

## 数据科学实践课程项目说明

本实践课程内容为开放性命题，旨在提供一个实践性的机会，让学生运用所学的数据科学技能解决一个真实世界的问题。在本次项目中，您将有机会展示您在数据收集、数据处理、特征工程、模型选择和建模、数据可视化等方面的能力，并将结果有效地呈现和沟通。最终将由教师（40%）、助教（30%）、学生（30%）在报告当日共同打分。

评分细则如下：

### **完整（40分）：**

项目目标与问题陈述清晰明确，包含必要的背景信息和目标定义。

数据收集和处理全面，包括数据源的选择、数据清洗和预处理的步骤。

特征工程和分析充分，包括特征选择、转换和创建，以及数据的可视化和摘要统计。

模型选择、建模和评估全面，包括合适的模型选择、参数设置和评估指标的使用

### **创意（20分）：**

提供了独特和创新的解决方案或方法，展示了对问题的独到见解。

尝试了新颖的数据处理、特征工程或模型选择方法，超出了课程中介绍的基本技术。

提出了有趣的观点、假设或问题，展示了对数据科学领域的思考和深入理解。

### **技术（20分）：**

实施技术方案的准确性和有效性，包括数据处理、特征工程和模型建模的正确性。

使用合适的技术工具和编程语言，保证代码的可读性和可复用性。

模型评估方法和指标的选择和应用准确无误，能够充分评估模型的性能和结果。

### **表达（20分）：**

结果呈现清晰、结构良好，包括报告的组织、段落和标题的使用。

使用恰当的图表、图像和表格来支持结果和结论，确保信息的易理解性。

字面表达准确、语法正确，避免拼写和语法错误，使得报告易于阅读和理解。

# 课程报告给分细则

报告时长 15 分钟，10 分钟讲述，3 分钟现场问答，2 分钟打分

## 内容（30 分）：

目标明确：报告中明确表达了目标和主题，并清楚地说明了要传达的核心信息。

结构清晰：报告按照逻辑和结构进行组织，包括背景、主体和结论等部分。

内容准确性：提供了准确、详尽和可靠的信息，包括相关的实验设定和结果数据。

## 演示技巧（10 分）：

演讲流畅性：演讲者语言流畅、清晰，能够顺畅地表达思想和观点，没有明显中断或语法错误。

姿态和声音：演讲者的姿态自然、自信，并能够以适当的音量和节奏传递信息。

观众参与：演讲者与观众建立联系，包括目光接触、姿态转变和观众参与互动等。

## 可视化呈现（40 分）：

图表和图像：使用适当的图表、图像和表格来支持演讲内容，并确保它们清晰可读、准确有序。

设计和布局：幻灯片设计简洁、美观，布局合理，文字和图像的排列有序、易于理解。

视觉吸引力：使用恰当的颜色、字体和样式，以增强幻灯片的视觉吸引力和专业性。

## 时间管理（10 分）：

符合时间限制：演讲者能够在规定的时间内完成演讲，不超过或严重不足时间限制。

平衡内容：演讲者能够合理分配时间，确保每个部分的内容得到充分呈现和讲解。

## 问题回答（10 分）：

回答准确性：演讲者能够准确回答观众提出的问题，并给出清晰、有逻辑性的解释。

沟通能力：演讲者与观众之间能够进行积极有效的沟通，表达清晰、明确的观点和回答。