

2025年第四届“创新杯”（原钉钉杯）大学生大数据挑战赛题目

初赛B：道路路面维护需求综合预测

一、赛题背景:

道路维护是保障交通系统安全、高效运行的关键任务，也是城市基础设施管理中的重要环节。随着现代交通网络规模的持续扩大和车辆保有量的快速增长，道路养护需求与有限的人力物力、财力资源之间的矛盾日益凸显。传统的道路维护决策通常依赖人工经验或单一指标（如路面破损程度），缺乏对多源数据的综合分析，容易导致资源浪费或维护滞后。例如，部分路段因过度维护造成资金浪费，而另一些高隐患路段却因未及时检测而引发交通事故；同时，不同区域的道路特征（如交通流量、气候条件、材料类型）差异较大，单一维护标准难以适应复杂场景。因此，如何基于多维度数据（如路面状况指数、交通荷载、气候环境、历史维护记录等）构建科学预测模型，并结合预测结果优化资源分配策略，成为交通管理部门亟待解决的核心问题。

本赛题提供某地区真实道路检查数据集，包含路段ID、路面状况指数（PCI）、道路类型、年平均日交通量、沥青类型、上次大修年份、平均降雨量、车辙深度、国际粗糙度指数（IRI）等多维度信息，以及目标标签（是否需要维护）。参赛队伍需通过数学建模与数据分析技术，设计高效、精准的预测模型和资源分配方案，为道路维护决策提供科学依据。

二、样本解读:

Column Name名称	Description特征解释
Segment ID	路段ID道路路段的唯一标识符
PCI	PCI路面状况指数（0=最差，100=最佳）
Road Type	道路类型（主要、次要、Barangay）
AADT	年平均日交通量
Asphalt Type	沥青类型：沥青混合料分类（如致密、开级配、SMA）
Last Maintenance	上次大修的最后一次维护年份
Average Rainfall	平均降雨量该地区的年平均降雨量（毫米）
Rutting	车辙深度（mm）
IRI	IRI国际粗糙度指数（m/km）
Needs Maintenance	需要维护目标标签：如果需要紧急维护，则为1，否则为0

三. 解决问题:

初赛任务A: 路面维护需求预测回归分析问题

使用路段ID、路面状况指数（PCI）、道路类型、年平均日交通量、沥青类型、上次大修年份、平均降雨量、车辙深度、国际粗糙度指数（IRI）等特征进行数学模型，预测某条道路是否需要维护。要求输出模型的准确率、召回率、F1值，并分析各特征对预测结果的贡献程度。

初赛任务B: 特征重要性分析与维护策略优化问题

在上一题的基础上，使用数学模型对道路维护需求进行预测，并分析各特征的重要性。进一步，根据特征重要性排序，提出一种基于关键特征的道路维护优先级划分策略（例如：高优先级维护路段、中优先级维护路段、低优先级维护路段）。要求输出优先级划分规则及对应的路段数量分布。

题目声明：本赛事所有赛题仅授权 2025年第四届“创新杯”（原钉钉杯）大学生大数据挑战赛参赛队伍使用，任何组织及个人未经组委会书面授权，严禁用于校内竞赛，篡改、复制等侵权行为。