实验1：WEKA安装与配置

**1 实验目的**

1. 基于WEKA搭建数据挖掘任务执行环境。

**2 实验平台与工具**

1. Windows、Linux操作系统
2. 数据分析与挖掘系统WEKA

**3 实验内容**

1. 在WEKA主页<https://www.cs.waikato.ac.nz/ml/weka/> 下载并且安装WEKA；
2. 解释WEKA应用程序中Visualization菜单、Tools的4个子菜单、Explorer、Experimenter、KnowledgeFlow、Workbench的基本功能。



**4. 规则与要求**

（1）独立完成，严禁相互抄袭（如有发现抄袭和被抄袭均判为0分），以及从网络上直接摘抄别人的观点和总结（该行为将影响报告成绩）。

（2）实验报告符合学术写作的排版要求，请参考群文件中的“报告模板.docx”和“参考文献格式.docx”的排版格式。

（3）实验报告内容详实，采用图文混合的方式叙述安装和配置过程。

Tip：Win+Shift+S 在Windows中可以快速截屏。

（4）报告文件见附件，提交报告时请以附件形式插入到超星作业中。

实验报告

报告标题：WEKA安装与配置

学号：21190630

姓名：黄艺杰

日期：2023年4月6日

# 一、实验环境

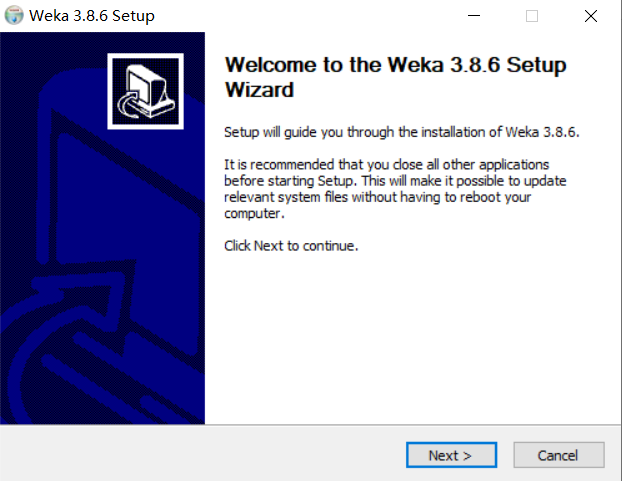
1. 操作系统：Windows10

2. 软件（含版本号）：weka-3-8-6-azul-zulu-windows

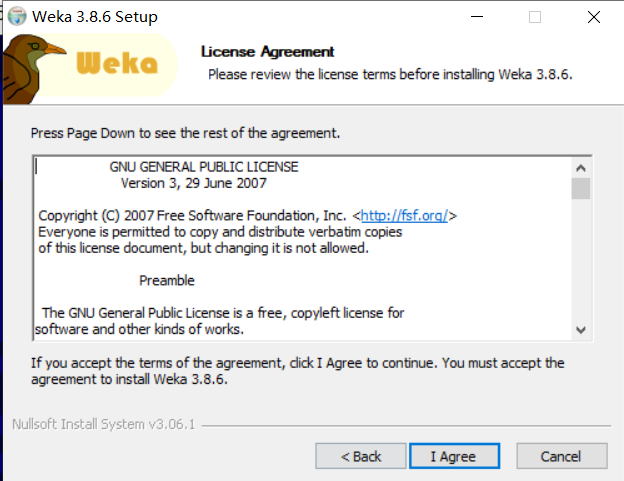
# 二、实验内容及其完成情况

1. 下载并安装Weka

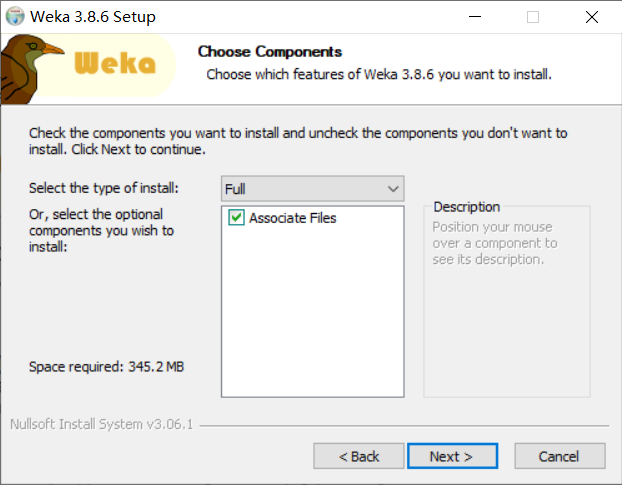
点击安装包开始安装



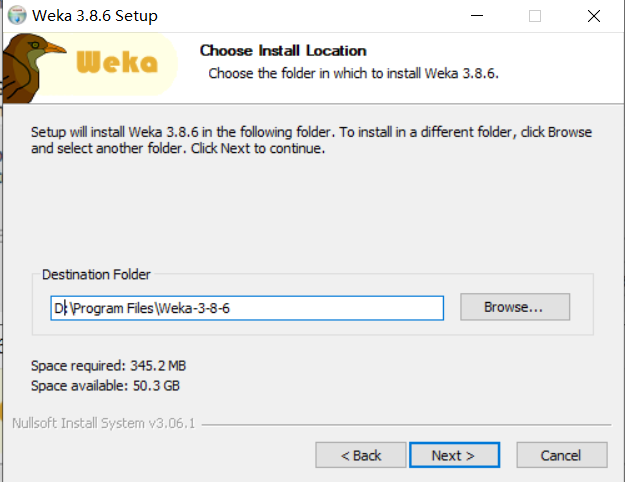
点击同意GNU GPL协议



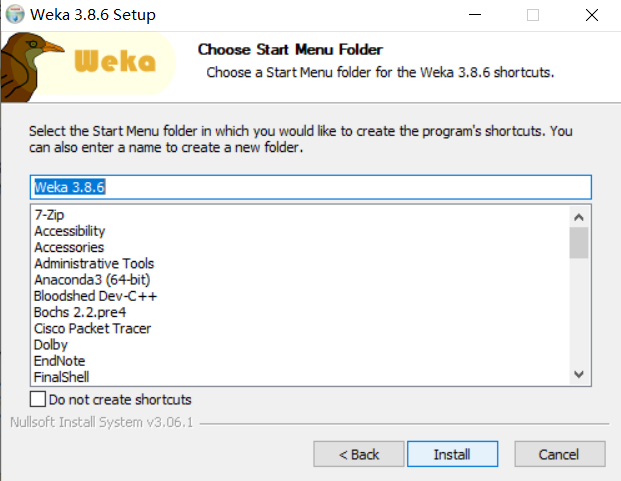
下一步是选择安装组件。选项有Full 、Minimal和custom三项，默认为Full，由于完全安装也不占多大空间，选择默认的Full选项，因此保持默认选项不动，单击Next按钮



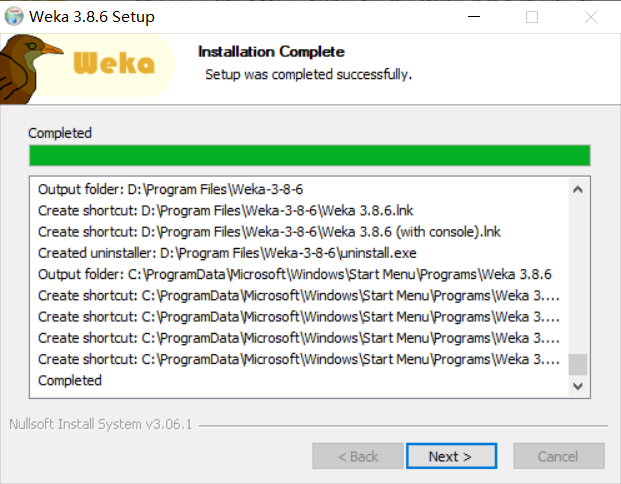
下一步是选择安装路径，我将路径修改成D盘



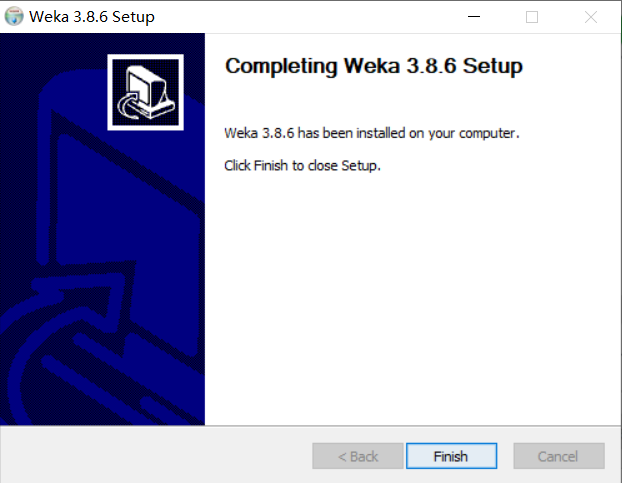
最后就是选择开始菜单文件夹名称，使用默认的就行，点击安装



等待安装



最后安装成功



点击图标进入Weka主界面



2. 解释WEKA应用程序中Visualization菜单、Tools的4个子菜单、Explorer、Experimenter、KnowledgeFlow、Workbench的基本功能

（1）Visualization可视化菜单下有5个选项，分别为：

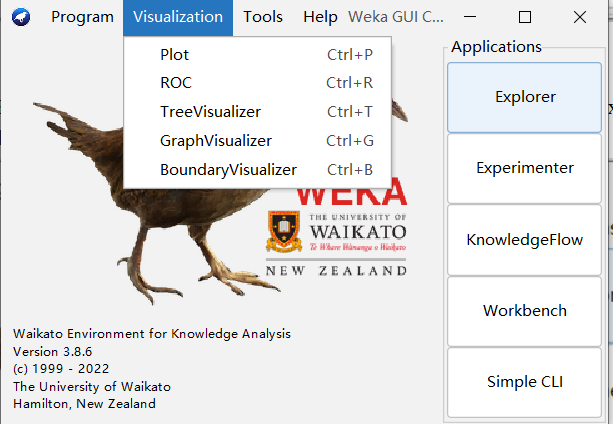
Plot：散点图，绘制数据集的 2D 散点图

ROC：受试者工作特征曲线，显示预先保存的 ROC 曲线

TreeVisualizer：树结构可视化，显示有向图，例如，决策树

GraphVisualizer： 图结构可视化，显示 XML、BIF 或 DOT 格式的图，例如，贝叶斯网络

BoundaryVisualizer：边界可视化，显示二维空间中分类器决策边界的可视化



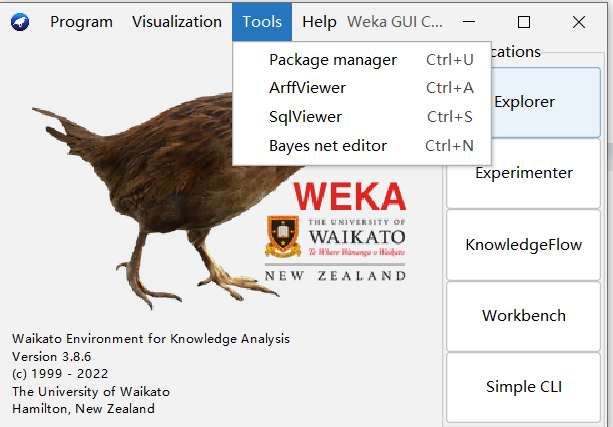
1. Tools工具的4个子菜单分别是：

Package manager：包管理器，Weka包管理系统的图形接口

ArffViewer：ARFF文件查看器，以电子表格形式查看ARFF文件的MDI应用

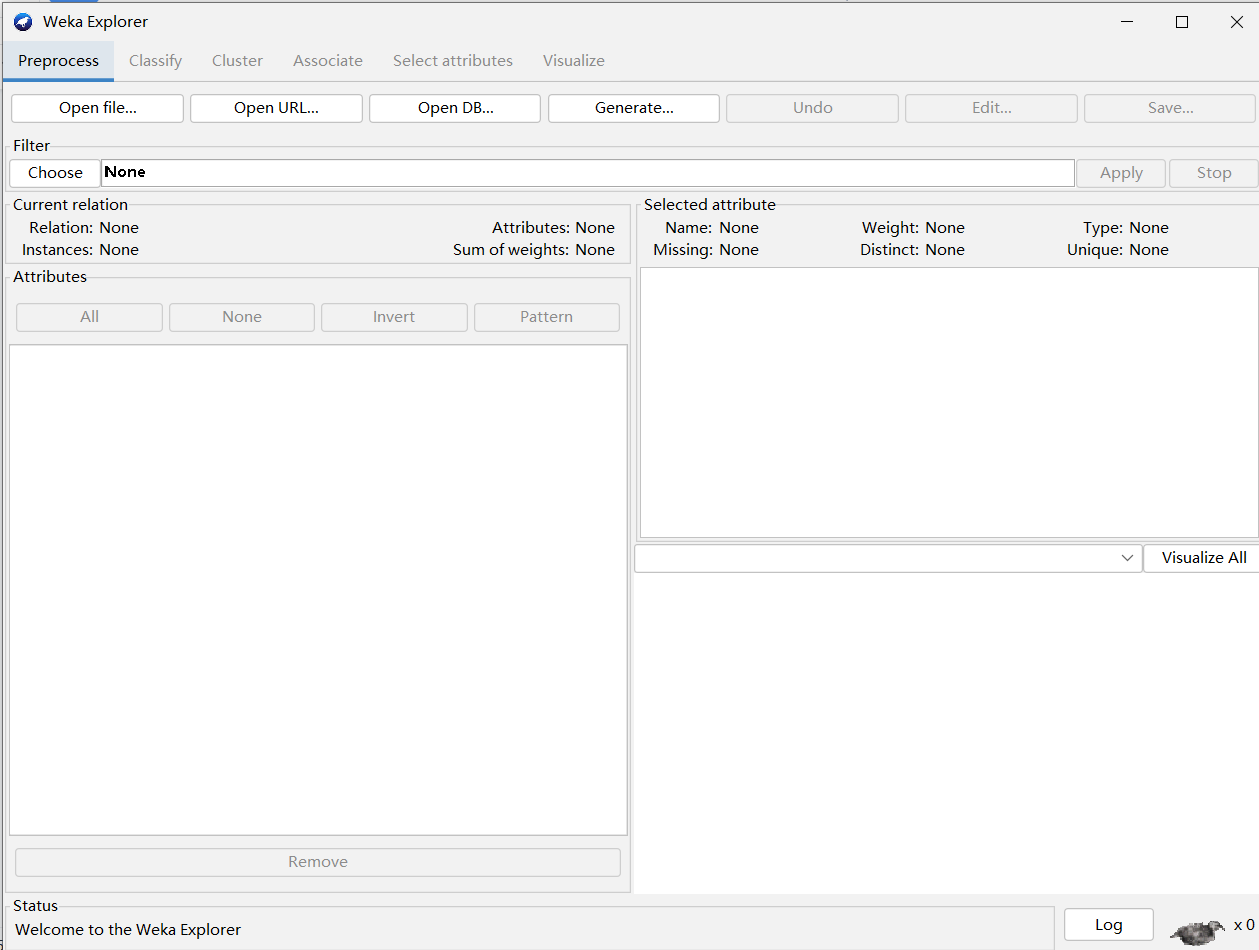
SqlViewer：SQL查看器，一个SQL 工作表单，通过JDBC查询数据库

Bayes net editor： 贝叶斯网络编辑器，一个用于编辑、可视化和学习贝叶斯网络的应用

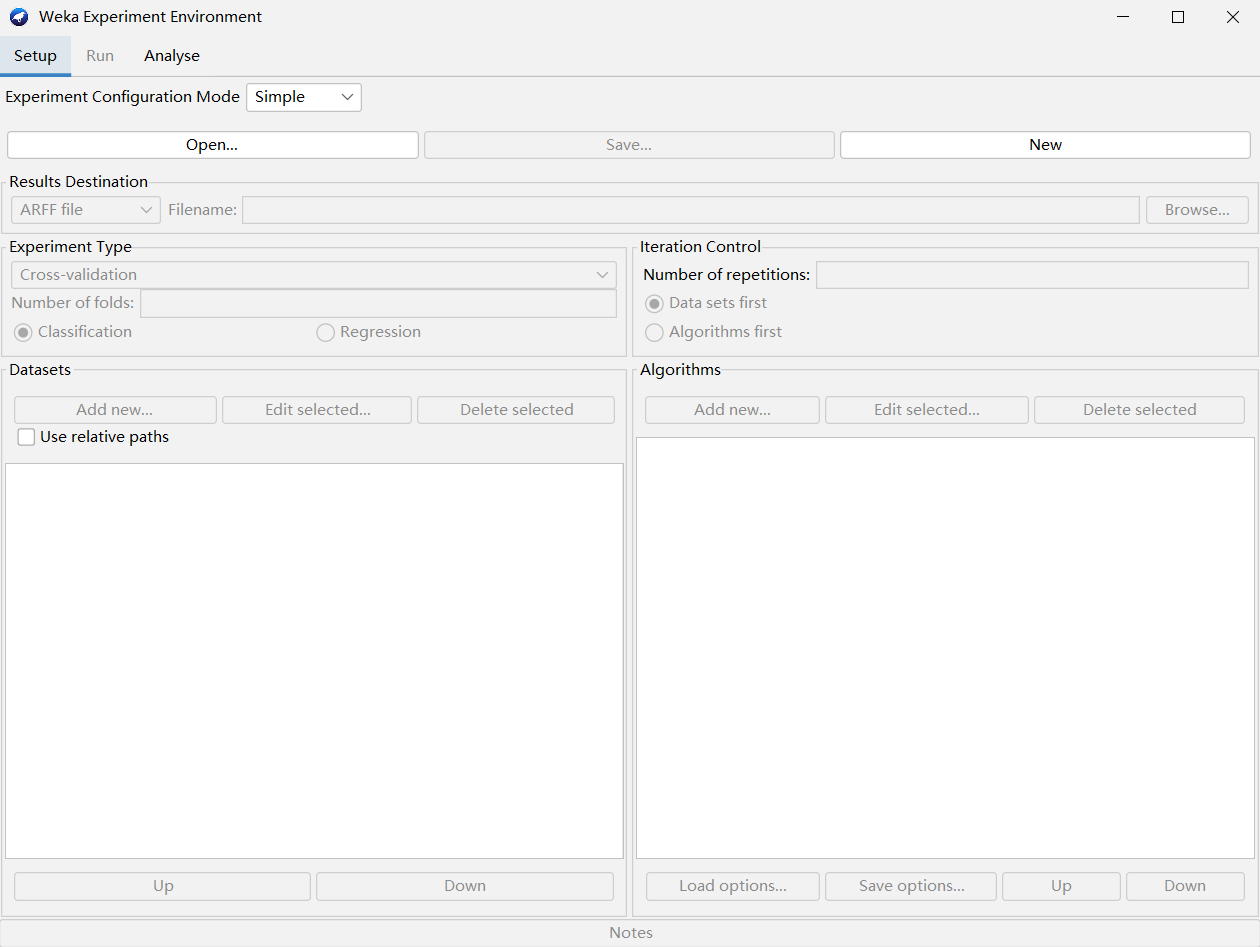


（3）Explorer、Experimenter、KnowledgeFlow、Workbench的基本功能

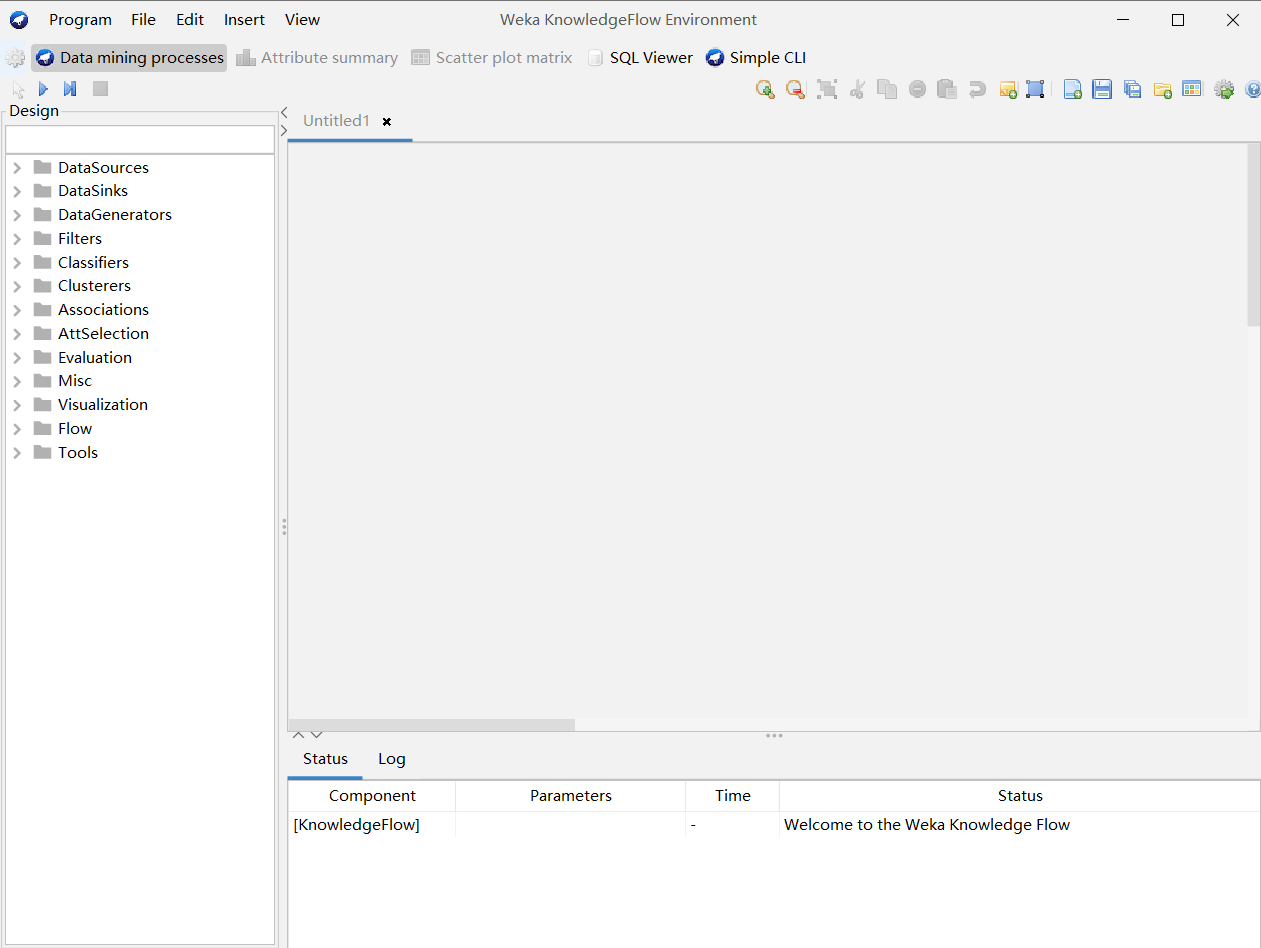
Explorer：探索者界面，用来进行数据实验、挖掘的环境，它提供了分类，聚类，关联规则，特征选择，数据可视化等等功能



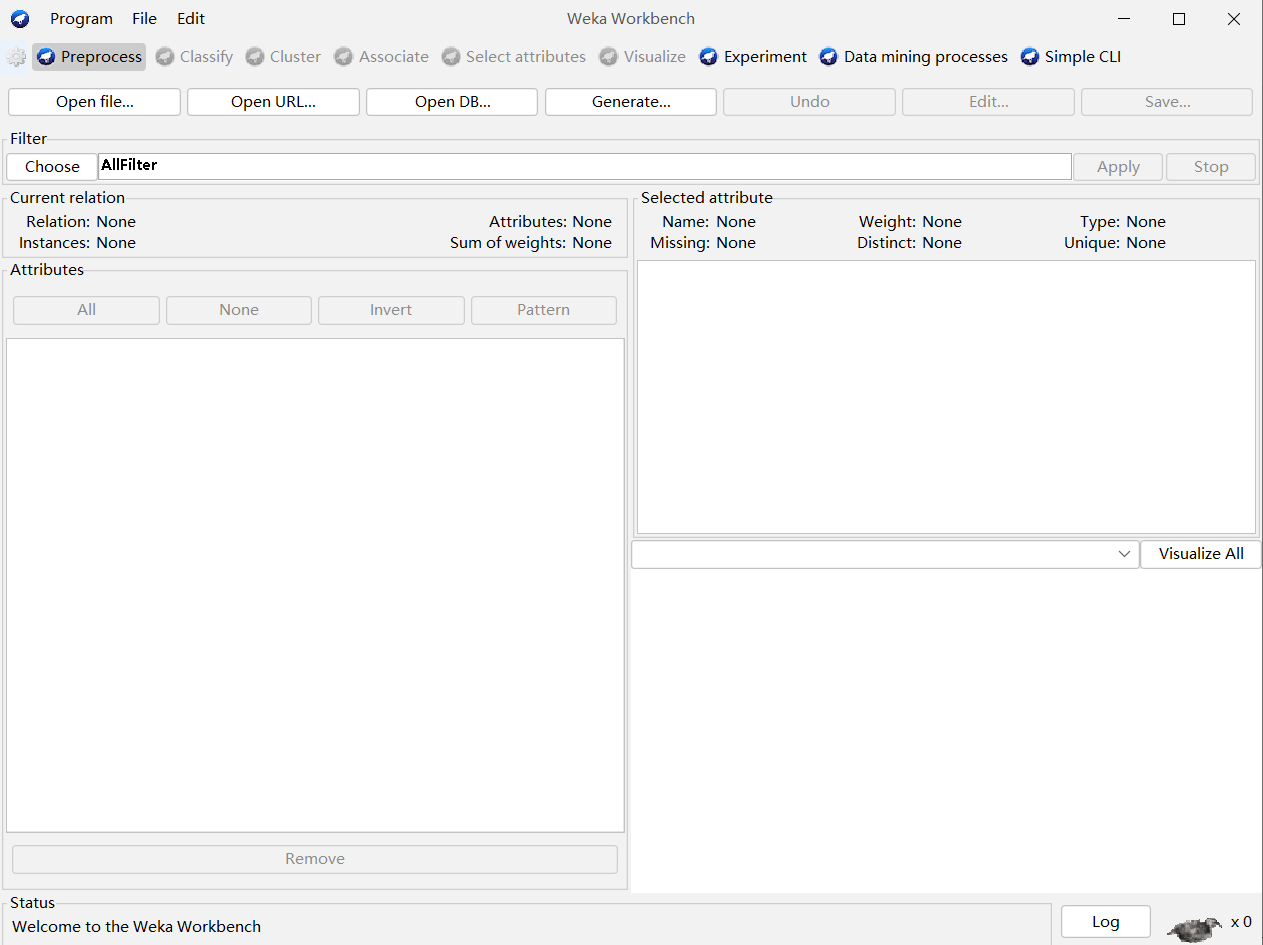
Experimenter：实验者界面，Weka提供的实验者工作环境，用户可以比较不同的学习方案。尽管探索者界面也能通过交互完成这样的功能，但通过实验者界面，用户可以让处理过程实现自动化。实验者界面更加容易使用不同参数去设置分类器和过滤器，使之运行在一组数据集中，收集性能统计数据，实现重要的测试实验



KnowledgeFlow：知识流界面，可以使用增量方式的算法来处理大型数据集，用户可以定制处理数据流的方式和顺序。知识流界面允许用户在屏幕上任意拖曳代表学习算法和数据源的图形构件，并以一定的方式和顺序组合在一起。也就是，按照一定顺序将代表数据源、预处理工具、学习算法、评估手段和可视化模块的各构件组合在一起，形成数据流。如果用户选取的过滤器和学习算法具有增量学习功能，那就可以实现大型数据集的增量分批读取和处理



Workbench： 工作台界面，提供了一个多合一的应用程序，该应用程序包含了前面各节中描述的所有主要的WEKA GUI（Graphical User Interface，图形用户界面）

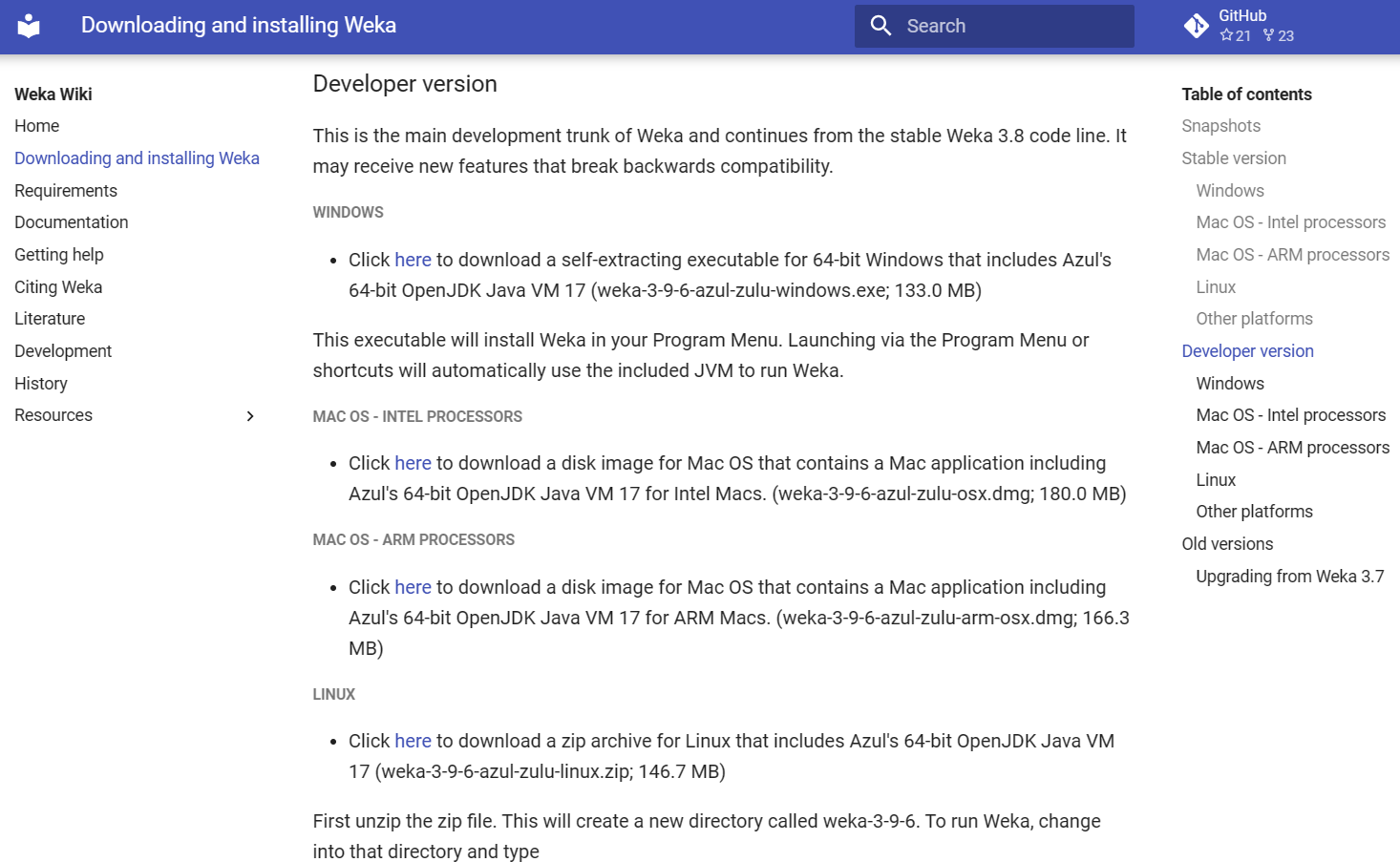


# 三、实验总结

问题1：

1. 问题描述

这里因为是老师提供的安装包，所以安装过程较为顺利，从官网上来看，老师给的应该是开发者版本，如果是其他版本需要安装jdk并配置环境变量，同时配置jdk路径后可能出现闪退的情况



1. 问题分析（可能的原因、难点、挑战）

系统找不到环境变量

1. 解决方案

将配置的jdk和jre放在path路径的最前面