平时作业2：实现LSH算法  
作业要求：

1. 实现LSH算法（从Shingling到Locality-Sensitive Hashing完整步骤）
2. 用一个文档数据集对LSH的效果进行测试，设置k=10，相似性阈值0.8，通过调整M、b、r分析LSH的假阴性和假阳性。
3. 撰写完整的实验报告（1~3页）
4. 代码和实验报告的语言风格不限

作业命名：平时作业2+学号+姓名.zip（包括代码、数据和实验报告）

邮件主题：平时作业2+学号+姓名

截至日期：2024.3.24 24:00前提交

邮箱：BDSysu2024Spring@163.com