

附件 3

人工智能大模型赛项竞赛规则

一、竞赛背景

人工智能技术带来的运算能力、数据分析能力、规律发掘能力，使得数据挖掘、预测、决策优化等变得前所未有的精准，特别是以 ChatGPT 为代表的大模型技术正加速千行百业数字化与智能化转型，是新质生产力发展的核心推动力。人工智能大模型赛事旨在鼓励参赛者充分学习大模型技术、使用大模型技术解决专业问题，提升大模型技术思维能力，提高各专业学生对新时代人工智能技术的认知和兴趣。

二、竞赛器材

比赛器材不限，由参赛队伍自备，要求具备人工智能大模型训练、推理、展示能力，包含但不限于内置麦克风、音箱、视觉摄像头等软硬件设备，内存容量不得超过 32GB，显存容量不得超过 24GB，比赛过程中，需显示训练日志等过程性内容。

三、竞赛规则

参赛团队基于有限的硬件资源及标准的软件环境，使用人工智能大模型技术，训练出具有一定学科属性的人工智能对话大模型，该模型经训练后需要能准确回答相关知识。最终考核参赛团队训练出的大模型准确性及对大模型相关技术的掌握与应用能力。

（一）竞赛任务

比赛任务分为数据整理、大模型调试训练及系统评测、功能演示三部分，总分为 100 分。

1.数据整理

比赛现场各团队通过抽签的方式抽取所要构建的学科大模型任务，之后参赛团队使用自备的参赛设备进行数据的搜集、整理、生成。数据质量对最终大模型训练效果影响较大，各团队需要对数据进行一定的处理后以 Excel 文件的方式提交至评测平台。

(1) 数据量（6 分）

数据量 ≥ 300 条，得 6 分； $100 \leq$ 数据量 < 300 条，得 5 分； $50 \leq$ 数据量 < 100 条，得 2 分；数据量 < 50 条，得 0 分。

(2) 数据质量（4 分）

数据重复率 $< 1\%$ ，得 4 分； $1\% \leq$ 数据重复率 $< 3\%$ ，得 2 分； $3\% \leq$ 数据重复率 $< 5\%$ ，得 1 分；数据重复率 $\geq 5\%$ ，得 0 分。

2.大模型调试训练及系统评测

基于对任务和已整理好的数据，使用自备的比赛器材，进行大模型机型调试训练，参赛队伍需要理解模型参数的意义及对应调整的效果影响，掌握大模型技术知识，对代码进行调整编写，最终生成大模型文件。

(1) 大模型文件上传（5 分）

上传文件至评测平台且完成评测的，得 5 分。

(2) 大模型性能评测（70 分）

比赛过程中，参赛队伍每 20 分钟可上传大模型文件至平台，评测平台对大模型文件进行性能评测。大模型评测平台基于“融

合智创”多维度学科数据集，该数据集涵盖人文、社科、理工等近 80 个专业数据主题，具有完整可复现的学科大模型评测能力，支持大语言模型一站式评测，将为参赛过程中参赛队伍所产出的大语言模型提供统一评测支持。

3.功能演示

(1) 文字对话（10 分）

现场随机抽取三个相关问题，参赛队伍使用文字与大模型对话，需让大模型完成连续回答，中间如有问题回答错误，该任务评测便终止。答对 1 个，得 2 分；答对 2 个，得 6 分；全部答对，得 10 分。

(2) 语音对话（5 分）

现场随机抽取一个相关问题，参赛队伍使用语音与大模型对话。参赛队伍现场录制视频并通过平台上传，要求视频、声音清晰。识别准确，得 3 分；播报成功，得 2 分。

(二) 比赛时间

180 分钟

(三) 排名规则

得分高的团队排名在前，若出现同分情况，则提交时间在前的团队排名在前。

(四) 其他事项

1.参赛队伍不得使用网络，不得使用手机等通讯设备。基础数据现场提供。

2.参赛队伍自备笔记本电脑，并提前安装好视频录制软件。