关卡3:

openGauss的DB4AI特性应用

openGauss的DB4AI特性应用

任务一：重新编译数据库环境，并安装Madlib

1. 登录tpch数据库，查看schema，存在madlib表示安装完成，将执行结果截图：

gsql -d tpch -p 5432

\dn

任务二：使用SVM对TPCH的orders表进行分类问题预测。

1. 使用SVM对TPCH的orders表进行分类，使用默认的参数进行训练，分类的条件为o\_totalprice < 100000'，查看训练模型，将执行结果截图：

打开列显示：

\x on

查看模型order\_svm。

select \* from order\_svm;

关闭列显示：

\x off

2. 使用模型，进行预测，将结果与原表进行关联查询，将执行结果截图。

SELECT \*, o\_totalprice < 100000 AS actual FROM orders JOIN orders\_pred USING (o\_orderkey) ORDER BY o\_orderkey;

任务三：使用linregr\_train对TPCH的orders表进回归预测。

1. 使用linregr\_train对TPCH的orders表进回归预测，将执行结果截图：

打开列显示：

\x on

查看模型order\_svm。

select \* from orders\_linregr;

关闭列显示：

\x off

2. 预测，并对比结果，

挑战一：使用Madlib中的其他功能，完成某个数据挖掘工作（如决策树）。

1. 使用Madlib中的其他功能，完成某个数据挖掘工作，将执行结果截图：

实践思考题1：分类模型与回归模型有何不同？

实践思考题2：什么是SVM算法？