### 计算机科学与工程学院 学院毕业设计(论文)中期报告

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 学生姓名 |  | | | | 指导教师 | | 李锡祚 | | | 职称 | | 教授 | |
| 专业 |  | | | | | | | | | 班级 | | |  |
| 毕业设计（论文）题目 | | | 基于协同过滤的音乐推荐系统的设计与实现 | | | | | | | | | | |
| 题目来源 | 为所在专业自拟题目 | | | 题目类型 | | 工程实践型 | | | 是否校外题目 | | 否 | | |
| 公司名称  （地址） |  | | | | | | | | 校外指导教师姓名（职称） | |  | | |
| 毕业设计（论文）进展情况，遇到哪些困难，如何解决（加A4纸打印的附页，字数在2000字以上）  一 、毕业设计（论文）进展情况  1）完成了协同过滤相关论文的翻译工作, 并进行查阅资料明确了设计方向。  2）对音乐推荐系统有了较为详尽的了解和研究，明确了毕设的完成过程。  3）完成了使用python网络爬虫获取音乐数据集（包括音乐文件，音乐封面，歌词，歌手，评论等内容）。  4）完成了提取歌曲的mfcc特征值并进行保存，同时对mfcc特征值进行了K-MEANS聚类，得到了相同种类的歌曲。  5）分析并设计了数据库，编写了sql代码来创建数据表等。  6）编写了后台代码，可以给前端提供支持。    二 、已经完成的工作详细叙述：  1、 协同过滤  协同过滤（简称CF）是根据已有数据来推测出未知数据。在协同过滤这个算法中会从海量的数据中找到与你品味比较类似（ 即相似度达到指定范围）的数据，而这些数据会成为你的邻居，系统会根据你的这些邻居为你推荐你心仪的物品。启发式的协同过滤算法可以分为基于用户的协同过滤算法和基于项目的协同过滤算法。  基于用户(user-based)的协同过滤算法是通过，挖掘用户的历史行为数据，发现用户对商品或内容的偏好，并对这些偏好进行度量和打分。之后根据不同用户对相同商品或内容的态度以及偏好程度，来计算用户之间的相似度关系。基于用户的协同过滤，主要计算的是用户与用户之间的相似度，只需要找出相似用户喜欢的物品，并预测出目标用户对对应物品的评分，就能够找到评分最高的物品推荐给用户，这样能够挖掘用户的隐藏属性。  基于项目的协同过滤推荐的基本原理是，使用所有用户对物品或者信息的偏好，发现物品和物品之间的相似度，然后根据用户的历史偏好信息，将类似的物品推荐给用户。       本应用主要采用的是第二种方法。  2、MFCC特征值  在语音处理领域里，梅尔频率倒谱(mel-frequency cepstrum简称MFC)表示一个语音的短时功率谱，是一个语音的对数功率谱在频率的一个非线性梅尔刻度上进行线性余弦转换所得，而所有的梅尔频率倒谱系数(Mel-frequency cepstral coefficients  简称MFCC)共同的组成一个MFC。       本应用通过调用python的包来提取歌曲的mfcc特征值。主要代码如下：       def getMusicMfcc(filePath):             Y, sr = librosa.load(filePath, sr=22050\*2, duration=15)             return librosa.feature.mfcc(y=y, sr=sr, n\_mfcc=13)  3、K-MEANS 聚类     对提取到的mfcc特征值使用K-MEANS聚类算法进行聚类，主要代码如下：  [IDX, C, SUMD] = kmeans(mfcc\_feature(:, 1:16796), K, 'Distance', DISTANCE, 'Display', 'iter', 'Start', 'plus', 'MaxIter', 300);  disp(' Kmeans聚类完毕...}');  music\_list = cell(K, 2);  music\_labels = mfcc\_feature(:, 16797);  for k = 1:K  index = find(IDX == k);  music\_label = zeros(size(index, 1), 1);  for each=1:size(index, 1)  music\_label(each) = music\_labels(index(each));      end  music\_list{k, 1} = music\_label;  music\_list{k, 2} = length(music\_list{k, 1});  end    4、编写后台代码     分为用户模块，歌曲模块，评论模块，歌单模块，歌手模块等。     用户模块提供用户界面的登录，注册，个人信息查询等。     歌曲模块包括查找歌曲，获取歌曲信息，对歌曲进行模糊搜索等。     评论模块包括添加评论，查找评论，删除评论等。     歌手模块包括查找歌手，获取歌手信息等。    三 、后期要完成的工作  1）在老师的帮助下完善毕业设计与论文。  2）完成剩余部分编码工作，对系统进行优化。  3）参照他人的界面，设计自己的前端。    四、中期总结  在整个过程中遇到过许多的困难，让我更加认识到翻阅资料的重要性，许多错误和知识点其实已经有人提供了很好的解决办法了，并不需要我去闭门造车。 | | | | | | | | | | | | | |
| **英文翻译** | | □ 优 □ 良 □ 中 □ 差 | | | | | | | | | | | |
| **工作进度** | | □ 提前完成 □ 按计划完成 □ 没有完成 | | | | | | | | | | | |
| **工作态度** | | □ 认真 □ 一般 □ 不认真 | | | | | | | | | | | |
| **工作质量** | | □ 优 □ 良 □ 中 □ 差 | | | | | | | | | | | |
| 指导教师意见    签字  年  月  日 | | | | | | | | | | | | | |
| 系（专业）主任意见    签字  年  月  日 | | | | | | | | 学院意见    签字  年  月  日 | | | | | |

大连民族大学教务处制