

蓝鲸数据挖掘平台用户手册

版本 v1.1

1	引言	1
	1.1 目的	1
	1.2 平台简介	1
	1.3 适用范围	1
	1.4 功能模块	1
2	快速开始	2
	2.1 进入平台	2
	2.2 注册并登录	2
	2.2.1 邮箱或手机号注册登录	2
	2.2.2 微信扫码登录	3
3	用户操作指南	4
	3.1 软件下载和使用	4
	3.1.1 软件下载	4
	3.1.2 软件操作	4
	3.2 平台资源	5
	3.2.1 平台数据资源	6
	3.2.2 平台案例资源	8
	3.3 自建资源	8
	3.3.1 自建数据资源	8
	3.3.2 自建案例资源	10
	3.4 团队管理	11
	3.4.1 团队创建	11
	3.4.2 用户管理	12
	3.5 资源管理	13
	3.5.1 授权数据	13
	3.5.2 授权案例	14

目录

1 引言

1.1 目的

本手册的目的是对蓝鲸数据挖掘平台的操作进行说明。

1.2 平台简介

蓝鲸数据挖掘平台是基于 Orange 开发的一款机器学习和数据可视化工具, 采用图形化方式建模,内置分类预测、回归预测、聚类分析、关联规则、统计和 降维等算法模型。

1.3 适用范围

本手册适用于所有蓝鲸数据挖掘平台的使用者,包括教师、学生、科研人员 以及有数据分析需求的从业人员。

1.4 功能模块

蓝鲸数据挖掘平台的功能模块如图 1-1 所示。

用户角色 功能模块	游客	免费用户 (注册并登录)	企业用户 (开通企业控制台)
软件下载	使用软件 <mark>基础</mark> 功能 • 数据导入 • 数据预处理 • 数据建模 • 模型评估	使用软件 <mark>所有</mark> 功能 • 数据导入 • 数据预处理 • 数据建模 • 模型评估 • 引用蓝鲸资源	使用软件 <mark>所有</mark> 功能 • 数据导入 • 数据预处理 • 数据建模 • 模型评估 • 引用蓝鲸资源
平台资源	无	使用"蓝鲸"平台 <mark>部分</mark> 资源 • 免费数据 • 免费案例	使用"蓝鲸"平台 <mark>所有</mark> 资源 • 免费数据+专业数据 • 免费案例+专业案例
自建资源	无	无	 研发资源并授权给成员(学生) 研发数据 授权数据 研发案例 授权案例
教学教研	无	无	团队及成员管理 创建团队(班级) 加入或移除成员(学生) 给成员授权资源(数据或案例)

图 1-1 蓝鲸平台功能模块

2 快速开始

2.1 进入平台

用浏览器进入蓝鲸平台官网 https://bw.dashenglab.com, 界面如图 2-1。

🎃 藍鲸	首页产品介绍资源广场下载案例学习手册
	<section-header></section-header>

图 2-1 蓝鲸平台首页

2.2 注册并登录

要使用蓝鲸平台的数据和资源,须注册并登录平台。平台支持两种注册方式: (1)邮箱或手机号注册;(2)微信扫码注册。

2.2.1 邮箱或手机号注册登录

操作步骤:在蓝鲸平台官网导航栏点击【资源广场】,点击右上角【注册】, 如图 2-2;点击【立即注册】,按要求输入信息,注册完成后即可登录,如图 2-3。



图 2-2 进入资源广场注册

Ď	4	20						
密码登录	_	微信扫码						
机号								
机号								
			Ø					
立即登录								
<u>:重置</u> 仰注册								
	密码登录 4机号 4机号 2 重重置 吹注册	 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・					

图 2-3 注册账号

2.2.2 微信扫码登录

操作步骤:进入【资源广场】,点击右上角【注册】,选择【微信扫码】即可快速登录,如图 2-4。

Ď	⇒ 20					
密码登录	微信扫码					
☑ Hey (●), 欢迎进入 ♀ 若已注册,请关联您的账号!	.l I					
建议绑定密码账	号,防止失联更安全					
关联邮	箱或手机号					
稍后关联						

图 2-4 微信扫码登录

3 用户操作指南

3.1 软件下载和使用

蓝鲸软件是一款机器学习与数据可视化工具,它将算法封装在组件中,用户 无需编程,只需通过简单的拖拽操作就可以快速创建数据挖掘工作流。用户可在 蓝鲸官网免费下载蓝鲸软件。

3.1.1 软件下载

操作步骤:点击蓝鲸首页导航栏的【下载】,根据计算机系统选择相应的版本,下载安装即可,如图 3-1、图 3-2。



图 3-1 在平台首页点击"下载"

🂽 蓝鲸	首页 产品介绍 资源广场 下载 案例学习 手册	
	蓝鲸数据挖掘平台	
	Windows Mac	

图 3-2 选择版本

3.1.2 软件操作

蓝鲸软件的核心功能模块包括数据预处理、数据可视化、数据建模、模型评估等,各模块下的详细功能如图 3-3 所示。

Construction Construction<			
数据预处理	数据可视化	数据建模	模型评估
• 数据记录选择	• 直方图	• 回归分析	• 混淆矩阵
• 数据特征选择	 箱线图 	• 决策树	• ROC曲线
• 数据标准化	 树图 	• 逻辑回归	• 误差分析
• 缺失值处理	 散点图 	• 神经网络	
• 降维	 维恩图 	• 随机森林	
• 数据特征编辑	 地图 	• 支持向量机	
• 数据合并	• 序列图	• 时间序列预测	
• 离散化	 螺旋图 	• 关联规则	
• 异常值处理	• 相关图	 聚类 	

图 3-3 蓝鲸软件功能

关于软件详细功能的教程可在蓝鲸网站查看,如图 3-4。

<u></u>	蓝鲸	首页	品介绍	资源广场	下载	案例学习	手册				
	组件手册										
	★主页 / 组件手册 / 可	视化 / 箱线图									
	▼ 组件手册		箱	线图							
	▶ 蓝鲸										
	▶ 数据			±±							
	▼ 可视化			보호							
	箱线图										
	分布图		用图	形展示属性值的	分散情况。						
	散点图		输入								
	马赛克图		90.06	初八致路来							
	▶ 模型		输出	】: 择的数据:在新	街街園中选中	中的实例					
▶ 评估				箱绊图组供可以	「香香粉」な	6公本結:2	从而识别教捉使山的局	◎塗信武重复信	判断教促的信本和雇害	门及查委日纪教程的协信	ь
	▶ 无监督		位数	、尾长、异常值	1、分布区间	间等信息。	7711日年1037年1137年1137月	中中国动业规国。	A THE REPORT OF MELTING AND A THE PARTY OF A THE PA	NAX HE TO A DESCRIPTION	ь т

图 3-4 在蓝鲸平台首页查看软件教程

3.2 平台资源

蓝鲸平台提供物流、电商、供应链等多个行业的数据和案例资源,供教学或 科研使用。平台资源包括免费资源和专业资源,免费资源所有注册用户均可使用, 专业资源需联系销售购买。

3.2.1 平台数据资源

操作步骤:点击蓝鲸首页导航栏的【资源广场】,点击【数据】即可查看数据集列表,并且可以筛选需要的数据类别,如图 3-5。



图 3-5 查看数据集列表

点击任一数据集,进入数据详情页,点击【加入我的】,如图 3-6。

Ô.	藍鲸数据	首页	软件下载	资源广场	我的收藏	我的资源	企业控制台		137180194
				波士顿房价数 醫務 加入我的 ?	y据集 Free	创建时间	J] : 2021-01-28	♥ 数据大小:68	
	数据简介								
	该数据集包含美国人口普查局收集的有关波士	顿地区住房	价格的信息,這	5月于回归任务。数	居集共有 506 个实	例,每个实例包	括13 个特征变量和1个目标变量。		
	数据详情								
	每条数据包含房屋以及房屋周围的详细信息 1.标值,片约记罪率; 2.占地面积超过25,000平方英尺的住宅用地 3.每个城镇非零售业务的比例; 4.Charles River虚拟变量(如果是问道,则 5.一氧化氮浓度(每千万份); 6.每间住宅的中均问题; 7.1940年以前建造的自住单位比例; 8.加以取画版注锁的五个就业中心;	1。变量信息 比例; 为1;否则为0)	如下:);						

图 3-6 将数据集"加入我的"

之后打开蓝鲸软件,点击【蓝鲸】模块的【数据】组件并登录,如图 3-7。

▲ *	编辑 视图 组件 选项 帮助 账号		
30	蓝鲸		
1		■ 蓋輸浏览器 — □ ×	
۴	а ж. 1)HT	
	数据		
	可视化		
-	模型	密码登录 微信扫码	
×× ×	评估		
*	无监督	* 输入邮箱或手机号	
	图像分析	输入邮箱或手机号	
~	时间序列		
	生物信息学	* 请输入密码	
	关联分析	输入密码	
		立即登录	
		2	
			~
14.42	人坦林以且二甘光明		
石烊-	"T租什以亚不共说明。 丁作遗示例 或打开 古德市		
	ALTERNATION PROVIDENT		

图 3-7 点击"数据"组件并登录

登录后双击打开画布上的【数据】组件即可使用已加入的数据资源,需要选择数据集和数据表,如图 3-8。



图 3-8 打开"数据"组件选择数据集

至此已完成了数据的导入。"数据"组件输出端可以连接到数据表格、特征 选择、缺失值处理等组件,具体的软件操作教程可在蓝鲸网站查看。

3.2.2 平台案例资源

操作步骤:点击蓝鲸首页导航栏的【资源广场】,点击【案例】即可查看案 例资源列表,并且可以筛选需要的案例类别,点击需要的案例并【加入我的】, 此处和 3.2.1 步骤相似。

之后打开蓝鲸软件,点击【蓝鲸】模块的【案例】组件并登录。登录后双击 打开画布上的【案例】组件,选择需要查看的案例并点击【下载】。和数据资源 不同的是,蓝鲸软件中打开的案例是工作流的形式,如图 3-9。



图 3-9 打开"案例"组件选择需要的案例进行下载

3.3 自建资源

蓝鲸资源广场除了平台资源外还支持用户上传自建资源,可以方便地将自己的数据或案例上传到平台,并授权给他人使用(例如教师授权给学生以实现教学目的)。数据研发功能支持上传.sqlite、.csv、.xlsx格式的数据文件,案例研发功能仅支持.ows格式的工作流文件。

3.3.1 自建数据资源

操作步骤:点击【资源广场】-【企业控制台】,点击团队名称进入团队,点击左侧【教研管理】下的【数据研发】,点击【新建数据集】,如图 3-10;之后根据提示创建数据集并完善数据集信息即可,如图 3-11、图 3-12。

★ > > > > > > > > > > > > > > > > > >									
□ 秋研管理 ^ ● 数据研发 新建者									
数据研发									
案例研发	# 数据名称	数据大小	类别	上传日期	操作				
	1 111			2021-04-12	(失效				

图 3-10 数据研发

会 蓝鲸数排	E	救研管理 / 姜	如居研发				
回 教研管理	^	2 数据	研发				新建数据集
数据研发					上传数据	×	
家例研发		#	数据名称	数据大小			操作
目 资源管理	~	1	111		* 第一步, 创建数据		
み 用户管理	~	2	333		间初日八支江站台村小		
		3	11111		T-#	完美教展 取 当	0.953
		4	11111			2021104114	()#X)

图 3-11 创建数据

🎃 藍鲸数据	¢	研管理 / 数据研发	19
回 教研管理	^	■ 数据研发 / 1111	
数据研发			
案例研发		 数据名称: 1111 	
回 资源管理	~	数据文件: 占 请上传sqlite/.ssv/.slsz文件	
み 用户管理	~	质量评级:	
		主题标签: 金融 医疗与健康 物流 电子商务 经济 交通运输 零售 汽车 供应链 社会公共 生产制造 其他	
		数据商介:	

图 3-12 完善数据集信息

点击【保存】,提示创建成功,即完成了数据集的创建,如图 3-13。

教研管理 / 数据研发		❷ 修改成功		1 8 😣
€数据研发 /	波士顿房价			参改 删除
数据名称:	波士顿房价	数据文件:	下载数据占	
质量评级:	****	数据简介:	11111111111	
主题标签:	经济	资源来源:	免费	
封面logo:				
详细描述:	22222222222			

图 3-13 保存创建的数据集

3.3.2 自建案例资源

操作步骤:点击【资源广场】-【企业控制台】,点击团队名称进入团队,点击左侧【教研管理】下的【案例研发】,点击【新建案例】,如图 3-14;之后根据 提示创建案例并完善案例信息即可,如图 3-15。

▲ 藍鲸数据	教	研管理/案例研发				
回 教研管理	^	🗙 案例研发				新建案例
数据研发		4	소드ril	上法口题	+9.//-	
案例研发		# 条例石标	衆加		SHI F	

图 3-14 教研管理-案例研发-新建案例

·	居	教研管理 / 案例研发		
回 教研管理	^	★ 案例研发	_	新建率例
数据研发		a an third and	创建案例	×
案例研发		# 案例名称	· 第─步, 创建案例	Berr-
回 资源管理	~	1 222	请输出案例名称	
& 用户管理	~	2 444		3035
			下一步,完	語案例 取消

图 3-15 教研管理-案例研发-完善案例信息

点击【保存】,提示创建成功,即完成了案例的创建,如图 3-16。

Ç,	藍鯨数据	4	教研管理 / 案例研发		♥ 案例修改成功		13718019418
Ľ	平台管理	×	★ 案例研发 /	基于时间序列的分仓商品需求预测			修改 删除
	教研管理	^	案例名称:	基于时间序列的分仓商品需求预测	案例文件:	下載案例と	
	案例研发		质量评级:	****	案例简介:	国内某大型电商平台积累了海量的交易数据,通过数据挖掘 技术,能够为商家提供供应链决策支持,大幅降低运营成	
	资源管理	^				中,使开用户体验。本条例以过去来一年的海里关系行为或 据以及商品信息数据为依据,运用时间序列ARIMA模型,精 准刻画商品需求的变动规律,对未来一周的全国和区域分仓 每日票本电量进行研测	
	授权数据				Maximum de Mari	- 190001 - 1900001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001 - 190001	
	授权案例		王懿标益:	智尤	資源米源:	专业	
\$	资源发布	^	封面logo:				
	数据发布						
	案例发布						
R	用户管理	Ŷ	详细描述:	1.本案例所用到的预测方法是自回归移动平均(Af	RIMA)模型,以及会用到	评价指标选取和使用的相关理论基础;	
2	团队审核	~		 2.本案例要分析的数据是监察数据挖掘系统目带的 求,进行特征选择。 	数据资源——"智能分仓」	tem_feature2 数据表。需要对原始数据进行清洗,结合本任务的需	
				 3.首先筛选出畅销商品,通过数据可视化的方式(4.通过自相关系数以及偏自相关系数的变化趋势, 	序列图等)观测商品需求 以及差分的方式实现序列	的变化趋势,判断序列的平稳性; 的平稳化处理;构建商品需求预测ARIMA模型。	
4	ē	•		 5.根据构建好的模型对某款商品进行未来一周(75) 6.通过多种途径,观察预测精度,判定模型的预测 	天)的需求量预测。 性能。		

图 3-16 保存创建的案例

3.4 团队管理

平台资源和自建资源都可以授权给其他用户使用,例如教师授权给班级的学生,实现教学和教研的目的。蓝鲸平台的企业用户可通过企业控制台实现团队管理和资源管理的功能。

3.4.1 团队创建

若用户为教师类别,教师便可申请创建团队,作为团队指导员的身份邀请学 生加入团队,并对学生进行管理。

操作步骤:点击【资源广场】-【企业控制台】-【申请开通企业控制台】-【创建新团队】,创建成功后需要选择创建好的团队并点击【申请加入】,申请完 成后等待审核即可,如图 3-17 至 3-19。

۵.	蓝鲸数据	首页	软件下载	资源广场	我的收藏	我的资源	企业控制台
	您尚未开通任何企业						
	在企业版中, 您可以:						
	1) 企业版拥有更专业的案例库和教学资源;						
	2) 适应团队管理,可灵活共享案例授权;						
	3) 根据教研需求,可灵活研发自有案例;						
	4) 综合管理团队成员以及学员信息。						
	申请开通企业控制台						

图 3-17 开通企业控制台

▲	首页	软件下载	资源广场	我的收藏	我的资源	企业控制台
请选择您的团队身份						
			♀ 没有团队?去创	建新团队		
		*团队名称:	请选择您要开通的	的团队	~	
		手机号码:	请输入您的手机等	码		
		申请理由:	试用 已购买	其他		
			申	青加入		

图 3-18 完善团队信息

·	首页	软件下载	资源广场	我的收藏	我的资源	企业控制台
请选择您的团队身份						
			♀没有团队?去	创建新团队		在这里选择刚
	l	*团队名称:	蓝鲸小队		~	刚创建的团队
		手机号码:	13800001234			
		申请理由:	试用 已购	买 其他		点击申请加入
			Ħ]请加入		

图 3-19 申请加入创建的团队

3.4.2 用户管理

创建团队审核完成后,团队指导员可以添加其他教师和学员加入团队。

操作步骤:点击【资源广场】-【企业控制台】,单击团队名称,点击【用户 管理】-【成员管理】-【团队中导入用户】-【管理团队用户点击此处前往】,进 入界面后便可添加教师或成员,如图 3-20、图 3-21。

图 3-20 从团队导入用户

从团队中导入用户	×
如果要管理团队用户请点击此处前往	
请输入用户名称进行搜索	Q
教师学员	批量添加
名称	

图 3-21 添加用户

3.5 资源管理

资源管理指的是用户在创建自己的团队后,可为团队成员授权数据和案例资 源的使用权限。

3.5.1 授权数据

操作步骤:点击【资源管理】-【授权数据】,单击想要授权的数据集,如图 3-22 所示;点击【授权用户】,选择授权团队成员的姓名,点击【授权】,团队成员即可使用该数据,如图 3-23、图 3-24。

註								
□ 教研管理	~	授权数据						
□ 资源管理	^						_	
授权数据		# 数据名称	数据大小	类别	授权数量	到期日期		
授权案例		1 银行营销数据集	3.89M	金融	2/5	2022-04-12	Â	

图 3-22 资源管理-授权数据集

🏊 藍鲸数排	8 3	意源管理 / 授权数据			
■ 教研管理	~	银行营销数据集 / 授权查看			授权用户取
■ 资源管理	^		5.5		
授权数据		姓名	角色	授权日期	到期日期
授权案例				暂无数据	

图 3-23 资源管理-授权数据集-授权用户

银行营销数据集 授权	输入昵称查询 Q X
姓名	角色
1908524092	教师
15333866021	学生
< 1 >	第1-2条/共2条记录
*勾选列表中的用户进行授权	授权取消

图 3-24 选择授权用户-授权

3.5.2 授权案例

操作步骤:点击【资源广场】-【企业控制台】,单击团队名称,点击【资源 管理】-【授权案例】,单击授权的数据集,点击【授权用户】,如图 3-25、图 3-26; 选择要授权的用户,点击【授权】,如图 3-27。

🍌 蓝鲸数据	201	S調管理 / 招	受权案例			
回 教研管理	~	授权案例	ภ			
■ 资源管理	^					
授权数据		#	案例名称	类别	授权数量	到期日期
授权案例		1	基于关联规则的超市顾客购物行为 分析	零售	0/5	2022-04-12

图 3-25 资源管理-授权案例

🏊 蓝鲸数	据	资源管理 / 授权案例			
□ 教研管理	~	基于关联规则的超市顾客购物行为分析 / 授权直看 授权用户			
■ 资源管理	^				
授权数据		姓名	角色	授权日期	到期日期
授权案例				暂无数据	

图 3-26 资源管理-授权案例-授权用户

基于关联规则的超市顾客	购物 输入昵称以查询 Q X
□ 姓名	角色
1908524092	教师
15333866021	学生
< 1 >	第1-2条/共2条记录
*勾选列表中的用户进行授权	授权取消

图 3-27 选择授权用户-授权