

数据处理方法及装置

本申请要求于 2024 年 8 月 28 日提交中国专利局、申请号为 202411191435.9、发明名称为“数据处理方法及装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本说明书实施例涉及机器学习技术领域，特别涉及数据处理方法及装置。

背景技术

随着计算机和互联网技术的发展，深度学习技术在越来越多的场景中得以应用。深度学习通常集中在易于人类理解的数字媒体文件上，如文本、音频和图像；现有的深度学习模型大多在媒体数据领域运行，例如，语言模型通过 token 预测在文本生成和理解方面表现出色；音频处理模型通过结合文本和语音，实现了高级语音识别和语音到语音的翻译；图像处理模型通过预测下一个像素来生成图像。这些模型在各自领域内取得了显著的成果。但是这恰恰忽略了数字世界中无处不在的原生二进制数据。字节作为所有数字数据、设备和软件的基础，从计算机处理器到日常电子产品中的操作系统，都是依附于字节完成的处理。而现有技术中的模型预测过程忽略了数据最本质的要素，导致模型预测能力以及速度并未得到更全面的开发，因此亟需一种有效的方案以解决上述问题。

发明内容

有鉴于此，本说明书实施例提供了一种数据处理方法。本说明书一个或者多个实施例同时涉及一种数据处理装置，一种计算设备，一种计算机可读存储介质以及一种计算机程序产品，以解决现有技术中存在的技术缺陷。

根据本说明书实施例的第一方面，提供了一种数据处理方法，包括：

获取目标预测任务的多媒体数据，并将所述多媒体数据转换为二进制字节序列；

将所述二进制字节序列输入至多媒体处理模型，其中，所述多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元；

通过所述线性投影单元将所述二进制字节序列转换为字节序列向量，并利用所述序列解码单元对所述字节序列向量进行处理，获得序列特征；

利用所述字节解码单元对所述序列特征进行处理，获得预测字节，并基于所述预测字节构建所述目标预测任务对应的目标任务数据。

可选地，所述将所述多媒体数据转换为二进制字节序列，包括：

按照所述目标预测任务关联的转换策略确定所述多媒体数据对应的码点信息；

根据所述码点信息和所述转换策略中配置的编码规则对所述多媒体数据进行处理，获得所述多媒体数据对应的二进制字节序列。

可选地，所述通过所述线性投影单元将所述二进制字节序列转换为字节序列向量，包括：

将所述二进制字节序列输入至所述线性投影单元，通过所述线性投影单元将所述二进制字节序列中包含的二进制字节转换为字节向量；

根据每个二进制字节对应的字节向量构建所述二进制字节序列对应的初始字节序列向量；

将所述初始字节序列向量转换为一维向量，并将所述一维向量映射至二进制字节空间，获得字节序列向量。

可选地，所述利用所述序列解码单元对所述字节序列向量进行处理，获得序列特征，包括：

将所述字节序列向量输入至所述序列解码单元，通过所述序列解码单元对所述字节序列向量进行自回归预测处理，获得序列特征；

其中，所述利用所述字节解码单元对所述序列特征进行处理，获得预测字节，包括：

将所述序列特征输入至所述字节解码单元，通过所述字节解码单元对所述序列特征进行自回归重建处理，获得预测字节。

可选地，所述基于所述预测字节构建所述目标预测任务对应的目标任务数据，包括：

确定所述转换策略对应的逆转换策略，并根据所述逆转换策略对所述预测字节进行解析；

根据解析结果确定所述预测字节对应的预测码点信息，并根据所述预测码点信息生成所述目标预测任务对应的目标任务数据。

可选地，所述将所述二进制字节序列输入至多媒体处理模型步骤执行之前，还包括：

获取样本数据和所述样本数据对应的样本标签，将所述样本数据转换为样本二进制字节序列，以及将所述样本标签转换为标签字节；

将所述样本二进制字节序列输入至初始多媒体处理模型进行处理，获得训练预测字节；

基于所述训练预测字节和所述标签字节对所述初始多媒体处理模型进行训练，直至获得满足训练停止条件的所述多媒体处理模型。

根据本说明书实施例的第二方面，提供了一种数据处理装置，包括：

获取模块，被配置为获取目标预测任务的多媒体数据，并将所述多媒体数据转换为二进制字节序列；

输入模块，被配置为将所述二进制字节序列输入至多媒体处理模型，其中，所述多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元；

转换模块，被配置为通过所述线性投影单元将所述二进制字节序列转换为字节序列向量，并利用所述序列解码单元对所述字节序列向量进行处理，获得序列特征；

构建模块，被配置为利用所述字节解码单元对所述序列特征进行处理，获得预测字节，并基于所述预测字节构建所述目标预测任务对应的目标任务数据。

根据本说明书实施例的第三方面，提供了一种计算设备，包括：

存储器和处理器；

所述存储器用于存储计算机可执行指令，所述处理器用于执行所述计算机可执行指令，该计算机可执行指令被处理器执行时实现上述数据处理方法的步骤。

根据本说明书实施例的第四方面，提供了一种计算机可读存储介质，其存储有计算机可执行指令，

该指令被处理器执行时实现上述数据处理方法的步骤。

根据本说明书实施例的第五方面，提供了一种计算机程序产品，包括计算机程序或指令，该计算机程序或指令被处理器执行时实现上述数据处理方法的步骤。

本实施例提供的数据处理方法，为了能够提高模型对数字系统的全面理解和模拟，从而提高对多媒体数据的处理能力以及效率，可以在获取到目标预测任务的多媒体数据后，先将多媒体数据转换为二进制字节序列；此时可以将二进制字节序列输入至多媒体处理模型，且多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元；在此基础上，可以通过线性投影单元将二进制字节序列转换为字节序列向量，之后利用序列解码单元对字节序列向量进行处理，获得序列特征；实现在字节序列维度挖掘结构关系。再利用字节解码单元对序列特征进行处理，获得预测字节，实现在字节维度挖掘字节关系，最后即可基于预测字节构建目标预测任务对应的目标任务数据，实现通过模型进行二进制字节维度的处理，可以简化不同数据类型之间的集成过程，提升了模型在跨模态任务中的适用性和性能。

附图说明

- 图 1 是本说明书一个实施例提供的一种数据处理方法的示意图；
- 图 2 是本说明书一个实施例提供的一种数据处理方法的流程图；
- 图 3 是本说明书一个实施例提供的一种数据处理方法的处理过程流程图；
- 图 4 是本说明书一个实施例提供的一种数据处理装置的结构示意图；
- 图 5 是本说明书一个实施例提供的一种计算设备的结构框图。

具体实施方式

在下面的描述中阐述了很多具体细节以便于充分理解本说明书。但是本说明书能够以很多不同于在此描述的其它方式来实施，本领域技术人员可以在不违背本说明书内涵的情况下做类似推广，因此本说明书不受下面公开的具体实施的限制。

在本说明书一个或多个实施例使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本说明书一个或多个实施例。在本说明书一个或多个实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本说明书一个或多个实施例中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

应当理解，尽管在本说明书一个或多个实施例中可能采用术语第一、第二等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本说明书一个或多个实施例范围的情况下，第一也可以被称为第二，类似地，第二也可以被称为第一。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

此外，需要说明的是，本说明书一个或多个实施例所涉及的用户信息（包括但不限于用户设备信息、用户个人信息等）和数据（包括但不限于用于分析的数据、存储的数据、展示的数据等），均为经用户授权或者经过各方充分授权的信息和数据，并且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准，并提供有相应的操作入口，供用户选择授权或者拒绝。

在本说明书中，提供了一种数据处理方法，本说明书同时涉及一种数据处理装置，一种计算设备，一种计算机可读存储介质以及一种计算机程序产品，在下面的实施例中逐一进行详细说明。

实际应用中，深度学习模型主要聚焦于处理易于人类理解的数字媒体数据，如文本、音频和图像。这些模型虽然在各自领域取得了显著成果，但在处理原生二进制数据方面存在明显不足。特别是：文本处理-语言模型通过 token 预测在文本生成和理解方面表现出色，但它们依赖于对文本的 token 化过程，无法直接处理二进制数据。音频处理-音频处理模型虽然能够处理音频数据，但同样局限于特定数据类型，无法有效处理和模拟二进制数据。图像处理-图像处理模型虽然能够预测图像中的下一个像素，但在处理长字节序列和复杂二进制数据时表现不佳。因此亟需一种有效的方案以解决上述问题。

参见图 1 所示的示意图，本实施例提供的数据处理方法，为了能够提高模型对数字系统的全面理解和模拟，从而提高对多媒体数据的处理能力以及效率，可以在获取到目标预测任务的多媒体数据后，先将多媒体数据转换为二进制字节序列；此时可以将二进制字节序列输入至多媒体处理模型，且多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元；在此基础上，可以通过线性投影单元将二进制字节序列转换为字节序列向量，之后利用序列解码单元对字节序列向量进行处理，获得序列特征；实现在字节序列维度挖掘结构关系。再利用字节解码单元对序列特征进行处理，获得预测字节，实现在字节维度挖掘字节关系，最后即可基于预测字节构建目标预测任务对应的目标任务数据，实现通过模型进行二进制字节维度的处理，可以简化不同类型数据之间的集成过程，提升了模型在跨模态任务中的适用性和性能。

参见图 2，图 2 示出了根据本说明书一个实施例提供的一种数据处理方法的流程图，具体包括以下步骤。

步骤 S202，获取目标预测任务的多媒体数据，并将所述多媒体数据转换为二进制字节序列。

本实施例提供的数据处理方法可以应用于任意场景下的模型应用，比如文本处理场景（摘要提取、文本分类、翻译），图像处理场景（图像裁剪、图像生成、图文生成），音频处理场景（消除杂音、音频生成、音频转换）等。本实施例在此不做任何限定。本实施例以数据处理方法在文本处理场景中的应用为例对数据处理方法进行说明，其他场景的描述均可参见本实施例中相同或相应的描述内容，本实施例在此不做任何限定。此外，本实施例提供的数据处理方法还可以应用于跨模态的场景。

具体的，目标预测任务具体是指多媒体数据处理场景下触发的任务，该任务可以是文本处理场景中的文本处理任务，如摘要提取、文本分类、翻译等；也可以是图像处理场景中的图像处理级人物，如图像裁剪、图像生成、图文生成等；还可以是音频处理场景中的音频处理任务，如消除杂音、音频生成、音频转换等。相应的，多媒体数据即为触发目标预测任务时得到的需要输入模型进行处理的数据，如文本、图像、视频、音频等。相应的，二进制字节序列具体是指对多媒体数据进行字节编码后得到的二进制字节组成的序列，其为计算机处理的最基础单位，通过将多媒体数据转换为二进制字节进行模型处理，可以有效的提高预测效率。同时能够对模型进行最有效的开发，实现模型从最基础单位上进行多媒体数据分析。

基于此，为了能够提高模型对数字系统的全面理解和模拟，从而提高对多媒体数据的处理能力以及

效率，可以在获取到目标预测任务的多媒体数据后，先将多媒体数据转换为二进制字节序列；此时可以将二进制字节序列输入至多媒体处理模型，且多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元；在此基础上，可以通过线性投影单元将二进制字节序列转换为字节序列向量，之后利用序列解码单元对字节序列向量进行处理，获得序列特征；实现在字节序列维度挖掘结构关系。再利用字节解码单元对序列特征进行处理，获得预测字节，实现在字节维度挖掘字节关系，最后即可基于预测字节构建目标预测任务对应的目标任务数据。

进一步的，在对多媒体数据进行二进制字节序列构建时，为了能够实现跨模态且更精准的完成模型预测，可以按照预设的转换策略完成。本实施例中，具体实现方式如下：

按照所述目标预测任务关联的转换策略确定所述多媒体数据对应的码点信息；根据所述码点信息和所述转换策略中配置的编码规则对所述多媒体数据进行处理，获得所述多媒体数据对应的二进制字节序列。

具体的，转换策略具体是指用于将关联目标预测任务的多媒体数据按照指定编码方式转换为二进制字节序列的策略，该策略可以确定多媒体数据中每个字符的码点信息，以及编码规则，其中，编码规则用于对多媒体数据进行编码处理，以得到二进制字节组成二进制字节序列。

基于此，在获取到目标预测任务的多媒体数据后，可以先按照目标预测任务关联的转换策略确定多媒体数据对应的码点信息；在此基础上，可以根据码点信息和转换策略中配置的编码规则对多媒体数据进行处理，以实现将多媒体数据转换为多个二进制字节，从而组成多媒体数据对应的二进制字节序列，以便后续输入模型进行预测处理。

举例说明，用户通过客户端提交问题文本“A 古诗的作者是谁”，在接收到该问题文本后，可以先确定每个字单元在在 Unicode 编码中的对应码点。在此基础上，可以采用 UTF-8（8-bit Unicode Transformation Format，万国码）对每个字单元进行编码处理。在 UTF-8 编码中，每个字单元对应的码点会被转换为设定长度字节的序列。通过计算 UTF-8 编码的字节序列，并进行字节分配，可以得到问题文本“A 古诗的作者是谁”对应的十六进制字节序列 {A\xe5\x8f\xa4\xe8\xaf\x97\xe7\x9a\x84\xe4\xbd\x9c\xe8\x80\x85\xe6\x98\xaf\xe8\xb0\x81}，在此基础上，将其转换为二进制字节序列，即可用于后续模型处理。

综上，通过采用转换策略进行多媒体数据的转换，可以实现将多媒体数据按照设定规范转换为统一格式的进制字节序列，从而实现后续模型可以精准的预测。

步骤 S204，将所述二进制字节序列输入至多媒体处理模型，其中，所述多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元。

步骤 S206，通过所述线性投影单元将所述二进制字节序列转换为字节序列向量，并利用所述序列解码单元对所述字节序列向量进行处理，获得序列特征。

步骤 S208，利用所述字节解码单元对所述序列特征进行处理，获得预测字节，并基于所述预测字节构建所述目标预测任务对应的目标任务数据。

具体的，在上述得到多媒体数据对应的二进制字节序列后，可以将二进制字节序列输入至包含线性

投影单元、序列解码单元和字节解码单元的多媒体处理模型。以实现多媒体处理模型可以通过进行二进制字节级别的预测输出预测字节，以构建出目标预测任务对应的目标任务数据，以便后续使用。

其中，线性投影单元具体是指用于将二进制字节序列中包含的二进制字节映射为向量表达的单元。相应的，序列解码单元具体是指构建序列维度特征的解码器，字节解码单元具体是指构建字节维度特征的解码器，用于输出预测字节。相应的，字节序列向量具体是指对二进制字节序列进行向量映射后得到的序列向量。序列特征具体是指从字节序列向量中提取到的序列级别的特征，预测字节具体是指针对目标预测任务对应的多媒体数据进行处理后得到的模型输出结果，且该结果为二进制字节，用于构建目标任务数据。相应的，目标任务数据具体是指对多媒体数据进行模型处理后得到的结果，比如目标预测任务为文本生成任务，则目标任务数据即为文本生成结果，再比如目标预测任务为音频消除杂音任务，则目标任务数据即为消除杂音的音频。

也就是说，在多媒体处理模型对二进制字节序列进行处理时，为了能够挖掘二进制字节级别之间的字节关系，可以先通过线性投影单元将二进制字节序列转换为字节序列向量，在得到二进制字节序列对应的向量表达后，即可进行后续的预测处理，而为了提高模型预测能力，实现对二进制字节进行充分的数据关系分析，可以先利用序列解码单元对字节序列向量进行处理，获得对应序列维度的序列特征；再利用字节解码单元对序列特征进行处理，即可完成预测处理，从而获得预测字节；此后基于预测字节构建目标预测任务对应的目标任务数据，即可便于下游使用。比如将生成的文本反馈给用户，或者基于生成的文本回答用户问题等。本实施例在此不做任何限定。

实际应用中，多媒体处理模型可以采用分层 Transformer 架构，将字节序列分割成片段以提高计算效率。模型可以包括三个主要组件，分别为线性投影单元、序列解码单元和字节解码单元，从而完成字节级别的预测。

进一步的，在将二进制字节序列输入至多媒体处理模型后，为了能够实现在二进制字节维度完成预测，需要先进行字节到向量的转换。本实施例中，具体实现方式如下：

将所述二进制字节序列输入至所述线性投影单元，通过所述线性投影单元将所述二进制字节序列中包含的二进制字节转换为字节向量；根据每个二进制字节对应的字节向量构建所述二进制字节序列对应的初始字节序列向量；将所述初始字节序列向量转换为一维向量，并将所述一维向量映射至二进制字节空间，获得字节序列向量。

具体的，字节向量具体是指将二进制字节中包含的二进制字节进行向量转换后得到的向量表达。相应的，初始字节序列向量具体是指将每个二进制字节对应的字节向量进行拼接后得到的向量表达。二进制字节空间具体是指将一维向量映射到字节空间，实现通过字节空间映射出的字节序列向量体现多媒体数据的数据内在关联关系，从而便于模型进行预测。

基于此，在将二进制字节序列输入至多媒体处理模型后，可以先通过线性投影单元对序列进行向量映射，此时可以通过线性投影单元将二进制字节序列中包含的二进制字节转换为字节向量；在此基础上，可以先根据每个二进制字节对应的字节向量构建二进制字节序列对应的初始字节序列向量；而后可以将初始字节序列向量转换为一维向量，再将一维向量映射至二进制字节空间，即可得到对应多媒体资源的

字节序列向量，以便后续进行预测使用。

实际应用中，线性投影单元在对二进制字节序列进行映射处理时，可以将每个二进制字节序列作为一个矩阵，每个字节可以通过独热编码表示为 257 维向量，包括所有 256 个字节值和一个<eop>（序列结尾）标记，此后将序列转换为一维向量，再映射到二进制字节空间，即可得到字节序列向量。

5 综上，通过在模型中嵌入线性投影单元，用于将二进制字节序列映射至二进制字节空间，可以实现得到字节序列向量的情况下，依旧能够表征多媒体数据中数据关系，进而使得模型可以精准进行预测。

更进一步的，在得到字节序列向量后，即可结合模型中的序列解码单元和字节解码单元完成字节预测。本实施例中，具体实现方式如下：

10 将所述字节序列向量输入至所述序列解码单元，通过所述序列解码单元对所述字节序列向量进行自回归预测处理，获得序列特征；将所述序列特征输入至所述字节解码单元，通过所述字节解码单元对所述序列特征进行自回归重建处理，获得预测字节。

基于此，在得到线性投影单元映射后的字节序列向量后，可以将字节序列向量输入至序列解码单元，通过序列解码单元对字节序列向量进行自回归预测处理，实现在序列级别进行结构预测，从而获得序列特征；此后再将序列特征输入至字节解码单元，通过字节解码单元对序列特征进行自回归重建处理，进
15 而实现在字节维度进行二进制字节预测，从而获得预测字节，以便后续再将预测字节转换为目标任务数据，以用于下游任务处理。

实际应用中，序列解码单元可以实现自回归地预测后续序列的特征，学习数据的结构，而字节解码单元可以根据片段特征自回归地重建片段内的字节序列，从而得到预测字符便于后续使用。

20 综上，通过采用序列解码器从序列维度进行特征预测，以及再根据预测结果从字节维度进行预测字节生成，从而可以保证模型从二进制维度输出更加精准预测结果。

在上述得到预测字节后，考虑到最后得到的结果需要反馈给用户或者业务端使用，因此还需要将预测字节转换为目标任务数据。本实施例中，具体实现方式如下：

25 确定所述转换策略对应的逆转换策略，并根据所述逆转换策略对所述预测字节进行解析；根据解析结果确定所述预测字节对应的预测码点信息，并根据所述预测码点信息生成所述目标预测任务对应的目标任务数据。

具体的，逆转换策略具体是指将预测字节转换为目标任务数据的策略，即通过逆转换策略可以将预测字节转换为文本、图像或者音频等。相应的，预测码点信息具体是指预测字节对应的码点信息。

30 基于此，在模型输出预测字节后，可以先确定转换策略对应的逆转换策略，在此基础上，可以根据逆转换策略对预测字节进行解析；实现根据解析结果确定预测字节对应的预测码点信息，此后即可根据预测码点信息生成目标预测任务对应的目标任务数据，以便下游使用。

沿用上例，在得到问题文本“A 古诗的作者是谁”对应的二进制字节序列后，可以将二进制字节序列输入至多媒体处理模型；首先，通过多媒体处理模型中的线性投影单元对二进制字节序列进行映射处理，以得到字节序列向量 ZLc；其次，可以将字节序列向量 ZLc 输入至多媒体处理模型中的序列解码器，通过序列解码器自回归地预测序列特征，从而得到序列向量 Lc。最后，再将序列向量 Lc 输入至多媒体处

理模型中的字节解码器，通过字节解码器自回归地重建序列内的字节序列，从而可以生成预测字节。在得到预测字节后，再按照上述编码的逆过程对预测字节进行转换，即可得到问题文本“A 古诗的作者是谁”对应的答案文本“甲”，并将其反馈给用户即可。

5 综上所述，通过包含线性投影单元、序列解码单元和字节解码单元的多媒体处理模型对多媒体数据进行处理，可以实现模型从而字节维度给出预测结果，从而有效的提高模型预测精度和处理效率。同时可以占用更少的计算资源。

此外，为了能够使得多媒体处理模型可以从字节维度完成预测，需要在训练阶段使得模型具备该方面的能力。本实施例中，模型的训练包括：

10 获取样本数据和所述样本数据对应的样本标签，将所述样本数据转换为样本二进制字节序列，以及将所述样本标签转换为标签字节；将所述样本二进制字节序列输入至初始多媒体处理模型进行处理，获得训练预测字节；基于所述训练预测字节和所述标签字节对所述初始多媒体处理模型进行训练，直至获得满足训练停止条件的所述多媒体处理模型。

15 具体的，样本数据和样本标签具体是指训练多媒体处理模型的样本，通过将其转换为二进制序列，可以实现模型学习二进制字节的预测能力，从而提高模型对字节级别数据的挖掘能力。相应的，标签字节即为样本标签对应的二进制字节。训练预测字节具体是指模型输出的预测字节。其中，训练停止条件具体是指停止训练多媒体处理模型的条件，包括但不限于损失值比较条件、验证条件或者迭代次数条件，本实施例在此不做任何限定。

20 基于此，在模型训练阶段，可以先获取样本数据和样本数据对应的样本标签，此后可以将样本数据转换为样本二进制字节序列，以及将样本标签转换为标签字节；在此基础上，即可进行模型的训练，通过将样本二进制字节序列输入至初始多媒体处理模型进行处理，即可获得训练预测字节；在此基础上，结合训练预测字节和标签字节对初始多媒体处理模型进行训练，直至获得满足训练停止条件的多媒体处理模型即可。

25 本实施例提供的数据处理方法，为了能够提高模型对数字系统的全面理解和模拟，从而提高对多媒体数据的处理能力以及效率，可以在获取到目标预测任务的多媒体数据后，先将多媒体数据转换为二进制字节序列；此时可以将二进制字节序列输入至多媒体处理模型，且多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元；在此基础上，可以通过线性投影单元将二进制字节序列转换为字节序列向量，之后利用序列解码单元对字节序列向量进行处理，获得序列特征；实现在字节序列维度挖掘结构关系。再利用字节解码单元对序列特征进行处理，获得预测字节，实现在字节维度挖掘字节关系，最后即可基于预测字节构建目标预测任务对应的目标任务数据，实现通过模型进行二进制字节维度的处理，30 可以简化不同数据类型之间的集成过程，提升了模型在跨模态任务中的适用性和性能。

下述结合附图 3，以本说明书提供的数据处理方法在图像数据处理场景中的应用为例，对所述数据处理方法进行进一步说明。其中，图 3 示出了本说明书一个实施例提供的一种数据处理方法的处理流程图，具体包括以下步骤。

步骤 S302，获取目标预测任务的多媒体数据，按照目标预测任务关联的转换策略确定多媒体数据对

应的码点信息。

步骤 S304, 根据码点信息和转换策略中配置的编码规则对多媒体数据进行处理, 获得多媒体数据对应的二进制字节序列。

5 步骤 S306, 将二进制字节序列输入至多媒体处理模型, 其中, 多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元。

步骤 S308, 通过线性投影单元将二进制字节序列中包含的二进制字节转换为字节向量。

步骤 S310, 根据每个二进制字节对应的字节向量构建二进制字节序列对应的初始字节序列向量。

步骤 S312, 将初始字节序列向量转换为一维向量, 并将一维向量映射至二进制字节空间, 获得字节序列向量。

10 步骤 S314, 通过序列解码单元对字节序列向量进行自回归预测处理, 获得序列特征。

步骤 S316, 将序列特征输入至字节解码单元, 通过字节解码单元对序列特征进行自回归重建处理, 获得预测字节。

步骤 S318, 确定转换策略对应的逆转换策略, 并根据逆转换策略对预测字节进行解析。

15 步骤 S320, 根据解析结果确定预测字节对应的预测码点信息, 并根据预测码点信息生成目标预测任务对应的目标任务数据。

本实施例提供的数据处理方法, 为了能够提高模型对数字系统的全面理解和模拟, 从而提高对多媒体数据的处理能力以及效率, 可以在获取到目标预测任务的多媒体数据后, 先将多媒体数据转换为二进制字节序列; 此时可以将二进制字节序列输入至多媒体处理模型, 且多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元; 在此基础上, 可以通过线性投影单元将二进制字节序列转换为字节序列向量, 之后利用序列解码单元对字节序列向量进行处理, 获得序列特征; 实现在字节序列维度挖掘结构关系。再利用字节解码单元对序列特征进行处理, 获得预测字节, 实现在字节维度挖掘字节关系, 最后即可基于预测字节构建目标预测任务对应的目标任务数据, 实现通过模型进行二进制字节维度的处理, 可以简化不同数据类型之间的集成过程, 提升了模型在跨模态任务中的适用性和性能。

25 与上述方法实施例相对应, 本说明书还提供了数据处理装置实施例, 图 4 示出了本说明书一个实施例提供的一种数据处理装置的结构示意图。如图 4 所示, 该装置包括:

获取模块 402, 被配置为获取目标预测任务的多媒体数据, 并将所述多媒体数据转换为二进制字节序列;

输入模块 404, 被配置为将所述二进制字节序列输入至多媒体处理模型, 其中, 所述多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元;

30 转换模块 406, 被配置为通过所述线性投影单元将所述二进制字节序列转换为字节序列向量, 并利用所述序列解码单元对所述字节序列向量进行处理, 获得序列特征;

构建模块 408, 被配置为利用所述字节解码单元对所述序列特征进行处理, 获得预测字节, 并基于所述预测字节构建所述目标预测任务对应的目标任务数据。

一个可选的实施例中, 所述获取模块 402 进一步被配置为:

按照所述目标预测任务关联的转换策略确定所述多媒体数据对应的码点信息；根据所述码点信息和所述转换策略中配置的编码规则对所述多媒体数据进行处理，获得所述多媒体数据对应的二进制字节序列。

一个可选的实施例中，所述转换模块 406 进一步被配置为：

5 将所述二进制字节序列输入至所述线性投影单元，通过所述线性投影单元将所述二进制字节序列中包含的二进制字节转换为字节向量；根据每个二进制字节对应的字节向量构建所述二进制字节序列对应的初始字节序列向量；将所述初始字节序列向量转换为一维向量，并将所述一维向量映射至二进制字节空间，获得字节序列向量。

一个可选的实施例中，所述转换模块 406 进一步被配置为：

10 将所述字节序列向量输入至所述序列解码单元，通过所述序列解码单元对所述字节序列向量进行自回归预测处理，获得序列特征；

其中，所述构建模块 408 进一步被配置为：

将所述序列特征输入至所述字节解码单元，通过所述字节解码单元对所述序列特征进行自回归重建处理，获得预测字节。

15 一个可选的实施例中，所述构建模块 408 进一步被配置为：

确定所述转换策略对应的逆转换策略，并根据所述逆转换策略对所述预测字节进行解析；根据解析结果确定所述预测字节对应的预测码点信息，并根据所述预测码点信息生成所述目标预测任务对应的目标任务数据。

一个可选的实施例中，所述装置还包括：

20 训练模块，被配置为获取样本数据和所述样本数据对应的样本标签，将所述样本数据转换为样本二进制字节序列，以及将所述样本标签转换为标签字节；将所述样本二进制字节序列输入至初始多媒体处理模型进行处理，获得训练预测字节；基于所述训练预测字节和所述标签字节对所述初始多媒体处理模型进行训练，直至获得满足训练停止条件的所述多媒体处理模型。

本实施例提供的数据处理装置，为了能够提高模型对数字系统的全面理解和模拟，从而提高对多媒体数据的处理能力以及效率，可以在获取到目标预测任务的多媒体数据后，先将多媒体数据转换为二进制字节序列；此时可以将二进制字节序列输入至多媒体处理模型，且多媒体处理模型包含线性投影单元、序列解码单元和字节解码单元；在此基础上，可以通过线性投影单元将二进制字节序列转换为字节序列向量，之后利用序列解码单元对字节序列向量进行处理，获得序列特征；实现在字节序列维度挖掘结构关系。再利用字节解码单元对序列特征进行处理，获得预测字节，实现在字节维度挖掘字节关系，最后
25 即可基于预测字节构建目标预测任务对应的目标任务数据，实现通过模型进行二进制字节维度的处理，可以简化不同数据类型之间的集成过程，提升了模型在跨模态任务中的适用性和性能。

上述为本实施例的一种数据处理装置的示意性方案。需要说明的是，该数据处理装置的技术方案与上述的数据处理方法的技术方案属于同一构思，数据处理装置的技术方案未详细描述的细节内容，均可以参见上述数据处理方法的技术方案的描述。

图 5 示出了根据本说明书一个实施例提供的一种计算设备 500 的结构框图。该计算设备 500 的部件包括但不限于存储器 510 和处理器 520。处理器 520 与存储器 510 通过总线 530 相连接，数据库 550 用于保存数据。

5 计算设备 500 还包括接入设备 540，接入设备 540 使得计算设备 500 能够经由一个或多个网络 560 通信。这些网络的示例包括公用交换电话网（PSTN, Public Switched Telephone Network）、局域网（LAN, Local Area Network）、广域网（WAN, Wide Area Network）、个域网（PAN, Personal Area Network）或诸如因特网的通信网络的组合。接入设备 540 可以包括有线或无线的任何类型的网络接口（例如，网络接口卡（NIC, network interface controller））中的一个或多个，诸如 IEEE802.11 无线局域网（WLAN, Wireless Local Area Network）无线接口、全球微波互联接入（Wi-MAX, Worldwide Interoperability for
10 Microwave Access）接口、以太网接口、通用串行总线（USB, Universal Serial Bus）接口、蜂窝网络接口、蓝牙接口、近场通信（NFC, Near Field Communication）。

在本说明书的一个实施例中，计算设备 500 的上述部件以及图 5 中未示出的其他部件也可以彼此相连接，例如通过总线。应当理解，图 5 所示的计算设备结构框图仅仅是出于示例的目的，而不是对本说明书范围的限制。本领域技术人员可以根据需要，增添或替换其他部件。

15 计算设备 500 可以是任何类型的静止或移动计算设备，包括移动计算机或移动计算设备（例如，平板电脑计算机、个人数字助理、膝上型计算机、笔记本计算机、上网本等）、移动电话（例如，智能手机）、可佩戴的计算设备（例如，智能手表、智能眼镜等）或其他类型的移动设备，或者诸如台式计算机或
个人计算机（PC, Personal Computer）的静止计算设备。计算设备 500 还可以是移动式或静止式的服务器。

其中，处理器 520 用于执行如下计算机可执行指令，该计算机可执行指令被处理器执行时实现
20 上述数据处理方法的步骤。

上述为本实施例的一种计算设备的示意性方案。需要说明的是，该计算设备的技术方案与上述的数据处理方法的技术方案属于同一构思，计算设备的技术方案未详细描述的细节内容，均可以参见上述数据处理方法的技术方案的描述。

本说明书一实施例还提供一种计算机可读存储介质，其存储有计算机可执行指令，该计算机可执行指令被处理器执行时实现上述数据处理方法的步骤。
25

上述为本实施例的一种计算机可读存储介质的示意性方案。需要说明的是，该存储介质的技术方案与上述的数据处理方法的技术方案属于同一构思，存储介质的技术方案未详细描述的细节内容，均可以参见上述数据处理方法的技术方案的描述。

本说明书一实施例还提供一种计算机程序产品，包括计算机程序或指令，该计算机程序或指令被处理器执行时实现上述数据处理方法的步骤。
30

上述为本实施例的一种计算机程序产品的示意性方案。需要说明的是，该计算机程序产品的技术方案与上述的数据处理方法的技术方案属于同一构思，计算机程序产品的技术方案未详细描述的细节内容，均可以参见上述数据处理方法的技术方案的描述。

上述对本说明书特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下，

在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外，在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中，多任务处理和并行处理也是可以的或者可能是有利的。

所述计算机指令包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可
5 执行文件或某些中间形式等。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U 盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM, Read-Only Memory)、随机存取存储器(RAM, Random Access Memory)、电载波信号、电信信号以及软件分发介质等。需要说明的是，所述计算机可读介质包含的内容可以根据专利实践的要求进行适当的增减，例如在某些地区，根据专利实践，计算机可读介质不包括电载波信号和电信信号。

10 需要说明的是，对于前述的各方法实施例，为了简便描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本说明书实施例并不受所描述的动作顺序的限制，因为依据本说明书实施例，某些步骤可以采用其它顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定都是本说明书实施例所必须的。

在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述的部分，可以参见其它
15 实施例的相关描述。

以上公开的本说明书优选实施例只是用于帮助阐述本说明书。可选实施例并没有详尽叙述所有的细节，也不限制该发明仅为所述的具体实施方式。显然，根据本说明书实施例的内容，可作很多的修改和变化。本说明书选取并具体描述这些实施例，是为了更好地解释本说明书实施例的原理和实际应用，从而使所属技术领域技术人员能很好地理解和利用本说明书。

权 利 要 求

- 1.一种数据处理方法，其特征在于，包括：
获取目标预测任务的多媒体数据，并将所述多媒体数据转换为二进制字节序列；
将所述二进制字节序列输入至多媒体处理模型，其中，所述多媒体处理模型包含线性投影单元、序
5 列解码单元和字节解码单元；
通过所述线性投影单元将所述二进制字节序列转换为字节序列向量，并利用所述序列解码单元对所
述字节序列向量进行处理，获得序列特征；
利用所述字节解码单元对所述序列特征进行处理，获得预测字节，并基于所述预测字节构建所述目
标预测任务对应的目标任务数据。
- 10 2.根据权利要求1所述的数据处理方法，其特征在于，所述将所述多媒体数据转换为二进制字节序
列，包括：
按照所述目标预测任务关联的转换策略确定所述多媒体数据对应的码点信息；
根据所述码点信息和所述转换策略中配置的编码规则对所述多媒体数据进行处理，获得所述多媒体
数据对应的二进制字节序列。
- 15 3.根据权利要求1所述的数据处理方法，其特征在于，所述通过所述线性投影单元将所述二进制字
节序列转换为字节序列向量，包括：
将所述二进制字节序列输入至所述线性投影单元，通过所述线性投影单元将所述二进制字节序列中
包含的二进制字节转换为字节向量；
根据每个二进制字节对应的字节向量构建所述二进制字节序列对应的初始字节序列向量；
20 将所述初始字节序列向量转换为一维向量，并将所述一维向量映射至二进制字节空间，获得字节序
列向量。
- 4.根据权利要求1所述的数据处理方法，其特征在于，所述利用所述序列解码单元对所述字节序列
向量进行处理，获得序列特征，包括：
将所述字节序列向量输入至所述序列解码单元，通过所述序列解码单元对所述字节序列向量进行自
25 回归预测处理，获得序列特征；
其中，所述利用所述字节解码单元对所述序列特征进行处理，获得预测字节，包括：
将所述序列特征输入至所述字节解码单元，通过所述字节解码单元对所述序列特征进行自回归重建
处理，获得预测字节。
- 5.根据权利要求2所述的数据处理方法，其特征在于，所述基于所述预测字节构建所述目标预测任
30 务对应的目标任务数据，包括：
确定所述转换策略对应的逆转换策略，并根据所述逆转换策略对所述预测字节进行解析；
根据解析结果确定所述预测字节对应的预测码点信息，并根据所述预测码点信息生成所述目标预测
任务对应的目标任务数据。
- 6.根据权利要求1至5任一项所述的数据处理方法，其特征在于，所述将所述二进制字节序列输入

至多媒体处理模型步骤执行之前，还包括：

获取样本数据和所述样本数据对应的样本标签，将所述样本数据转换为样本二进制字节序列，以及将所述样本标签转换为标签字节；

将所述样本二进制字节序列输入至初始多媒体处理模型进行处理，获得训练预测字节；

5 基于所述训练预测字节和所述标签字节对所述初始多媒体处理模型进行训练，直至获得满足训练停止条件的所述多媒体处理模型。

7.一种数据处理装置，其特征在于，包括：

获取模块，被配置为获取目标预测任务的多媒体数据，并将所述多媒体数据转换为二进制字节序列；

输入模块，被配置为将所述二进制字节序列输入至多媒体处理模型，其中，所述多媒体处理模型包

10 含线性投影单元、序列解码单元和字节解码单元；

转换模块，被配置为通过所述线性投影单元将所述二进制字节序列转换为字节序列向量，并利用所述序列解码单元对所述字节序列向量进行处理，获得序列特征；

构建模块，被配置为利用所述字节解码单元对所述序列特征进行处理，获得预测字节，并基于所述预测字节构建所述目标预测任务对应的目标任务数据。

15 8.一种计算设备，其特征在于，包括：

存储器和处理器；

所述存储器用于存储计算机可执行指令，所述处理器用于执行所述计算机可执行指令，该计算机可执行指令被处理器执行时实现权利要求 1 至 6 任意一项所述方法的步骤。

9.一种计算机可读存储介质，其特征在于，其存储有计算机可执行指令，该计算机可执行指令被处

20 理器执行时实现权利要求 1 至 6 任意一项所述方法的步骤。

10.一种计算机程序产品，其特征在于，包括计算机程序或指令，该计算机程序或指令被处理器执行时实现权利要求 1 至 6 任意一项所述方法的步骤。