**B题 智能评阅算法效果的综合评价**

2024年初，中国明确提出加快发展新质生产力推动高质量发展，人工智能被确立为培育新动能的核心领域。考试评卷智能化在人才选拔领域成为人工智能技术落地的重要手段。

当前主流网上评卷系统仅实现客观题自动批改，主观题仍需人工多评，耗费大量人力。随着AI技术进步，基于大模型、手写识别、自然语言处理等技术的智能评卷系统取得突破：AI通过深度学习评分标准，能独立完成稳定高效评分，结合人工复核与数据分析，形成“人机协同”管理模式，既保证质量又提升效率。

某实验室率先研究AI辅助评卷创新模式：通过两种智能算法背对背进行AI评分，与人工一评结果交叉验证。这种“一人工+双AI”的协同机制，标志着智能评卷从理论探索进入规模化应用阶段，为教育考试数字化改革提供了可复制范本。

附件1wl.xlsx中给出了两类人工智能阅卷与人工阅卷的成绩结果，其中子对象1为填空题型，其他均为简答题；子对象1到子对象4对应的满分分值分别为6、9、14、16。请建立数学模型，解决下列问题：

**问题1** 对于附件1所给的数据，请分析人工评阅和两类人工智能算法评阅数据的分布特点。

**问题2** 以附件1为例，选择不同评价角度，构建“智能评阅算法”的评价指标体系，设计智能评阅算法的综合评价模型，利用所给数据对两类人工智能算法给出评价。

**问题3** 附件2中给出了其他几个科目的抽样数据（各科目数据解释和附件1类似），请在问题2基础上，针对学科维度展开评阅效果评价对比。

**问题4** 如果现在要使用上述两类人工智能算法，请结合附件3设计一种使用方案，并对所设计方案进行分析和评价。

**备注：**（1）附件3提供了科目题号、分值和误差阈值，供参考；（2）指标体系的层级和各层下的指标、定义、量化公式在论文中要清晰明确；（3）建议设计并在论文中插入指标体系可视化思维导图（思维导图在线链接（可查看编辑，放入图的标题中或者图释中））。