
협업필터링 기법을 이용한 모바일 광고 추천 시스템

김은숙* · 윤성대*

*부경대학교

Using collaborative filtering techniques Mobile ad recommendation system

Eun-suk Kim* · Sung-dae Yoon*

*Pukyong National University

E-mail : myway85@hanmail.net

요 약

최근 모바일 시장이 급속도로 성장함에 따라 현대인들은 컴퓨터가 가지는 여러가지 제약들을 극복하여 원하는 정보를 빠르게 얻을 수 있는 수단으로 모바일 콘텐츠 사용이 늘고 있다 그러나 광범위한 콘텐츠의 추천으로 콘텐츠의 선택에 있어 많은 시간이 소요된다 이러한 문제점을 해결하기 위해 사용자가 원하는 콘텐츠를 예측하여 정확하게 추천해 줄 수 있는 시스템이 필요하다

본 논문에서는 이러한 사용자들의 요구에 맞추어 원하는 콘텐츠를 제공하기 위하여 협업필터링을 이용하여 추천의 선택 횟수를 높일 수 있는 방법을 제시한다 첫번째 단계에서 대분류로 카테고리를 구분하여 목표고객과 사용자간의 유사도를 구하고 최근접 이웃을 구성하여 대분류 카테고리간 선호도 예측값을 구하여 가장 높은 대분류 카테고리를 목표고객에게 추천한다 두 번째 단계에서 소분류 카테고리 간 선호도 예측값을 구하여 가장 높은 소분류 카테고리를 목표고객에게 추천한다 실험에서 대분류 카테고리 기반 협업필터링으로 모바일 콘텐츠를 추천하고 소분류 카테고리 기반 협업필터링으로 모바일 콘텐츠를 추천해 두 가지 방법의 결과를 비교하여 소분류 카테고리 기반 협업필터링의 방법이 선택 횟수가 높다는 것을 검증하였다.

ABSTRACT

Due to recent rapid growth of mobile market, the modern people increasing make use of mobile contents as a means to obtain the desired information quickly by overcoming various restraints of a computer. The wide range of recommended contents, however, takes much time in selection of contents. To resolve such issues, a system that predicts the contents desired by the user and makes an accurate recommendation is necessary.

In this paper, in order to provide the desired contents in line with the user demands, a method to increase select the number of recommendation using cooperative filtering is proposed. In the first step, the categories are formulated with super-classes and the similarity between the target customer and users is found, and the nearest-neighbors are constituted to find the preference predictions between super-classes, and the super-class with the highest resulting value is recommended to the target customer. In the second step, the preference predictions between sub-classes are found and the sub-class with the highest value is recommended to the target customer. In the experiment, mobile contents are recommended through super-class-based cooperative filtering, and then the mobile contents are recommended through sub-class-based cooperative filtering, and sub-class collaborative filtering method to select a high number of verification.

Key Word

협업필터링, 모바일 콘텐츠 추천, 카테고리 기반

I. 서론

최근 모바일 사용이 급격히 증가됨에 따라 걸어 다니면서 e-mail을 보내고 동영상을 볼 수 있는 이동성, 많은 친구들을 하나로 묶어 언제든 대화할 수 있는 소통, 원하는 정보를 어디서든 접근할 수 있는 접속성 등의 혜택을 소비자들이 누릴 수 있게 되었다. 또한 소비자 개인의 특성을 보다 정확하게 파악해서 최적의 서비스를 제공하려는 개인화의 흐름이 중요시 되고 있다[1]. 이러한 개인화의 흐름에 맞추어 개인화 추천 서비스가 제공되고 있는데 개인화 추천 서비스란 구매 행동 분석을 토대로 사용자의 요구에 맞는 서비스 및 상품을 추천해주는 시스템이자[2], 사용자들이 선호하는 서비스 및 상품을 쉽게 찾을 수 있도록 도와주는 정보 필터링 기술이기도 하다

모바일 콘텐츠 광고 시장은 다양한 사용자 층의 확산으로 인해 미디어로서의 기본적인 가치 평가가 되고 있으며, 낮은 비용(온라인 콘텐츠 대비 50% 수준)으로 광고주가 목표 고객에게 직접 접근할 수 있어 더욱 가치가 높게 평가되고 있다[4].

모바일 콘텐츠 시장 규모가 커짐에 따라 경쟁도 치열해졌다. 경쟁이 치열해진 만큼 철저한 시장조사와 빠르고 정확한 정보의 습득을 바탕으로 사용자의 선호도에 따른 콘텐츠를 추천하는 시스템을 개발하는 일이 중요하다고 볼 수 있다

본 연구에서는 사용자가 원하는 모바일 콘텐츠를 추천하는데 있어 사용자가 원하는 콘텐츠를 추천하기 위한 방법으로 아이템을 소분류로 세분화하여 콘텐츠를 추천하여 목표고객의 콘텐츠 선택의 횟수가 높은 시스템을 제시하고 실험을 통해 타당성을 검토한다

II. 관련연구

2.1 추천 시스템(Recommendation System)

추천 시스템은 고객이 관심을 가지는 상품에 관한 정보나 인구통계학적 정보, 과거 구매행동 분석을 토대로 고객의 요구에 맞는 상품을 추천해주는 시스템을 말한다[4].

추천 시스템은 현재 Amazon.com, CDnow.com 등의 전자상거래 사이트에서 적용되고 있으며[5], 동영상 추천, 음악 추천, 도서 추천 등의 다양한 분야에서 이용이 되고 있는 추세이다

2.2 협업 필터링 기법(Collaborative Filtering)

(1) 협업 필터링 추천 시스템의 개요

협업필터링 추천 방법이란 많은 사용자들로부터 얻은 기호정보(taste information)에 따라 사용자들의 관심사들을 자동적으로 예측하게 해주는 방법이다. 고객들의 선호도와 관심 표현을 바탕으로 선호도 관심에서 비슷한 패턴을 가진[6] 고객들을 식별해 내는 기법이다. 비슷한 취향을 가진 고객들에게 서로 아직 구매하지 않은 상품들은 교차 추천하거나 분류된 고객의 취향이나 생활 형태에 따라 관련 상품을 추천하는 형태의 서비스를 제공하기 위해 사용된다

협업 필터링 기법은 사용자 기반 협업 필터링(user-based collaborative filtering)과 아이템 기반 협업 필터링(item-based collaborative filtering)으로 나눌 수 있다[3].

(2) 협업 필터링 추천 방법의 문제점

기존 협업 필터링의 문제점은 사용자 기반 협업 필터링의 경우 사용자와 아이템의 수가 많아지면 연산량이 증대되어지고[3], 선호도가 입력되지 않은 상품의 수가 많아짐에 따라 신뢰성이 다소 낮아진다는 문제점이 있다[7].

III. 설계 및 구현

본 연구에서는 목표고객에게 원하는 모바일 콘텐츠를 추천하기 위한 방법으로 카테고리 대분류로 추천하는 방법보다, 카테고리 소분류로 추천하였을 경우 목표고객이 원하는 콘텐츠를 선택하는 횟수가 높은 시스템을 제안한다

3.1 카테고리 기반 협업필터링 실험

사용자와 목표고객 Data를 이용하여 대분류 카테고리 기반 협업 필터링의 콘텐츠 추천 방법과 소분류 카테고리 기반 협업 필터링의 콘텐츠 추천 방법을 구현해보고, 목표 고객이 콘텐츠를 선택하는 횟수를 비교하여 평가한다.

1단계 : 평가치 매트릭스 준비

목표고객에게 모바일 콘텐츠를 추천하기 위하여 5개의 카테고리에 대한 목표고객과 20명의 고객들의 평가치를 가져온다

표 1 . 목표고객 평가치

대분류 사용자	대분류 1	대분류 2	대분류 3	대분류 4	대분류 5
Target	7	0	5	4	0

표 2 . 사용자 평가치

대분류 사용자	대분류 1	대분류 2	대분류 3	대분류 4	대분류 5
A	7	1	3	0	0
B	3	0	1	0	1
C	5	0	2	1	3
D	4	1	2	0	6
E	2	0	5	2	7
:					
R	6	0	0	1	0
S	0	4	6	0	1
T	3	1	4	1	3

표1, 표2는 목표고객과, 사용자들이 콘텐츠를 선택 했을 때 각 카테고리 별로 누적된 데이터 횟수를 나타낸다.

2단계 : 최근접 이웃 구성

목표고객과 다른 사용자들의 유사도를 피어슨상관계수 방법을 사용하여 구하고 K-NN기법으로 표 3과 같이 최상위 5명으로 최근접 이웃을 구성한다

표 3 . 목표고객과 사용자간 유사도

사용자	유사도
A	0.78
B	0.655
I	0.82
O	0.79
R	0.73

3단계 : 대분류 카테고리 유사도 계산

K-NN기법으로 선정된 5명의 이웃들이 본 대분류 카테고리를 대상으로 대분류 카테고리간의 유사도를 계산한다.

표 4 . 대분류 카테고리간 유사도

목표고객 사용자	대분류1	대분류2	대분류3	대분류4	대분류5
대분류1	1	1.47	2.64	1.44	0.63
대분류2	1.47	1	1.15	0.82	0
대분류3	2.64	1.15	1	0.47	0.41
대분류4	1.44	0.82	0.47	1	0
대분류5	0.63	0	0.41	0	1

표4는 목표고객과 최근접 이웃으로 선정된 5명과의 대분류 카테고리간 유사도를 나타낸다

4단계 : 대분류 선호도 예측 값 계산

대분류 카테고리에 대한 선호도 예측 값을 계산하여 값이 가장 높은 대분류 카테고리를 추천한다

표 5 . 선호도 예측값

대분류 사용자	대분류 1	대분류 2	대분류 3	대분류 4	대분류 5
Target	7	5.62	5	4	6.21

표5와 같이 목표고객이 선택하지 않은 대분류 중 선호도 예측 값이 높은 대분류에 해당하는 콘텐츠를 추천한다.

5단계 : 소분류 카테고리 정보 준비

선호도 예측값이 가장 높은 대분류 카테고리에 해당하는 목표고객, 사용자의 소분류 카테고리의 선택 정보를 그림1, 그림2와 같이 준비한다.

	대분류 1			대분류 2			대분류 3			대분류 4			대분류 5		
	소분류 1	소분류 2	소분류 3	소분류 1	소분류 2	소분류 3	소분류 1	소분류 2	소분류 3	소분류 1	소분류 2	소분류 3	소분류 1	소분류 2	소분류 3
target	6	0	3	0	0	0	2	4	5	4	5	5	3	2	1

그림 1 . 목표고객 소분류 카테고리 선택 정보

	대분류 1			대분류 2			대분류 3			대분류 4			대분류 5		
	소분류 1	소분류 2	소분류 3	소분류 1	소분류 2	소분류 3	소분류 1	소분류 2	소분류 3	소분류 1	소분류 2	소분류 3	소분류 1	소분류 2	소분류 3
A	10	4	1	3	0	0	0	5	2	0	0	0	0	0	0
B	1	2	6	0	0	0	4	0	1	0	0	0	1	3	6
C	4	0	1	0	0	0	2	1	6	1	2	5	1	6	0
D	1	0	4	1	5	2	0	7	0	0	0	0	4	6	3
E	5	0	1	0	0	0	0	5	1	3	1	4	3	1	2
F	0	0	0	0	0	0	2	4	8	2	1	0	2	5	0
G	0	0	0	2	7	5	0	2	0	1	1	4	1	4	1
:															
Q	9	0	4	5	3	0	0	0	0	3	2	4	6	5	3
R	8	3	5	0	0	0	0	0	0	0	0	0	2	4	3
S	0	0	0	7	3	3	3	5	4	0	0	0	0	8	3
T	8	5	5	4	2	5	3	8	1	1	5	2	5	8	3

그림 2 . 사용자 소분류 카테고리 선택 정보

6단계 : 소분류 카테고리간 유사도

이웃 고객들이 본 대분류 5에 해당하는 소분류 카테고리와 목표고객이 본 소분류 카테고리를 대상으로 소분류 카테고리간 유사도를 표6과 같이 계산하여 나타낸다.

표 6 . 소분류 카테고리간 유사도

목표고객 사용자	소분류1	소분류2	소분류3
소분류1	1	2.25	2.31
소분류2	2.25	1	3.56
소분류3	2.31	3.56	1

7단계 : 소분류 선호도 예측 값 계산

계산된 유사도 테이블 각각의 유사도 값을 가중치로 사용하여 목표 고객에게 추천하고자 하는 콘텐츠 카테고리에 대한 선호도 예측 값을 계산하여 가장 높은 소분류 카테고리를 추천한다

표 7 . 소분류 선호도 예측 값

	소분류1	소분류2	소분류3
Target	1.45	1.77	2.4

표7과 같이 목표 고객에게 소분류에 해당하는 콘텐츠를 추천함을 알 수 있다

3.2 실험 평가

사용자 20명의 데이터를 사용하여 모바일 피부과 진료 항목 광고 150개를 8일간 대분류, 소분류로 나누어 진행 했을 때 목표고객 선택 값이 일자가 경과함에 따라 어떻게 변화 하는지 관찰한다.

그림3은 협업필터링을 이용한 대분류 콘텐츠 추천 방법과, 소분류 콘텐츠 추천 방법을 이용했을 때 목표고객의 선택값의 변화를 그래프로 표현한 것이다.

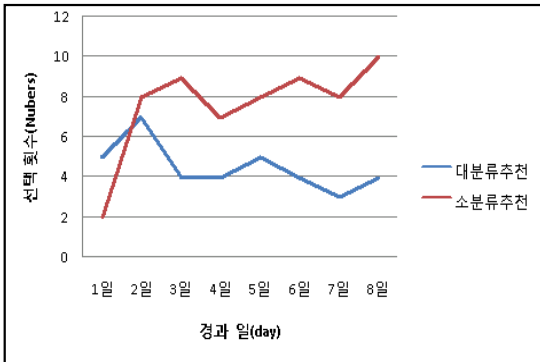


그림 3 . 날짜경과에 따른 목표고객 선택 횟수

소분류 카테고리 기반 협업 필터링의 추천 방법이 대분류 카테고리 기반 협업 필터링 방법으로 콘텐츠 추천을 진행한 경우보다, 목표고객에게 추천 되는 콘텐츠의 추천 범위가 좁아지므로 일자가 지남에 따라 목표고객이 콘텐츠를 선택하는 횟수가 높게 나타난다는 것을 알 수 있다.

IV. 결론 및 향후 연구

전세계 모바일 콘텐츠 시장이 늘어남에 따라 사용자의 요구도 높아지고 있다. 사용자에게 광범위한 콘텐츠를 추천하는 것보다 목표고객이 원하는 콘텐츠를 추천해주는 시스템이 중요한 역할을 할 것이다

본 논문에서는 카테고리 협업 필터링 기반의 모바일 진료과목 추천과정에서 기존의 대분류 카테고리 기반 협업 필터링의 추천방법의 문제점 중 추천 범위가 넓은 문제점을 해결하고자 소분류 카테고리 기반의 협업필터링 방법을 제안하여 목표고객에게 추천해본 결과 목표고객의 콘텐츠 선택 횟수가 높았다

향후 연구 과제로는 제시한 대분류 소분류의 분류 방법을 보다 다양한 분야에 적용하고 정확성을 높이기 위한 방법에 대한 연구가 필요할 것이다

참고 문헌

[1] Pottoniee, O., "Ubiquitous Personalization: a Smart Card Based Approach, Gemplus Research Labs, BP 100, 138811 Gemenos Cedex, 2000.

[2] Mulvenna, M., Sarabjot, S. A. and Alex, G. B., "Personalization on the Net Using Web Mining; Interoduction", Communications of the ACM, Vol. 43, No. 8, 2000.

[3] JL Herlocker, JA Konstan, J Riedl, A Bochers, "An algorithmic frame work for performing-collaborative filtering", Association for Computing Machinery, pp. 230-237, 1999.

[4] 김은규, 조인휘, "카테고리 협업 필터링 기반의 모바일 콘텐츠 추천 시스템 설계, 한양대학교, pp. 1-16, 2012.

[5] 최현석, 배효철, 서정진, 윤경로, "감성 및 상황 정보 융합 기반의 확장된 협업 필터링 기법을 이용한 음악추천시스템", 한국방송공학회 학술발표대회pp. 82-84, 2011.

[6] 강민정, "스마트폰 이용자의 어플리케이션 구매를 위한 추천 시스템, 홍익대학교, 2011.

[7] 김재경, 안도현, 조운호, "개인별 상품 추천 시스템, WebCF-PT 웹마이닝과 상품계층도를 이용한 협업필터링", 경영정보학연구, 제 15권 제 1호, pp.63-79, 2005.