

题目编号：SH-06

多模态智能评测系统技术研究 比赛方案

一、发榜单位

科大讯飞股份有限公司

二、题目名称

多模态智能评测系统技术研究

三、题目介绍

APA 研究表明人类在获取信息时，大概 70%的信息通过视觉，20%通过听觉，剩余通过其他感官获取，人类处理信息是多模态的，通过多模态信息交互完成对场景的理解，如一条 15 秒短视频包含视觉动作、听觉解说、弹幕文本等多维度信息，通过跨模态关联提取实现深度内容理解。而人工智能技术模拟人类多感官协同认知，早期的人工智能单模态模型（如 CNN、LSTM 等）仅模拟单一感官，独立处理视频、图像、文本、语音等单模态数据，仅能完成独立任务（如图像分类、文本分类等），导致信息割裂。随着获取数据形态的质变和人工智能技术的不断突破，人工智能经历从感知到认知的范式革命，Transformer 架构驱动的多模态融合（如 CLIP 的图文语义关联）复现了人类“联觉”能力，实现从“感官割裂”到“全局认知”的突破。随着人工智能技术赋能千行百业，单一模态的局限性

和片面性已无法满足行业复杂场景的需求，如在医疗、教育、金融、工业等领域，单一的视频或语音分析可能无法全面评估一个人的表现或状态，需要结合多种模态数据评测来提高准确性和可靠性。比如医疗中的病情评估，需要结合医学影像（视觉）、患者主诉（文本）、语音记录来判断病情进展；教育中的口语考试，仅依赖语音识别无法区分“背诵流畅”与“即兴表达”（需结合视频中的眼神和肢体语言）；金融中的欺诈，传统单模态生物识别（如人脸）易受攻击，需结合人脸识别（视觉）、声纹验证（语音）、虹膜、操作行为分析的多模态认证降低金融欺诈率；工业中的质检，仅依赖视觉易受光线干扰，需融合红外热成像（触觉延伸）与设备振动数据（音频）可提升缺陷检出率。

随着大模型技术的不断突破和各行各业积极探索大模型技术场景应用，为了应对行业场景的复杂性和多样性，需利用大模型的强大语义理解和生成能力、不断进化能力、以及协同单模态 AI 能力来研究和开发有效的面向不同行业的多模态智能评测技术，提高专业性、准确性和鲁棒性。虽不同行业涉及模态数据存在差异性 & 评测指标的不同，但每个行业都面临着人才的招聘需求，模态数据是一致的。数据显示，到 2025 年，更多行业将把 AI 面试纳入人才招聘流程，AI 面试的市场占有率预计将增长至 55%。目前对于初步面试，通过多模态分析评测求职者的语言表达、表情变化和逻辑能力等指标，为企业节省

高达 70%的时间成本，同时由于算法偏见和适配性等问题让 AI 面试评测存在专业性、准确性和全面性不足问题，那么提高 AI 面试中面试评测的专业性、准确性和全面性，是当前企业 AI 面试面临的技术突破。请围绕企业 AI 面试进行多模态智能评测系统技术研究。

四、参赛对象

本题目只设学生赛道。

参赛对象为 2025 年 6 月 1 日以前正式注册的全日制非成人教育的各类高等院校在校专科生、本科生、硕士研究生、博士研究生（不含在职研究生），参赛人员年龄在 40 周岁以下，即 1985 年 6 月 1 日（含）以后出生。

同一作品不得同时参加第十九届“挑战杯”全国大学生课外学术科技作品竞赛（以下简称第十九届“挑战杯”竞赛）其他赛道的评比。

参赛对象可以团队或个人形式参赛，每个团队不超过 10 人，每件作品可由不超过 3 名指导教师进行指导。可以跨专业、跨学校、跨单位、跨地域组队，但同一团队所有成员均应符合本赛道相关年龄、身份要求。每件作品只可由 1 所高等院校作为参赛主体提交申报。

五、答题要求

1. 作品要求

参赛者需设计并开发一套面向企业 AI 面试的多模态智能评

测系统，具体要求如下：

（1）场景覆盖不限于：人工智能、大数据、物联网、智能系统等至少 3 个技术领域的典型岗位面试场景；

（2）多模态数据分析评测：至少整合语音（语言逻辑、情感语调、语速）、视频（微表情、肢体语言）、文本（应答内容）等多维度数据，构建动态量化评测体系；包含至少 5 项核心能力指标（如专业知识水平、技能匹配度、语言表达能力、逻辑思维能力、创新能力、应变抗压能力等）；

（3）智能反馈不限于：①支持生成可视化评测报告，包含能力雷达图、关键问题定位及反馈建议（如“回答缺乏 STAR 结构”、“眼神交流不足”）帮企业做出决策；②给求职者提供个性化改进建议；

（4）创新性要求不限于：①支持真实的交互体验，模拟不同面试官的性格和风格，增加面试的真实性；②探索 AI 技术突破（如基于大模型的动态追问（如生成个性化追问逻辑链）、跨模态数据关联分析）；③鼓励结合心理学、教育学理论设计评测维度，增强反馈的科学性；④探索人机协同的混合决策机制，兼顾效率提升与人文关怀的双重需求；

（5）数据合规性：若使用真实面试数据，需符合隐私保护与伦理规范，确保用户隐私数据（如视频、语音）的加密存储与合规使用；

（6）其他场景和行业延伸（可选）：向公务员选拔、职业

技能认证等高公信力场景拓展，推动技术的社会化普惠应用。

2. 作品提交形式

作品形式应包括如下内容：

（1）材料文档：内容包括但不限于作品设计实现方案、PPT 作品介绍、10 分钟以内的演示视频等；

（2）系统模块：作品的源代码及可执行程序(如有)，提供程序在测试服务器上的部署说明，确保可在测试服务器上顺利运行。

六、作品评选标准

以百分制对参赛作品进行打分，根据分数决定团体奖名次，具体主要包括以下内容：

（1）作品完整性：30 分

研发并完成面向企业 AI 面试的多模态智能评测系统，作品完整、实现数据采集→分析→反馈闭环，没有明显缺陷，为 30 分。如未完成，按照完成度给予评定，最多不超过 20 分；

（2）技术创新性：20 分

根据参赛作品技术创新点（如多模态算法融合、技术架构、动态交互技术等）进行打分；

（3）用户体验：20 分

根据参赛作品的用户体验（如界面友好性、交互流畅度、沉浸感等）进行打分；

（4）实用价值：30 分

根据参赛作品的实用性（如多模态评测的专业性、准确性和全面性、数据驱动决策反馈、个性化建议、支持多类行业岗位等）进行打分。

七、作品提交时间

2025 年 5 月—8 月 10 日，各高校应组织学生参赛，安排专业人员给予指导，为参赛团队提供支持保障。

8 月 15 日前，各参赛团队通过大赛申报系统提交作品，具体要求详见作品提交方式。

2025 年 8 月底前，由大赛组委会会同发榜单位共同完成初审，确定入围终审擂台赛的晋级作品和团队。

2025 年 9 月，发榜单位安排专门团队提供帮助和指导，各晋级团队完善作品，冲刺攻关参加终审擂台赛，角逐“擂主”。

八、参赛报名及作品提交方式

（一）报名方式

1. 参赛选手登录“挑战杯”官网 2025.tiaozhanbei.net，在“揭榜挂帅”擂台赛报名入口注册账号，登录大赛申报系统在线填写报名信息。报名信息提交后，下载打印系统生成的报名表。

2. 申报人在报名表对应位置加盖所在学校公章。

3. 将盖章版报名表扫描件上传至报名系统，等待系统审核。请参赛选手注意查看审核状态，如审核不通过，需重新提交。

4. 系统开放报名时间为 2025 年 5 月 30 日—6 月 30 日，逾

期后系统将自动关闭报名功能。

（二）作品提交方式

参赛团队请将参赛作品（包括但不限于作品设计实现方案文档、PPT、10 分钟以内演示视频、源码、系统部署说明、测试数据等）统一打包压缩提交至大赛申报系统，压缩包命名方式为：申报人所在单位-申报人姓名-作品名称-联系电话（例如：XX 大学-张 XX-XX 方案-手机号）。如遇压缩包较大无法发送，请将压缩包上传至相对安全云盘，并将云盘链接和提取码、以及云盘文件截图（含上传时间）提交至大赛申报系统。

九、赛事保障

本单位将为此次比赛组建专业指导教师团队，介绍技术细节要求，针对比赛过程中的疑问定期进行解答。比赛过程中，参赛团队如需本单位提供与项目相关的其他必须帮助，请提前与本单位联系，本单位将在许可范围内给予参赛团队帮助。

十、设奖情况及奖励措施

1. 设奖情况

面向学生赛道，根据评分规则，综合评定参赛团队。原则上设“擂主”1 个（从“特等奖”中评选），特等奖 5 个，一等奖 5 个，二等奖 5 个，三等奖 5 个，最终授奖数量可视作品申报数量和质量情况报组委会同意后动态调整。

2. 奖励措施

面向学生赛道，本单位将结合项目实际情况，拟在特等奖基础上额外奖励“擂主”团队 10 万元；奖励特等奖每个团队 10000 元；奖励一等奖每个团队 5000 元；奖励二等奖每个团队 3000 元；奖励三等奖每个团队 2000 元。

对于选择本题目的学生可优先安排实习绿色通道。

3. 奖金发放方式

比赛结束后，本单位比赛专班工作人员与获奖团队取得联系，填写奖金申请表，待获奖团队提供银行卡详细信息后 1 个季度内，将奖金一次性发放至获奖团队提供的银行卡中。

十一、比赛专班联系方式

1. 专家指导团队

顾问专家：江老师，联系电话：13696542465

顾问专家：郑老师，联系电话：17681133583

负责比赛期间技术指导保障。

2. 赛事服务团队

联络专员：陈老师，联系电话：18118159492

联络专员：朱老师，联系电话：13965057275

负责比赛期间组织服务及后期相关赛务协调联络。

3. 联系时间

比赛期间工作日（9:00-17:00）

附：发榜单位简介

科大讯飞股份有限公司成立于 1999 年，是亚太地区知名的智能语音和人工智能上市民营企业。自成立以来，一直从事智能语音、计算机视觉、自然语言处理、认知智能等人工智能核心技术研究并保持国际前沿水平。科大讯飞积极推动人工智能源头核心技术研发和产业化落地，致力于“让机器能听会说，能理解会思考，用人工智能建设美好世界”。2008 年公司在深圳证券交易所挂牌上市。作为中国人工智能“国家队”，科大讯飞承建了中国唯一的认知智能全国重点实验室和语音及语言信息处理国家工程研究中心，同时是中国语音产业联盟理事长单位、中科院人工智能产学研创新联盟理事长单位、长三角人工智能产业链联盟理事长单位。2024 年，科大讯飞作为第一完成单位的“多语种智能语音关键技术及产业化”项目获得 2023 年度国家科学技术进步奖一等奖。科大讯飞坚持“平台+赛道”的发展战略。基于人工智能核心技术，科大讯飞多年来持续赋能教育、医疗、金融、汽车、城市、运营商、工业等行业赛道并取得广泛成效，持续助力民生和产业高质量发展。2023 年 5 月 6 日，科大讯飞发布讯飞星火大模型。2024 年 10 月 24 日，基于首个国产万卡算力集群训练的讯飞星火 4.0 Turbo 正式发布，国产超大规模智算平台“飞星二号”启动。2025 年 1 月，讯飞星火 X1 及首个具备端到端语音到语音同传能力的星火语音同传大模型，讯飞星火持续赋能千行百业。