

# 扣分细则

- 扣分细则

- 1 Clustering (25)

- 正确答案

- $$\{x^{(3)}, x^{(4)}, x^{(5)}\}$$
$$\{x^{(1)}, x^{(2)}\},$$

- 满分要求:完整展示四次迭代, 并每一次迭代的数据计算正确
    - 扣分点: 如只计算三次, 不能表明“算法已迭代”, 扣5分左右。用非欧式距离来计算, 扣1分左右。

- 2 PCA (25分)

- 2.1 (20分)

- 正确答案

- $$(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}).$$

- 满分要求: 答案正确, 过程正确
    - 扣分点: 答案不正确的话, 会结合计算过程扣5-10分

- 2.2 (5分)

- 正确答案

- $$: [-2\sqrt{2} \quad 0 \quad 0 \quad 2\sqrt{2}]$$

- 满分要求: 答案正确
    - 以下几种情况只可能出现一种
      - 扣分点1: 因为使用**未去中心化**数值进行计算, 导致计算错误, 扣1分左右。
      - 扣分点2: 第一问计算错, 第二问一定是计算错的, 扣5分
      - 扣分点3: 计算出二维的映射结果, 扣5分

- 3 Learning to rank (50分)

- 3.1 (15分)

- 正确答案: 正确说出三种算法的区别与联系, **必须有自己的文字描述**。
    - 扣分点: 可以粘贴ppt内容, 但要是**只**粘贴ppt, 没有自己的文字描述的话, 会扣分5分左右

- 3.2 (10+2\*5=20分)

- data transformation, function transformation, (10分)

- 答题关键点:

- 需要在数据处理上表现出“处理成对的数据”, 并为这些对分配+1, -1这样的值

- 展现出对于“三分类”的处理
  - data transformation与function transformation很相关，你可以在data transformation把AB, AC and BC单拎出来说明“符合这些关系的数据对的y值为+1”。或者在function transformation部分描述清楚，某些数据需要加loss
  - 扣分点：两个关键点少一个扣5分
- how to predict, (5分)
  - 根据 $f(x)$  排序
  - 扣分点：答错扣5分
- and how to evaluate the ranking results. (5分)
  - DCG/NDCG/pairwise disagreement都可以
  - 扣分点：
    - 有的同学采用其他方法，也是可以的，但是一定要描述清楚。有一些同学的公式中无法看出清晰的判断关系，（把初始数据和pairwise转换后的数据混淆了）是不能得分的。
    - 单独几个同学直接没写这一部分答案，大家之后答题的时候一定要把题看完整。
- 3.3 (15 分)
  - 正确答案：给loss加权重，使相关文档少的query损失权重更大（只要一个定性的关系就可以），并且给出公式，然后简要说明。
  - 扣分点：思路不满足“使相关文档少的损失权重更大”，不得分
  - 扣分点：不写公式，与修改权重关系不清晰等，意思对了，但是贴合不紧密，扣5-10分。