

我随即在uci的其他数据集上进行了实验，得到结果如下：

数据集	iris	breast_cancer	wine
聚类数目	3	2	3
NMI	0.758	0.464	0.428

可见聚类结果的好坏也会受到数据集特性的影响。

总结 && 感想

通过实现K-means聚类算法，并在鸢尾花数据集上进行实验，我得到了较高的NMI值，这说明聚类结果与真实标签的一致性较高。这表明K-means算法在这个数据集上表现良好，能够有效地将样本分配到不同的聚类中心。然而，需要注意的是，聚类结果的好坏也受到数据集特性的影响，对于其他数据集可能会有不同的表现。因此，在实际应用中需要根据具体情况选择适合的聚类算法和评估指标。

这次复现K-means聚类算法的实验让我对聚类算法有了更深入的理解。通过编写代码和运行实验，我能够清楚地看到算法是如何工作的，这也加深了我在课上学习到的理论知识。