**《软件需求分析说明书》**

**项 目 名 称：基于机器学习大数据评文本分析的电商推荐系统**

**成 员 名 单：李星煜、李星灿、吴泽林**

**导 师：袁宇丹**

**工 程 领 域：软件工程**

**中国科学技术大学软件学院**

**一、引言**

（一）研究背景

据第45次网络发展报告显示，截至2020年3月28日，我国网民规模为9.04亿，互联网普及率达到64.5%，并且截至2020年3月，我国网络购物用户规模达到7.10亿，越来越多的用户正参与到网络购物中去，并且通过对商品进行评论来表达自己对商品的喜好程度，如此便会产生一系列的商品评论信息。

电子商务蓬勃发展，但是在其繁荣发展的背景下也隐藏诸多问题：

（1）消费者抉择困难。各类商品品牌款式繁多，各有特色，消费者在选择选择购买商品时，很难从上百种，甚至上千种商品中，选择出一个满足自己的需求的商品。

（2）购物模式的局限。网购模式下，消费者不能实实在在体验到商品，消费者获取商品信息的方式有两种，商家的宣传广告，商家提供的照片和特点描述往往是有过滤有筛选的信息，这种手段下，消费者所获得的信息并不全面可靠。第二种，从以往购买过此商品的用户评论中获取，从上千条评论中寻找关键信息显然十分消耗时间，因此从海量评论中获取商品信息也不现实。

（3）评论信息冗余。大数据评论下，海量的文本数据中描述的商品特征往往集中在少数几个，评论的雷同率相当高，对于消费者来说会造成大量时间浪费。

（二）研究意义

为了解决消费者选择困难的问题，并且节约消费者的时间，在购物网站中根据购物者行为信息来对其进行个性推荐是十分有必要的。此外，商品评价信息具有很大的价值，用户经常根据商品评价信息做出决策，商家也可根据商品评价信息调整销售策略。在此背景下，如何挖掘商品评论数据体系背后蕴藏的信息，并根据个人需求进行推荐十分具有研究意义，并逐渐成为一个重要的研究课题。

**二、需求概述**

（一）目标

本项目实现的电商推荐系统，包含其他购物网站拥有的功能，比如注册、登录、浏览商品、将商品加入购物车等。此外还包含个性化推荐功能，用户登录后，系统就会根据用户购物车信息、浏览历史信息以及与当前用户相似的用户信息来为当前用户推荐商品，以此来方便用户购买商品，提高用户购物效率。

（二）研究内容

本项目的主要研究内容：

（1）运用JSP、CSS、JavaScript等开发前台页面

（2）利用SSM开发电商系统

（3）研究并制定个性化推荐策略。

**三、功能需求**

（一）课题内容

本课题要实现一个电商推荐系统，在此系统中用户可以进行注册、登录、浏览商品、将商品加入购物车、下订单等所有购物系统都存在的功能，同时系统也会为用户进行个性化推荐商品。其中个性化推荐策略分为按照商品相似度推荐、按照用户浏览历史推荐、按照用户相似度推荐三种，以此来推荐出用户可能感兴趣或者希望购买的商品。

（二）系统需求分析

系统中包含两个用例角色，一个是用户，另一个是系统。

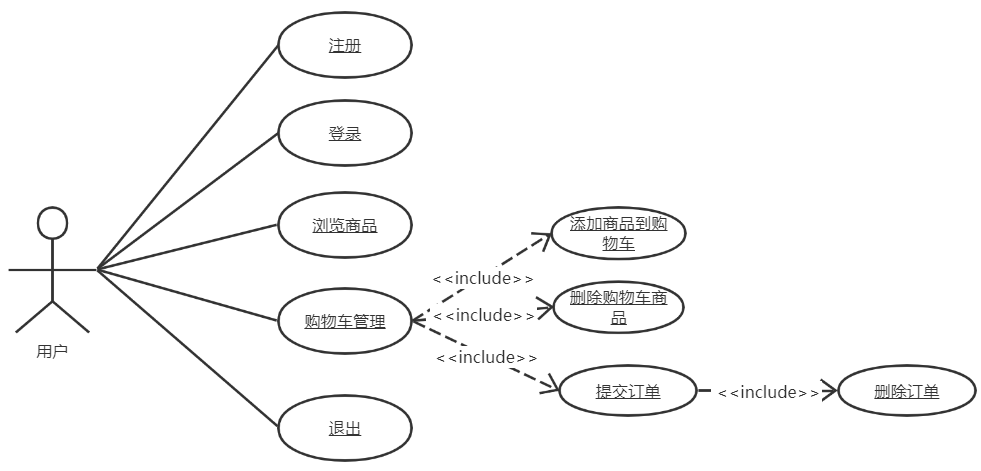
（1）用户用例

(a) 浏览商品，无论用户登录与否，都能执行浏览商品的操作；

(b) 注册新账号，当用户想要执行登录操作必须有注册好的账号；

(c) 用户登录，用户登录之后可以执行将商品加入购物车操作以及下订单的操作；

(d) 购物车管理，包含用户向购物车中添加商品、删除商品、提交订单。



（2）系统用例

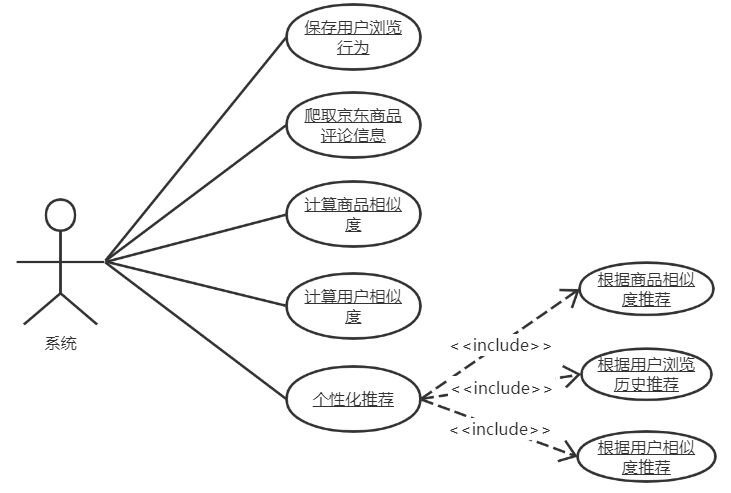
(a) 保存用户浏览行为，此用例是为了保存商品点击量以及用户的浏览历史；

(b) 爬取京东商品评论信息，此部分爬取京东商品信息以及评论是为计算商品相似度做铺垫；

(c) 计算商品相似度，便于根据用户购物车内商品进行推荐；

(d) 计算用户相似度，便于根据用户相似度给当前用户推荐商品；

(e) 个性化推荐商品，包含按照购物车内商品推荐、用户浏览历史推荐、相似用户推荐。



**四、非功能需求**

（一）性能需求

（1） 响应时间

系统的响应时间应至少与现有各购物商城相当。这要求我们必须建立结构合理的索引以及设计高效准确的查询、搜索算法。

（2）推荐的商品必须合理

系统推荐给用户的商品应该尽可能是用户想要购买的商品，这涉及到计算商品相似度算法，如果此算法足够准确，那么给用户推荐的商品大概率会被用户接受。

（3）灵活性

系统要有好的接口，便于以后的扩充，以适应信息源的增加或者减少。

（二）运行环境需求

编译器：Eclipse

编程语言：Java

数据库：MySQL

服务器：Tomcat

（三）安全性需求

电商推荐系统应能保障用户信息的安全，预防一些常见的网络攻击，能够应对一些非法链接的攻击。

（四）可维护性

系统的编码应遵循规范，便于维护。

（五）可靠性

该项目中的服务器应该能够支持大量用户同时访问的需求，并且对于计算商品相似度算法设计也应该考虑到大量商品评论数据的处理情况。