

数据挖掘可视化系统_操作说明书 V1.0

1. 简介

欢迎使用数据挖掘可视化系统！使用本系统前，请您务必阅读该操作手册。

为了使您尽快掌握，并且有效利用数据挖掘可视化系统的操作方法，特为您准备此手册。

系统分为三大模块：数据集模块，原始数据初始化模块，数据挖掘可视化模块。

主要功能是：

1. 提供多种数据集供用户选择。
2. 根据用户选择的数据集，使用不同方式可视化原始数据。
3. 利用不同算法对数据进行数据挖掘，并进行模型的评估以及可视化。

系统架构图如下所示：



2. 系统安装

2.1 运行环境

硬件环境：联网的电脑

软件环境：

可正常运行的浏览器：星愿，Chrome，Safari,Microsoft Edge 等

2.2 安装

本系统采用浏览器/服务器模式， 系统在服务器端部署完毕， 成功运行后， 用户仅需凭借联网状态下的电脑或移动笔记本， 提供浏览器访问系统网站即可正常使用本系统。

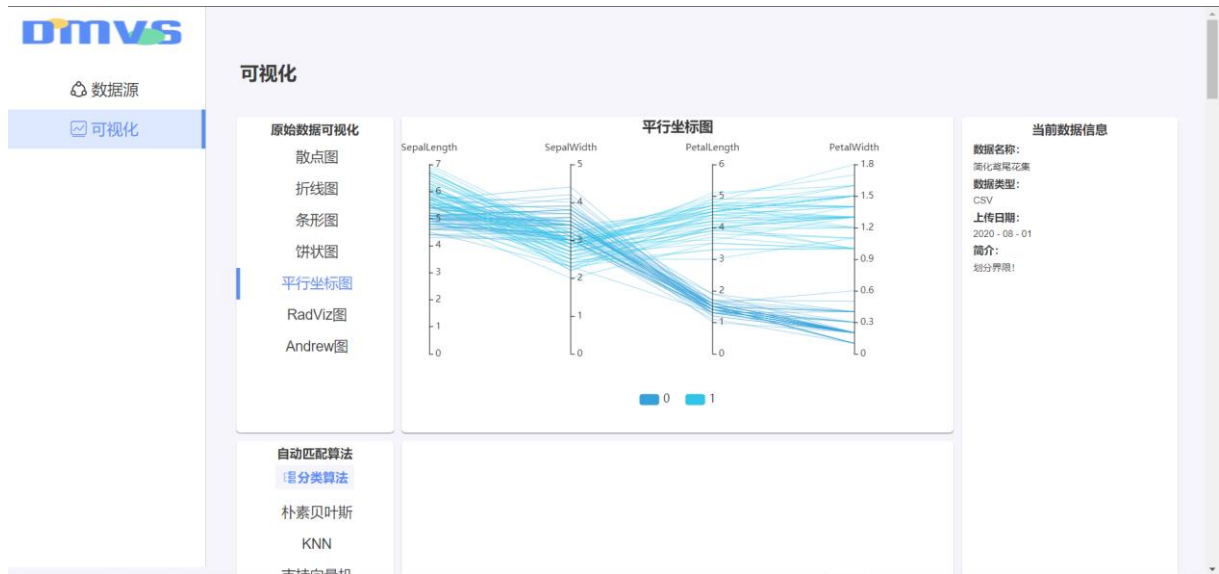
3. 操作步骤

3.1 数据集选择



打开网站，出现的是 显示选择数据集的页面，有多种数据集供用户选择。左上角可以选择切换为可视化页面，右上角可以搜索数据源名称，左边边框可以切换到可视化界面。

3.2 原始数据可视化



在切换到可视化界面后，用户可以选择多种方式对刚才选择的数据进行可视化。包括散点图，折线图等多种形式。右侧是对数据集的介绍。

3.3 数据预处理和模型训练

The screenshot shows the DMVS data preprocessing and model training interface. On the left, there is a sidebar with a '数据源' (Data Source) button and a '预处理' (Preprocessing) button. The main area is titled '预处理' (Preprocessing) and contains a list of preprocessing methods: '对数几率回归' (Logistic Regression), '回归决策树' (Regression Decision Tree), and '回归决策树' (Regression Decision Tree). The '对数几率回归' (Logistic Regression) method is selected and displayed in the center. Below the list, there is a '表格展示' (Table Display) section with three tabs: '提交并预处理' (Submit and Preprocess), '进行训练' (Train), and '进行预测' (Predict). The '进行训练' (Train) tab is selected, and a table of data is displayed. The table has 6 columns: 'id', 'Sepal.Length', 'Sepal.Width', 'Petal.Length', 'Petal.Width', and 'Class'. The data is as follows:

id	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Class
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0
5	5.4	3.9	1.7	0.4	0
6	4.6	3.4	1.4	0.3	0
7	5.0	3.4	1.5	0.2	0
8	4.4	2.9	1.4	0.2	0
9	4.9	3.1	1.5	0.1	0
10	5.4	3.7	1.5	0.2	0

用户点击提交和预处理，和进行训练后，系统会智能选定算法对数据进行训练并建立模型。向下拉动即可查看全部数据。

3.4 模型的评估以及数据挖掘可视化



稍等几许，便可以看到数据挖掘算法的可视化结果以及模型评估参数的大小。