

人工智能实践课程项目一

音乐推荐系统

1. 学习目标

通过完成该项目，学生将学习如何将学到的理论知识应用于实践中。每个小组需要使用相同的数据集来实现自己的推荐系统。学生们需要在有限的时间内通过团队协作完成实验，撰写报告并准备演讲展示。

2. 项目简介

本课程项目的任务是通过给定的数据集实现一个音乐推荐系统，每个小组可以自由的选择要采用的方法与工具来获得最好的推荐结果。

据报道 [1]，中国有超过 9.77 亿人每周都听音乐，而 66% 的人通过流媒体来听音乐。为了给用户提供更好的体验，如何为用户推荐喜爱的音乐就变得非常重要。本项目使用的数据集来自 Last.fm 音乐网站 [2]，数据集在 2011 推荐系统会议(ACM RecSys)中发布。

3. 音乐数据集 [3]

3.1 统计数据

1892 个用户(user)

17632 个歌手(artist)

12717 对好友关系

92834 个用户歌手关系组合，如元组: [user, artist, listeningCount]

11946 种风格标签(tag)

186479 个标签标记，如元组: [user, tag, artist]

3.2 文件

artists.dat 文件包含听过的歌手信息

tags.dat 文件包含数据集中用到的 tag

user_artists.dat 文件包含 user 听过的 artist 以及听过的次数

user_friends.dat 文件包含用户间的好友关系

user_taggedartists.dat 和 user_taggedartists-timestamps.dat 文件包含 user 对 artists 标记 tag 的行为，文件同时包含标记的时间戳(timestamp)

4. 项目提交

项目提交截止时间为 5 月 31 日 12:00，提交内容包含源代码，报告。

4.1 代码

你可以自由选择要使用的程序语言（如：Python、Java 等），第三方库（如：Scikit-learn 等），团队合作工具（如：Github 等）。

4.2 报告

项目报告提交格式为 PDF，你可以自由选择撰写工具（如：腾讯文档、LaTeX、MS Word 等）。

报告不限字数，每个小组提交一个报告并在报告中注明组员。

一个较好的项目报告一般包含以下几项：

- 项目背景
介绍一下你的项目的背景，意义等。
- 相关工作
通过阅读相关文献，介绍一下当下推荐系统的发展，有哪些常用的推荐算法，这些算法在不同推荐中的应用。别人的相关工作有没有什么缺点，对你的音乐推荐项目有没有什么帮助。
- 实验方法
详细介绍你项目选用的推荐方法，如协同过滤、矩阵分解、近邻、深度学习、基于内容的推荐等，以及为什么选择该方法。
- 实验结果
介绍一下你选用的评价标准（如：推荐的精确度、准确度、AOC、点击率等）及实验的结果。
- 总结
总结并讨论项目的结果与意义，有哪些方面还可以在未来改进。
- 参考文献

项目中用到的参考文献。

4.3 评分标准

音乐推荐系统课程项目分值为 30 分，项目中如果使用了任何非小组成员编写的代码，文本等均需注明出处，如果发现拷贝或抄袭将按照学校有关规定处理。

具体评分标准如下表：

推荐方法的合理性	6 分
推荐方法的新颖性	6 分
实验结果	6 分
实验报告撰写	6 分
演讲展示	6 分

参考文献

- [1] M. Savage, "英媒：逾 9 亿中国人每周听音乐，热爱现场的乐迷比例比美国多," 澎湃新闻, 2016. [Online]. Available: https://www.thepaper.cn/newsDetail_forward_1429142_1.
- [2] Last.fm, [Online]. Available: <https://www.last.fm/>.
- [3] GroupLens, "HetRec 2011," [Online]. Available: <http://files.grouplens.org/datasets/hetrec2011/hetrec2011-lastfm-2k.zip>. [Accessed 18 04 2020].