

标识：RICH-Datlab

版本：1.0

技术文件

DatLab 开发者平台手册

北京睿企

二〇二一年三月

修订记录

版本	修订人	修订时间	修订内容
----	-----	------	------

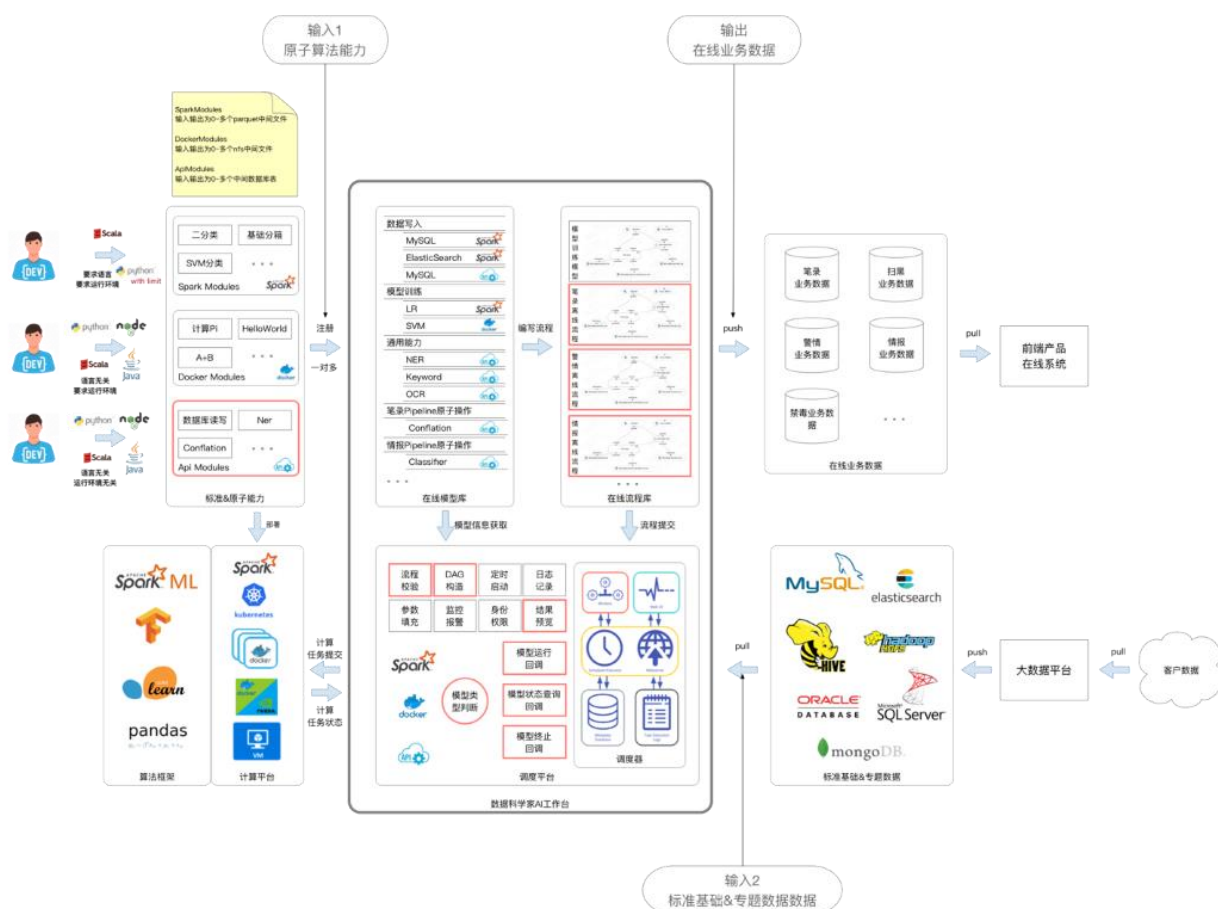
目录

1 概述.....	2
2 架构图.....	3
3 Datlab 平台服务模块接口开发规范.....	3
3.1 接口说明.....	3
3.2 接口定义.....	3
4 数据服务接口.....	6
4.1 接口说明.....	6
4.2 接口定义.....	6
5 基础服务注册为平台《模块》.....	9
5.1 打开注册接口页面.....	9
5.2 填写注册信息.....	10
5.3 验证.....	11
6 查询模块信息.....	12
7 修改模块信息.....	13

1 概述

数据和服务（算法）的通用调度平台 Datlab 本身可以提供各种功能性模块，但同时也提供用户可以根据接口设计规范，开发出属于自己的功能模块，在流程实验中进行使用。以下对于模块接口的设计规范，以及模块的注册使用进行详细说明

2 架构图



3 Datlab 平台服务模块接口开发规范

3.1 接口说明

- 所有希望可以被平台调度的服务模块接口需要按照 3.2 的接口定义实现
- 接口的输入输出均为 0~多张符合 mysql 协议的数据库表

3.2 接口定义

接口约束

接口编号: apiId	每一个服务需要有自己唯一的接口编号
调用 URI	每一个服务有自己的 URI
调用方法	必须为 POST

Content-Type	必须为 application/json
幂等	除了数据库写入服务，所有接口服务在输入不变的情况下，重复调用接口输出内容不应改变
数据表描述方式：DBInfo	每一张数据库表的描述方式为 <pre>{ "host": "数据库地址", "port": "数据库端口号", "user": "数据库用户名", "password": "数据库密码", "database": "数据库名称", "tableName": "数据库表名称" }</pre>

请求参数

名称	类型	默认值	是否必须	描述
parameters	Map<String, Object>	{}	是	接口参数，其中的 key-value pair 会根据不同的接口需求，传输不同的内容
inputs	List<Map<String, Object>>	[]	是	每一个输入对应的数据库表信息
outputs	List<Map<String, Object>>	[]	是	每一个输出对应的数据库表信息, 说明输出表未进行初始化
experimentId	String		是	这次运行的实验 id

返回数据

字段	类型	是否必须	说明
requestId	String	是	每次请求的唯一 id 通过该 id 可以追溯服务内部请求细节
experimentId	String		这次运行的实验 id, 和输入一致即可
apiId	String	是	服务编号，由服务开发者定义
version	String	是	服务版本，由服务开发者定义
code	String	是	返回状态码，" 0" 表示正常
message	String	是	异常附加信息

data	Object	否	执行过程信息, 可以将必要的过程信息打印出来

调用示例

地址	http://ip:port/api/algo/ner
请求体	<pre>{ "parameters": { "threshold": 0.5, "mode": "bert" }, "inputs": [{ "name": "input1", "dbInfo": { "host": "数据库地址", "port": "数据库端口号", "user": "数据库用户名", "password": "数据库密码", "database": "数据库名称", "tableName": "数据库表名称" } }], "outputs": [{ "name": "output1", "dbInfo": { "host": "数据库地址", "port": "数据库端口号", "user": "数据库用户名", "password": "数据库密码", "database": "数据库名称", "tableName": "数据库表名称" } }] }</pre>

返回示例

```
{
  "requestId": "aa4e55e5-ccb2-4f6d-ac8b-39abf74ff72",
  "version": "v1",
  "apiId": "B001",
  "code": "0",
  "message": "success",
  "data": [
```

```

{
  "content": {
    "score": [
      0.8964171409606934
    ]
  }
}
]
}

```

4 数据服务接口

4.1 接口说明

- 直接通过数据进行交互的接口，由 3 中的接口代为进行分页处理，进行统一的调度
- 接口的输入输出均为 0~多份数据信息

4.2 接口定义

接口约束

接口编号：apiId	每一个服务需要有自己唯一的接口编号
调用 URI	每一个服务有自己的 URI
调用方法	必须为 POST
Content-Type	必须为 application/json
幂等	所有接口服务在输入不变的情况下，重复调用接口输出内容不应改变
数据表描述方式：DataInfo	<p>每一份数据库的描述方式为 Column 名称为数据库字段的驼峰格式 比如数据库中叫做 entity_name 那么对应请求名称为 entityName</p> <pre> { "name" : "数据标识", "data" : [{ "column1" : "value1", "column2" : "value2" }, { "column1" : "value1", </pre>

	<pre> "column2" : "value2" }] } </pre>
--	---

请求参数

名称	类型	默认值	是否必须	描述
parameters	Map<String, Object>	{}	是	接口参数，其中的 key-value pair 会根据不同的接口需求，传输不同的内容
inputs	List<DataInfo>	[]	是	每一个输入对应的数据
experimentId	String		是	这次运行的实验 id

返回数据

字段	类型	是否必须	说明
requestId	String	是	每次请求的唯一 id 通过该 id 可以追溯服务内部请求细节
experimentId	String		这次运行的实验 id, 和输入一致即可
apiId	String	是	服务编号，由服务开发者定义
version	String	是	服务版本，由服务开发者定义
code	String	是	返回状态码，"0" 表示正常 异常时 "API 名称" + "错误编码" 来表示，API 名称大写，编码保持 5 位数，如 OCR00001 NER00002 KEYWORD00003 ROLE00004 RELATION00005 CIRMETHOD00006
message	String	是	异常附加信息
data	List<DataInfo>	是	返回结果

调用示例

地址	http://ip:port/api/algo/ner
请求体	<pre> { "parameters": { "threshold": 0.5, "mode": "bert" }, "inputs": [{ "name": "input1", "data": [{ "column1": "value1" }, { "column1": "value1" }] }], "experimentId": "1" }</pre>

返回示例

```

{
  "requestId": "aa4e55e5-ccb2-4f6d-ac8b-39abf74ff72",
  "version": "v1",
  "apiId": "B001",
  "code": "0",
  "message": "success",
  "data": [
    {
      "name": "output1"
      "data": [
        {
          "column1": "value1",
          "code": "0"
        },
        {
          "column1": "value1",
          "code": "0"
        }
      ]
    }
  ]
}
```



```
]
}
]
}
```

5 基础服务注册为平台《模块》

5.1 打开注册接口页面

访问

http://worker_ip:31510/modules/add

如图：

模块基本信息

*模块名称

*模块对应接口

*familyid

*版本号

*页面层级路径

*模块说明

模块参数信息

*参数名称	*参数值类型	参数枚举值	默认值	*displayname	*help Text	*操作
暂无数据						
+ 添加						

输入参数信息

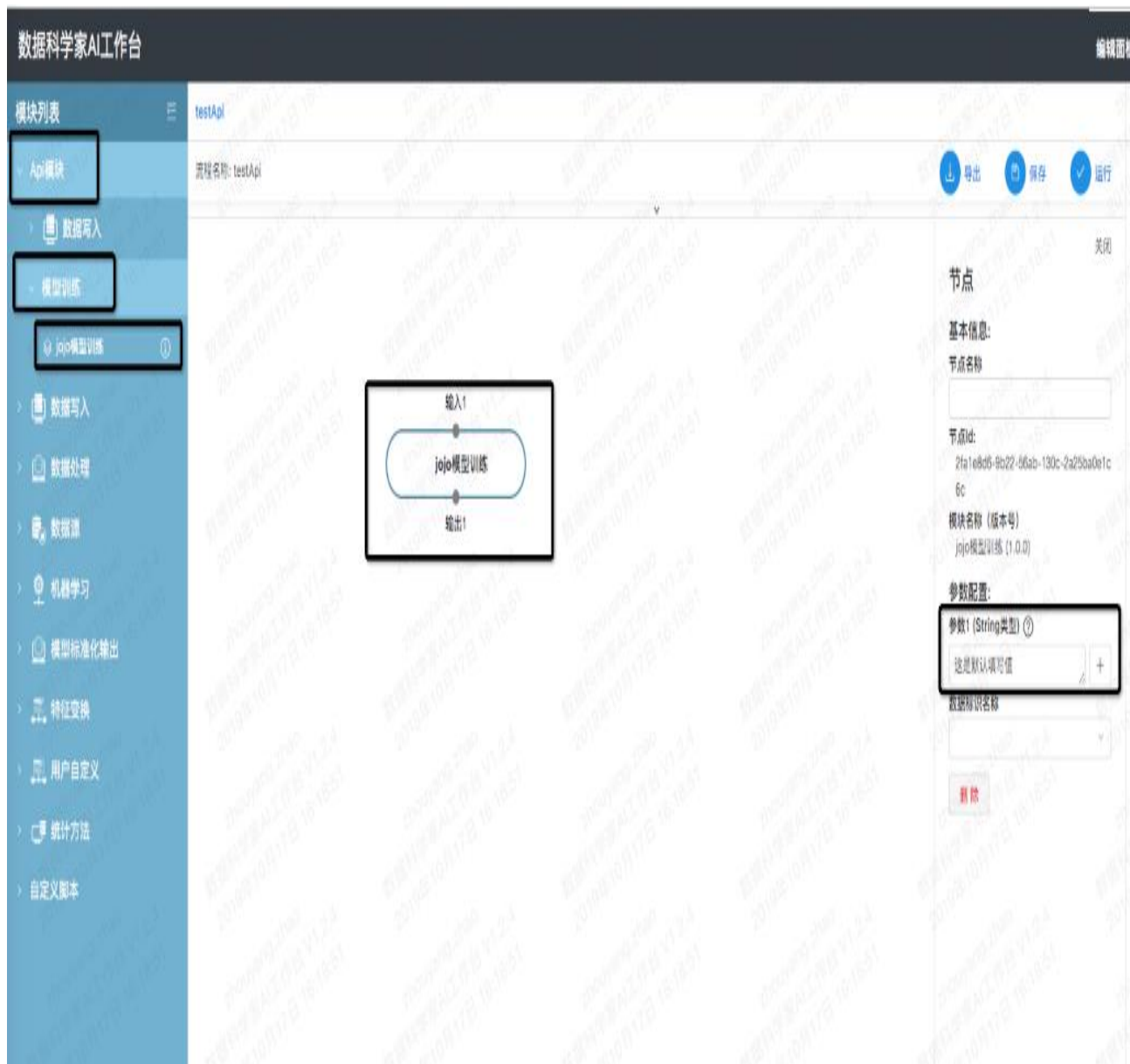
*参数名称	*displayname	*help Text	*操作
暂无数据			

5.2 填写注册信息

按照提示进行模块信息填写
部分模块基本信息填写说明，填写完成后进行保存。

字段	说明
模块对应接口	如果 API 部署在 k8s 集群中，ip 部分直接填写 WORKER_IP 即可 如 http://WORKER_IP:12345/path
familyId	模块功能代码，如 OCR，NER，Keyword 同样功能的模块 familyId 一样即可
页面层级路径	必须以”Api 模块”打头 如 Api 模块/模型训练, 在使用时可按照路径找到该模块
模型参数信息	在模块参数信息下，点击添加，按照提示填写参数信息
输入信息	模块输入，给出每个输入的名称和描述，说明，输入的名称是代码中获取输入表信息的唯一标识
输出信息	模块输出，给出每个输出的名称和描述，说明，输出的名称是代码中获取输入表信息的唯一标识

如果构造一个两个输入，一个输出的模块，1 个参数的模块如下：

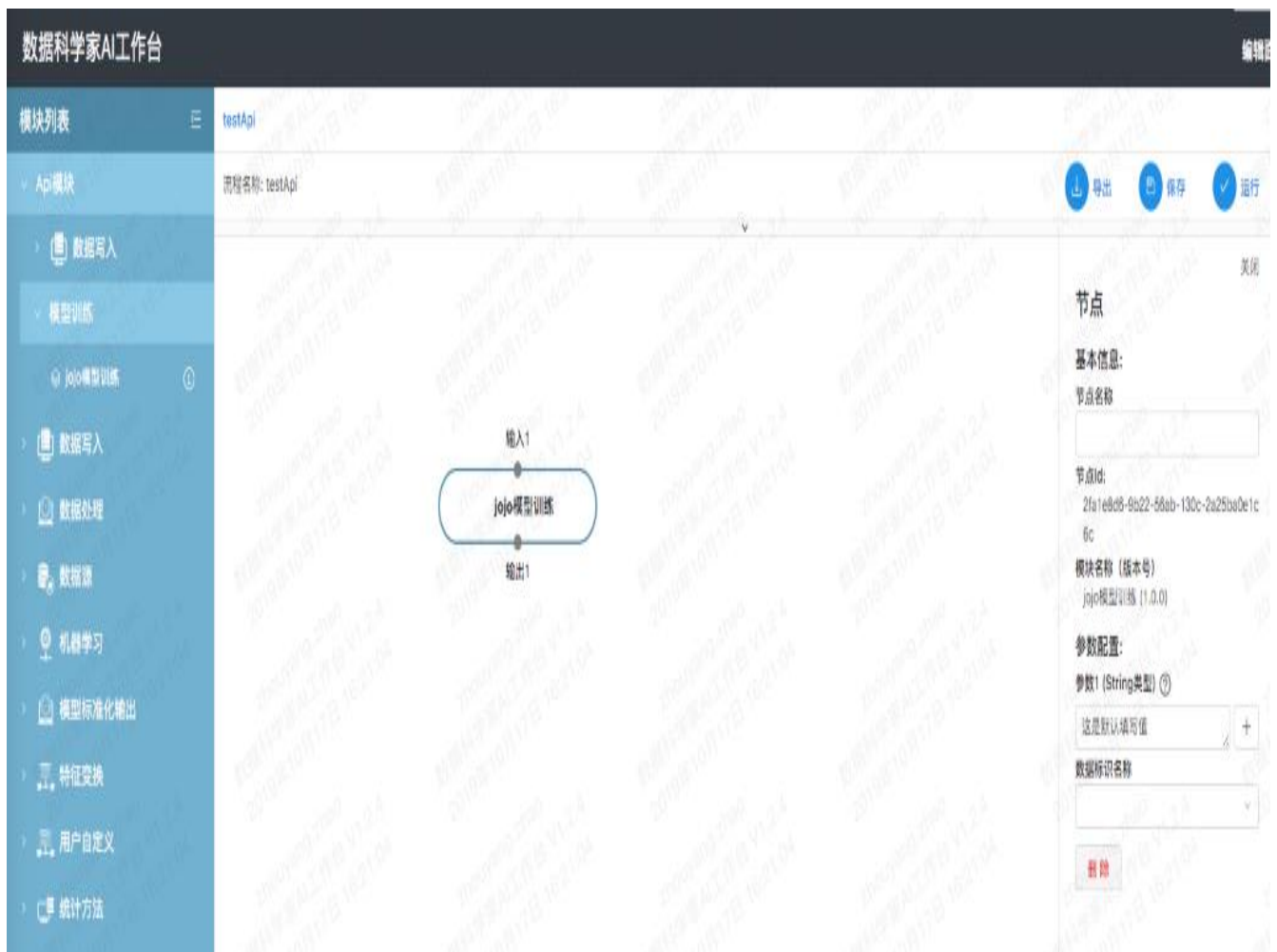


5.3 验证

访问

http://worker_ip:31510/

通过左侧拖拽模块列表即可验证



6 查询模块信息

访问

http://worker_ip:31510/modules/list 页面

模块名称	描述	页面路径	版本号	类型	注册时间	注册人	操作
蓝信信息发送模块V2	蓝信消息发送模块	Api模块/蓝信/	1.0.0	Api		delta	查看 编辑
QB主体归一化合并	qingbao_subject_normalize_merge_handleV1.0	Api模块/cia-Pipeline	1.0	Api		delta	查看 编辑
重新生成xzqhdn	重新生成xzqhdn	Api模块/警情Pipeline/临时数据同步模块	1.0.0	Api		hanzhong.zhang	查看 编辑
情报线索预警	情报线索预警	Api模块/cia-Pipeline	3.0	Api		delta	查看 编辑
情报人员预警	情报人员预警	Api模块/cia-Pipeline	3.0	Api		delta	查看 编辑
情报事件预警	情报事件预警	Api模块/cia-Pipeline	3.0	Api		delta	查看 编辑
python脚本执行	该模块用于执行python脚本	Api模块/固化模块	v1	Api		delta	查看 编辑
情报事件预警	情报事件预警	Api模块/cia-Pipeline	2.0	Api		delta	查看 编辑
情报人员预警	情报人员预警	Api模块/cia-Pipeline	2.0	Api		delta	查看 编辑

7 修改模块信息

按照 6 进行模块查询完之后，点击编辑按钮，按照提示进行编辑工作。

模块管理 > 模型注册 > 编辑

模块基本信息

*模块名称

蓝信信息发送模块V2

*模块对应接口

http://WORKER_IP:32114/api/lanxin/sciSendMessage

*familyid

lanxin_sendmessageV2

*版本号

1.0.0

*页面层级路径

Api模块/蓝信/

*模块说明

蓝信消息发送模块

模块参数信息

*参数名称	*参数值类型	参数枚举值	默认值	*displayname	*help Text	*操作
cycles	Integer		10	cycles	的最高循环失败数，超出则模块失败，否则继续请求	编辑
maxFailCount	Integer		30	maxFailCount	的总失败次数，超出则模块失败，否则继续请求	编辑