

협업 필터링을 이용한 개인화 전자제품 추천 시스템

김성권*, 박두순*

*순천향대학교 컴퓨터소프트웨어공학과
e-mail : novaray@naver.com

The Personalized Electrical Goods Recommendation System using Collaborative Filtering

Sung-Kwon Kim*, Doo-Soon Park*

*Dept. of Computer Software Engineering, SoonChunHyang University

요 약

최근 전자제품들이 많이 출시되고 있다. 전자제품이 많고 특징도 매우 다양하며, 사용법도 복잡하다. 이런 이유로 쇼핑몰에서 전자 제품들을 직접 고르는 무척이나 힘들다. 그래서 사용자가 쇼핑몰에 들어갔을 때 사용자의 성향에 따라 사용자한테 가장 알맞은 전자 제품들을 추천받고 싶다. 사용자의 성향을 나이, 성별, 지역, 소득기준, 외제/국산에 따라 협업 필터링 방법으로 전자 제품을 추천하는 시스템을 제안한다.

1. 서론

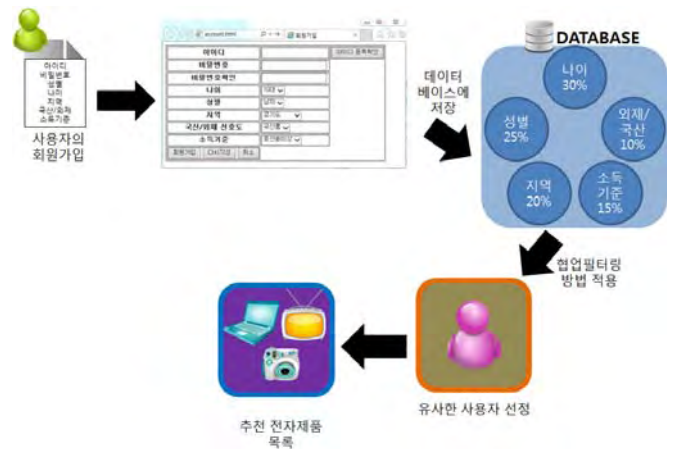
요즘 전자제품들은 많은 기능이 담겨있고 스마트한 전자제품이라고 해서 속히 말하는 똑똑한 전자제품들이 많이 출시되고 있다. 그래서 최근에는 전자제품을 평소에 잘 다루던 사람이라면 새로 나오는 제품이나 스마트 전자제품을 구매해도 알맞게 전자제품을 선택 할 수 있지만 잘 모르던 사람이 어떤 특정 전자제품을 구매하고자 할 때에는 그 제품이 기능이 무엇인지, 큰 어려움이 따르기 마련이다. 어떤 전자제품이 어떤 기능을 갖고 있고, 그 기능이 구매자가 필요시하는 기능인지 잘 모르는 경우가 발생한다. 또한 최근에는 한 전자제품 안에 한 기능만 있는 것이 아니라 여러 가지를 접목한 전자제품이 많기에 사용자가 원하는 전자제품을 찾는 것이 어렵다.

본 논문에서는 구매자가 해당 가전제품에 대한 추천을 사용자의 성향인 나이, 성별, 지역, 소득기준, 외제/국산에 따라 협업 필터링(collaborative filtering)을 이용하여 가전제품을 추천하는 시스템을 제안한다.

2. 전자제품 추천 시스템의 구성

추천시스템을 이용하려면 해당 전자제품에 대한 연령대별 만족도와 해당 전자제품의 복잡도가 필수적으로 필요하다. 개인화 요인으로 쓰일 것이기 때문이다.

(그림 1)은 추천 시스템 구성도이다. 회원 가입을 할 때 나이, 성별, 지역, 소득기준, 외제/국산 등을 작성하게 하고 그것에 따라 가중치를 부여해서 선호도가 유사한 사람을 찾아 그 사람이 추천한 전자 제품을 추천하는 시스템이다.



(그림 1) 추천 시스템 구성도

협업 필터링 기법으로 아이템을 추천하는 과정은 세 단계로 나누어진다.

단계 1. 사용자 간 유사도 계산

1단계에서는 추천 대상이 되는 사용자와 다른 사용자들 사이의 유사도를 산출하는 작업이 수행된다. 이러한 사용자 간 유사도에는 피어슨 상관계수(Pearson correlation coefficient)이 사용하도록 한다. 피어슨 상관계수를 이용하여 사용자 간 유사도를 산출하는 식은 다음 식과 같다.

$$\omega(A, B) = \frac{\sum_{i=1}^q (R_{A,i} - \overline{R_A})(R_{B,i} - \overline{R_B})}{\sqrt{\sum_{i=1}^q (R_{A,i} - \overline{R_A})^2 \sum_{i=1}^q (R_{B,i} - \overline{R_B})^2}}$$

상기 식에서 $\omega = (A, B)$ 는 사용자 A와 사용자 B의 유사도이고, i 는 사용자 A와 사용자 B가 공통으로 평가한 상품의 인덱스(index)이다. $R_{A,i}$ 는 상품 i 에 대한 사용자 A의 평가점수이고, $R_{B,i}$ 는 상품 i 에 대한 사용자 B의 평가점수이다. $\overline{R_A}$ 는 사용자 A의 평가점수 평균값이고, $\overline{R_B}$ 는 사용자 B의 평가점수 평균값이다.

단계 2. 이웃 선택

1단계에서 추천 대상이 되는 사용자와 다른 모든 사용자들 간의 유사도가 산출되고 나면, 이 유사도를 기반으로 추천 대상 사용자와 가장 유사한 N명의 이웃을 2단계에서 선택한다.

단계 3. 추천목록의 생성

마지막 3단계에서는 앞 단계에서 선택된 이웃들의 평가점수를 기반으로 추천 대상 사용자의 평가점수를 예측한 후 추천목록을 생성한다. 평가치 $R_{A,i}$ 는 고객 A의 아이템 i 에 대한 평가치로 고객 A의 최근접 이웃의 평가치들을 가중 평균하여 다음과 같이 예측한다.

$$R_{A,i} = \overline{R_A} + \frac{\sum_{j=1}^k \omega(A, j)(R_{j,i} - \overline{R_j})}{\sum_{j=1}^k |\omega(A, j)|}$$

여기서 $\overline{R_j}$ 는 고객 A의 최근접 이웃인 고객 j 의 이용 가능한 평가치들의 평균값을 뜻한다. 목표고객이 평가하지 않은 아이템들의 평가치를 예측한 후 Top-N 기법으로 추천목록을 생성한다[1]. 즉, 목표고객의 예측된 평가치 중에서 수치가 높은 상위 N개의 아이템을 목표고객에 대한 추천목록으로 생성한다[2][3][4].

본 논문에서의 협업 필터링 추천 방법은 회원가입을 할 때 등록한 나이, 성별, 지역, 소득기준, 외제/국산을 기반으로 각각의 가중치를 나이(30%), 성별(25%), 지역(10%), 소득기준(15%), 외제/국산(20%)으로 부여한다. 가중치 중에서 나이는 10~ 90대의 분들이 선호하는 전자 제품이 다를 것이라고 예상해서 구분하였고, 지역은 대한민국 안의 8개의 도와 7개의 시로 구별해서 선택 할 수 있게 하였다. 또한, 외제/국산은 외제품과 국산품 중 무엇을 더 선호하는지 선택하게하며, 소득기준은 중산층 이상, 이하로 나누어 세 부담이 증가하는 연봉 5500만원을 기준으로 이상/이하로 나누게 한다[5].

성향이 유사한 사용자들을 그룹화하고 그룹군의 평점 정보를 가지고 전자 제품을 추천하게 된다.

3. 전자제품 추천 시스템의 구현

본 논문에서의 전자제품 추천 시스템은 먼저 사용자가

간단하게 회원가입을 먼저 거쳐야 한다. 회원가입 절차는 (그림 2)와 같다.

아이디	<input type="text"/>	아이디 중복확인
비밀번호	<input type="password"/>	
비밀번호확인	<input type="password"/>	
나이	10대 ▾	
성별	남자 ▾	
지역	경기도 ▾	
국산/외제 선호도	국산품 ▾	
소득기준	중산층이상 ▾	
<input type="button" value="회원가입"/> <input type="button" value="다시작성"/> <input type="button" value="취소"/>		

(그림 2) 회원가입 절차

다음 (그림 3)은 데이터베이스에 저장 되어있는 정보는 사용자들의 회원가입을 통해서 얻은 정보이다. 이 데이터베이스에서 회원가입 되어있는 유저들의 일부분을 보여주고 있다.

ID	NAME	GENDER	AGE	AREA	PRODUCT	MONEY
nova03	김성권	남	20	경기도	국산	1
bat1234	박주현	남	20	서울특별시	외제	2
leets12	이태수	남	30	충청남도	국산	2
lleyh86	이영희	여	30	전라북도	국산	2
changpeace	이창현	남	10	경상남도	외제	1
parkgj	박관주	남	10	경기도	국산	1
kimmk	김민기	남	20	인천광역시	국산	1
jinzyman	조수연	남	20	경기도	외산	1

(그림 3) 데이터베이스에 저장되어 있는 값

전자제품 정보가 들어있는 데이터베이스가 필요하다. 다음 그림들은 각 카테고리 별로 나뉘어져 있는 그림이다.

(그림 4)는 20대가 산 전자제품 품목들이고, (그림 5)는 여자가 산 전자제품 품목, (그림 6)은 지역이 '경기도'인 사람이 산 전자제품 품목을 볼 수 있다.

ID_NUM	AGE	PRODUCT
1	20	SRS-X1
2	20	MW235SK
7	20	25490-KR
8	20	BL142

(그림 4) 20대가 산 전자제품 품목 일부

ID_NUM	GENDER	PRODUCT
4	여	BKP-H31DP
9	여	FD-08BL
11	여	HD-4646
12	여	PM9631

(그림 5) 여자가 산 전자제품 품목 일부

ID_NUM	AREA	PRODUCT
1	경기도	SRS-X1
6	경기도	RQ-1145
8	경기도	BL142
12	경기도	FD-08BL

(그림 6) 경기도 지역에서 산 전자제품 품목 일부

다음으로 이제 앞서 가중치를 준 것을 바탕으로 성향이 비슷한 유저를 찾아야 한다. (표 1)은 사용자와 가장 유사한 사용자를 찾기 위해 추천받을 유저와 많은 사용자들을 각 가중치에 따라 계산한 다음, 내림차순으로 정렬한 결과이고 상위 3명의 사용자를 뽑아낼 수 있다.

유저 번호	나이	성별	지역	소속기준	외제/국산	가중치 합	비교 유저 번호
1번 유저	0.27	0.3	0.074	0.15	0.2	0.994	94
1번 유저	0.24	0.3	0.088	0.15	0.2	0.978	37
1번 유저	0.24	0.3	0.08	0.15	0.2	0.97	95
1번 유저	0.21	0.3	0.1	0.15	0.2	0.96	20
1번 유저	0.21	0.3	0.095	0.15	0.2	0.955	93
1번 유저	0.3	0.3	0.08	0.075	0.2	0.955	30
1번 유저	0.3	0.3	0.08	0.075	0.2	0.955	92
1번 유저	0.3	0.3	0.1	0.15	0.1	0.95	59
1번 유저	0.27	0.3	0.08	0.075	0.2	0.925	65
1번 유저	0.3	0.3	0.06	0.15	0.1	0.91	24
1번 유저	0.18	0.3	0.074	0.15	0.2	0.904	82
1번 유저	0.27	0.3	0.08	0.15	0.1	0.9	19
1번 유저	0.27	0.3	0.054	0.075	0.2	0.899	34
1번 유저	0.27	0.3	0.054	0.075	0.2	0.899	83
1번 유저	0.3	0.15	0.095	0.15	0.2	0.895	75
1번 유저	0.15	0.3	0.095	0.15	0.2	0.895	86
1번 유저	0.24	0.3	0.095	0.15	0.1	0.885	3
1번 유저	0.24	0.3	0.068	0.075	0.2	0.883	46
1번 유저	0.3	0.15	0.08	0.15	0.2	0.88	60
1번 유저	0.27	0.3	0.03	0.075	0.2	0.875	74
1번 유저	0.21	0.3	0.08	0.075	0.2	0.865	5

(표1) 가중치들의 일부 결과 표

가중치를 계산한 표에서 1번 유저에서 가중치의 값이 제일 높은 것을 순차적으로 3개를 뽑아낸다. (그림 7)은 1번 유저한테서 성향이 비슷한 유저 94번, 37번, 95번 유저한테서 전자제품 3개를 추천하는데 Top 3 유저한테서 한 개씩의 전자제품을 골라내서 사용자에게 추천을 해주는 시스템이다.

(그림 7)은 데이터베이스에서 쿼리문을 써서 Top 3에 드는 94, 37, 95번 유저를 추출한 결과이다.

ID	ID_NUM	NAME	PURCHASE
jung91	94	정명석	T1407W8
Closer	37	원미영	YN-101
limda88	95	임다솔	CDD-9037T

(그림 7) Top 3 유저 추출

이제 Top 3안에 든 유저들의 전자제품 정보를 추출해야 한다. 전자제품 정보가 들어있는 데이터베이스를 이용해서 추천 받고자 하는 사용자에게 전자제품 정보와 사진들을 함께 보여준다. (그림 8)는 전자제품 정보가 들어있는 데

이터베이스에서 쿼리문을 이용해 해당 전자제품의 정보들을 추출해 낸 결과이고 (그림 9)은 추천 결과를 띄운 화면이다.

PURCHASE	OPTION1	OPTION2	OPTION3
T1407W8	물살세기 3단계	물높이 단계 10단계	에너지효율등급 2등급
YN-101	연속가습시간 9시간	최대 분무량 250cc	소비전력 35W
CDD-9037T	건조전용	용량 9인용	가정용

(그림 8) 전자제품 정보 데이터베이스



(그림 9) 전자제품 추천 결과 화면

4. 결론

본 논문에서는 협업 필터링 기법을 이용해서 개인화 추천 시스템을 구현하였다.

향후 연구과제로는 본 논문에서 사용한 개인화 요인 이외에 많은 전자제품 중에서 사용자에게 알맞은 전자제품들을 추천할 수 있게끔 사용자가 필요로 하는 전자제품들을 찾을 수 있게 개인화요소를 더 추가하거나 다른 방법을 모색하여 사용자에게 알맞은 전자제품들을 추천할 수 있게 하는 것이 필요한 것으로 보인다.

Acknowledgments

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the C-ITRC(Convergence Information Technology Research Center) (IITP-2015-IITP-2015-H8601-15-1009) supervised by the IITP(Institute for Information & communications Technology Promotion).

참고문헌

- [1] 박석두, "추천시스템의 성능 향상을 위한 단계적 협업 필터링의 설계 및 구현," 아주대학교 일반대학원 석사 논문, 2007. 1

- [2] 전병국, 안현철, “사용자 리뷰 마이닝을 결합한 협업 필터링 시스템: 스마트폰 앱 추천에의 응용,” 지능 정보 연구, 21권 2호, pp1-18, 2015. 6
- [3] 김선호, 박두순, “개인의 감정과 협업필터링을 이용한 개인화 영화 추천 시스템. 한국정보처리학회 추계학술 발표 논문집, pp1176-1178, 2013. 11
- [4] 박석두, “단계적 협업필터링을 이용한 추천시스템의 성능 향상,” 한국지능정보시스템학회 춘계학술대회 논문집, pp218-225, 2007. 5
- [5] 중산층 살린다더니 세 부담 늘려 -
<http://www.newspim.com/view.jsp?newsId=20150120000320> 2015, 1