

2024年（第17届）中国大学生计算机设计大赛人工智能挑战赛赛题

大语言模型专项挑战赛

（版本：V20240310.02）

一、挑战内容

（一）赛题背景

人工智能作为新一轮产业变革的核心驱动力，正在深刻改变人类的生产生活方式，成为引领科技发展的关键要素。国务院印发《新一代人工智能发展规划》指出：到 2030 年人工智能理论、技术与应用总体达到世界领先水平，成为世界主要人工智能创新中心，人工智能核心产业规模超过 1 万亿元，带动相关产业规模超过 10 万亿元。中国将举全国之力，建设人工智能强国，重点是人工智能人才的培养。

2022 年底以来，以 ChatGPT 为代表的 AIGC 异军突起，在全球范围内掀起了一场关于大模型的竞速赛，大模型有目共睹的强大能力掀开了人工智能通用化的序幕，其赋予各行各业解决场景效率优化问题的能力，推动产业变革，从而提质降本增效，促进产业转型焕新，升级发展。

当前，AI 技术已步入了大模型时代，并成为了全球创新的焦点，大模型将成为未来的基础设施，如何将大模型有效地应用于具体场景并提高生产力是业界共同需要探索的方向，随着 AI 大模型爆发，大模型迭代和训练所需的算力呈指数级增长，增速将超过单芯片算力增长速度。同时，单个 AI 超算规模也受到功耗、土地、散热等因素制约。在此形势下，算力供给与需求的缺口持续放大，云端算力面临成本以及供给压力，无法承载 AI 算力需求，算力走向边缘成为大势所趋。边缘算力在成本、时延、隐私上具有天然优势，可以作为云端算力的补充，将 AI 的多样化、海量复杂需求进行本地化预处理，并将其导向大模型，实现 AI 触及万千场景，在边缘算力的加持下，大模型部署在边缘侧与移动端将成为必然趋势，由此将推动边缘 AI 普及，而边缘 AI 作为智能应用创新的关键使能技术，将驱动智能驾驶、智能监测、智能制造、智能家居、虚拟现实等领域的应用快速落地。

基于以上事实,人工智能挑战赛重点考察参赛选手能够基于大模型开展创新应用开发的能力,包括基于商用大模型 API 应用创新、开源大模型的部署、微调定制、基于业务场景的大模型融合创新应用等技术能力,要求选手具备使用人工智能开发平台、深度学习训练框架和推理框架并适当结合主流大模型能力开展工程应用创新的能力。通过举办本赛项,可以提升人工智能相关人才的专业素养,能够在人工智能迅速发展,创新应用场景广泛落地的趋势下,具备宏观视角并能迁移所学知识,在更广阔的人工智能领域内有所作为,从而为人工智能的发展储备和输送新鲜血液。

(二)任务内容

针对大语言模型在商业分析、教育与学习辅助以及企业知识管理等场景应用需求,要求参赛选手完成大语言模型的调用或本地部署以及基于 LLM API 完成助手类应用作品,包括:

(1) 调用商用大语言模型或本地私有化部署大语言模型。

(2) 设计一款基于大语言模型(LLM)的问答助手类应用,该应用可实现在给定的文档(包括但不限于商业分析、教育与学习辅助以及企业知识管理等场景)上进行检索、理解和交互,协助使用者高效阅读和查询。

(2) 在边缘侧设备(综合算法与场景创新实验箱)实现文档的解析与转化、用户交互界面设计,同时调用 LLM API 提取文档中的关键信息和主体,识别重要段落、句子和术语,并将其归纳和总结,为用户提供基于文档的问答。

作品在以上任务的基础上,可以进一步拓展,运用人工智能的相关技术,对大模型进行优化、作品功能创新开发等任务。

最终,参赛队将应用端部署到实验箱上进行测试展示。

二、挑战规则

(一)软件环境

基于大语言模型及边缘端设备进行相关 AI 软件开发。

(二)任务要求

1. 规定动作:选手需要设计一款基于大语言模型(LLM)的问答助手类应用,

该应用可实现在给定的中文文档（包括但不限于商业分析、教育与学习辅助以及企业知识管理等场景）上进行检索、理解和交互，协助使用者高效阅读和查询。规定动作包括：在边缘侧设备（综合算法与场景创新实验箱）实现文档的解析与转化、用户交互界面设计，同时调用 LLM API 提取文档中的关键信息和主体，识别重要段落、句子和术语，并将其归纳和总结，为用户提供基于文档的问答；选手可自行部署本地大模型或选择合适的国内商用大模型作为该应用的 LLM 能力底座；

2. 拓展动作 01（自行部署本地大模型或国内商用大模型均可）：本赛题鼓励选手在人工智能大模型应用功能或业务场景上进行创新（应用端需部署在边缘侧，尽可能调用边缘侧硬件传感器，如摄像头、语音麦克风等）；以上拓展动作的交互功能需集成进规定动作的交互界面中。
3. 拓展动作 02（仅限选手自行部署的本地大模型）：参赛选手可对大模型（仅限选手自行部署的本地大模型）进行优化，针对中国传统文化知识领域（包括且不限于经典古籍、古典诗词、民族刺绣）测试的方向包括 1. AGI 文生文模型；2. AGI 文生图模型。以上拓展动作的交互功能需集成进规定动作的交互界面中。
4. 现场测试：为确保公平，参赛选手需要将参赛作品的客户端部署在边缘侧（综合算法与场景创新实验箱），同时本赛题现场测试时，将针对所有选手采用统一的文档进行测试。
5. 说明：学生参赛作品的 LLM 能力底座可以选择调用国内商用大语言模型或私有化部署的大语言模型两种方式，本赛项鼓励采用私有化部署方式，如采用私有化部署方式，需自行准备算力硬件并携带至比赛现场，算力设备硬件规格参数不高于：显存 16G，内存 64G，不允许将私有化模型部署在云端（为避免使用代理转发技术，一切调用互联网侧 API 的行为皆作为调用商业大模型处理）。

(三) 计分规则

本赛题现场测试按如下规则排序：规定动作测试排名×50%+ 拓展动作展示评分排名×50%，得本赛项现场测试排名。

注：所有计分规则，以最终专家委员会审定为准。

三、设备支持

(一)支持单位名称

江苏天穹实业有限公司

(二)支持单位联系方式

郑钥宁 15380774810

(三)支持单位承诺

1. 现场测试中，支持单位将提供统一实验箱，参赛队可以提前部署测试；
2. 支持单位提供相关实验箱硬件技术参数技术咨询服务。

(四)支持单位设备环境介绍

(1) 现场测试实验箱

国赛现场测试将使用天穹综合算法与场景创新实验箱：



图 综合算法与场景创新实验箱示意图

(2) 现场测试实验箱参数

实验箱中配备摄像头、触摸屏、麦克风、喇叭等模块，所有模块已在平台内部连接，无需额外硬件操作，实验箱采用标准的音频/视频输入输出模块、传感器输入输出模块，保证项目和实验运行结果的一致性。在箱体内部已经规划出传感器实验区、音视频输出区、音视频输入区，结构清晰、功能完整。产品中配备开源传感器接口和 4 个 USB 标准接口，用于采集外部数据，同时也可以扩展其他开源传感器以及其他 USB 设备。

模块	分类
----	----

主控	Jetson Orin Nano核心处理器
尺寸	450X330X185mm
净重	9.5kg
外壳材料	金属铝制外箱
显示屏	13.3寸HDMI Touch触摸式显示屏
喇叭	内置USB声卡喇叭
舵机云台	XSG90 PWM 180二自由度云台
RGB摄像头	采用1080P USB2.0免驱摄像头，mjpeg输出，100度广角镜头
麦克风	360度麦克风
环境传感器	BME680 支持温度、挥发性有机物浓度、相对湿度、气压、海拔读取
土壤湿度传感器	测量土壤中水分数值
红外传感器	XD80NK红外光电开关传感器
按钮模块	触发GPIO电平高低
光敏传感器	测量环境光照强度
NFC读卡器	搭配5张NFC卡
超声波传感器	SR04高精度超声波测距避障传感器

四、其他附加说明

此部分说明为针对各赛题的统一说明。

(一)赛程与最终排名

1. 省/区域赛排名：省/区域赛可参考本文说明实施，也可以在公平、公正、公开原则下，结合参赛师生与当地承办单位的实际情况，充分考虑赛程与选拔的可行性另行制定，参赛队在参加省/区域赛时必须遵守省/区域赛的比赛与计分规则，本赛项国赛组委会不干预、不参与省/区域赛的排序与选拔，但可以在省/区域赛前提供相关测试，供省/区域赛评判参考；
2. 赛前准备：是指各参赛队在学校或其他地点，在国赛之前进行的各项准备工作，赛前准备由各参赛队和所在学校自行安排组织；

3. 赛前测试：是指国赛之前，如果该赛项安排了现场测试，那么给予参赛队熟悉场地、适应场地的测试环节，该环节是否安排，由承办单位视现场条件决定，并赛前通知；该环节允许指导教师与参赛学生共同参与；在该环节中，只能尽可能模拟现场正式比赛的情况，不保证与比赛测试当天、当时的现场各方因素完全相同，这些因素包括：①光照、温湿度等环境的变化，②某些赛项会在比赛前临时调整赛场布置，随机摆放道具等，③比赛测试场地与赛前测试场地不是同一块场地，使用的道具不是同一套道具，使用的设备同型号但不是同一台设备，④其他因素；若比赛没有现场测试环节，则不安排赛前测试环节；
4. 现场测试：是指正式计入成绩的比赛测试环节；每队进行现场测试占用的时间由各赛项单独拟定，但一般不超过 20 分钟（需要长时间计算或展示的赛项除外）；现场测试一般安排在演示答辩之前，但也有可能在演示答辩之后，以承办单位现场条件与最终安排为准；现场测试环节，参赛队必须服从现场安排，不舞弊，不破坏现场秩序；指导教师不得参与现场测试，不得在现场测试时进行场外指导；
5. 演示答辩：是指直接面向评委进行作品展示、宣讲，回答评委提问，并计入成绩的比赛环节；该环节共 20 分钟，一般 10 分钟用于参赛队的作品演示与成果汇报，10 分钟用于评委提问与参赛队回答问题，两部分时间均不超过 10 分钟；如该赛项没有现场测试环节，而必须在演示答辩环节演示作品的，作品演示与成果汇报最多不超过 15 分钟，总时间不超过 20 分钟；该环节需要参赛学生充分准备，在演示现场遇到任何软硬件问题，其维修处理时间都计入演示汇报时间，不得超时，且不安排再次演示答辩；演示答辩环节，参赛队必须服从现场安排，不破坏现场秩序；指导教师不得参与演示答辩，不得在演示答辩时进行场外指导；
6. 有现场测试：最终参赛队排名 = 现场测试排名 × 65% + 演示答辩排名 × 35%；
无现场测试：最终参赛队排名 = 演示答辩排名 × 100%。

(二)参赛队自行保管独立使用设备的

1. 参赛队和指导教师是所使用设备的共同第一负责人，参赛队或指导教师

收到设备后，必须第一时间检查设备是否完备可用，如有问题，支持单位必须及时给予支持，协助其进行设备调试；

2. 设备使用过程中的任何故障、损坏，请参赛队直接与支持单位联系，原则上，人为因素造成的故障损坏，由参赛队负责相关维修费用；由于设计、制造缺陷导致的损坏、故障，由支持单位负责免费维修；
3. 因设备故障、维修造成对备赛的影响，都视为正常比赛事件，比赛不会因个别队伍的设备问题进行推迟或调整；

（以下仅限有现场测试环节的赛项）

4. 支持单位可以提供多套相同的场地、道具同时并行进行现场测试，计分规则相同，参赛队必须在备赛时就考虑其中可能出现的各种差异，在现场测试中服从现场安排，并不得在正式测试或重测时就不同场地、道具上的差异拒绝比赛或提出更换场地、道具的要求；
5. 现场测试过程中发生设备故障（支持单位因素造成的，非自行编写的软件、自行改装、或部分比赛规则中规定的可救援问题，且非 6. 所述情况），那么在测试中，或测试结束后 30 分钟内，且在同场地的下一个测试队开始前，由参赛队向主裁判提出重测申请，由主裁判与设备支持单位确定实属设备故障，方可进行重新测试；
6. 两队或多队同时参与且相互影响的对抗赛不安排重测，所有故障、意外都视为正常比赛事件，成绩经裁判确认后有效；参赛队应在备赛时充分考虑可能出现的各种状况，提前做好应急预案与防范措施；
7. 重测安排在同场地所有队伍测试结束之后，按申请先后依次进行；重测只安排一次，且必须服从现场场地安排；重测形式与正常现场测试中的形式（包括测试轮数、计分方式、道具是否随机摆放等）完全相同（光照、温湿度、时间、必须更换同型号设备等不可抗改变因素除外）；重测必须全部重新进行，不得对单独环节进行单独重测；重测后，取重测成绩作为最终成绩，除非 5. 所述故障依然存在，取两次最好成绩作为最终成绩。

(三)大赛统一提供测试设备的

（以下仅限有现场测试环节的赛项）

1. 支持单位可以提供多套相同型号的设备、场地、道具同时并行进行现场

测试，计分规则相同，参赛队必须在备赛时就考虑其中可能出现的各种差异，在现场测试中服从现场安排，并不得就不同场地上的差异拒绝比赛或提出更换设备、场地、道具的要求；

2. 现场测试过程中发生设备故障（支持单位因素造成的，非自行编写的软件、自行改装、或部分比赛规则中规定的可救援问题，且非 3. 所述情况），主裁判与支持单位确认后，除了当前受影响的队伍可参加重测外，可追溯之前的使用队伍是否也受到影响到，若有，则通知之前已测试的队伍，由之前已测试的队伍自行决定是否参加重测；
3. 两队或多队同时参与且相互影响的对抗赛，允许参赛队在正式测试前，在不损坏设备、场地、道具，且不影响其他队伍的前提下，通过运行测试程序等手段，检查设备的完备性，时间不超过 5 分钟，一旦参赛队确认设备可用，对抗赛开后，所有故障、意外都视为正常比赛事件，且不安排重测，成绩经裁判确认后有效；参赛队应在备赛时充分考虑可能出现的各种状况，提前做好应急预案与防范措施；
4. 重测安排在同场地所有队伍测试结束之后；两次以上重测需大赛组委会批准；重测必须服从场地、设备、道具和测试顺序等现场安排；重测形式与正常现场测试中的形式（包括测试轮数、计分方式、道具是否随机摆放等）完全相同（光照、温湿度、时间、必须更换同型号设备等不可抗改变因素除外）；重测必须全部重新进行，不得对单独环节进行单独重测；重测后，取最后一次重测成绩作为最终成绩，除非最后一次重测中 2. 所述故障依然存在，取最近两次测试的最好成绩作为最终成绩。

(四)其他说明

1. 比赛中如有现场测试，场地、环境、道具、设备等，以赛场实况为准；现场测试前，所有参赛队必须对场地、环境进行全面检查，一旦测试开始，不得以场地、环境、布局、道具等问题提出重测要求；
2. 本文件内容如有更新，以最新发布版本为准；
3. 因疫情等各种因素导致各种情况的改变与调整，以组委会最后通知或现场安排为准；
4. 本文件由中国大学生计算机设计大赛人工智能挑战赛组织方负责解释。

五、修订记录

- V20240221.01: 发布;
- V20240310.02: 对拓展动作进行了细节描述修改。

以下空白