

2019년도
학사학위논문

협업필터링을 활용한 화장품 추천 시스템

Cosmetic Recommendation System Using
Collaborative Filtering

2019년 11월 22일

순천향대학교 공과대학
컴퓨터소프트웨어공학과

김원진, 신현우

2019년도
학사학위논문

협업필터링을 활용한 화장품 추천 시스템

Cosmetic Recommendation System Using
Collaborative Filtering

2019년 11월 22일

순천향대학교 공과대학
컴퓨터소프트웨어공학과

김원진, 신현우

협업필터링을 활용한 화장품 추천 시스템

Cosmetic Recommendation System Using Collaborative Filtering

지도교수 박 두 순

이 논문을 공학사학위 논문으로 제출함

2019년 11월 22일

순천향대학교 공과대학
컴퓨터소프트웨어공학과

김원진, 신현우

김원진, 신현우의 공학사학위논문을
인준함

2019년 11월 22일

심 사 위 원 박 두 순 인

순천향대학교 공과대학

컴퓨터소프트웨어공학과

차 례

그림 차례	
표 차례	
초 록	
제 1 장 서 론	1
제 2 장 관련 연구	2
2.1 JSP(Java Server Page).....	2
2.2 자바스크립트 (JavaScript).....	3
2.3 추천 시스템.....	3
2.4 협업필터링(Collaborative Filtering).....	4
2.5 아파치스파크(Apache Spark).....	5
2.5.1 MF (Matrix Factorization).....	7
2.5.2 ALS (Alternating Least Squares).....	8
제 3 장 협업필터링을 활용한 화장품 추천시스템 설계	9
3.1 협업필터링을 활용한 화장품 추천시스템 데이터.....	9
3.2 협업필터링을 활용한 화장품 추천시스템 설계	13
제 4 장 협업필터링을 활용한 화장품 추천시스템 구현	14
4.1 협업필터링을 활용한 화장품 추천시스템 구현 환경	14
4.2 협업필터링을 활용한 화장품 추천시스템 구현	15
제 5 장 결 론	18
[참 고 문 헌]	19
[부 록]	20

그림 차례

[그림 2-1] JSP 실행 구성도	2
[그림 2-2] 추천시스템의 필요성	4
[그림 2-3] 사용자 기반 협업필터링	5
[그림 2-4] 아이템 기반 협업필터링	5
[그림 2-5] 아파치 스파크	6
[그림 2-6] 평점 행렬을 사용자와 아이템 행렬로 분해	7
[그림 2-7] 예측 평점 행렬을 계산.....	8
[그림 3-1] DB 구상도	9
[그림 3-2] 카테고리 테이블	10
[그림 3-3] 브랜드 테이블	11
[그림 3-4] 상품 테이블	11
[그림 3-5] 멤버 테이블	12
[그림 3-6] 리뷰 테이블	12
[그림 3-7] 협업필터링을 활용한 화장품 추천 시스템 구성도.....	13
[그림 4-1] 메인 화면	15
[그림 4-2] 분류별 인기순위 목록	15
[그림 4-3] 로그인 화면	16
[그림 4-4] 회원 가입 화면	16
[그림 4-5] 화장품 상세 페이지	16
[그림 4-6] 화장품 리뷰 등록	17
[그림 4-7] 화장품 추천 목록	17

표 차 례

<표 4-1> 협업필터링을 활용한 화장품 추천 시스템 구현 환경	14
---	----

초 록

최근 화장품 종류가 다양해 지고 수가 증가하고 있어 사용자들이 화장품을 선택하는데 많은 시간을 할애하고 어려움을 겪고 있다. 추천시스템을 활용하여 사용자들에게 화장품을 추천해 줌으로 인해 사용자들이 자신에게 맞는 화장품을 쉽고 편리하게 접할 수 있다.

실제로 추천 시스템은 아마존, 넷플릭스, 페이스북, 유튜브 등 다양한 곳에서 활용되고 있으며 아마존은 1/3이 주문이 추천시스템으로 이루어지고 넷플릭스는 3/4가 추천시스템을 이용하고 있다.

추천시스템은 베스트셀러기반, 인구통계정보기반, 최소질의 대상 상품결정 방법, 정보필터링 방법 등으로 나뉘고 그중에서 정확도가 제일 높은 정보필터링 방법은 다시 콘텐츠 자체 내용을 기반으로 추천해 주는 내용기반(Content-based)과 사용자들로부터 얻은 기호정보에 따라 사용자들의 관심사를 자동으로 예측해 주는 방법인 협업필터링(Collaborative Filtering)으로 나누어진다.

협업필터링은 다시 사용자-사용자 간의 유사도를 기준으로 하는 경우는 사용자 기반(User-Based), 아이템-아이템 간의 유사도를 기준으로 하는 경우는 아이템 기반 (Item-Based)으로 나뉜다[1].

본 논문에서는 협업필터링(User-based collaborative filtering)을 이용하여 사용자들의 화장품 성향을 분석 후 유사한 성향을 가진 사용자들이 선호하는 화장품을 추천해 준다.

제 1 장 서 론

과거에는 대부분 매장에서 직접 화장품을 구매하여 판매자에게 화장품을 추천받거나 정보를 제공받아 체험을 해 본 후 화장품을 구매하였으나 최근에는 TV, 컴퓨터, 각종 스마트 기기를 통하여 언제 어디서나 화장품에 대한 정보를 쉽게 접할 수 있다. 그러나 사용자 입장에서는 다양해지는 화장품 종류와 사전 지식 부족으로 인하여 자신에게 맞는 화장품을 고르는데 어려움이 생긴다.

추천시스템을 이용하여 사용자들에게 화장품을 추천해 줌으로써 이러한 문제를 해결 할 수 있다. 추천 시스템에는 베스트셀러기반 추천방법, 인구통계정보 기반 추천방법, 최소질의대상 상품결정방법, 정보필터링 방법 등이 있다. 정보필터링 방법은 다른 기법들에 비해 정확도가 높아 일반적으로 사용되고 있으며 제품간의 연관성을 기반으로 추천하는 콘텐츠 기반 필터링(Content-Based filtering)과 사용자들로부터 얻은 기호정보에 따라 사용자들의 관심사들을 자동적으로 예측하게 해주는 방법인 협업필터링(Collaborative Filtering)으로 구분된다. 협업 필터링 기법은 일반적으로 고객들이 동적인 평가결과를 보이는 상품군에 대한 상대적으로 높은 예측력을 보이는 장점을 가지고 있다[1].

본 논문에서는 협업필터링(Collaborative Filtering)을 이용하여 사용자의 성향을 분석해 유사한 성향을 가진 사용자들이 선호하는 화장품을 사용자에게 추천해 주었다.

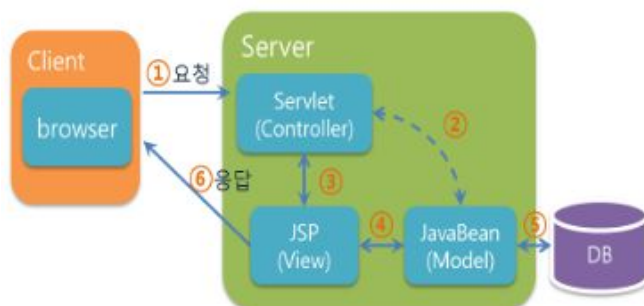
제 2 장 관련연구

본 장에서는 화장품 추천시스템을 설계하고 구현 하는데 사용하고, 필요한 기술이 JSP(Java Server Page), 자바스크립트(JavaScript), 아파치 스파크(Apache Spark), 추천시스템, 협업필터링 기법에 대하여 설명한다[1][4].

2.1 JSP(Java Server Page)

JSP는 HTML 내에 자바 코드를 삽입하기 위하여 웹 서버에서 동적으로 웹페이지를 생성하여 웹 브라우저에 돌려주는 언어이다[4]. 자바 서버 페이지는 실행 시에는 자바 Servlet으로 변환된 후 실행된다. Servlet과 마찬가지로 Java 언어를 기반으로 하지만 ASP, PHP와 같이 서버에서 실행되는 스크립트 언어 방식인 웹페이지를 작성해서 Servlet의 장점은 그대로 갖추고 있다. 하지만 서블릿과는 달리 HTML 표준에 따라 작성되므로 웹 디자인하기에 편리하다[4].

JSP의 장점으로서는 JAVA가 제공하는 기능 사용가능, 세밀한 제어를 통한 성능발휘, 뛰어난 보안성, 플랫폼에 독립적, 분산 네트워크 환경에 적합 등 다양한 장점이 있다. JSP 실행 구성도는 [그림 2-1]과 같다[5].



[그림 2-1] JSP 실행 구성도

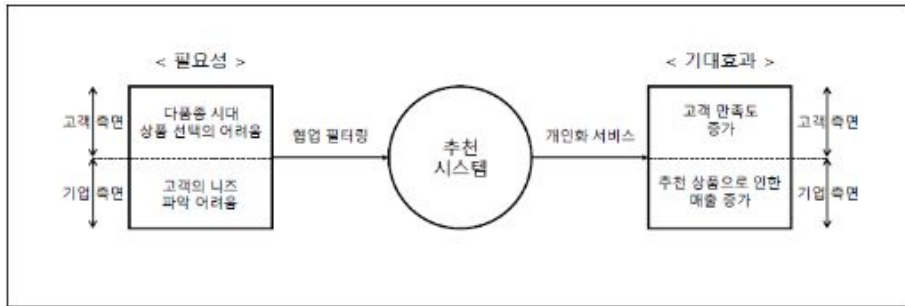
2.2 자바스크립트(JavaScript)

자바스크립트(JavaScript)는 객체 기반의 스크립트 프로그래밍 언어이다. 이 언어는 웹 브라우저 내에서 주로 사용되며, 다른 프로그램의 내장 객체에도 접근할 수 있는 기능이 있다. 자바스크립트의 일반적인 용도는 웹페이지에 기능을 더해 HTML 웹페이지를 동적이고 살아있게 하는 것이다. 자바 스크립트는 컴파일 과정이 없으므로 다른 언어와 비교했을 때 빨리 스크립트 코드를 작성할 수 있게 해주며 기존 C 나 자바와 달리 굉장히 단순한 구조와 원칙을 가지고 있으므로 쉽게 배우고 이해할 수 있고 객체 지향형 프로그래밍과 함수형 프로그래밍을 모두 표현할 수 있다는 장점이 있다[4].

2.3 추천시스템

추천 시스템은 사용자가 선호할 만한 아이템을 추측함으로써 여러 가지 항목 중 사용자에게 적합한 특정 항목을 선택(information filtering)하여 제공하는 시스템을 일컫는다. 여기서 ‘필터링’이란 여러 가지 항목 중 적당한 항목을 선택하는 기술을 말하는 IT 용어이다[6].

추천시스템의 종류는 그 구현방식에 따라 베스트셀러기반 추천방법, 인구통계 정보기반 추천방법, 최소질의대상 상품결정방법, 정보필터링 방법 등으로 구분된다. 그중에서도 정보필터링방법은 타 기법들보다 정확도가 높아 가장 일반적으로 적용되는데, 이는 다시 콘텐츠기반 필터링(content-based filtering) 기법과 협업 필터링(collaborative filtering) 기법으로 나누어 진다[1]. 추천시스템이 필요한 이유는 [그림 2-2]과 같다[7].



[그림 2-2] 추천시스템의 필요성

2.4 협업필터링 (Collaborative Filtering)

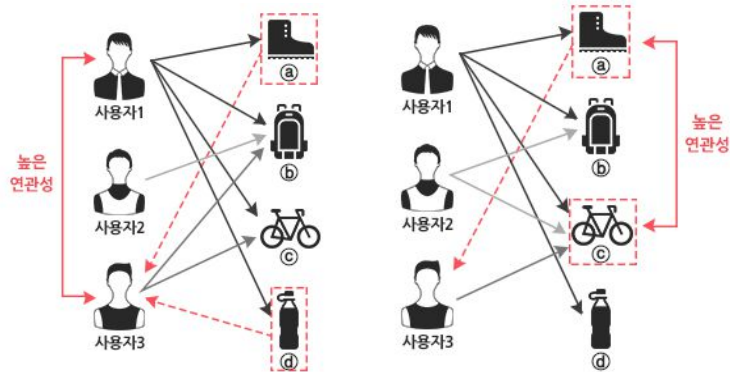
협업 필터링(collaborative filtering)은 많은 사용자로부터 얻은 기호정보 (taste information)에 따라 사용자들의 관심사들을 자동으로 예측하게 해주는 방법이다. 협필터링(collaborative filtering)은 크게 메모리 기반(Memory-based), 모델기반(Model-based), 하이브리드 기반(Hybrid)으로 나뉘어진다. 이중 가장 흔하게 사용되는 것은 메모리 기반 협업필터링 기법이다[1].

메모리기반 협업필터링 기법은 유사도를 기반으로 동작하며 메모리 기반 협업 필터링은 사용자기반(User-Based)과 아이템 기반(Item-Based)로 나뉘어진다.

아이템 기반 협업필터링은 [그림 2-4]과 같이 내가 이전에 구매했던 항목을 기반으로 아이템들의 유사도를 계산해 가장 유사도가 비슷한 아이템을 추천해 주는 방법이다. 아마존에서 판매의 35%가 아이템 기반 협업필터링을 활용한 추천으로 발생한다[8].

사용자 기반 협업필터링은 [그림 2-3]과 같이 사용자의 성향을 분석한 후 해당 사용자와 비슷한 성향을 가진 사용자들이 어떠한 상품을 선호하는지 분석해 추

천해 주는 방법이다. 실제 넷플릭스에서는 고객들의 영화 취향 데이터를 기반으로 사용자 기반 협업필터링을 활용하여 영화를 추천함으로써 구매의 2/3가 추천으로 발생한다[8].



[그림 2-3] 사용자기반 협업필터링(오)

[그림 2-4] 아이템기반 협업필터링(왼)

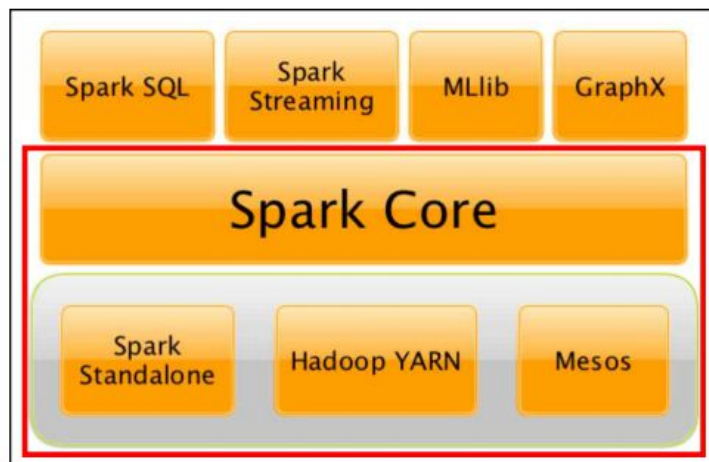
2.5 아파치 스파크 (Apache Spark)

스파크는 세련된 방법으로 클러스터의 여러 노드로 프로그램을 분배하고 그 위에서 동작하는 프로그램을 개발할 수 있도록 개발된 오픈 소스 프레임워크이다. 빅데이터 워크로드에 주로 사용되는 분산처리 시스템이다. 빠른 성능을 위해 인 메모리 캐싱과 최적화된 실행을 사용하고 일반 배치처리, 스트리밍분석, 머신러닝, 그래프 데이터베이스 및 임시 쿼리를 지원한다. 스파크의 데이터셋과 데이

터프레임을 통한 추상화는 개발자가 데이터 처리 파이프라인의 어느지점에서라도 데이터를 클러스터의 메모리로 저장할 수 있도록 해주며 인메모리 캐시 기능은 반복적으로 이뤄지는 크고 작은 모든 데이터 처리에 이상적이다[9][11].

기존의 하둡을 통해 끌어오는 데이터들은 시간 소요가 크기 때문에 실시간으로 분석해야 하는 업무에서 어려운 부분이 있어 새로 개발되었으며 스파크는 유연성과 맵리듀스에 대한 확장성을 훨씬 빠른 속도로 제공합니다. 데이터가 메모리에 저장돼 있을 때는 아파치 하둡보다 100배 빠르고, 디스크에 저장돼 있을 때는 10배 빠르다[10].

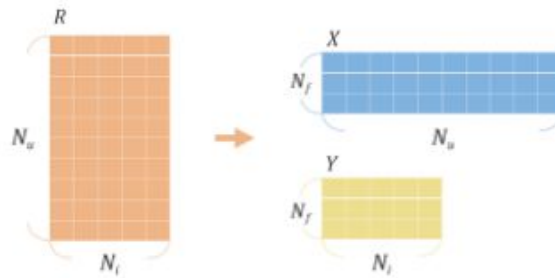
실제 은행, 통신업체, 게임업체, 정부를 비롯해 애플, 페이스북, IBM, 마이크로소프트와 같은 모든 주요 IT 기업들이 아파치 스파크를 사용한다. 스파크 실행 구성도는 [그림 2-5]와 같다[13].



[그림 2-5] 아파치 스파크

2.5.1 MF (Matrix Factorization)

MF는 행렬 인수 분해라는 수학적 방법으로 접근한다. 타겟 변수가 있는 지도 학습 기반 기계학습이며 회귀와 분류 모두 하는 방법이다. 이는 행렬은 두개의 하위 행렬로 분해할 수 있으며, 다시 곱해져서 원래 행렬과 동일한 크기의 단일 행렬이 될 수 있는 성질에 기인한 것이다. SGD(Stochastic Gradient Descen) 방법, ALS(Alternative Least Square) 또는 마코프체인 방법으로 훈련되는 시스템이다. 변수들을 저차원 공간에 매핑함으로써 변수의 상호작용을 모델링 한다[13]. 다음 [그림 2-6]은 평점 행렬을 사용자와 아이템 행렬로 분해한 그림이다.



[그림 2-6] 평점 행렬을 사용자와 아이템 행렬로 분해

R : 사용자가 아이템에 남긴 평점을 행렬로 나타냄

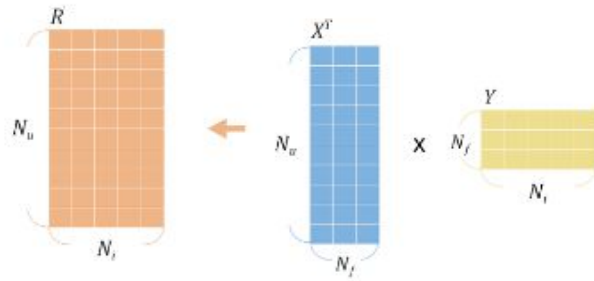
N_u : 사용자 수

N_i : 아이템 수

X : 사용자 잠복인자 행렬

Y : 아이템 잠복인자 행렬

N_f : Matrix Factorization 학습 시에 정하는 임의의 차원 수



[그림 2-7] 분해된 행렬을 다시 곱하여 예측 평점 행렬을 계산

위 [그림 2-7]과 같이 분해된 행렬을 다시 곱하여 예측 평점 행렬을 구한 후 Loss Function을 이용하여 예측 평점 행렬의 오차가 최대한 작아지도록 한다.

$$\min_{x^*, y^*} \sum_{u,i} c_{ui} (p_{ui} - x_u^T y_i)^2 + \lambda (\sum_u \|x_u\|^2 + \sum_i \|y_i\|^2)$$

X와 Y를 곱하여 얻은 예측 평점 행렬의 오차가 최대한 작아지도록 수식을 구성

2.5.2 ALS (Alternating Least Squares)

ALS는 먼저 사용자 혹은 아이템의 Latent Factor 행렬을 아주 작은 랜덤 값으로 초기화하고 둘 중 하나를 상수처럼 고정시켜 Loss Function을 Convex Function으로 만든다[13].

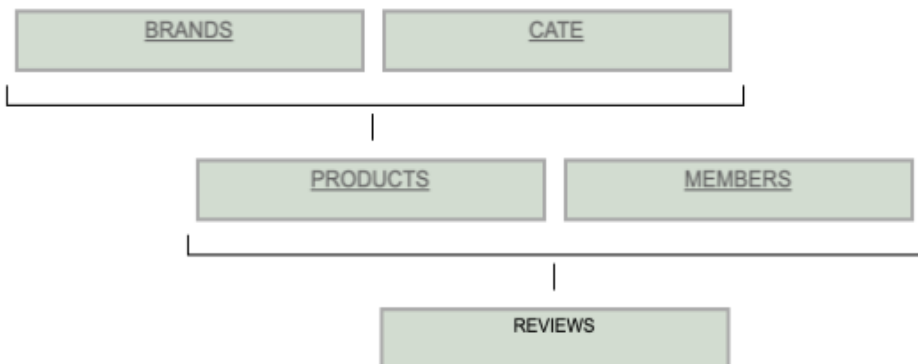
그리고 이를 미분한 다음, 미분 값을 0으로 만드는 사용자 혹은 아이템의 Latent Factor 행렬을 계산하는 과정을 사용자 한번, 아이템 한번 반복하면서 최적의 X, Y를 찾아내는 것입니다[13].

제 3 장 협업필터링을 활용한 화장품 추천시스템 설계

본 장에서는 협업필터링을 활용한 화장품 추천 시스템에 대해서 설계한다. 설계를 하기 위한 데이터 준비과정, 전체 시스템 구조와 화장품 추천 시스템의 구성도 및 문제점, 해결 과정을 설명한다.

3.1 협업필터링을 위한 데이터

협업 필터링을 활용한 화장품 추천 시스템을 개발하기 위해서는 우선 화장품을 보여주기 위하여 화장품 정보가 필요하다. 사용자에게 화장품에 대한 정보를 제공하고 다른 사용자들의 화장품 성향을 파악하기 위해 화장품에 대한 정보가 필요하여 JAVA를 이용하여 화장품 전체, 화장품 상세, 브랜드, 리뷰 등과 같은 html, JSON 파일을 파싱하여 [그림 3-1] 과 같이 나누어 저장하였다.



[그림 3-1] DB 구조도

카테고리 테이블 생성 후 화장품 사이트에서 HTML로 되어있는 페이지를 JAVA로 파싱하여 카테고리에 관한 정보를 [그림 3-2]와 같이 DB에 저장한다. 이때 카테고리는 대분류, 중분류, 소분류로 나누어 사용해야 하기 때문에 상위분류를 나타내는 칼럼과 현재 분류를 나타내는 칼럼을 생성하여 저장한다.

Results Explain Describe Saved SQL History				
NO	CATE_NO	NAME	SUPER_NO	CLASS
4	0002	메이크업	2	2
5	0003	바디케어	2	2
6	0004	파마케어	2	2
7	0005	향수	2	2
8	0001	화장품	3	3
9	0002	다크스맥	3	3
10	0006	클렌징	3	3
11	0007	선크림	3	3
12	0004	대모 로스메릭	3	3
13	0001	베이스	4	3
14	0005	네일케어	4	3
15	0006	립케어크림	4	3
16	0007	아이메이크업	4	3
17	0004	로션/크림	5	3
18	0005	워터스크림	5	3
19	0006	미스트/오일	5	3
20	0007	핸드/풋	5	3
21	0008	립케어	5	3
22	0008	샴푸/린스	6	3
23	0007	트리트먼트/팩	6	3
24	0006	스타일링/세션스	6	3
25	0005	얼모/원/크림	6	3
26	0003	여성향수	7	3
27	0004	남성향수	7	3
28	0001	스킨케어	8	4
29	0002	로션	8	4
30	0003	메이크업/세럼	8	4
31	0004	크림	8	4
32	0005	미스트	8	4

[그림 3-2] 카테고리 테이블

Results	Expain	Describe	Stored	SQL	History					
NO	ID	PASSWORD	NAME	BIRTHDATE	REGDATE	TYPE	GENDER			
M1000000500014	shin1	1111	이혁	01/01/2019	18-NOV-19 02:47:31.000000 AM	1	F			
M000000176530	sa****	1111	이혁	01/01/1975	09-NOV-19 02:48:46.885993 PM	-	F			
M000000176541	lhwk****	1111	이혁	01/01/1975	10-NOV-19 04:21:11.351249 AM	-	F			
M000000176513	w013055****	1111	이혁	01/01/2005	10-NOV-19 04:48:57.945697 AM	-	F			
M000000176420	dsejyho****	1111	이혁	01/01/1975	09-NOV-19 01:53:34.456580 PM	-	F			
M0000001763730	gjy****	1111	이혁	01/01/1985	10-NOV-19 05:15:56.676760 AM	-	M			
M0000001763621	coom****	1111	이혁	01/01/1975	10-NOV-19 05:18:15.338067 AM	-	M			
M0000001763563	claw****	1111	이혁	01/01/1985	10-NOV-19 05:16:17.120567 AM	-	F			
M0000001763452	wuzhenya****	1111	이혁	01/01/1975	10-NOV-19 05:18:15.432761 AM	-	F			
M0000001762216	msoe****	1111	이혁	01/01/1975	10-NOV-19 04:26:44.155475 AM	-	F			
M0000001761818	wbinus****	1111	이혁	01/01/1975	10-NOV-19 12:47:51.581377 PM	-	F			
M0000001668680	ewu****	1111	이혁	01/01/1996	10-NOV-19 03:40:27.500784 AM	-	F			
M0000001668199	loves****	1111	이혁	01/01/1996	10-NOV-19 01:22:59.198639 AM	-	F			
M0000001667695	dlst****	1111	이혁	01/01/1986	10-NOV-19 05:02:10.086662 AM	-	F			
M0000001666154	kongde****	1111	이혁	01/01/1996	10-NOV-19 05:15:53.194826 AM	-	F			
M0000001665226	plm****	1111	이혁	01/01/1986	09-NOV-19 04:09:43.982780 PM	-	F			
M0000001665189	sky****	1111	이혁	01/01/1986	09-NOV-19 06:53:41.679105 PM	-	F			
M0000001664641	jje****	1111	이혁	01/01/1975	10-NOV-19 01:23:00.546680 AM	-	M			
M0000001664018	skagne****	1111	이혁	01/01/1996	10-NOV-19 03:59:41.847000 AM	-	M			
M0000001663548	hongda****	1111	이혁	01/01/1996	09-NOV-19 01:57:01.211901 PM	-	M			
M0000001663342	ljdylq****	1111	이혁	01/01/1996	09-NOV-19 02:14:45.680704 PM	-	M			
M0000001663273	cljtk****	1111	이혁	01/01/1986	10-NOV-19 04:41:16.977750 AM	-	F			
M0000001662671	02****	1111	이혁	01/01/1996	09-NOV-19 04:28:35.233346 PM	-	F			
M0000001660160	acac****	1111	이혁	01/01/1996	09-NOV-19 09:07:14.208043 AM	-	M			
M0000001660159	sunng****	1111	이혁	01/01/1996	09-NOV-19 05:23:11.523629 AM	-	F			
M0000001659458	chdrja****	1111	이혁	01/01/1986	09-NOV-19 06:20:06.016500 AM	-	F			
M0000001659360	anot****	1111	이혁	01/01/1975	09-NOV-19 06:23:17.454820 AM	-	M			
M0000001659193	shu****	1111	이혁	01/01/1986	10-NOV-19 05:03:22.601048 AM	-	M			
M0000001658183	yjnu****	1111	이혁	01/01/1986	09-NOV-19 04:30:03.716085 PM	-	F			

[그림 3-5] 멤버 테이블

키보 타입 생성 등 히자프 사이트에서 ICON으로 되어있는 페이지를 JAVA

[illegible]

[그림 3-6] 리뷰 테이블

3.2 협업필터링을 활용한 화장품 추천 시스템 설계

[그림 3-7] 협업필터링을 활용한 화장품 추천 시스템 구성도이다



[그림 3-7] 협업필터링을 활용한 화장품 추천 시스템 구성도

사용자기반 협업필터링을 기반으로 사용자가 입력한 화장품의 평점을 분석해 해당 사용자와 유사한 성향을 가진 사용자들이 선호하는 화장품을 추천해주는 방식을 선택하였다. 이때 기존 사용자들에게는 추천이 잘 이루어지나 신규 가입자들에 대해 성향분석을 할 수 없어 추천해줄 수 없는 문제가 발생하였다.

이 문제를 해결하기 위해 신규 가입자들에게 최초 가입 시 자신이 이용해본 화장품들에 대한 평점을 입력받아 화장품에 대한 사용자 성향을 분석할 수 있을 때 화장품을 추천해 주게 하였다.

제 4 장 협업필터링을 활용한 화장품 추천시스템 구현

본 장에서는 협업필터링을 활용한 화장품 추천시스템 구현환경에 관해 기술하고 협업필터링을 활용한 화장품 추천시스템의 실제 구현된 웹 페이지와 간단한 설명을 기술한다.

4.1 협업필터링을 활용한 화장품 추천 시스템 구현 환경

< 표 4-1 > 은 협업필터링을 활용한 화장품 추천 시스템의 구현환경을 나타낸 표 이다.

<표 4-1> 협업필터링을 활용한 화장품 추천 시스템 구현 환경

시스템 구성요소	시스템 세부 내용
Server OS	MAC Java JDK 8.0 Tomcat 8.5
Database	Oracle
Web Language	HTML, JSP
API	Apache Spark
Develop Tools	Eclipse

Mac OS를 기반으로 Java JDK 8.0, Tomcat8.5을 사용하였고 Database는 Oracle을 사용하였다. Eclipse에서 JSP와 Servlet을 이용하여 협업필터링을 활용한 화장품 추천 시스템을 개발하였다.

4.2 협업필터링을 활용한 화장품 추천 시스템 구현

[그림 4-1]과 같이 메인에서는 화장품을 보여주지 않고 사이트에 관한 설명과 로그인을 유도하는 글을 띄워준다



[그림 4-1] 메인 화면

[그림 4-2]와 같이 화장품의 목록 창은 로그인을 하지 않아도 제품 메뉴에서 화장품에 대한 정보를 찾아볼 수 있다.



[그림 4-2] 분류별 인기순위 목록

[그림 4-3]은 로그인 화면이다.



[그림 4-3] 로그인 화면

[그림 4-4]와 같이 회원가입시 이미 사용중인 아이디가 있거나 비밀번호와 비밀번호 확인이 다를시 사용자에게 알려주며 내용을 전부 입력해야 작동한다.



[그림 4-4] 회원 가입 화면

[그림 4-5], [그림 4-6]과 같이 화장품의 상세정보 및 다른 사용자의 리뷰를 확인 할 수 있고 화장품에 대한 리뷰를 작성할 수 있다.



[그림 4-5] 화장품 상세 페이지



[그림 4-6] 화장품 리뷰 등록

신규 사용자, 기존 사용자가 리뷰를 작성하면 DB에 데이터를 저장하고 사용자, 상품, 평점 데이터를 불러와 평점행렬을 생성한다. 평점 행렬을 [그림 2-6]과 같이 사용자와 아이템 Latent Factor 행렬로 분해한 다음 아주 작은 랜덤 값으로 초기화 한다. 그 다음 둘 중 하나를 상수처럼 고정시켜 Loss Function을 Convex Function으로 만들고 이를 미분한 다음, 미분 값을 0으로 만드는 사용자 혹은 아이템의 Latent Factor 행렬을 계산한다. 이 과정을 사용자 한번 아이템 한번 반복하면서 최적의 값을 찾아낸다. 이를 통해 얻어진 최종 선호도 예측을 [그림 4-7]과 같이 화면에 출력해 준다[14].



[그림 4-7] 화장품 추천 목록

제 5 장 결 론

최근 인터넷과 스마트 기기의 발달로 사용자들은 어디서든 쉽게 아이템에 대한 데이터를 얻을 수 있다. 그러나 아이템에 대한 수가 급격하게 증가함에 따라 해당 분야에 대한 지식이 없는 경우 아이템을 선정하는 데 어려움을 겪는 경우가 많다. 이에 따라 최근 기업에서는 사용자들의 요구를 분석하여 아이템에 대한 추천을 제공하는 데 중점을 두고 있다. 추천시스템을 이용하면 사용자 개개인에게 필요에 맞는 아이템을 제공해 주기 때문에 기업은 고객을 유지하고 경쟁력을 확보시킬 수 있고 사용자들은 자신에게 적합한 아이템을 추천받아 사용할 수 있다.

본 논문에서는 여러 사이트의 상품과 사용자 데이터를 저장하고 분산처리 시스템인 아파치 스파크를 이용하여 많은 양의 데이터를 학습하고 학습을 통해 얻은 결과를 기준으로 사용자들의 성향을 분석하여 더욱 정확한 추천이 이루어지도록 하였다.

[참 고 문 헌]

- [1] 종합 평점과 다기준 평점을 선택적으로 활용하는협업필터링 기반 하이브리드 추천 시스템 2018
- [2] 협업필터링시스템의 발전 및 활용 방안 연구 2018
- [3] 오피니언 마이닝을 이용한 키워드 여론분석 시스템 2014.10
- [4] 위키백과 - <https://ko.wikipedia.org/wiki>
- [5] Beom's Blog - <http://myblog.opendocs.co.kr/archives/tag/jsp-컨테이너>
- [6] 콘텐츠 추천 알고리즘의 진화 - http://www.kocca.kr/insight/vol05/vol05_04.pdf
- [7] 도리의 디지털라이프 - <http://blog.skby.net/협업-필터링-collaborative-filtering/>
- [8] 협업 필터링을 활용한 맞춤 여행 정보 제공 모델 구축 - http://www.tourbigdata.kr/down/s3_bronze01_SA246.pdf
- [9] LG CNS - <https://blog.lgcns.com/940?category=515147>
- [10] 에이콘아카데미 공식블로그 - <https://m.blog.naver.com/acornedu/221083892521>
- [11] 샌디 라이자, 유리 레이저슨, 셀 오언, 조시 월스 , 2018,
9가지 사례로 익히는 고급 스파크 분석, Location : 한빛미디어
- [12] IT WORLD - <http://www.itworld.co.kr/news/107192>
- [13] Apache Spark - <https://jaceklaskowski.gitbooks.io/mastering-apache-spark/spark-overview.html>
- [14] 갈아먹는 추천 알고리즘 - <https://yeomko.tistory.com/5?category=805638>