

附件 2

2022 年广东省人工智能开发员职业技能竞赛 技术工作文件

2022 年广东省人工智能开发员职业技能竞赛组委会
2022 年 10 月

目 录

一、技术描述.....	3
二、试题与评判标准.....	5
三、评分流程及考核细则.....	8
四、竞赛场地、设施设备安排.....	8
五、工具材料安排及清单.....	9
六、项目特殊说明.....	9
七、安全健康和防疫要求.....	10

一、技术描述

（一）项目概要

人工智能训练师是使用智能训练软件，在人工智能产品实际使用过程中进行数据库管理、算法参数设置、人机交互设计、性能测试跟踪及其他辅助作业的人员。

近年来，新冠疫情时有起伏，为了减少感染风险，佩戴口罩已经成为人们进出公共场所、办公区域的必备防护措施。因此，需要有效的识别系统用于检查人在被管制的区域是否戴口罩。为了完成这项任务，训练深度学习模型需要大量的遮挡人脸数据集来检测。本次比赛采用了某个园区门口采集到的人脸数据，作为训练数据集。

本次比赛对选手的技能要求主要包括：掌握人工智能基础算法和相关理论知识；熟悉人工智能算法开发流程；能够使用 **OpenCV** 的 **Python** 版本及其他 **Python** 基础工具库，实现数据读取和可视化、数据集划分、数据基础处理以及数据增强等常用操作；通过对模型结构的分析，能够使用 **TensorFlow** 深度学习框架定义卷积神经网络并输出算法结构；通过定义基础操作，实现模型训练、参数调优、结果保存等步骤；利用训练好的模型完成算法测试及结果输出，并能解决实际任务。

大赛设置职工组和学生组，赛题内容均参考人工智能训练师职业技能三级标准，以个人赛方式由参赛选手独立完成考核内容。

(二) 基本知识与能力要求

相关要求		权重比例 (%)
1	数据预处理	30
基本知识	<ul style="list-style-type: none"> ◆数据预处理的必要性 ◆数据预处理的方法选择 ◆数据预处理的常用操作 	
工作能力	<ul style="list-style-type: none"> ◆读取、分析、可视化数据 ◆划分数据集 ◆实现数据基础处理 ◆实现数据增强方法 	
2	算法定义及训练	50
基本知识	<ul style="list-style-type: none"> ◆分析当前业务场景的特点 ◆选择算法模型结构 ◆优化或调整相应算法 ◆自定义超参数并实现算法迭代训练 	
工作能力	<ul style="list-style-type: none"> ◆模型结构定义 ◆模型结构输出 ◆超参数设置 ◆启动算法训练 ◆分析算法训练效果 ◆根据模型表现进行优化 	
3	算法测试及结果输出	20
基本知识	<ul style="list-style-type: none"> ◆使用训练好的算法进行结果测试 ◆测试结果通过量化指标和可视化的形式进行展示 ◆根据算法测试结果分析可优化的步骤 	
工作能力	<ul style="list-style-type: none"> ◆加载权重并启动算法测试 ◆根据业务场景，输出量化指标 ◆测试结果可视化 ◆结果分析及调优 	
合计		100

二、试题与评判标准

(一) 试题范围

模块内容	评分点	评分描述	评价内容
数据预处理 (30%)	1	读取数据并输出可视化结果，包括原图及标准框	可视化指定数据
	2	将数据集原图和标签形成对应关系列表	打印输出原图和标签的对应列表
	3	按比例划分训练集测试集，并打乱图片顺序	分别打印输出训练集和测试集的混序列表
	4	数据基础处理，主要涉及尺寸裁剪、平移、旋转、缩放、镜像	可视化处理后的数据
	5	数据增强处理，主要涉及灰度变换、色彩抖动、模糊处理、多场景模拟	可视化处理后的数据
算法定义与训练 (50%)	6	根据模型结构图，定义卷积神经网络中的部分操作	检查网络定义方式是否正确
	7	输出模型结构	打印输出模型结构
	8	超参数配置	检查超参数配置是否合理
	9	基础操作定义	检查基础操作定义是否正确
	10	训练模型，输出迭代结果	打印输出每次迭代结果
算法测试及结果输出 (20%)	11	保存权重文件	检查是否存在权重文件
	12	读取权重文件	检查读取方式是否正确
	13	执行测试程序并输出结果	可视化测试结果
	14	利用训练好的模型完成推理任务	输出推理任务结果

(二) 比赛时间及试题具体内容

1. 比赛时间安排：本次比赛共分为初赛和决赛两个阶段。初赛考核内容为理论，以答卷（闭卷）方式进行，具体形式为客观题（判断题、单选题、多选题），满分 100 分，考试时间 120 分钟；决赛考核内容为实际操作，具体形式为人脸识别项目实操，满分 100 分，考试时间 360 分钟。

2. 试题：详见样题大纲。

（三）评判标准

1. 分数权重：初赛理论部分满分 100 分，其中，判断题占比 20%，单选题占比 50%，多选题占比 30%。决赛实践部分满分 100 分，其中，数据预处理环节占比 30%，算法定义与训练环节占比 50%，算法测试及结果输出环节占比 20%。

2. 评判方法：

初赛理论部分，判断题和单选题直接按参考答案评分，多选题错选或漏选均不得分，选手需使用答题卡作答，机器阅卷。

实际操作部分，根据输出结果和代码补充结果进行评分，不涉及主观评分，具体参照试题答案及评分标准。实际操作部分的裁判员可设置为“一对多”模式，即 1 名裁判员评阅多名选手的试卷，并将分数记录于给定位置。

3. 成绩并列：当理论部分出现选手成绩并列时，则按多选题得分排名；当实践部分出现选手成绩并列时，则按“算法定义与训练环节”、“数据预处理环节”、“算法测试及结果输出环节”顺序的得分情况进行排名；当总成绩出现并列时，则按实践操作成绩排名，若仍不能分出先后，则按“算法定义与训练环节”、“数据预处理环节”、“算法测试及结果输出环节”顺序的得分情况进行排名。

三、评分流程及考核细则

初赛：初赛设置为理论考试环节（满分 100 分），以理论知识闭卷考试的形式进行，职工组前 60 名选手进入决赛，学生组

前 60 名选手进入决赛。

决赛：

(1) 职工组：

决赛为实际操作（满分 100 分）。各参赛人员总成绩，按照理论知识占 30%、实践操作占 70%加权计算，个人总成绩为理论成绩*30%+实操成绩*70%。当排名出现成绩相同时，按实际操作成绩排名；若仍不能分出先后，则按“算法定义与训练环节”、“数据预处理环节”、“算法测试及结果输出环节”顺序的得分情况进行排名。最终排名按个人总成绩排列。

(2) 学生组：

决赛为实际操作（满分 100 分）。各参赛人员总成绩，按照理论知识占 30%、实践操作占 70%加权计算，个人总成绩为理论成绩*30%+实操成绩*70%。当排名出现成绩相同时，按实际操作成绩排名；若仍不能分出先后，则按“算法定义与训练环节”、“数据预处理环节”、“算法测试及结果输出环节”顺序的得分情况进行排名。最终排名按个人总成绩排列。

现场除参赛选手外，还应设置现场指引人员、技术支持人员、裁判组、仲裁组等工作人员。

1. 现场指引人员负责引导选手熟悉考场环境、考前安全检查、解决选手除考试内容外的其他问题。

2. 技术支持人员负责解决选手考试内容相关问题，如平台和赛题内容等，并辅助裁判员完成赛后的阅卷评分工作。

3. 裁判组由若干裁判员和裁判长组成，除评分工作外，还需担任监考员角色。考试过程中，若有选手提出平台或赛题相关问题，需由监考员传递至技术支持人员，待商议确定后再同步至参赛选手。评分阶段，根据实际参赛人数，每名裁判员将对若干选手进行实践操作打分。若有异议，则向裁判长发出审核需求，待裁判长与技术支持人员确认后，再同步至裁判员进行打分。

4. 仲裁组负责受理比赛中出现的所有申诉进行仲裁。仲裁组需认真负责地受理选手申诉，并将处理意见尽快反馈给当事人。组委会仲裁组的裁决为最终裁决，参赛人员不得因对处理意见不服而停止竞赛，否则视弃权处理。

比赛期间，参赛选手应严格遵守考试纪律，严禁违反考试相关规定，一经发现将取消比赛资格。裁判组成员应严格执行回避制度，负责竞赛评判工作，保证竞赛全程的公平、公正、公开。其他工作人员应严格遵守比赛制度，做好现场及技术支持工作。参赛人员对不符合竞赛规定的工具和设备，有失公正的评审、计分以及对工作人员的违规行为等，可提出申诉。选手申诉须在竞赛成绩公布三日以内以书面形式向组委会仲裁组提出，过期即视为放弃申诉。

四、竞赛场地、设施设备安排

（一）赛场规格要求

实际操作考试考场为标准机房，每间机房可同时容纳 40-50 人进行比赛。

（二）基础设施清单

实际操作考试的赛场提供设施、设备清单表如下，相关设备均无需选手准备。

序号	名称	数量	技术规格
1	电脑	1台/选手	电脑需支持本地文件编辑功能。 硬件配置：CPU：i5 以上；内存：8G 以上；硬盘：50G 以上；网卡：无要求。
2	谷歌浏览器	1个/选手	最新版本

五、工具材料安排及清单

比赛过程均在现场提供的设备上完成，无需选手自带工具或材料，禁止选手携带任何通讯设备进入赛场，禁止选手从赛场带出任何工具或材料。

六、项目特殊说明

参赛人员不可携带任何通讯设备（手机、个人电脑等）进入考场，不允许使用 USB 接口传输数据、预制模板等。

七、安全健康和防疫要求

为贯彻落实国家与广东省关于做好新冠肺炎疫情常态化防控工作的有关要求，本次大赛将建立以下疫情防控机制。

1. 科学规划活动方案及场地布局。

尽量选择通风良好的地点举办赛事及启动仪式等相关活动，并按防控要求合理规划大赛场地分区（测温区、活动区、应急医疗服务区、临时隔离间等）、布局及人员动线，各分区要有醒目的指引标识，要提出流量管控方案。大赛现场设置足够的洗手设备或免洗手消毒液。

2. 控制人员规模和密度。

根据场地规模控制大赛参与人数，参赛人员原则上控制在场所最大容纳人数的70%以内，实行预约分流、分批错峰入场等管控措施，引导大赛参与人员有序进出，保持1米以上社交距离，尽量避免握手、拥抱等密切接触行为。

3. 做好大赛工作人员健康防护。

活动举办单位负责安排专人负责健康监测工作，做好大赛工作人员日常体温监测，如出现发热（ $\geq 37.3^{\circ}\text{C}$ ）、干咳等症状应及时报告并进行排查。活动现场大赛工作人员分发的一次性医用口罩、消毒剂等防疫物品，按照每工作4-6小时更换一次的频率进行发放。大赛工作人员自觉做好个人卫生防护，每日通过“粤康码”等小程序如实进行健康申报，科学佩戴口罩，加强卫生意识。

4. 落实参赛人员扫码测温及核酸检测。

做好参赛人员体温监测，通过红外测温枪等设备确保体温检测准确性。对参赛人员进行健康扫码或登记，可通过实名登记、预约入场等方式对参赛人员进行人员管理，查验为红码或出现发热（ $\geq 37.3^{\circ}\text{C}$ ）、干咳等症状的人员，严禁参加与大赛相关的活动。原则上，竞赛初赛及决赛签到时，所有参与竞赛的选手、教练、裁判及工作人员等，需持48小时内核酸阴性结果（以粤康码显示为准）。

5. 设置临时隔离观察间。

根据参加活动人员和活动场所实际情况设置一定数量的临

时医学观察点，对于初测体温 $\geq 37.3^{\circ}\text{C}$ 人员的实行单独隔离观察，单独隔离观察间用于不需要在医院隔离的具有发热等症状人员的隔离观察。观察点要设在相对独立、通风良好的房间，需配备 1-2 名工作人员，负责体温检测和发热人员的管理，并配备红外测温仪、水银温度计、一次性医用外科口罩、消毒纸巾、医用乳胶手套、快速手消毒剂，84 消毒剂等物品。