

“贴心云衣柜——基于机器学习和社交网络的服装搭配推荐” 结题报告

计算机科学与技术系，2016 级，李鑫烨，曹佳涵，金亦凡，李光耀

指导老师：卜磊，副教授

摘要：本项目 4 名成员均为计算机系 2016 级本科生。指导老师为卜磊副教授。当下穿衣搭配是一个很热门的话题，本项目开发了一款基于机器学习和社交网络的服装搭配推荐 APP——贴心云衣柜，它可以为用户生成个性化的穿衣搭配，还可以通过对相似风格的用户间进行相互推荐。

关键词：Android；服装搭配推荐；机器学习；协同过滤；社交网络

1. 选题背景

1.1. 背景

在物质生活水平不断提高的现代社会，人们对于服装搭配的要求也在一天天地提高：不仅要做到舒适得体，还要用不同的穿搭应对不同的场合。可以说，服装搭配已经成为了现代人出门之前必须要解决的一个问题。然而与此同时，随着生活节奏的进一步加快，思考如何搭配着装和选择购买新衣物的时间也愈加缺乏。在这种情况下，我们不难发现，一个智能的、个性化的，能应对上述诸多问题的穿衣搭配推荐 APP，必然能够吸引一个很大的用户群体，在 APP 市场中占有一席之地。

1.2. 知识储备

Android APP 的开发会用到 Android 的应用，服务器端需要用 java 语言进行开发，java 语言我们小组成员有一定基础，但 Android 开发此前未曾接触过，我们需要从零开始自学起。另外，本项目的核心是使用机器学习算法进行个性化推荐，因此算法的设计与改进也十分重要，我们需要掌握机器学习的决策树算法和协同过滤的隐语义模型算法，并不断进行调参，算法的实现和优化也是一大重点。

这一技术路线需要掌握的知识 and 理论有：Android 开发、java 开发、机器学习等等。这些也是项目中需要学习的知识。

2. 项目情况

2.1. 成员组成

项目成员均为计算机科学与技术系 2016 级本科生。

李鑫烨：项目负责人，动手能力较强，擅长做实验和协调各成员的工作。主要负责学习算法决策树的实现和各个模块的连接、督促成员的工作和联系导师等。

曹佳涵：执行力强，注重理论学习和动手实践相结合。主要负责推荐算法协同过滤的实现，以及后端数据库的搭建。

金亦凡：具有很强的创新思维。主要负责客户端架构，前端界面及各功能模块的实现。

李光耀：思维清晰，知识面广。主要负责前端功能实现和客户端与服务器端的 socket 通信。

2.2. 导师指导情况

项目初期，卜磊老师给我们介绍了当前 APP Store 的内容在穿衣个性化推荐上的缺失的现状，在此基础上，他提出了关于这个 APP 的构想和最终功能的预期，并提供给我们完成该项目所需要自主学习的相关知识，给我们初步定下了项目的目标。

项目进展过程中，老师时常关注着项目进展，并安排了一位研究生学姐为我们提供帮助。大二上学期末，在老师的指导下，我们进行了模块划分和分工。中期答辩前，老师对我们的答辩材料给出了非常有用的建议。中期答辩后，老师为项目接下来的进展指明了方向。

结项答辩前，老师对项目成果进行了评估，并就结项事宜进行了详细的说明和指导。

整个项目过程，老师非常认真和仔细地提供帮助和关心，使得项目能够顺利地进展。

3. 创新特色

3.1. 需求分析

在项目开发的前期准备时，开发小组访问了 APP 商店，发现市场中已经存在很多有关穿衣搭配的 APP，但是这些 APP 的推荐功能较为基础，它们只是简单地处理了用户现有的衣物的搭配问题，并没有推荐系统的处理。事实上，用户更加青睐于个性化的 APP，期望 APP 能够符合自己的品味。更进一步的需求是满足大多数人都具有的融入集体的心理，在社交网络中分享自己的风格，并找到同道中人。

以网易云音乐为例，该款 APP 的成功之处并不在于其与其他音乐 APP 相同或相似的音乐播放功能或庞大的音乐库，而是在于它特有的个性化音乐推荐功能，以及完善的社交功能，包括音视频分享功能和评论功能等。用户可以通过强大的推荐功能找到更多符合自己听歌风格的歌曲，丰富个人歌单；同时用户还能够评论、分享自己喜欢的歌曲，参与到社区互动，也可以找到听歌风格相近的用户集群。这种用户体验，是一个单纯的音乐播放器所无法提供的。

因此，在分析大多数用户的需求、并与市面上的同类型或同性质的 APP 进行横纵向比较之后，开发小组明确了开发方向，即开发一款这样的 APP：它既有普通穿衣搭配 APP 的功能（包括衣物管理系统、服装搭配收藏系统以及基于用户现有衣物和天气等客观信息的搭配推荐系统），又有类似于网易云音乐的个性化推荐系统和社交系统。用户能够收到 APP 为其提供的符合该用户个性的一套或多套穿衣搭配（内含的衣物单品可能不在用户现有衣物中），也能够参加社交活动，查看好友的衣物搭配，同时在社交网络中寻找到与其风格类似的其他用户。

3.2. 创新功能

本项目开发的 APP 的核心功能有二：第一，根据天气等一些外在条件，在现有的衣物中为用户推荐一套适当的衣物搭配；第二，建立一个社交网络，根据该用户的穿衣风格，从社交网络中寻找与其风格相近的其他用户，并将这些用户或他们的衣物搭配收藏推荐给该用户。项目以这两个功能为核心进行架构。

除此之外，该 APP 还有一些优先级较低但同样十分必要的功能，例如天气、个人衣柜、服装搭配收藏、新衣物导入，以及一些社交网络的基本功能（好友系统、朋友圈等）。

3.2.1. 项目框架

从功能列表中，我们能够大致将所有功能分为三个大类：演示、存储和计算。演示是 APP 最基础的功能，用户的各种个人信息、为用户推荐的各种信息以及其他种种用户会想看到的信息需要在手机界面上向用户演示；用户的个人信息，例如衣物收藏等，在每次启动 APP 时均需要进行加载，故需要选择一个合适的存储方式；最后，为用户推荐信息需要进行大量的计算。

鉴于用户对 APP 的轻量化需求，开发小组决定开设一台服务器，将所有计算工作以及部分存储工作转移至服务器，以达到轻量化的目标。用户集的信息存储在服务器中，本用户的一些信息（例如衣物收藏）则存储在本地；推荐衣物搭配及用户的计算过程放在服务器端，通过 Socket 架设接口传输推荐请求和推荐结果。客户端的主要任务则是演示。

3.2.2. 推荐算法

在对应用于两大核心功能的推荐算法的选择上，开发小组的考虑如下：首先，对于基于已有衣物向用户进行穿衣搭配推荐这一功能，设想运用机器学习的方法。

考虑服装搭配推荐这一场景下，所选用机器学习算法在用户已有衣物中进行挑选搭配并推荐。现实生活中衣物数量往往十分有限，因此算法需要适用较小的训练数据集规模。同时，算法需要灵活地挑选多件衣物单品以构成具有一定整体风格的服装搭配，在此过程中可能出现推荐衣物单品时，与其搭配的衣物尚未选取或进行推荐，表现为训练数据中的缺失值。因此选取算法需对缺失值不敏感。综合这些特点，最终实现在小规模数据集上表现较好且对缺失值不敏感的，可解释性较好的决策树算法。

在多用户推荐的推荐环境设定下，每个用户的数据由他对各类衣物的偏好程度组成，这个偏好程度由他的历史穿衣数据以及浏览记录数据计算得到。但我们真正要从数据中提取的，是每个用户对于不同穿衣风格的偏好，由于不同种类的衣物可能会属于同一种风格，因此衣物种类之间并不是完全独立的。在理想情况下，我们希望通过一种相互独立的属性，能将不同风格的用户划分开来，而这种属性并不是显式地体现在数据中的。另一方面，由于每个用户并不是对于每种衣物都有数据，因此用户-衣物矩阵是稀疏的，很多信息因此缺失，我们需要重新计算并预测新的用户-衣物矩阵来进行用户相似度计算，最终我们采用隐语义模型(Latent Factor Model)来求解问题。

3.3. 科学意义

在衣物搭配推荐上，我们的决策树算法可以根据用户的历史穿衣记录，实现令人满意的推荐。对于多个用户，我们的隐语义模型也可以挖掘用户对各种风格衣物的偏好程度，并计算相似用户，完成用户之间的相互推荐，增强 APP 的社交功能。我们的算法可以为衣物搭配方面的推荐提供一种新的思路，并实现了较好的实践。

3.4. 应用价值

从用户体验的角度分析，该 APP 能够提供当下市面上各种穿衣搭配 APP 所不能提供的一些用户体验。当用户想要在自己的衣柜中增添新衣物时，个性化的推荐功能可以更快地给用户指明方向；同时，在用户社群中找到风格类似的同道中人也能为用户带来极好的体验感，用户之间可以互相借鉴穿衣搭配，交流某种风格的穿搭心得等。

项目主要阶段以实现功能为主，在开发完成之后如果对整个项目进行优化，包括界面优化、

性能优化等，那么这款 APP 成品必定能够占据一定市场，吸引一定量的用户集群。

4. 进展成果

4.1. 进展情况

项目进展可分为：前期准备、前期实施、中期总结和规划、后期整合阶段。

前期准备：在和导师进行交流后，我们决定开发一个主体业务逻辑部署在服务器的 APP，客户端除了用户交互界面显示之外，只对输入图像进行识别和特征提取。我们了解到 Android 系统上的 APP 开发需要掌握 Java 语言并且服务器端算法实现采用 Java 语言也能获得较好的效果。看了框架进行基本的了解，并通过一些简单的 Demo 来了解 Android 系统上 APP 的开发。

前期实施：我们先从客户端入手，完成 APP 界面上基本的用户交互效应逻辑。对服务器端代码也进行了大致的架构和划分，服务器的运行逻辑即为通过 socket 模块接收用户的响应，并调用响应的算法模块进行计算，并更新信息到数据库。曹佳涵完成协同过滤算法模块的实现，李光耀负责客户端 socket 模块的实现，金亦凡完成客户端上数据库模块，李鑫烨负责完成对服务器端代码的进一步整合和测试。

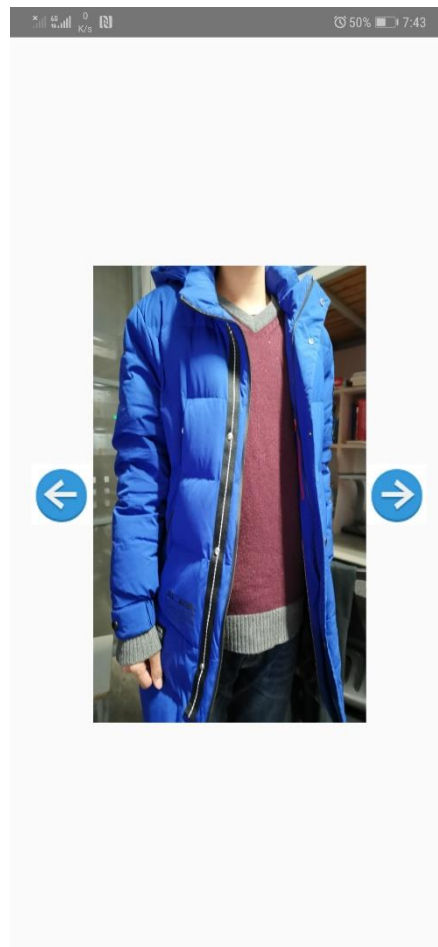
中期总结和规划：目前我们自己实现的衣物识别和特征提取，存在准确率低、速度慢的问题，继续运用图形学的分析方法难以解决问题，考虑到衣物识别并不是我们项目的重点，我们决定调用已有的 API 进行深度网络的训练以实现衣物识别和特征提取。同时我们目前仅仅实现了协同过滤算法，将相似用户匹配到一起而未能构建社交圈，因此在相似用户匹配基础上添加社交模块是我们努力的一个方向。

后期整合：在项目后期，主要进行社交模块已有框架与现有 APP 的整合，调优决策树与协同过滤算法，进行系统性的性能测试与调优，针对应用场景数据特点进行针对性的改进，稳定服务器与 APP 之间的联系，提供 APP 稳定性。

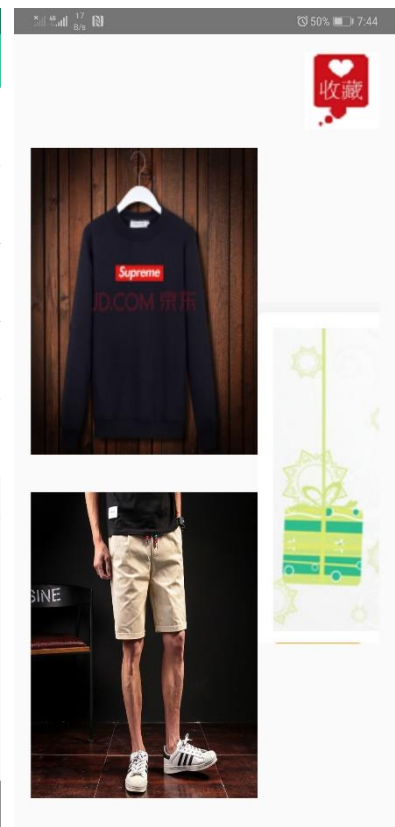
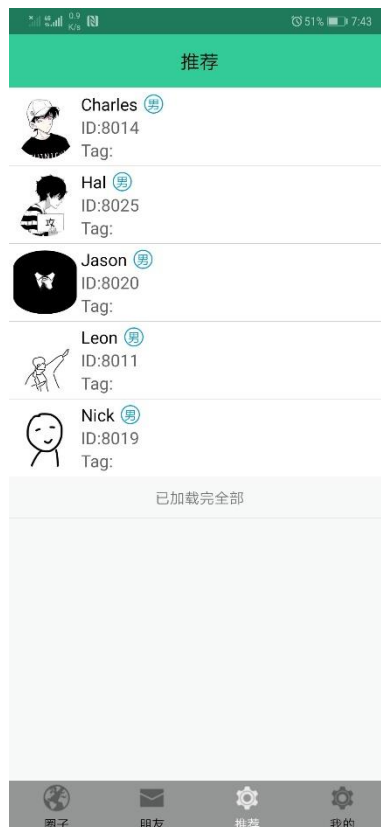
4.2. 创新成果

我们项目开发的 APP 实现了两大核心功能：第一，根据天气等一些外在条件，在现有的衣物中为用户推荐一套适当的衣物搭配；第二，建立一个社交网络，根据该用户的穿衣风格，从社交网络中寻找与其风格相近的其他用户，并将这些用户或他们的衣物搭配收藏推荐给该用户。

初始界面和衣柜：



社交朋友圈，用户推荐和衣物搭配推荐：



5. 收获体会

5.1. 总结

在本次项目实践中，我们项目小组对于机器学习算法以及软件项目开发有了更加深刻的理解。我们开发了一款穿衣搭配 APP——贴心云衣柜，用来满足人们对于节省穿衣搭配时间、以及进行社交活动的需求。在项目开发过程中，我们也学习了 Android Studio 的使用，后台服务器及 Socket 的搭建，以及系统化进行项目开发的软件工程方法。当前版本的贴心云衣柜实现了所有核心功能和绝大部分的辅助功能。在之后的后续开发过程中，我们会进一步对项目进行优化，例如增加运行流畅度、优化算法精准度、添加一些更加强大的辅助功能等。

5.2. 各成员体会

李鑫烨

在创新项目推进过程中，我们经历了一个完整的工程全过程。从需求分析、评估可行性和制定技术路线到将复杂问题模块分解，小组成员不再仅仅着眼于技术细节，而观察到了整个工程开发的全貌，领会到了架构的重要性。同时工程也是小组成员集体智慧的结晶，在开发的过程中我们深刻体会到了团队协作的重要性。创新项目在专业课程之外，为我们拓宽了对成熟应用的技术的视野。同时对卜老师在项目过程中的悉心指导表示衷心感谢！

曹佳涵

在一开始我们定下项目框架的时候，还没有明确具体用哪种算法来实现推荐，在项目的进行过程中，我学习了很多机器学习的推荐算法，阅读了很多相关的论文，最终我重新思考问题本身，发现其数据稀疏，特征之间相互影响，不完全独立的特点，而隐语义模型十分适合这类问题，最终实现出来的效果也非常令人满意。此外在实现服务器端数据库的时候，也让我对数据库和 java 的应用更加熟练。整个项目的实现过程让我学习到了很多，也收获了很多。

金亦凡

在项目刚刚起步时，我们还未对各种开发框架和方法有过研究。在项目进展的过程中，我学习了各种 Android Studio 开发知识，一些工具（例如数据库等）的应用，以及系统化审视整个项目的方法。一开始对于一个小白来说，这是个艰难的过程，但在经历过这个过程之后，我也有了经验与体会，体会到了 Android 前端开发的乐趣。同时，跟项目小组其他成员的合作也增强了我的团队意识和沟通能力。通过这次项目我受益匪浅。

李光耀

通过参加这次创新项目，我提升了自己各方面能力，包括动手能力、知识水平、团队合作意识等。在项目前期，由于对诸多领域一无所知，我会表现得比较焦虑。随着每周一次的小组讨论，使我对这个大型项目有了更深的理解，同时对整个系统进行架构、责任功能分解，使得模块具体化。大型软件项目的开发就被这样分解下来。不得不说这种探索的过程提高了我解决复杂问题的信心。通过这个 APP 开发，我积累下了宝贵的经验，掌握了很多开发工具的使用，学习到了很多课堂不曾涉及的知识，非常感谢这次创新项目。

参考文献

- [1] <https://itunes.apple.com/cn/app/chuan-yi-zhu-shou-jiao-ni/id577700263?mt=8>
- [2] <http://www.cs.cmu.edu/%7Etom/pubs/MachineLearning.pdf>
- [3] https://en.wikipedia.org/wiki/Decision_tree
- [4] 项亮编著, 2012 年 6 月,《推荐系统实践》第三章——推荐系统冷启动问题
- [5] https://en.wikipedia.org/wiki/Collaborative_filtering#Data_sparsity
- [6] Meyer, C. D. (2000), Matrix Analysis and Applied Linear Algebra
- [7] <https://console.heweather.com/my/service>
- [8] 郭霖编著, 2016 年 12 月,《第一行代码——Android Studio》
- [9] <https://github.com/TommyLemon/Android-ZBLibrary>
- [10] Badrul Sarwar, Badrul Sarwar, Joseph Konstan, John Riedl, WWW '01 Proceedings of the 10th international conference on World Wide Web, Item-based collaborative filtering recommendation algorithms
- [11] Rodolphe Jenatton, Nicolas L. Roux, Antoine Bordes, Guillaume R. Obozinski, Advances in Neural Information Processing Systems 25 (NIPS 2012), A latent factor model for highly multi-relational data