

《数据库应用技术》考核说明 2024

成绩构成说明:

- 闭卷笔试 (60%)
- 课程论文(15%) 钉钉提交 doc 文献综述
- 项目代码实践(15%) 钉钉提交 doc 实验报告/ 前 6 章 SQL 脚本
- 平时&考勤(10%)

笔试: 6月14日(周五)下午至晚上随堂进行笔试(闭卷), 试卷分值分布如下:

机器学习概念理解(10%) 关系数据库概念与数据库设计(40%) SQL 查询语句编写(50%)

课程论文: 选择机器学习相关**算法**主题, 查阅中英文学术文献 3 篇以上(知网或谷歌学术下载), 撰写文献综述, 字数不限

项目代码实践:

1. SQL Server 教材前 6 章脚本(要求经过调试运行)

2. 从以下题目中选择 **1** 个题目进行实验, 撰写实验报告。



实验报告应至少包括: 1 问题分析、2 代码或实验过程分析、3 结果比较或效果验证

1. 找到一份训练数据, 使得岭回归/套索回归的效果优于最小二乘线性回归
2. 绘制逻辑回归损失函数, 验证该损失函数是凸函数
3. 对同一份分类数据, 同时训练感知机/逻辑回归/支持向量机, 绘制三直线
4. 绘制约束优化函数的优化过程
5. 手写数字识别: 将数字写在纸上, 拍照进行识别。加深网络层数实验
6. 使用神经网络拟合曲线
7. 在MNIST手写识别数据集上训练自编码器
8. 基于SVD的推荐算法
9. 高斯混合模型
10. 隐马尔可夫模型: 智能输入法, 拼音输入法提示
11. 马尔可夫随机场: 图像分割
12. 贝叶斯网络: 航班延误预测

第 4 题的约束优化函数如下:

例 1.1.1 设平面上有两个线段 u_1u_2 和 v_1v_2 . 试求这两个线段上相距最近的两个点 u^* 和 v^* .

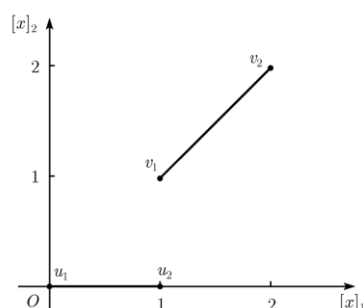
这个问题可以写成一个最优化问题. 事实上, 线段 u_1u_2 和 v_1v_2 上的点分别可表为

$$u = \alpha u_1 + (1 - \alpha) u_2, \quad \alpha \in [0, 1] \quad (1.1.1)$$

和

$$v = \beta v_1 + (1 - \beta) v_2, \quad \beta \in [0, 1]. \quad (1.1.2)$$

显然, 点 u 和 v 的距离是变量 α 和 β 的函数



第 6 题的数据如下:

```
X = np.linspace(0, 20, num=200)
```

```
y = X + np.sin(X)*2 + np.random.normal(size=X.shape)
```

