

第 5 章



飞桨 EasyDL 结构化数据 AI 应用建模

主要内容：

- EasyDL 中表格数据预测的建模、训练、校验、部署及测试的基本方法；
- EasyDL 中时序数据预测的建模、训练、校验、部署及测试的基本方法。



微课视频

5.1 表格数据预测建模

本节在简要介绍 EasyDL 结构化数据中提供的表格数据预测模型有关基本概念、适用场景和 AI 应用建模一般流程的基础上,按照创建模型、数据准备、模型训练、模型校验、模型发布、接口测试六大步骤,阐述表格数据预测模型建模及其应用测试的基本方法。

5.1.1 基本流程

表格数据预测指的是通过机器学习技术从结构化(表格化)数据中发现潜在规律,从而创建机器学习模型,并基于机器学习模型处理新的数据,为业务应用生成预测结果。根据预测数据的不同,EasyDL 提供的表格数据预测模型可以分为如下 3 种类型。

(1) 回归模型。表格数据中的目标列是连续的实数范围。如在销量预测场景中,销量值可能是某个取值范围内的任意值,解决该问题的模型属于回归模型。

(2) 二分类模型。表格数据中的目标列是离散值,且只有两种可能的取值。如在精准营销场景中预测一个用户是否为潜在购买用户,其目标列仅存在“True”和“False”两种取值,解决该问题的模型属于二分类模型。

(3) 多分类模型。表格数据中的目标列是离散值,并具有有限的可能取值。如在用户分类场景中,根据用户的历史消费数据,将用户划分到不同消费偏好的类别中,解决该问题的模型属于多分类模型。

表格数据预测建模的基本流程如图 5-1 所示。



图 5-1 表格数据预测建模的基本流程

表格数据预测模型使用场景丰富多彩,既可用于工业领域故障诊断、状态预测,还可用于商业领域精准营销、客户管理等。

5.1.2 创建模型

这个阶段的主要任务是在 EasyDL 中,按照操作向导完成表格数据预测模型创建。打开 EasyDL 平台主页,如图 5-2 所示。



图 5-2 EasyDL 平台主页

打开平台后,先单击【AI 主站】按钮,然后单击页面中的【控制台】,进入如图 5-3 所示的百度 AI 登录界面。

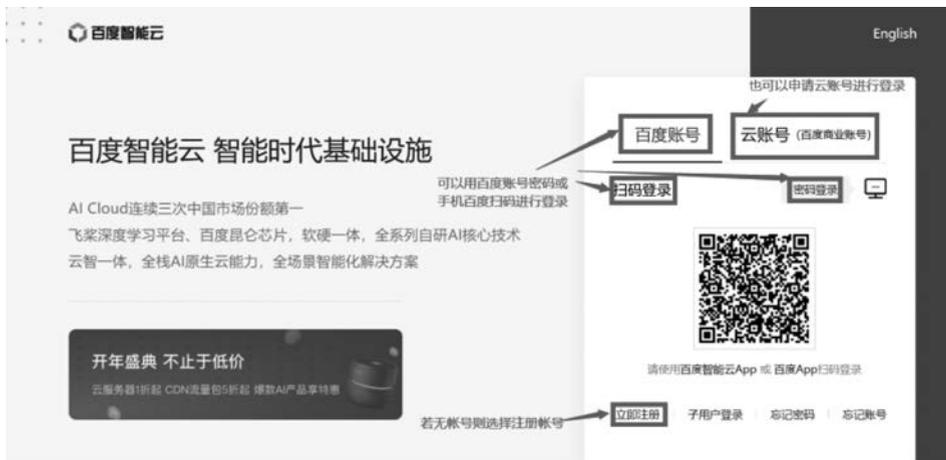


图 5-3 百度 AI 登录界面

单击【百度账号】或【云账号】登录。如无账号,选择注册一个账号,然后登录,进入 EasyDL 服务平台,如图 5-4 所示。

单击图 5-4 中【立即使用】按钮,显示选择模型类型操作界面,如图 5-5 所示。

选择【表格数据预测】,进入表格数据预测模型中心,单击导航栏【模型中心】→【我的模型】,进入如图 5-6 所示的表格数据预测模型操作界面。



图 5-4 EasyDL 服务平台



图 5-5 选择模型类型操作界面



图 5-6 表格数据预测模型操作界面

单击【创建模型】按钮,进入表格数据预测建模信息配置页面,如图 5-7 所示。



图 5-7 表格数据预测建模信息配置页面

完成创建后,模型列表中显示表格数据预测模型创建结果,如图 5-8 所示。



图 5-8 表格数据预测模型创建结果

5.1.3 数据准备

这个阶段的主要任务是提供表格数据预测模型训练所需要的数据集。单击左侧导航栏【模型中心】→【数据总览】,在右侧显示的【我的数据总览】中,单击【创建数据集】,启动训练用数据集创建,如图 5-9 所示。

在弹出的操作界面中填写【数据集名称】,完成表格数据预测模型训练数据集命名,如图 5-10 所示。

完成数据集准备后,单击 EasyDL 控制台左侧导航栏【模型中心】→【数据总览】,即可查看用于表格数据预测模型训练的数据集(此时数据集中并无数据),如图 5-11 所示。

在【我的数据总览】中选择刚创建的数据集,单击操作链接【导入】,进入如图 5-12 所示的配置导入数据集界面。

在【导入方式】列表框中选择“上传 CSV 文件”或“上传压缩包”,弹出如图 5-13 所示的上传数据提示信息对话框。



图 5-9 启动训练用数据集创建



图 5-10 表格数据预测模型训练数据集命名



图 5-11 查看用于表格数据预测模型训练的数据集



图 5-12 配置导入数据集界面



图 5-13 上传数据提示信息对话框

单击【已阅读并上传】按钮,在弹出的文件打开对话框中选择所需 CSV 文件或压缩包。待文件上传完毕,单击【确认并返回】,进入【我的数据总览】。完成数据上传大约需要 10 分钟。上传完成状态的数据集信息如图 5-14 所示。



图 5-14 上传完成状态的数据集信息

需要注意的是,训练数据的质量决定了训练所得模型效果可达到的上限。数据上传后无法修改其内容。如果在导入训练数据后需要对其进行更改,必须重新导入。训练用数据文件目前仅支持 CSV 格式,数据文件内容至少包含两列,其中一列为要预测的值即目标列,其他列为属性列。EasyDL 规定数据总列数不得超过 1000 列,而且数据集的总行数不能超过 1000 万行。上传数据时,一次仅能上传一个文件,可以是一个 CSV 文件或由多个 CSV 文件压缩成的 ZIP 包(ZIP 包中的多个 CSV 文件必须使用相同的编码格式,都包含列名或都不包含列名,且列的顺序必须保持一致),单个上传文件大小不能超过 5GB,一个数据集包含的总文件大小不能超过 20GB。

注: EasyDL 技术迭代更新可能会导致上述实现过程的部分细节内容有所变化,读者可以参考最新使用文档进行具体操作的调整。

5.1.4 模型训练

这一阶段的主要任务是使用准备好的训练数据集对前期创建的模型进行训练。单击 EasyDL 控制台左侧导航栏中【训练模型】按钮,进入训练模型流程。主要操作包括选择模型、选择数据集、选择目标列、选择算法模型、选择部署方式等。表格数据预测模型训练参数设置如图 5-15 所示。

填写完毕模型训练信息后,单击【开始训练】按钮,启动模型训练。可以在【我的模型】中查看模型训练的进度,还可以勾选短信通知复选框,模型训练完成后会第一时间以短信的形式通知开发者模型训练相关信息。表格数据预测模型训练状态信息如图 5-16 所示。

对于“训练完成”的模型,可以查看其评估结果。单击【我的模型】按钮,进入模型列表页



图 5-15 表格数据预测模型训练参数设置



图 5-16 表格数据预测模型训练状态信息

面。单击待查看模型的【历史版本】，进入模型版本列表页面。单击待查看模型版本所在行的【完整评估结果】按钮，查看表格数据预测模型训练结果，如图 5-17 所示。



图 5-17 查看表格数据预测模型训练结果

5.1.5 模型校验

这一阶段的主要任务是在线检验完成训练的表格数据预测模型。单击【校验模型】按钮,进入模型校验页面。选择要校验的模型及其版本。单击【启动模型校验服务】按钮,启动表格数据预测模型校验,如图 5-18 所示。



图 5-18 启动表格数据预测模型校验

校验数据支持表单和 JSON 格式两种输入方式,可在两者之间任意切换。用户可以直接使用预置的数据进行预测,也可以修改后再进行预测。单击【预测】按钮,可以在右侧结果面板中查看训练所得表格数据预测模型的预测结果,如图 5-19 所示。



图 5-19 查看训练所得表格数据预测模型的预测结果

5.1.6 模型发布

这一阶段的主要任务是部署训练效果满意的表格数据预测模型。训练完成后,可将训

练效果满意的表格数据预测模型部署在公有云服务器、通用小型设备、本地服务器,也可以采用百度 AI 软硬一体方案。初学阶段,一般选择“公有云部署”方式。

训练完毕后可以在左侧导航栏中找到【发布模型】,依次进行“选择模型(拟部署模型)→选择部署方式(公有云部署)→选择版本(拟发布版本)→自定义服务名称→设置接口地址后缀→提交申请”等操作。设置模型发布相关信息操作界面如图 5-20 所示。



图 5-20 设置模型发布相关信息操作界面

当模型信息中显示“服务详情”时,表示已经发布成功。发布成功的模型信息如图 5-21 所示。



图 5-21 发布成功的模型信息

单击【服务详情】链接,显示发布模型的 API 接口地址相关信息,如图 5-22 所示。

单击【查看 API 文档】,可以进入表格数据预测模型 API 使用方法说明文档网页,如图 5-23 所示。

开发者可通过该文档了解表格数据预测模型 API 使用方法。



图 5-22 发布模型的 API 接口地址相关信息



图 5-23 表格数据预测模型 API 使用方法说明文档

5.1.7 接口测试

这一阶段的主要任务是测试云端部署的表格数据预测模型访问接口。模型调用测试前首先需要创建一个 EasyDL 结构化数据对应的模型应用。

进入百度智能云控制台(需要使用自己的账号和密码登录百度智能云)。在产品服务栏中选中【EasyDL 定制化训练平台】→【EasyDL 结构化数据】,进入 EasyDL 结构化数据应用中心,如图 5-24 所示。

EasyDL 结构化数据应用中心内单击导航栏【公有云部署】下的【应用列表】,在应用管理界面单击【创建应用】,填写【应用名称】【接口选择】【应用归属】【应用描述】等相关信息,完成表格数据预测模型应用配置,如图 5-25 所示。

单击【立即创建】按钮,完成表格数据预测模型应用的创建,应用列表中显示应用名称及参数信息,如图 5-26 所示。应用参数中 AppID、API Key、Secret Key 的取值至关重要,是后续 API 接口调用时鉴权的依据。



图 5-24 EasyDL 结构化数据应用中心

创建新应用

* 应用名称:

* 接口选择: 勾选以下接口, 使此应用可以请求已勾选的接口服务, 注意EasyDL结构化数据服务已默认勾选并不可取消。

- EasyDL
- 语音技术
- 文字识别
- 人脸识别
- 自然语言处理
- 内容审核
- UNIT
- 知识图谱
- 图像识别
- 智能呼叫中心
- 图像搜索
- 人体分析
- 图像增强与特效
- 智能创作平台
- EasyMonitor
- BML 数据管理接口
- BML数据管理
- 机器翻译

* 应用归属: 公司 个人

* 应用描述:

图 5-25 表格数据预测模型应用配置

应用名称	AppID	API Key	Secret Key	创建时间	操作
1 学生学分选修	23795030	5zE2XSVkS6upBW48z8izEH28	***** 显示	2021-03-14 16:20:07	报表 管理 删除

图 5-26 模型应用参数信息

公有云部署的表格数据预测模型 API 使用,可以按照如下几个步骤进行。

(1) 鉴权认证获取 API 访问令牌。打开 HTTP 调试工具软件 PostMan,新建一个 Request,完成如下设置。

请求方式: POST。

URL 地址: <https://aip.baidubce.com/oauth/2.0/token>。

URL 参数: grant_type = `client_credentials` & client_id = `创建应用的 API Key` & client_secret = `创建应用的 Secret Key`。

单击【Send】按钮,如无错误,PostMan 执行 POST 请求及返回信息如图 5-27 所示。

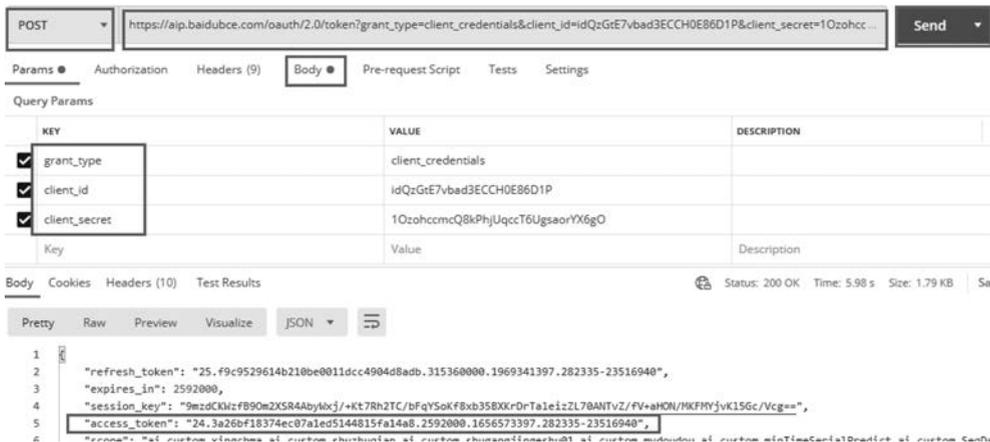


图 5-27 PostMan 执行 POST 请求及返回信息

在服务器返回的 JSON 字符串中,提取"access_token"键对应的取值,完成云端部署表格数据预测模型 API 访问令牌的获取。

(2) 使用令牌访问 API 进行表格数据预测。查看表格数据预测模型 API 调用文档,需要确认以下几个参数。

HTTP 请求方法: POST。

URL: https://aip.baidubce.com/rpc/2.0/ai_custom/v1/table_infer/spsspredict(公有云部署模型的访问接口地址)。

URL 参数: 模型 API 接口地址需附加参数"access_token",取值为上一步中通过 API Key 和 Secret Key 获取的 access_token。

Header 参数: 设置 Header 参数"Content-Type"取值为 application/json。

Body 参数: 请求正文,JSON 格式,包含提交云端部署模型进行预测的数据内容。表格

数据预测模型 API 请求的 Body 参数如表 5-1 所示。

表 5-1 表格数据预测模型 API 请求的 Body 参数

参 数	是否必选	类 型	说 明
include_req	否	boolean	返回结果是否包含特征数据, false 表示不包含, true 表示包含, 默认为 false
data	是	array	待预测数据, 每条待预测数据是由各个特征及其取值构成的键值对的集合

表格数据预测模型 API 返回参数为 JSON 字符串, 如表 5-2 所示。

表 5-2 表格数据预测模型 API 返回参数

字 段	是否必选	类 型	说 明
log_id	是	number	唯一的 logid, 用于问题定位
error_code	否	number	错误码, 当请求错误时返回
error_msg	否	string	错误描述信息, 当请求错误时返回
results	否	array(object)	预测结果数组

(3) PostMan 测试。根据上述表格数据预测模型 API 访问参数设置要求, 打开 HTTP 调试助手 PostMan, 新建 POST 请求, 填写表格数据预测模型 API 接口地址 (https://aip.baidubce.com/rpc/2.0/ai_custom/v1/table_infer/spsspredict), 并附带 URL 参数 access_token(第(1)步中获取的访问令牌)。【Header】参数中补充设置键值对 Content-Type: application/json。设置 POST 请求头部参数如图 5-28 所示。

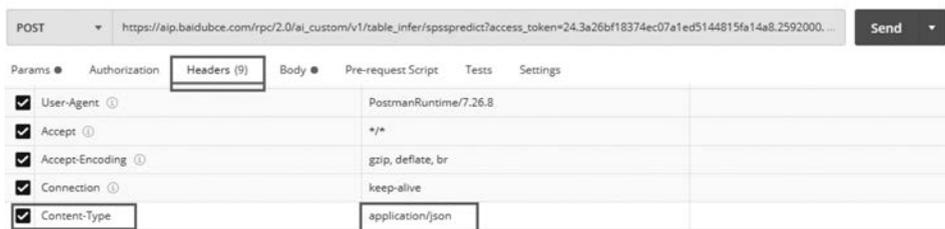


图 5-28 设置 POST 请求头部参数

单击【Body】选项, 选择 Body 参数格式为【raw】【JSON】, 设置 POST 请求 Body 参数格式如图 5-29 所示。



图 5-29 设置 POST 请求 Body 参数格式

按照 API 访问文档中 Body 参数设置要求, 填写 JSON 格式请求参数, 如图 5-30 所示。

单击【Send】按钮, 向云端部署的表格数据预测模型发起服务请求, 服务器返回的表格数据预测结果如图 5-31 所示。

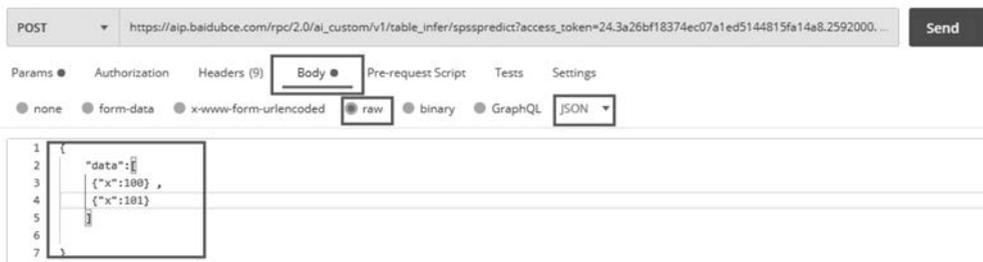


图 5-30 填写 JSON 格式请求参数

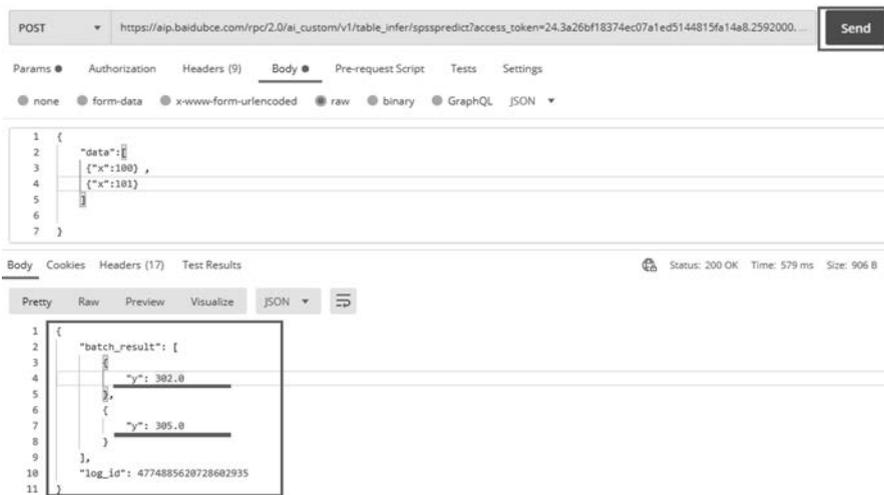


图 5-31 表格数据预测结果

测试结果表明,EasyDL 中创建并发布的表格数据预测模型,其 API 接口可以在第三方应用程序中直接应用,能够帮助开发者有效达成数据预测目标。



微课视频

5.2 时序数据预测建模

本节在简要介绍 EasyDL 结构化数据中提供的时序数据预测模型有关基本概念、适用场景和 AI 应用建模一般流程的基础上,按照创建模型、数据准备、模型训练、模型校验、模型发布、接口测试六大步骤,阐述时序数据预测模型建模及其应用测试的基本方法。

5.2.1 基本流程

时序预测指的是通过机器学习技术从历史数据中发现潜在规律,从而对未来的变化趋势进行预测。相较于表格数据预测使用的分类或回归模型,时序预测模型使用的训练数据中必须包含有效时序的特征,一般时序具有固定的频率,且在连续时间范围内的每个时间点上都有一个值。

时序预测模型基于包含时间特征的结构化数据进行建模,系统会基于用户上传的数据使用预置算法进行模型构建与训练。当完成模型训练后,系统不仅提供了常见的评估指标而且会生成可视化的预测序列效果图,用以验证模型性能。对于达到业务要求的时序预测模型,可以部署为在线服务,通过远程调用的方式对新的时间序列数据进行预测。

时序预测模型建模的基本流程如图 5-32 所示。

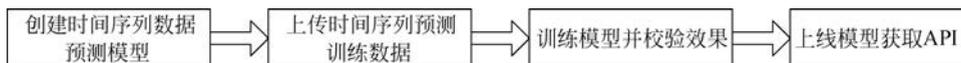


图 5-32 时序预测模型建模的基本流程

时序预测模型可以用于工业领域的维修预测、交通领域的交通流量预测,也可以用于商业领域的销量预测、价格预测等。

5.2.2 创建模型

这个阶段的主要任务是在百度 EasyDL 中,按照操作向导完成时序数据预测模型创建。打开 EasyDL 平台主页,如图 5-33 所示。



图 5-33 EasyDL 平台主页

单击【立即使用】按钮,显示选择模型类型操作界面,如图 5-34 所示。

选择【时序预测】,进入时序数据预测模型中心,如图 5-35 所示。

单击【创建模型】按钮,进入模型创建页面。依次填写“模型名称”“模型归属”“邮箱地址”“联系方式”“功能描述”等模型相关配置信息,完成时间序列预测建模信息配置,如图 5-36 所示。

单击【下一步】按钮完成模型创建。可在模型列表中查看新建的模型,如图 5-37 所示。



图 5-34 选择模型类型操作界面



图 5-35 时序数据预测模型中心



图 5-36 时间序列预测建模信息配置



图 5-37 模型列表中查看新建的模型

5.2.3 数据准备

这个阶段的主要任务是提供时序数据模型训练所需要的数据集。单击左侧导航栏【模型中心】→【数据总览】，在导航栏右侧显示的【我的数据总览】中，单击【创建数据集】按钮，启动时序数据预测模型训练数据集创建，如图 5-38 所示。



图 5-38 启动时序数据预测模型训练数据集创建

在弹出的操作界面中，填写时序数据预测模型训练数据集名称，完成时序数据预测模型训练数据集命名，如图 5-39 所示。

完成数据集准备后，单击 EasyDL 控制台左侧导航栏【模型中心】→【数据总览】，即可查看用于时序数据预测模型的数据集，如图 5-40 所示。

在【我的数据总览】中选择刚创建的数据集，单击【导入】，进入如图 5-41 所示的数据集导入方式配置界面。

选择【导入方式】为上传 CSV 文件或上传压缩包，单击【上传 CSV 文件】或【上传压缩包】，弹出如图 5-42 所示的上传数据提示信息对话框。



图 5-39 时序数据预测模型训练数据集命名



图 5-40 查看用于时序数据预测模型的数据集



图 5-41 数据集导入方式配置界面



图 5-42 上传数据提示信息对话框

单击【已阅读并上传】按钮，在弹出的文件打开对话框中选择所需文件或压缩包。待文件上传完毕，单击【确认并返回】按钮，进入【我的数据总览】，可查看用于时序数据预测模型训练的数据集导入进度，如图 5-43 所示。



图 5-43 查看用于时序数据预测模型训练的数据集导入进度

训练数据集的上传速度受个人使用计算机网络带宽所限，一般需要几分钟时间处理，当完成训练数据集上传后，【最近导入状态】由“正在导入”变为“已完成”。完成导入的数据集信息如图 5-44 所示。

单击【查看】操作链接，显示训练数据集详细信息，如图 5-45 所示。

需要注意的是，一个时序数据集可以包含一个或多个时间序列，目前时间序列仅支持 dd-mm-yyyy 的格式，其他时序数据文件要求和表格数据要求一致。

导入的数据文件可以是 CSV 文件或由 CSV 文件组成的压缩包文件。如果导入的是 CSV 文件，支持数据预览，如果是压缩包格式，则不支持预览。根据数据文件的实际情况进行列名设置。



图 5-44 完成导入的数据集信息

我的数据总览 > 【表格】销售预测/V1/查看

总行数 41 | 总列数 3 | 采样数 40 | 筛选显示列 (3/3) | 导入

datetime	area	sales_quantity
日期	类别	数值
1-1-2020	A	2000
2-1-2020	B	600
3-1-2020	A	2300
4-1-2020	B	550
5-1-2020	A	2100
6-1-2020	B	650
7-1-2020	A	2400

图 5-45 训练数据集详细信息

注：EasyDL 技术迭代更新可能会导致上述实现过程的部分细节内容有所变化，读者可以参考最新使用文档进行具体操作的调整。

5.2.4 模型训练

这一阶段的主要任务是利用导入的训练数据集训练创建的时序数据预测模型。单击【模型中心】中【训练模型】选项，首先进入训练模型参数配置环节，主要操作包括选择模型、选择数据集、选择时间列、选择时间间隔、选择目标列、滑动窗口大小、预测长度等。其中：“选择时间间隔”表示序列中相邻两个样本点的时间间隔；“滑动窗口大小”表示使用多少点数的历史数据生成预测数据；“预测长度”表示要预测的序列长度，该长度应小于滑动窗口大小。

一个典型的时序预测训练参数配置如图 5-46 所示。

配置完毕后，单击【开始训练】按钮启动训练任务。启动训练任务后，当显示“训练完成”状态时表示模型已完成训练。此类模型训练速度较快，具有 20 个特征列的 3 万条样本的表

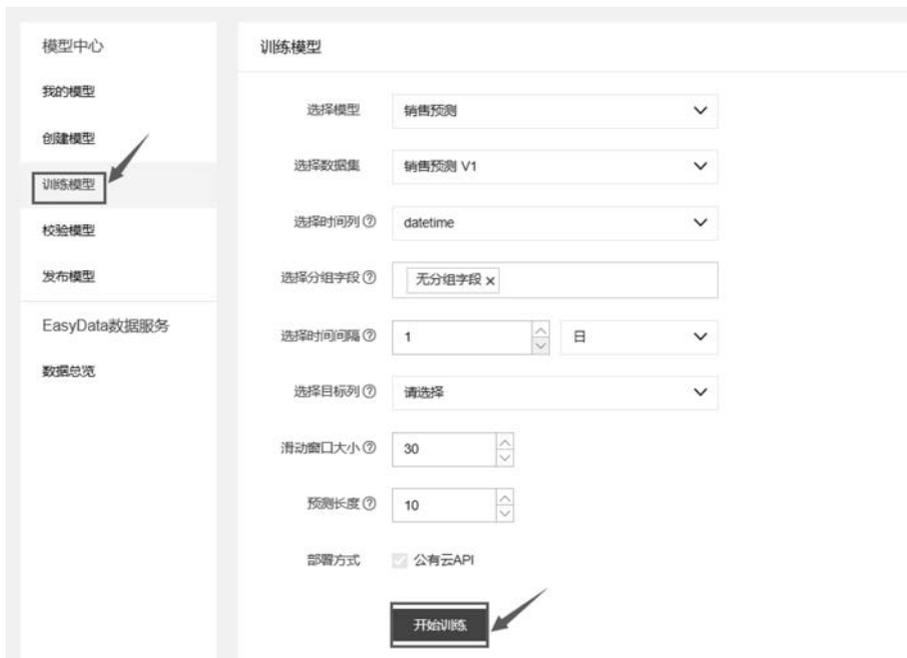


图 5-46 时序预测训练参数配置

格形式存储的时序数据,一般可在 30 分钟内训练完毕。训练中的时序数据预测模型信息如图 5-47 所示。



图 5-47 训练中的时序数据预测模型信息

当接收到训练完成通知短信后,刷新网页,可以查看时序数据预测模型训练结果信息,如图 5-48 所示。

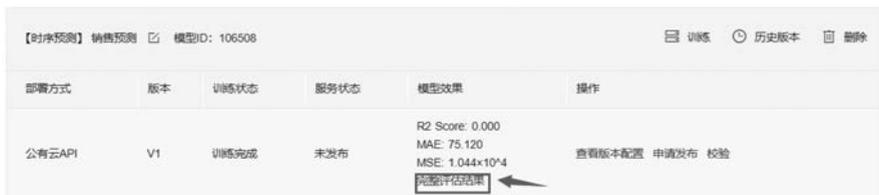


图 5-48 时序数据预测模型训练结果信息

5.2.5 模型校验

这一阶段的主要任务是在线检验完成训练的时序数据预测模型。单击左侧导航栏【校验模型】，进入模型校验页面。选择要校验的时序数据预测模型及其版本。单击【启动模型校验服务】按钮，启动时序数据预测模型校验，如图 5-49 所示。



图 5-49 启动时序数据预测模型校验

对于单序列模型，系统会自动生成校验数据，对于多序列模型，可以通过上传 CSV 文件来填充测试数据，但每次测试时只能包含一个序列的数据。单击【预测】按钮，可以在右侧结果面板中查看时序数据预测结果，如图 5-50 所示。

模型列表

选择模型: mymodel01 | 部署方式: 公有云API (目前仅支持) | 选择版本: V1

当前模型MAE(平均绝对误差) 6.530 评估报告

识别结果 如何优化效果?

预测数据		上传 csv 文件	复制json格式	预测
timestamp	value			
日期	数值			
1 2015-09-08 11:39:00	73			
2 2015-09-08 11:44:00	62			
3 2015-09-08 11:59:00	66			
4 2015-09-08 12:19:00	69			
5 2015-09-08 12:24:00	65			
6 2015-09-08 12:27:00	76			

预测结果

```

1  {
2    "tim... : [
3      "2015-09-09 15:33:00",
4      "2015-09-09 15:38:00",
5      "2015-09-09 15:43:00",
6      "2015-09-09 15:48:00",
7      "2015-09-09 15:53:00",
8      "2015-09-09 15:58:00",
9      "2015-09-09 16:03:00",
10     "2015-09-09 16:08:00",
11     "2015-09-09 16:13:00",
12     "2015-09-09 16:18:00",
13   ]
14  "val... : [
15     "0.057045099447"

```

预测 | 申请上线

图 5-50 时序数据预测结果

5.2.6 模型发布

这一阶段的主要任务是部署训练效果满意的时序数据预测模型。模型训练完成后,可将时序数据预测模型部署在公有云服务器、通用小型设备、本地服务器,也可以采用百度 AI 软硬一体方案。初学阶段,一般选择公有云部署方式。

训练完毕后可以在左侧导航栏中找到【发布模型】,依次进行“选择模型(拟部署模型)→选择部署方式(公有云部署)→选择版本(拟发布版本)→自定义服务名称→设置接口地址后缀→提交申请”等操作。设置时序数据预测模型发布相关信息操作界面如图 5-51 所示。



图 5-51 设置时序数据预测模型发布相关信息操作界面

当时序数据预测模型信息中显示“服务详情”时,表示已经发布成功。发布成功的时序数据预测模型信息如图 5-52 所示。



图 5-52 发布成功的时序数据预测模型信息

单击【服务详情】,显示如图 5-53 所示时序预测模型 API 接口地址信息。

单击【查看 API 文档】按钮,可以进入时序数据预测模型 API 使用方法说明文档网页,如图 5-54 所示。



图 5-53 时序预测模型 API 接口地址信息



图 5-54 时序数据预测模型 API 使用方法说明文档网页

开发者可通过该文档了解时序数据预测模型 API 使用方法。

5.2.7 接口测试

这一阶段的主要任务是测试云端部署的时序数据预测模型访问接口。模型调用测试前首先需要创建一个 EasyDL 结构化模型应用。进入百度智能云控制台(需要使用自己的账号和密码登录百度智能云),在产品服务栏中选中【EasyDL 定制化训练平台】,进入操作页面后单击【EasyDL 结构化数据】,进入 EasyDL 结构化数据应用中心,如图 5-55 所示。

在 EasyDL 结构化数据服务中心内单击导航栏【公有云部署】下的【应用列表】,在应用管理界面单击【创建应用】,填写【应用名称】【接口选择】【应用归属】【应用描述】等相关信息,完成时序数据预测模型应用配置,如图 5-56 所示。

单击【立即创建】按钮,应用列表中显示的时序数据预测模型应用 AppID、API Key、Secret Key 参数如图 5-57 所示。

公有云部署的时序数据预测模型 API 使用,可以按照如下几个步骤进行。



图 5-55 EasyDL 结构化数据应用中心

创建新应用

* 应用名称:

* 接口选择: 勾选以下接口, 使此应用可以请求已勾选的接口服务, 注意EasyDL结构化数据服务已默认勾选并不可取消。

- EasyDL
- 语音技术
- 文字识别
- 人脸识别
- 自然语言处理
- 内容审核
- UNIT
- 知识图谱
- 图像识别
- 智能呼叫中心
- 图像搜索
- 人体分析
- 图像增强与特效
- 智能创作平台
- EasyMonitor
- BML 数据管理接口
- BML数据管理
- 机器翻译

* 应用归属: 公司 个人

* 应用描述:

图 5-56 时序数据预测模型应用配置



图 5-57 时序数据预测模型应用 AppID、API Key、Secret Key 参数

(1) 鉴权认证获取 API 访问令牌。打开 HTTP 调试工具软件 PostMan, 新建一个 Request, 完成如下设置。

请求方式: POST。

URL 地址: <https://aip.baidubce.com/oauth/2.0/token>。

URL 参数: `grant_type=client_credentials&client_id=创建应用的 API Key&client_secret=创建应用的 Secret Key`。

单击【Send】按钮, 如无错误, PostMan 执行 POST 请求及返回信息, 如图 5-58 所示。



图 5-58 PostMan 测试 API 访问令牌的获取

在服务器返回的 JSON 字符串中, 提取“access_token”键对应的取值, 完成云端部署时序数据预测模型 API 访问令牌的获取。

(2) 使用令牌访问 API 进行时序数据预测。查看时序数据预测模型 API 调用文档, 需要确认以下几个参数。

HTTP 请求方法: POST。

URL: https://aip.baidubce.com/rpc/2.0/ai_custom/v1/table_infer/spsspredict (公有云部署模型的访问接口地址)。

URL 参数: 模型 API 接口地址需附加参数“access_token”(上一步中通过 API Key 和 Secret Key 获取)。

头部参数：设置 Content-Type 参数取值为 application/json。

Body 参数：请求正文,JSON 格式,包含提交云端部署模型进行预测的时间序列。时序数据预测模型 API 请求的 Body 参数如表 5-3 所示。

表 5-3 时序数据预测模型 API 请求的 Body 参数

参 数	是否必选	类 型	说 明
include_req	否	boolean	返回结果是否包含特征数据: false,不包含; true,包含
data	是	array	待预测数据,每条待预测数据是由各个特征及其取值构成的键值对的集合

Body 参数设置典型示例如下所示。

```
{
  "data":{
    "datetime":
    ["2015-09-0915:33:00","2015-09-0915:38:00","2015-09-0915:43:00"],
    "sales_quantity":
    ["10","15","20"]
  }
}
```

返回参数：时序数据预测模型 API 访问返回参数亦为 JSON 字符串,如表 5-4 所示。

表 5-4 时序数据预测模型 API 返回参数

字 段	是否必选	类 型	说 明
log_id	是	number	唯一的 logid,用于问题定位
error_code	否	number	错误码,当请求错误时返回
error_msg	否	string	错误描述信息,当请求错误时返回
results	否	array(object)	预测结果数组

(3) PostMan 测试。根据上述时序数据预测模型 API 访问参数设置要求,打开 HTTP 调试助手 PostMan,设置 HTTP 请求方式为 POST,填写时序数据预测模型 API 访问接口地址(URL): https://aip.baidubce.com/rpc/2.0/ai_custom/v1/table_infer/spsspredict。

设置 URL 参数为 access_token = ***** [先前获取访问令牌],并在 PostMan 【Headers】参数中补充键值对 Content-Type:application/json,设置 POST 请求头部参数如图 5-59 所示。

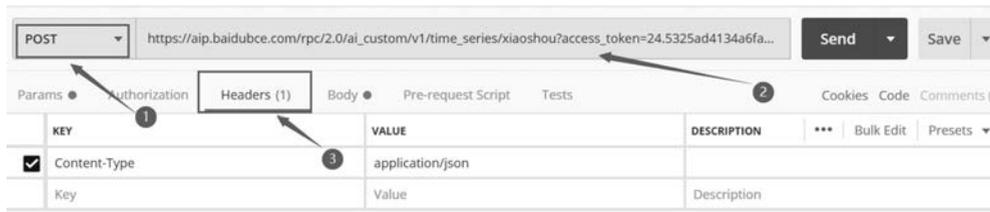


图 5-59 设置 POST 请求头部参数

单击【Body】选项,选择 Body 参数类型为【raw】【JSON】,设置 POST 请求 Body 参数格式如图 5-60 所示。

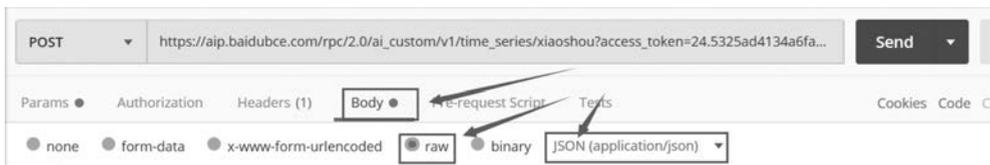


图 5-60 设置 POST 请求参数格式

按照 API 访问文档中 Body 参数设置要求,填写如下 JSON 格式请求参数。

```
{
  "data": {
    "col_0": [
      "2021-02-02 15:00",
      "2021-02-02 15:01",
      "2021-02-02 15:02",
      "2021-02-02 15:03",
      "2021-02-02 15:04",
      "2021-02-02 15:05",
      "2021-02-02 15:06",
      "2021-02-02 15:07",
      "2021-02-02 15:08",
      "2021-02-02 15:09",
      "2021-02-02 15:10",
      "2021-02-02 15:11",
      "2021-02-02 15:12",
      "2021-02-02 15:13",
      "2021-02-02 15:14",
      "2021-02-02 15:15",
      "2021-02-02 15:16",
      "2021-02-02 15:17",
      "2021-02-02 15:18",
      "2021-02-02 15:19",
      "2021-02-02 15:20"
    ],
    "col_1": [
      "0.0",
      "8.41471",
      "9.092974",
      "1.4112",
      "-7.568025",
      "-9.589243",
      "-2.794155",
      "6.569866",
      "9.893582",
      "4.121185",
      "-5.440211",
      "-9.999902"
    ]
  }
}
```

```

" - 5.365729",
"4.20167",
"9.906074",
"6.502878",
"- 2.879033",
"- 9.613975",
"- 7.509872",
"1.498772",
"9.129453"
] } }

```

单击【Send】按钮，向云端部署的时序数据预测模型发起服务请求，服务器返回的时序数据预测结果如图 5-61 所示。

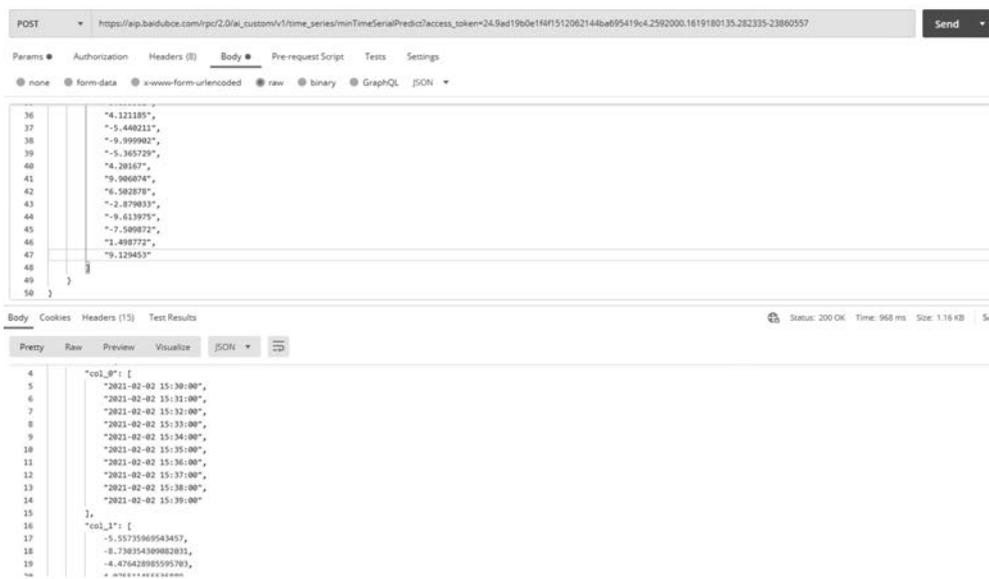


图 5-61 时序数据预测结果

对比申请数据和服务器返回数据，可见请求时序预测服务时，提交的时序数据的时间戳起始值为 2021-02-02 15:00:00，服务器返回预测结果的时间戳起始值为 2021-02-02 15:30:00，两者时间差与训练模型的滑动窗口值 30 一致，返回连续 10 个时间戳的预测结果，与训练模型时设置的预测长度值 10 一致。这说明服务器返回结果从格式角度看，完全符合训练模型设置。