Deepdetect总结

DeepDetect集成了Caffe，XGBoost和Tensorflow, 采用REST风格API和完全JSON化的输入／输出数据结构。

服务器信息：通过GET目录/info获取服务器信息；

服务管理：通过PUT（创建一个服务）、GET（获得一个服务状态）和POST（更新一个服务）目录/services进行服务的管理；

模型训练：通过POST（创建一个新的训练作业）、GET（获取一个训练作业的状态）和DELETE（取消一个训练作业）目录/train进行模型训练；

数据预测：通过POST（发送数据到服务）目录/predict进行数据预测。

Caffe为例：

/ info：查看服务信息

/ services：服务的创建和销毁。

/ train：训练数据

/ predict：预测数据

1、获得服务信息：

GET "http://localhost:8080/info"

2、创建一个图像分类的服务：

curl -X PUT "http://localhost:8080/services/myserv" -d "{

\"mllib\":\"xgboost\",

\"description\":\"example classification service\",

\"type\":\"supervised\",

\"parameters\":{

\"input\":{\"connector\":\"csv\"},

\"mllib\":{\"nclasses\":9}

},

\"model\":{\"repository\":\"/home/me/models/example\"}

}"

参数说明：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 参数 | | | | 类型 | 默认 | 说明 |
| description | | | | string | empty | 描述服务 |
| mllib | | | | string |  | 选择机器学习库caffe，xgboost,tensorflow |
| type | | | | string | supervised | 机器学习服务类型：监督和无监督 |
| Model | repository | | | string |  | 模型路径 |
| templates | | | string | templates | 模型模板库路径 |
| weights | | | string | empty | 预训练模型路径 |
| parameters | | input | connector | string |  | 数据类型 |
| mllib | nclasses |  |  | 类别数 |
| output |  |  |  |  |

3、删除服务

curl -X DELETE "http://localhost:8080/services/myserv?clear=full"

4、创建训练服务

curl -X POST "http://localhost:8080/train" -d "{"service":"imageserv","async":false,"parameters":{"mllib":{"gpu":true,"net":{"batch\_size":32},"solver":{"test\_interval":500,"iterations":30000,"base\_lr":0.001,"stepsize":1000,"gamma":0.9}},"input":{"connector":"image","test\_split":0.1,"shuffle":true,"width":224,"height":224},"output":{"measure":["acc","mcll","f1"]}},"data":["/media/img"]}"

参数说明：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | | | | **Type** | **Default** | **Description** |
| service | | | | string |  | 服务标识 |
| async | | | | Bool | true | 是否开始无阻挡训练(non-blocking training call) |
| data | | | | object | empty | 输入数据集进行训练，在某些情况下可以由输入连接器处理，一般不可选 |
| parameters | mllib | gpu | |  | true | 是否采用gpu，选择gpu个数采用gpid。 |
| net | batch\_size | int |  | 训练的batchsize |
| solver | test\_interval | | int |  | 测试阶段之间的迭代次数 |
| iterations | | int |  | 最大解释器迭代次数 |
| base\_lr | | real |  | 初始学习率 |
| stepsize | | int |  | 学习率下降之间的迭代次数 |
| gamma | | real |  | 学习率下降因子 |
| input | connector | |  |  | 连接器 |
| test\_split | | real | 0 | 测试分割数据集的一部分 |
| shuffle | | bool | false | 是否打乱数据集（分割之前） |
| width | | int | 227 |  |
| height | | int | 227 |  |
| output | measure | | array | empty | 要求的输出措施，“acc”：精度，“mcll”：多级日志丢失， “f1”：召回率 |