需求文档

1 引言

1.1 编写目的

在电子商务网站中，用户在购买服饰时，经常有这样的需求，例如用户准备购买一件喜欢的上衣的同时往往希望选择一条与之搭配的裤子。穿衣搭配是电商场景下服饰鞋包导购中非常重要的课题。

为了解决这个需求，本课设拟通过足够多的数据集，同时利用商品图像信息以及文字信息等综合来设计穿衣搭配系统。

为实现这个需求，通过需求分析、模型设计、程序测试等过程完成课程设计。该报告设计软件开发依据，软件验收的标准，促进课程项目顺利进行。

1.2 背景

本课设是图文算法结合的算法设计系统，同时，穿衣搭配算法设计需要比较大的数据集，所以本课设需要处理大量数据，同时，可能存在衣服搭配不合理的情况，比如上衣搭配了另一个上衣，类似这种预期不合理的情况都需要考虑。如果搭配数量不足或搭配数量过多的情况都需要进行取舍。

1.3 定义

搭配集：课程设计提供了的专家搭配集

商品集：所有商品的相关数据

用户历史购买集：用户的购买记录

1.4 参考资料

[1]<https://www.jianshu.com/p/78ada62c9953>

[2]<https://www.deepfashion.cn/index> （搭配素材）

[3]<https://wenku.baidu.com/view/a4244b2130126edb6f1aff00bed5b9f3f90f726e.html>

2 项目概述

2.1 需求概述

本课设使用三个模型，分别是：利用给定的专家搭配集通过计算不同商品之间的搭配度得到商品与其他商品的搭配度、基于商品标签的频次分析法、基于用户历史行为计算搭配度。利用这三种模型计算输入商品的搭配集。

2.2 运行环境

sklearn

python

2.3 条件与限制

由于图片数据量太大，需要有较高的硬件存储。由于高数据量运算，需要有较高的内存。

3 功能需求

该课设需要实现输入某个商品的id，通过模型计算获得该商品的搭配集。并通过界面显示出，搭配商品图片。

4 数据描述

**输入数据：**



**输出数据：**



**数据库描述**：商品基本信息数据(文本、图像)、用户历史行为数据和搭配套餐数据。如下：

**搭配套餐数据**：dim\_fashion\_match\_sets



**商品信息表**：dim\_items



**用户历史行为表**：user\_bought\_history



**数据的数据量和处理量**：

约88G向量文件

约2.5GB的图片压缩文件

5 性能需求

时间特性：需要通过数据处理（例如聚类）提高算法运行速率。以及算法上的一些改善，以提高运行速度。

容量需求：至少需要88G的硬盘存储空间