归一化类别概率相较于归一化类别概率所包含的信息量较高,本公开实施例中,可以基于第一模型输出非归一化类别概率,在其他可选的实施例中,也可以基于第一模型输出归一化类别概率。

[0094] 以已标注数据集中某一个已标注样本数据(第一样本数据)为例,在将第一样本数据输入第一模型之后,可以基于第一模型输出第一样本数据在N个类别的N个第一类别概率。例如,第一样本数据在第一类的第一类别概率是0.4,第一样本数据在第二类的第一类别概率是0.001,第一样本数据在第三类的第一类别概率是0.05,.....第一样本数据在第N类的第一类别概率是0.35,这样,就能确定出第一样本数据在各个类别的第一类别概率,其中,第一类别概率越大,则第一样本数据属于该类别的可能性越大,第一类别概率越小,则第一样本数据属于该类别的可能性越小。例如,第一样本数据在第一类的第一类别概率是0.4,第一样本数据在第二类的第一类别概率是0.001,则可以确定第一样本数据属于第一类的概率高于属于第二类的概率。

[0095] 在得到已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率之后,可以将N个第一类别概率按照从大到小的顺序进行排序,根据排序结果从N个第一类别概率中选取最大的K个第一类别概率。还是以已标注数据集中的第一样本数据为例,第一样本数据在第一类的第一类别概率是0.4,第一样本数据在第二类的第一类别概率是0.001,第一样本数据在第三类的第一类别概率是0.05,.....第一样本数据在第N类的第一类别概率是0.35,在将第一样本数据对应的N个第一类别概率按照从大大小进行排序之后,可以取前K个第一类别概率。以N是3000,K是20为例,可以对3000个第一类别概率按照从大到小的顺序进行排序,并选取最大的20个第一类别概率。

[0096] 由于当第一类别概率小于设定概率阈值时,第一样本数据属于该类别的可能性会很小,本公开实施例中,能够选取数值较高的第一类别概率,摒弃数值较低的第一类别概率,能够在保证输出的类别概率的准确性的基础上减少数据量,进而减少训练模型的计算量。在选取最大的K个第一类别概率之后,能够确定出最大的K个第一类别概率所对应的的K个第一预测标签,并基于已标注数据集、各个已标注样本数据的真实标签以及该K个第一预测标签,训练第二模型。

[0097] 本公开实施例中,可以基于第一模型对已标注数据集中的已标注样本数据进行预测,并输出各个已标注样本数据的第一类别概率以及各个已标注样本数据的第一预测标签,然后从第一模型输出的所有第一预测标签中选取概率最大的K个第一类别概率,以及所述K个第一类别概率所对应的K个第一预测标签。

[0098] 由于在基于第一模型对第二模型进行训练的过程中,需要将第一模型输出的第一预测标签保存至设定存储空间,需要基于该第一预测标签训练第二模型时,再从该设定存储空间调用第一预测标签,如果所存储的第一预测标签数量较大,可能会浪费设定存储空

间的内存资源。本公开实施例中,通过选取最大的K个第一类别概率所对应的K个第一预测标签训练第二模型,相较于直接基于第一模型输出的所有第一预测标签训练第二模型,第一方面,能够减少存储第一预测标签需要占用的内存空间;第二方面,由于数据量减少,在训练的过程中,如果需要基于第一预测标签计算第二模型的训练损失,能够提高数据的计算速度。

[0099] 在其他可选的实施例中,所述方法还包括:

[0100] 基于所述第一模型对未标注数据集进行处理,分别得到所述未标注数据集中各个未标注样本数据在M个类别的M个第二类别概率;

[0101] 从各个所述未标注样本数据对应有M个第二类别概率中选取最大的H个第二类别概率,并确定各个所述未标注样本数据的所述H个第二类别概率所对应的H个第二预测标签,其中,所述M和所述H为正整数,且所述H小于所述M;

[0102] 基于所述已标注数据集、所述未标注数据集、各个所述已标注样本数据的真实标签、各个所述已标注样本数据的K个所述第一预测标签、以及各个所述未标注样本数据的H个所述第二预测标签,训练所述第二模型。

[0103] 这里,第二类别概率的类型可以包括:非归一化类别概率和归一化类别概率。由于归一化类别概率相较于非归一化类别概率能够使各个类别之间的分别更加明显,本公开实施例中,可以基于第一模型输出归一化类别概率,在其他可选的实施例中,也可以基于第一模型输出非归一化类别概率。

[0104] 以未标注数据集中某一个未标注样本数据(第二样本数据)为例,在将第二样本数据输入第一模型之后,可以基于第一模型输出第二样本数据在M个类别的M个第二类别概率。例如,第二样本数据在第一类的第二类别概率是0.01,第二样本数据在第二类的第二类别概率是0.0001,第二样本数据在第三类的第二类别概率是0.45,.....第二样本数据在第N类的第二类别概率是0.35,这样,就能确定出第二样本数据在各个类别的第二类别概率,其中,第二类别概率越大,则第二样本数据属于该类别的可能性越大,第二类别概率越小,则第二样本数据属于该类别的可能性越小。例如,第二样本数据在第三类的第二类别概率是0.45,第二样本数据在第二类的第二类别概率是0.0001,则可以确定第二样本数据属于第三类的概率高于属于第二类的概率。

[0105] 在得到未标注数据集中各个未标注样本数据在M个类别的M个第二类别概率之后,可以将M个第二类别概率按照从大到小的顺序进行排序,根据排序结果从M个第二类别概率中选取最大的H个第二类别概率。还是以未标注数据集中的第二样本数据为例,第二样本数据在第一类的第二类别概率是0.01,第二样本数据在第二类的第二类别概率是0.0001,第二样本数据在第三类的第二类别概率是0.45,.....第二样本数据在第N类的第二类别概率是0.35,在将第二样本数据对应的M个第二类别概率按照从大大小进行排序之后,可以取前H个第二类别概率。以M是300,H是1为例,可以对300个第二类别概率按照从大到小的顺序进行排序,并选取最大的一个第二类别概率,并将最大的一个第二类别概率所对应的第二预测标签确定为第二样本数据的标签。

[0106] 本公开实施例中,可以基于第一模型对未标注数据集中的未标注样本数据进行预测,并输出各个未标注数据的第二类别概率以及各个未标注数据的第二预测标签,然后从第一模型输出的所有第二预测标签中选取概率最大的H个第二类别概率,以及所述H个第二

类别概率所对应的H个第二预测标签。通过增加未标注样本数据的第二预测标签,并基于第二预测标签对第二模型进行训练,扩充了第二模型的训练语料,能够提高数据的多样性,以提高所训练的第二模型的泛化能力。

[0107] 在其他可选的实施例中,所述基于所述已标注数据集、所述未标注数据集、各个所述已标注样本数据的真实标签、各个所述已标注样本数据的K个所述第一预测标签、以及各个所述未标注样本数据的H个所述第二预测标签,训练所述第二模型,训练所述第二模型,包括:

[0108] 将所述已标注数据集中的各个所述已标注样本数据输入所述第二模型,得到所述第二模型输出的第三预测标签;

[0109] 将所述未标注数据集中的各个所述未标注样本数据输入所述第二模型,得到所述第二模型输出的第四预测标签;

[0110] 利用预设损失函数,基于所述真实标签、各个所述已标注样本数据的K个所述第一预测标签、所述第三预测标签、各个所述未标注样本数据的H个所述第二预测标签、以及所述第四预测标签,确定所述第二模型的训练损失;

[0111] 基于所述训练损失,调整所述第二模型的模型参数。

[0112] 这里,预设损失函数用于衡量第二模型预测的好坏,本公开实施例中,通过将已标注样本数据输入第二模型预测得到第三预测标签,将未标注样本数据输入第二模型得到第四预测标签,并利用预设损失函数,基于真实标签、各个已标注样本数据的K个第一预测标签、第三预测标签、各个未标注样本数据的H个第二预测标签、以及第四预测标签,确定第二模型的训练损失,进而利用基于预设损失函数得到的训练损失调整第二模型的模型参数。

[0113] 本公开实施例中,第一方面,相较于直接基于第一模型输出的所有第一预测标签训练第二模型,能够减少存储第一预测标签需要占用的内存空间;第二方面,由于数据量减少,在训练的过程中,如果需要基于第一预测标签计算第二模型的训练损失,能够提高数据的计算速度;第三方面,通过增加未标注样本数据的第二预测标签,并基于第二预测标签对第二模型进行训练,扩充了第二模型的训练语料,能够提高数据的多样性,以提高所训练的第二模型的泛化能力;第四方面,还针对不同的损失计算任务使用新的预设损失函数,基于预设损失函数调整第二模型的模型参数,能够提升第二模型的性能。

[0114] 在其他可选的实施例中,该方法还包括:基于测试数据集对训练后的第二模型进行性能评估,得到评估结果;其中,测试数据集中测试数据的类型包括以下至少之一:文本数据类型、图像数据类型、业务数据类型和音频数据类型。这里,在得到训练后的第二模型之后,可以在测试数据集上评估其性能,逐步优化第二模型,直至找到最优的第二模型,例如最小化验证损失或最大化奖励的第二模型。这里,可以将测试数据集中的测试数据输入训练好的第二模型,经由该第二模型输出评估结果,然后将输出的评估结果与预设的标准进行比较,得到比较结果,并根据比较结果评估第二模型的性能,其中,测试结果可以为第二模型处理测试数据的速度或者精度。

[0115] 在其他可选的实施例中,所述基于所述真实标签、各个所述已标注样本数据的K个所述第一预测标签、所述第三预测标签、各个所述未标注样本数据的H个所述第二预测标签、以及所述第四预测标签,确定所述第二模型的训练损失,包括:

[0116] 基于所述真实标签和所述第三预测标签确定所述第二模型在所述已标注数据集

上的第一损失；

[0117] 基于各个所述已标注样本数据的K个所述第一预测标签和所述第三预测标签，确定所述第二模型在所述已标注数据集上的第二损失；

[0118] 基于各个所述未标注样本数据的H个所述第二预测标签和所述第四预测标签，确定所述第二模型在所述未标注数据集上的第三损失；

[0119] 基于所述第一损失、所述第二损失和所述第三损失的加权和，确定所述训练损失。

[0120] 这里，所述第一损失为真实标签和第三预测标签的交叉熵，第一损失的计算公式包括：

$$[0121] \quad loss^{(hard)} = -\sum_{i \in N} y_i' \log(y_i); \quad (1)$$

[0122] 公式(1)中， $loss^{(hard)}$ 表示所述第一损失；N表示所述已标注数据集的大小； y_i' 表示第i个维度的真实标签， y_i 表示第i个维度的第三预测标签，i为正整数。其中， y_i 的计算公式包括：

$$[0123] \quad y_i = -\frac{e^{Z_i}}{\sum_j e^{Z_j}}; \quad (2)$$

[0124] 公式(2)中， y_i 表示第i个维度的第三预测标签， Z_i 表示第i个维度的已标注数据的第一类别概率； Z_j 表示第j个维度的已标注数据的第一类别概率，i和j均为正整数。

[0125] 所述第二损失为各个已标注样本数据的K个第一预测标签与第三预测标签的交叉熵，第二损失的计算公式包括：

$$[0126] \quad loss^{(soft)} = -\frac{1}{T} \sum_{i \in ST_1} \hat{y}_i' \log(y_i); \quad (3)$$

[0127] 公式(3)中， $loss^{(soft)}$ 表示所述第二损失； \hat{y}_i' 表示第i个维度的第一预测标签； y_i 表示第i个维度第三预测标签；T表示预设温度参数； ST_1 表示第一预测标签的数量，可以等于K；i为正整数，这里，它所包含的分类信息就越多出的预测值更加平缓。其中， y_i 的计算公式包括：

$$[0128] \quad y_i = -\frac{e^{Z_i/T}}{\sum_j (e^{Z_j/T})} \quad (4)$$

[0129] 公式(4)中， y_i 表示第i个维度第三预测标签， Z_i 表示第i个维度的已标注数据的第一类别概率； Z_j 表示第j个维度的已标注数据的第一类别概率，T表示预设温度参数；i和j均为正整数。这里，预设温度参数的数值越大，输出的概率分布就越平缓，输出结果所包含的分类信息就越多，通设置预设温度参数，可以基于预设温度参数调整输出的概率分布的平缓度，进而调整输出结果所包含的分类信息量，能够提高模型训练的精确度和灵活性。

[0130] 所述第三损失为第二预测标签与第四预测标签的交叉熵，第三损失的计算公式包括：

$$[0131] \quad loss^{(hard_2)} = -\sum_{i \in M} y_i'' \log(y_i) \quad (5)$$

[0132] 公式(5)中， $loss^{(hard_2)}$ 表示所述第三损失； y_i'' 表示第i个维度的第二预测标签； y_i 表

示第*i*个维度的第四预测标签；*M*表示未标注数据集的大小；*i*为正整数。本公开实施例中，针对不同的损失计算任务使用新的预设损失函数，基于预设损失函数调整第二模型的模型参数，能够提升第二模型的性能。

[0133] 在其他可选的实施例中，所述基于所述第一损失、所述第二损失和所述第三损失的加权和，确定所述训练损失，包括：

[0134] 确定所述第一损失值与第一预设权重的第一乘积；

[0135] 根据所述第一预设权重确定损失权重，并确定所述第二损失值与所述损失权重的第二乘积；

[0136] 确定所述第三损失值与第二预设权重的第三乘积，所述第二预设权重小于或者等于所述第一预设权重；

[0137] 将所述第一乘积、所述第二乘积、以及所述第三乘积相加，得到所述训练损失。

[0138] 在其他可选的实施例中，所述训练损失计算公式包括：

$$[0139] \quad Loss = \alpha * loss^{(hard)} + (1 - \alpha) * loss^{(soft)} + \beta * loss^{(hard_2)}; \quad (6)$$

[0140] 公式(6)中，*Loss*表示所述第二模型的训练损失； $loss^{(hard)}$ 表示所述第一损失； $loss^{(soft)}$ 表示所述第二损失； $loss^{(hard_2)}$ 表示所述第三损失； α 为第一预设权重，且 α 大于0.5且小于1； β 为第二预设权重，且 β 小于等于 α 。本公开实施例中，一方面，针对不同的损失计算任务使用新的预设损失函数，基于预设损失函数调整第二模型的模型参数，能够提升第二模型的性能；另一方面，通过设置可调整的第一预设权重和第二预设权重，可以根据需要调整第一损失、第二损失和第三损失在训练损失中所占的比重，增加了模型训练的灵活性。

[0141] 在其他可选的实施例中，所述方法还包括：

[0142] 当所述训练损失在设定时长内的数值变化小于设定变化阈值时，停止训练所述第二模型。在其他可选的实施例中，还可以基于设定的验证集对第二模型的准确率进行验证，当准确率达到设定准确率时，停止训练第二模型，获得训练完成的目标模型。

[0143] 图2是根据一示例性实施例示出的另一种模型的训练方法的流程示意图，如图2所示，在基于第一模型(Teacher模型)训练第二模型(Student模型)的过程中，可以预先确定好第一模型，在已标注的训练数据集*L*上对第一模型进行微调，并保存微调后的第一模型，这里，可以将微调后的第一模型记为*TM*。这里，第一模型可以为预测精度高但计算速度慢的预训练模型，例如，Bert模型、Ernie模型、Xlnet模型等。

[0144] 在得到*TM*之后，可以使用*TM*对已标注数据集(转移集*T*)进行预测，分别得到已标注数据集中各个已标注样本数据在*N*个类别的*N*个第一类别概率，分别从各个已标注样本数据所对应的*N*个第一类别概率中选取最大的*K*个第一类别概率，并确定最大的*K*个第一类别概率所对应的*K*个第一预测标签，其中，*K*为超参数，例如，*K*可以等于20。

[0145] 本公开实施例中，还可以使用*TM*对未标注数据集*U*进行预测，分别得到未标注数据集中各个未标注样本数据在*M*个类别的*M*个第二类别概率，从各个未标注样本数据对应的*M*个第二类别概率中选取最大的*H*个第二类别概率，并确定最大的*H*个第二类别概率所对应的*H*个第二预测标签，其中，*H*可以等于1。这里，当第二类别概率为非归一化类别概率时，可以利用激活函数softmax对第二类别概率进行归一化处理。这样，就能够确定训练第二模型所需的数据。

[0146] 本公开实施例中,可以将已标注数据集中的各个已标注样本数据输入第二模型,得到第二模型输出的第三预测标签;将未标注数据集中的各个未标注样本数据输入第二模型,得到第二模型输出的第四预测标签;利用预设损失函数,基于真实标签、各个已标注样本数据的K个第一预测标签、第三预测标签、各个未标注样本数据的H个第二预测标签、以及第四预测标签,确定第二模型的训练损失;基于训练损失,调整第二模型的模型参数。

[0147] 本公开实施例中,第一方面,通过选取第一模型输出的最大的K个第一预测标签代替传统模型蒸馏中选取所有第一预测标签来训练第二模型,在不会对第二模型的性能产生影响的基础上,减少了内存消耗,并提升了第二模型训练速度;第二方面,充分利用未标注数据集,将未标注数据引入数据蒸馏的过程中,扩充了第二模型的训练语料,能够提高数据的多样性,以提高所训练的第二模型的泛化能力;第三方面,针对联合任务使用新的预设损失函数,基于预设损失函数调整第二模型的模型参数,能够提升第二模型的性能。

[0148] 本发明实施例还提供一种分类方法,可以采用经过训练的第二模型对待分类数据进行分类,可以包括如下步骤:

[0149] 步骤一、将待分类数据输入采用上述介绍的任一分类模型训练方法训练得到的第二模型,输出所述待分类数据在X个类别的X个类别概率。其中,X为自然数。

[0150] 步骤二、按照类别概率从大到小,确定X个类别概率中前预设数量个类别概率对应的类别标签。

[0151] 步骤三、将所述预设数量个类别标签确定为所述待分类数据的类别标签。

[0152] 可以根据实际需要确定待分类数据的类别标签数量,即预设数量,可以为一个或者多个。当预设数量为一个时,可以将类别概率最高的类别标签作为待分类数据的标签,当预设数量为多个时,可以按照类别概率从大到小的顺序,确定前多个类别概率,将这多个类别概率对应的类别标签,确定为待分类数据的类别标签。

[0153] 图3是根据一示例性实施例示出的分类模型的训练装置框图,如图3所示,该分类模型的训练装置300应用于电子设备,主要包括:

[0154] 第一确定模块301,配置为基于预先训练好的第一模型对已标注数据集进行处理,分别得到所述已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率;

[0155] 第一选取模块302,配置为分别从各个所述已标注样本数据所对应的N个第一类别概率中选取最大的K个所述第一类别概率,并确定各个所述已标注样本数据的所述K个第一类别概率所对应的K个第一预测标签,其中,所述K和所述N为正整数,且所述K小于所述N;

[0156] 第一训练模块303,配置为基于所述已标注数据集、各个所述已标注样本数据的真实标签以及各个所述已标注样本数据的K个所述第一预测标签,训练第二模型。

[0157] 在其他可选的实施例中,所述装置300还包括:

[0158] 第二确定模块,配置为基于所述第一模型对未标注数据集进行处理,分别得到所述未标注数据集中各个未标注样本数据在M个类别的M个第二类别概率;

[0159] 第二选取模块,配置为从各个所述未标注样本数据对应M个第二类别概率中选取最大的H个第二类别概率,并确定各个所述未标注样本数据的所述H个第二类别概率所对应的H个第二预测标签,其中,所述M和所述H为正整数,且所述H小于所述M;

[0160] 第二训练模块,配置为基于所述已标注数据集、所述未标注数据集、各个所述已标注样本数据的真实标签、各个所述已标注样本数据的K个所述第一预测标签、以及各个所述

未标注样本数据的H个所述第二预测标签,训练所述第二模型。

[0161] 在其他可选的实施例中,所述第二训练模块,包括:

[0162] 第一确定子模块,配置为将所述已标注数据集中的各个所述已标注样本数据输入所述第二模型,得到所述第二模型输出的第三预测标签;

[0163] 第二确定子模块,配置为将所述未标注数据集中的各个所述未标注样本数据输入所述第二模型,得到所述第二模型输出的第四预测标签;

[0164] 第三确定子模块,配置为利用预设损失函数,基于所述真实标签、各个所述已标注样本数据的K个所述第一预测标签、所述第三预测标签、各个所述未标注样本数据的H个所述第二预测标签、以及所述第四预测标签,确定所述第二模型的训练损失;

[0165] 调整子模块,配置为基于所述训练损失,调整所述第二模型的模型参数。

[0166] 在其他可选的实施例中,所述第三确定子模块,还配置为:

[0167] 基于所述真实标签和所述第三预测标签确定所述第二模型在所述已标注数据集上的第一损失;

[0168] 基于各个所述已标注样本数据的K个所述第一预测标签和所述第三预测标签,确定所述第二模型在所述已标注数据集上的第二损失;

[0169] 基于各个所述未标注样本数据的H个所述第二预测标签和所述第四预测标签,确定所述第二模型在所述未标注数据集上的第三损失;

[0170] 基于所述第一损失、所述第二损失和所述第三损失的加权和,确定所述训练损失。

[0171] 在其他可选的实施例中,所述第三确定子模块,还配置为:

[0172] 确定所述第一损失值与第一预设权重的第一乘积;

[0173] 根据所述第一预设权重确定损失权重,并确定所述第二损失值与所述损失权重的第二乘积;

[0174] 确定所述第三损失值与第二预设权重的第三乘积,所述第二预设权重小于或者等于所述第一预设权重;

[0175] 将所述第一乘积、所述第二乘积、以及所述第三乘积相加,得到所述训练损失。

[0176] 在其他可选的实施例中,所述装置300还包括:

[0177] 停止模块,配置为当所述训练损失在设定时长内的数值变化小于设定变化阈值时,停止训练所述第二模型。

[0178] 本发明实施例还提供一种分类装置,应用于电子设备,包括:

[0179] 分类模块,用于将待分类数据输入采用上述任一实施例提供的分类模型训练方法训练得到的第二模型,输出所述待分类数据在X个类别的X个类别概率;

[0180] 标签确定模块,用于按照类别概率从大到小,确定X个类别概率中前预设数量个类别概率对应的类别标签;

[0181] 类别确定模块,用于将所述预设数量个类别标签确定为所述待分类数据的类别标签。

[0182] 关于上述实施例中的装置,其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述,此处将不做详细阐述说明。

[0183] 图4是根据一示例性实施例示出的一种用于分类模型的训练装置1200,或者分类装置1200的框图。例如,装置1200可以是移动电话,计算机,数字广播终端,消息收发设备,

游戏控制台, 平板设备, 医疗设备, 健身设备, 个人数字助理等。

[0184] 参照图4, 装置1200可以包括以下一个或多个组件: 处理组件1202, 存储器1204, 电力组件1206, 多媒体组件1208, 音频组件1210, 输入/输出 (I/O) 的接口1212, 传感器组件1214, 以及通信组件1216。

[0185] 处理组件1202通常控制装置1200的整体操作, 诸如与显示, 电话呼叫, 数据通信, 相机操作和记录操作相关联的操作。处理组件1202可以包括一个或多个处理器1220来执行指令, 以完成上述的方法的全部或部分步骤。此外, 处理组件1202可以包括一个或多个模块, 便于处理组件1202和其他组件之间的交互。例如, 处理组件1202可以包括多媒体模块, 以方便多媒体组件1208和处理组件1202之间的交互。

[0186] 存储器1204被配置为存储各种类型的数据以支持在设备1200的操作。这些数据的示例包括用于在装置1200上操作的任何应用程序或方法的指令, 联系人数据, 电话簿数据, 消息, 图片, 视频等。存储器1204可以由任何类型的易失性或非易失性存储设备或者它们的组合实现, 如静态随机存取存储器 (SRAM), 电可擦除可编程只读存储器 (EEPROM), 可擦除可编程只读存储器 (EPROM), 可编程只读存储器 (PROM), 只读存储器 (ROM), 磁存储器, 快闪存储器, 磁盘或光盘。

[0187] 电力组件1206为装置1200的各种组件提供电力。电力组件1206可以包括电源管理系统, 一个或多个电源, 及其他与为装置1200生成、管理和分配电力相关联的组件。

[0188] 多媒体组件1208包括在所述装置1200和用户之间的提供一个输出接口的屏幕。在一些实施例中, 屏幕可以包括液晶显示器 (LCD) 和触摸面板 (TP)。如果屏幕包括触摸面板, 屏幕可以被实现为触摸屏, 以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界, 而且还检测与所述触摸或滑动操作相关的持续时间和压力。在一些实施例中, 多媒体组件1208包括一个前置摄像头和/或后置摄像头。当设备1200处于操作模式, 如拍摄模式或视频模式时, 前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0189] 音频组件1210被配置为输出和/或输入音频信号。例如, 音频组件1210包括一个麦克风 (MIC), 当装置1200处于操作模式, 如呼叫模式、记录模式和语音识别模式时, 麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器1204或经由通信组件1216发送。在一些实施例中, 音频组件1210还包括一个扬声器, 用于输出音频信号。

[0190] I/O接口1212为处理组件1202和外围接口模块之间提供接口, 上述外围接口模块可以是键盘, 点击轮, 按钮等。这些按钮可包括但不限于: 主页按钮、音量按钮、启动按钮和锁定按钮。

[0191] 传感器组件1214包括一个或多个传感器, 用于为装置1200提供各个方面的状态评估。例如, 传感器组件1214可以检测到设备1200的打开/关闭状态, 组件的相对定位, 例如所述组件为装置1200的显示器和小键盘, 传感器组件1214还可以检测装置1200或装置1200一个组件的位置改变, 用户与装置1200接触的存在或不存在, 装置1200方位或加速/减速和装置1200的温度变化。传感器组件1214可以包括接近传感器, 被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件1214还可以包括光传感器, 如CMOS或CCD图像传感器, 用于在成像应用中使用。在一些实施例中, 该传感器组件1214还可以包括加速度传感

器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0192] 通信组件1216被配置为便于装置1200和其他设备之间有线或无线方式的通信。装置1200可以接入基于通信标准的无线网络,如WiFi,2G或3G,或它们的组合。在一个示例性实施例中,通信组件1216经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,所述通信组件1216还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0193] 在示例性实施例中,装置1200可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述方法。

[0194] 在示例性实施例中,还提供了一种包括指令的非临时性计算机可读存储介质,例如包括指令的存储器1204,上述指令可由装置1200的处理器1220执行以完成上述方法。例如,所述非临时性计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0195] 一种非临时性计算机可读存储介质,当所述存储介质中的指令由移动终端的处理器执行时,使得移动终端能够执行一种分类模型的训练方法,所述方法包括:

[0196] 基于预先训练好的第一模型对已标注数据集进行处理,分别得到所述已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率;

[0197] 分别从各个所述已标注样本数据所对应的N个第一类别概率中选取最大的K个所述第一类别概率,并确定各个所述已标注样本数据的所述K个第一类别概率所对应的K个第一预测标签,其中,所述K和所述N为正整数,且所述K小于所述N;

[0198] 基于所述已标注数据集、各个所述已标注样本数据的真实标签以及各个所述已标注样本数据的K个所述第一预测标签,训练第二模型。

[0199] 或者使得移动终端能够执行一种分类方法,所述方法包括:

[0200] 将待分类数据输入采用上述任一实施例提供的分类模型训练方法训练得到的第二模型,输出所述待分类数据在X个类别的X个类别概率;

[0201] 按照类别概率从大到小,确定X个类别概率中前预设数量个类别概率对应的类别标签;

[0202] 将所述预设数量个类别标签确定为所述待分类数据的类别标签。

[0203] 图5是根据一示例性实施例示出的另一种用于分类模型的训练装置1300或者用于分类的装置1300的框图。例如,装置1300可以被提供为一服务器。参照图5,装置1300包括处理组件1322,其进一步包括一个或多个处理器,以及由存储器1332所代表的存储器资源,用于存储可由处理组件1322的执行的指令,例如应用程序。存储器1332中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外,处理组件1322被配置为执行指令,以执行上述分类模型的训练方法,所述方法包括:

[0204] 基于预先训练好的第一模型对已标注数据集进行处理,分别得到所述已标注数据集中各个已标注样本数据在N个类别的N个第一类别概率;

[0205] 分别从各个所述已标注样本数据所对应的N个第一类别概率中选取最大的K个所述第一类别概率,并确定各个所述已标注样本数据的所述K个第一类别概率所对应的K个第

一预测标签,其中,所述K和所述N为正整数,且所述K小于所述N;

[0206] 基于所述已标注数据集、各个所述已标注样本数据的真实标签以及各个所述已标注样本数据的K个所述第一预测标签,训练第二模型。

[0207] 或执行上述分类方法,所述方法包括:

[0208] 将待分类数据输入采用上述任一实施例提供的分类模型训练方法训练得到的第二模型,输出所述待分类数据在X个类别的X个类别概率;

[0209] 按照类别概率从大到小,确定X个类别概率中前预设数量个类别概率对应的类别标签;

[0210] 将所述预设数量个类别标签确定为所述待分类数据的类别标签。

[0211] 装置1300还可以包括一个电源组件1326被配置为执行装置1300的电源管理,一个有线或无线网络接口1350被配置为将装置1300连接到网络,和一个输入输出(I/O)接口1358。装置1300可以操作基于存储在存储器1332的操作系统,例如Windows Server™,Mac OS X™,Unix™,Linux™,FreeBSD™或类似。

[0212] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本发明的其它实施方案。本申请旨在涵盖本发明的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本发明的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本发明的真正范围和精神由下面的权利要求指出。

[0213] 应当理解的是,本发明并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本发明的范围仅由所附的权利要求来限制。

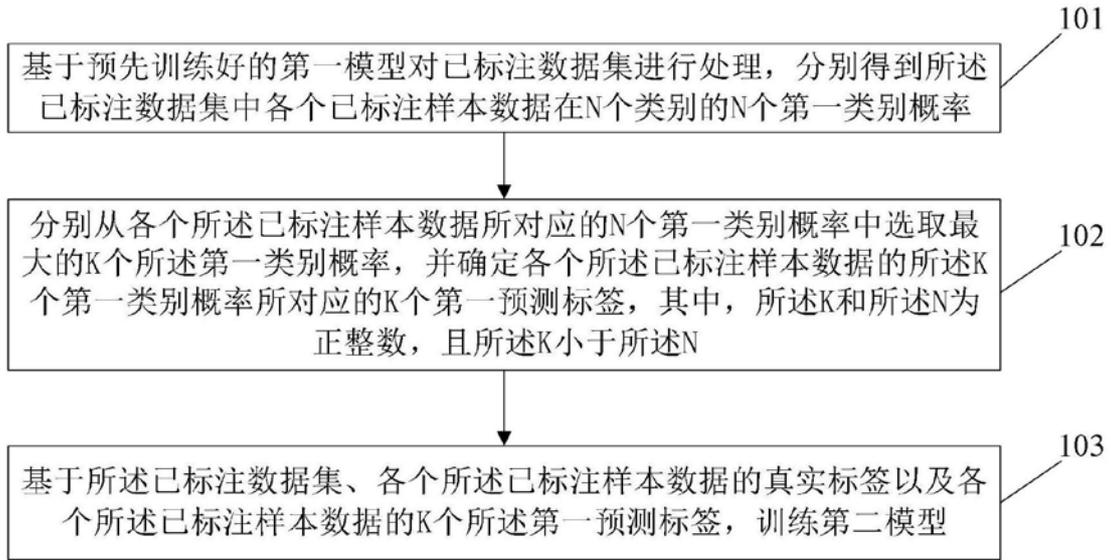


图1

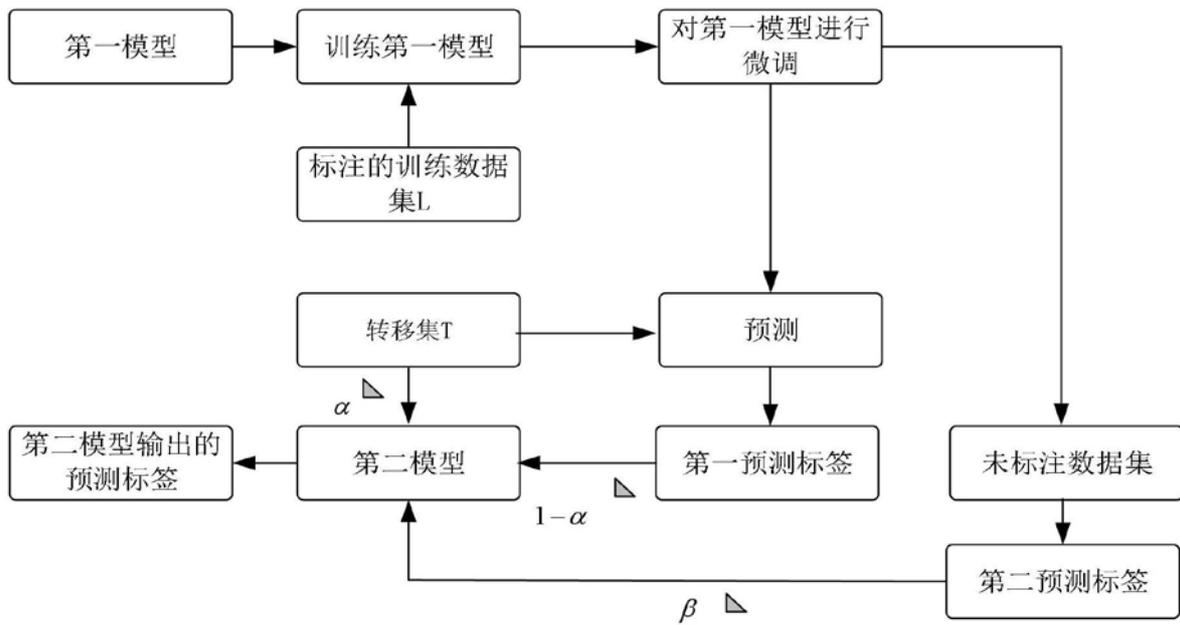


图2

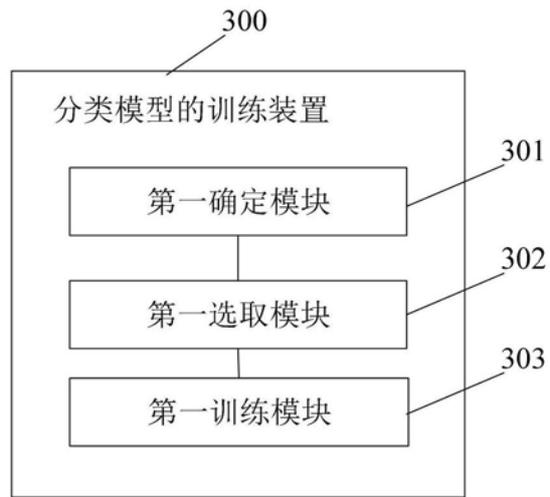


图3

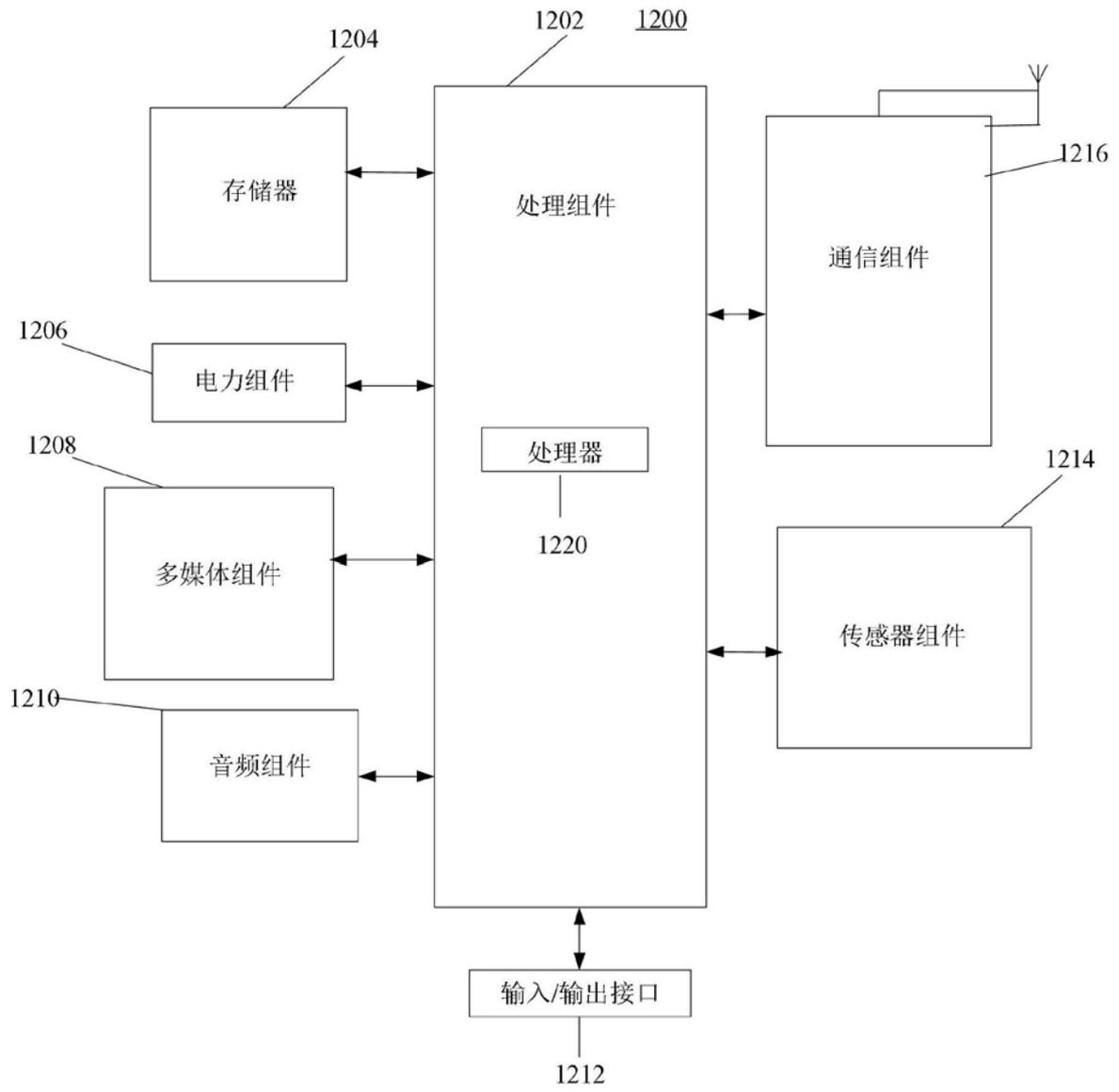


图4

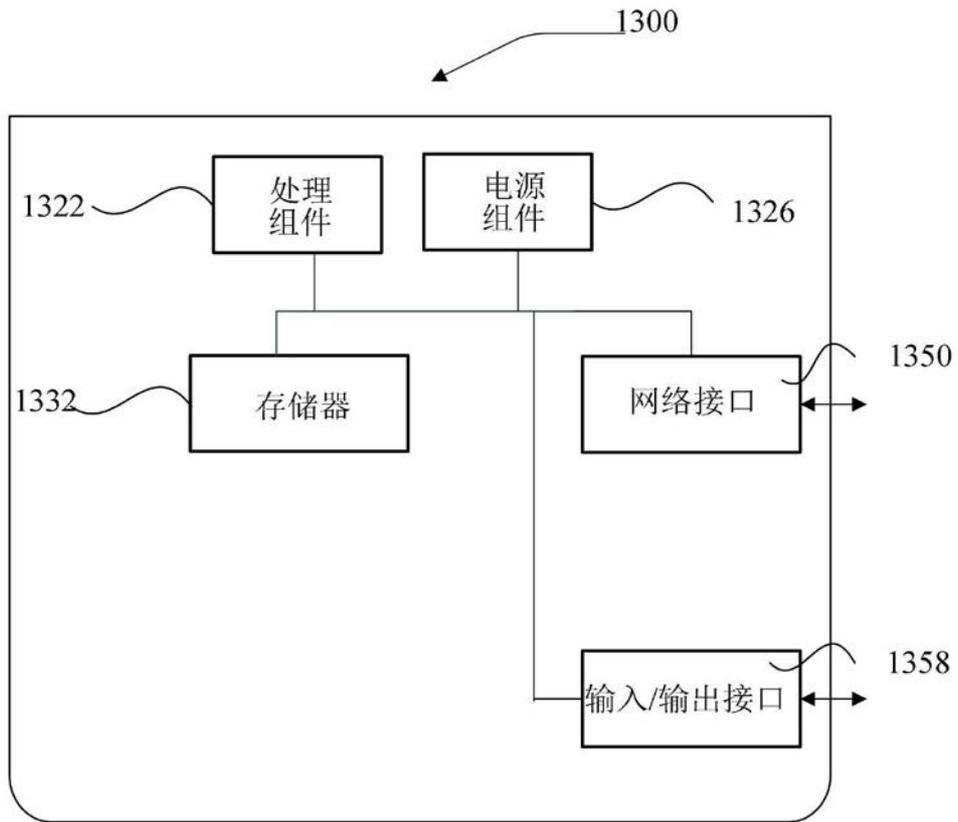


图5