



BRICS
Business Council



2023

金砖国家职业技能大赛 (金砖国家未来技能挑战赛)

省级选拔赛样题 (第一套)

BRICS-FS-26_人工智能计算机视觉应用

2023 年 5 月

目录

1 参赛形式.....3

2 竞赛内容.....3

3 项目模块.....3

 3.1 项目模块.....3

 3.2 任务内容.....4

模块一：人工智能环境搭建（15 分）4

 第 1 题：Python 开发环境搭建4

 第 2 题：人工智能开发环境搭建5

模块二：数据处理与分析（20 分）6

 第 3 题：图像数据清洗6

 第 4 题：图像预处理及可视化7

模块三：机器学习（30 分）8

 第 5 题：鸢尾花分类8

 第 6 题：房价预测10

模块四：计算机视觉应用（35 分）11

 第 7 题：图像识别之数字识别11

 第 8 题：目标检测之安全帽佩戴检测13

1 参赛形式

个人赛（1 人）。

2 竞赛内容

竞赛内容由 4 个模块组成，模块之间相互独立，模块错误不传递。

竞赛内容包含以下任务模块：

- 1) 人工智能环境搭建
- 2) 数据处理与分析
- 3) 机器学习
- 4) 计算机视觉应用

如果参赛选手不遵守职业健康安全环境要求，或使自己和其他选手面临危险，他们可能会被取消比赛资格。

参赛者完成竞赛后，由裁判组组织对选手提交结果进行评分。

3 项目模块

3.1 项目模块

人工智能计算机视觉应用赛项共 4 个模块，要求选手在 4 个小时内完成。具体项目模块名称和时间要求参照下表

模块	模块名称	权重	竞赛时长
模块一	人工智能环境搭建	15%	4h
模块二	数据处理与分析	20%	
模块三	机器学习	30%	
模块四	计算机视觉应用	35%	

3.2 任务内容

模块一：人工智能环境搭建（15 分）

第 1 题：Python 开发环境搭建

【功能说明】

在 Ubuntu 虚拟机中搭建 Python 所需基础开发环境，虚拟机中已安装 python 环境和包管理工具 pip，相关软件及依赖库的安装包放在虚拟机“/opt”目录下。

【任务要求】

本环节需要在 Ubuntu 系统中安装 Python 基础开发环境，安装完成后需根据答题区要求截图，并将截图保存到答题区对应位置。环境搭建具体要求如下：

1、根据提供的安装包安装 Python 环境，配置环境变量，在 Terminal 终端模拟器中输入“python”验证 python 安装成功。

2、根据提供的安装包安装 PythonIDE 工具 PyCharm，安装到虚拟机“/opt/pycharm2021”目录下，若目录不存在需自行创建。

3、根据提供的安装包安装 MySQL 数据库，安装到虚拟机“/opt/mysql”目录下，若目录不存在需自行创建，配置环境变量完成数据库安装，设置数据库账号为 root，密码为 123456。安装完成后在 Terminal 终端模拟器登录 MySQL 服务器。

【操作说明】

1、所有实操均在平台的虚拟机（Ubuntu 系统）中进行，虚拟机

内为断网环境。

2、本环节需要用到的安装包都存放在虚拟机“/opt”目录下，自行判断依赖库完整顺序完成环境搭建。

3、完成任务后，根据平台“答题区”对应内容要求，完成答题报告并点击“保存”按钮实时保存。

第 2 题：人工智能开发环境搭建

【功能说明】

Anaconda 是一个可用于科学计算的 Python 发行版，集成了非常多的第三方库（如 numpy、pandas、matplotlib）、可以根据不同的人工智能项目创建并管理不同的虚拟环境。

在 Ubuntu 虚拟机中安装 Anaconda 进行虚拟环境管理。虚拟机中已安装 python 环境和包管理工具 pip，相关软件安装包放在虚拟机“/opt”目录下。

【任务要求】

1、通过提供的安装包，将 Anaconda 安装到“/opt/anaconda3”目录下，若目录不存在则自行创建。

2、安装完成后，配置环境变量，在 Terminal 终端模拟器中验证并显示 Anaconda 版本信息。

3、使用 Anaconda 创建虚拟环境，并命名为“tf_env”，虚拟环境“tf_env”创建完成后，在 Terminal 终端模拟器中验证并显示虚

拟环境列表。

4、激活虚拟环境“tf_env”，在该环境中安装 TensorFlow 深度学习框架及其依赖库，完成搭建后，验证并显示当前虚拟环境中的软件及版本信息。

5、打开 Pycharm 编辑器，新建项目命名为“Project”并关联虚拟环境“tf_env”，在项目下创建 python 文件命名为“hello_tf”，在该 python 文件中导入 TensorFlow，并打印出版本信息。

【操作说明】

1、所有实操均在平台的虚拟机（Ubuntu 系统）中进行，虚拟机内为断网环境。

2、Anaconda 的安装包存放在虚拟机“/opt”目录下。

3、完成任务后，根据平台“答题区”对应内容要求，完成答题报告并点击“保存”按钮实时保存。

模块二：数据处理与分析（20 分）

第 3 题：图像数据清洗

【功能说明】

数据的质量很大程度决定了后续模型训练的准确性，对图像数据集进行清洗，输出清晰完整、符合标签类别的图片。

【任务要求】

读取所需数据集后，去除无法加载的异常图像、模糊图片、相似

图片、删除单通道图像等。



图 1：模糊图片示意图

【操作说明】

1、所有实操均在平台的虚拟机（Ubuntu 系统）中进行，虚拟机内为断网环境。

2、从虚拟机桌面“赛题/02_数据处理/03_图像数据清洗”路径下获取相关资料，结果保存至虚拟机桌面“提交文档/02 ”文件夹中。

3、完成任务后，根据平台“答题区”对应内容要求，完成答题报告并点击“保存”按钮实时保存。

第 4 题：图像预处理及可视化

【功能说明】

对图像数据集进行预处理操作，使得各种类图片数量分布均匀合理，符合要求。

【任务要求】

1、读取所需数据集后，对图像进行数据二值化、增强处理等。

2、根据给定的数据集，通过简单的图表展示数据集处理后的分布情况，并将图表通过 Django 框架渲染到指定页面。

【操作说明】

1、所有实操均在平台的虚拟机（Ubuntu 系统）中进行，虚拟机内为断网环境。

2、从虚拟机桌面“赛题/02 数据处理/04_图像预处理”路径下获取相关资料，结果保存至桌面“提交文档/02 ”文件夹中。

模块三：机器学习（30 分）

第 5 题：鸢尾花分类

【功能说明】

使用机器学习框架 sklearn 编写程序，实现鸢尾花的分类项目。

【任务要求】

在 machine_learning_iris.py 文件中实现鸢尾花分类功能，其中 load_data() 函数用于加载数据集，具体要求如下：

1、从 sklearn 对应库中直接导入鸢尾花数据集，训练集和测试集按 8:2 进行划分。

2、判读并选择合适的算法构建分类器。

3、使用训练集进行模型训练，使用测试集进行模型验证。

4、打印模型准确率在训练集及测试及上的准确率。

5、使用可视化库绘制散点图展示模型效果，要求传入参数为定

义好的分类器及待使用的数据集。

①、鸢尾花数据共有四个特征，选取‘花萼长度’和‘花萼宽度’作为两个特征值，标签值不变，重新训练一个模型。

②、绘制散点图展示模型划分结果（x 轴标签为花萼长度，y 轴标签为花萼宽度，标签字体大小为 13，横纵坐标轴的刻度范围均为对应数据的最小到最大值），点的大小设置为 30，颜色定义为 'g','b','r'。其余参数自行指定。

③、设置标题为‘鸢尾花分类’

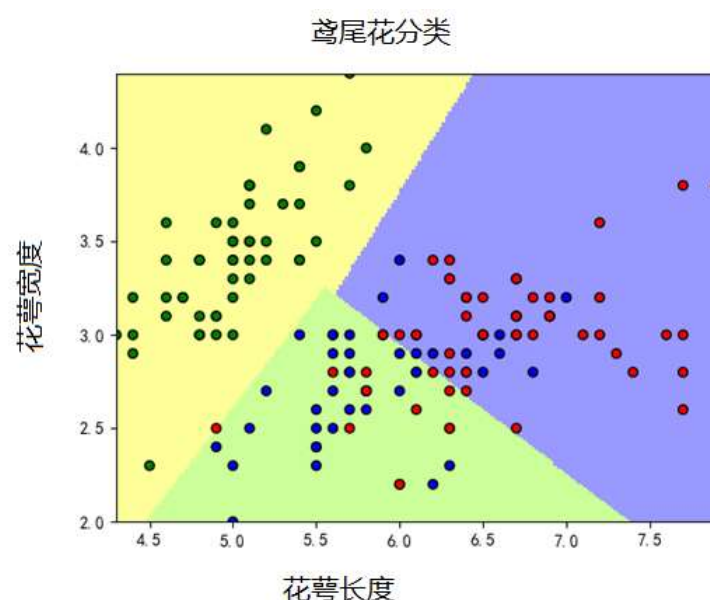


图 2：鸢尾花分类示意图

6、将图表通过 Django 框架渲染到指定页面。

【操作说明】

1、所有实操均在平台的虚拟机（Ubuntu 系统）中进行，虚拟机内为断网环境。

2、从虚拟机桌面“赛题/03_机器学习/05_鸢尾花分类”路径下获取相关资料，结果保存至虚拟机桌面“提交文档/03”文件夹中。

3、完成任务后，根据平台“答题区”对应内容要求，完成答题报告并点击“保存”按钮实时保存。

第 6 题：房价预测

【功能说明】

使用机器学习框架 sklearn 编写程序，实现房价预测项目。

【任务要求】

在 machine_learning_price.py 文件中实现房价预测功能，其中房价数据集存放在 house.csv，具体要求如下：

1、对 house.csv 数据进行数据字段分析，判断并选择合适的策略进行相关处理，如数据中存在的缺失值、异常值、数据标准化等，处理方式不限。

2、对 house.csv 数据中的字段进行相关性分析，计算不同数据特征之间的相关性，如皮尔逊系数，并选择合适的字段作为模型的数据特征。

3、基于机器学习框架 sklearn 进行数据划分，训练集和测试集划分比例为 8:2。

4、判读并选择合适的算法进行机器学习建模。

5、使用构建好的模型对测试集进行模型验证，并绘制模型预测值与真实值的折线对比图，将图表通过 Django 框架渲染到指定页面。

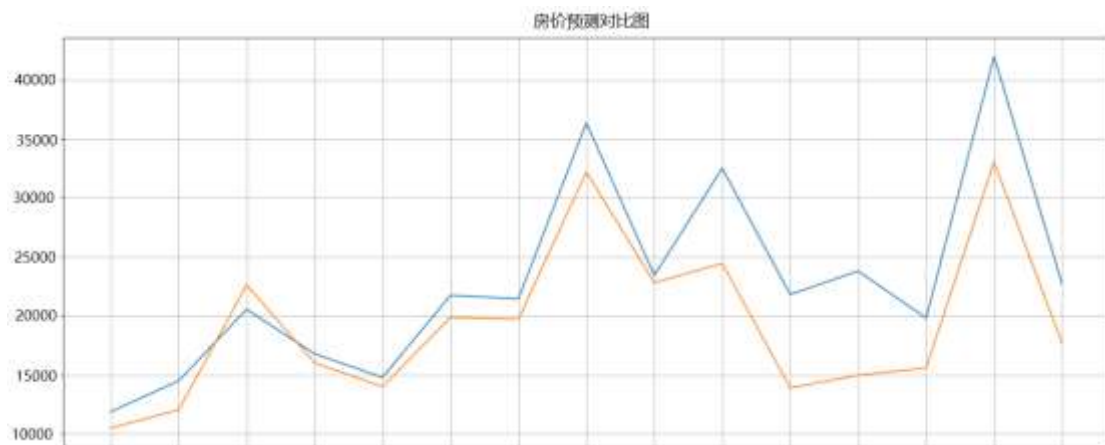


图 3：房价预测对比图

【操作说明】

- 1、所有实操均在平台的虚拟机（Ubuntu 系统）中进行，虚拟机内为断网环境。
- 2、从虚拟机桌面“赛题/03_机器学习/06_房价预测”路径下获取相关资料，结果保存至虚拟机桌面“提交文档/03”文件夹中。
- 3、完成任务后，根据平台“答题区”对应内容要求，完成答题报告并点击“保存”按钮实时保存。

模块四：计算机视觉应用（35 分）

第 7 题：图像识别之数字识别

【功能说明】

完成 LeNet-5 模型网络架构设计，编写代码，使用 TensorFlow 实现模型搭建，实现手写数字集分类，对数据进行数据清洗，数据预处理等操作，然后利用处理后的数据，完成模型的训练，实现数字图

像识别任务。

【任务要求】

在 num_train.py 文件中编写代码，实现图像数字识别任务，具体要求如下：

1、编写 load_data() 函数读取数据，进行数据增强，数据集划分操作：其中数据增强方式不限，数据训练集与测试集划分比例为 8:2。

2、编写 build_model() 函数构建深度学习模型和编译，要求模型在训练集的准确率不低于 95%，并将最终模型保存为 “model.h5”。

3、编写 plt_train() 函数将训练过程进行可视化展示，包括损失值变化和准确率变化，并将可视化图表通过 Django 渲染展示。

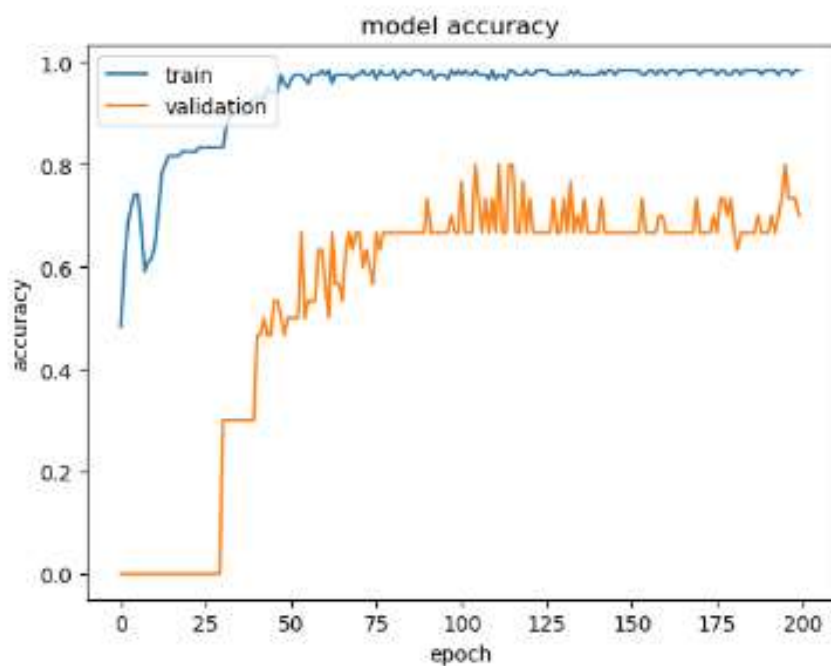


图 4：准确率变化示例图

4、调用保存好的模型实现图像识别功能。

【操作说明】

1、所有实操均在平台的虚拟机（Ubuntu 系统）中进行，虚拟机

内为断网环境。

2、从虚拟机桌面“赛题/04_计算机视觉应用/07_图像识别”路径下获取相关资料，结果保存至虚拟机桌面“提交文档/04”文件夹中。

3、完成任务后，根据平台“答题区”对应内容要求，完成答题报告并点击“保存”按钮实时保存。

第 8 题：目标检测之安全帽佩戴检测

【功能说明】

在多种工作场景中都需要佩戴安全帽进行头部防护工作，使用 TensorFlow 或 Keras 构建深度学习模型实现目标检测任务，根据任务提供数据集，检测出图片中人物是否佩戴安全帽并标框显示。

【任务要求】

在 `object_train.py` 中编写代码，实现安全帽佩戴检测功能，具体要求如下：

1、查看数据集，对数据集进行相应操作，如数据集划分、数据增强、数据尺寸设置等，方式不限。

2、选择适合的深度学习算法，构建深度学习模型，定义模型超参数、选择合适的模型评估函数、激活函数等。

3、进行模型训练并调优，使得模型在测试集上的准确率达到 70% 以上。

4、使用 Tensorboard 查看模型训练可视化过程。

5、构建模型预测函数，输入测试集查看模型的预测结果，并通过 Django 渲染结果展示。



图 5：安全帽佩戴检测示意图

【操作说明】

1、所有实操均在平台的虚拟机（Ubuntu 系统）中进行，虚拟机内为断网环境。

2、从虚拟机桌面“赛题/04_计算机视觉应用/08_目标检测”路径下获取相关资料，结果保存至虚拟机桌面“提交文档/04 ”文件夹中。

3、完成任务后，根据平台“答题区”对应内容要求，完成答题报告并点击“保存”按钮实时保存。

2023

金砖国家职业技能大赛 (金砖国家未来技能挑战赛)



金砖职赛微信号