건강관리를 위한 혼합 필터링을 이용한 개인화 식이영양 콘텐츠 추천

Personalized Dietary Nutrition Contents Recommendation using Hybrid Filtering for Managing Health

정경용*, 이영호**

상지대학교 컴퓨터정보공학부*, 가천의과학대학교 의료공학부**

Kyung-Yong Chung(dragonhci@hanmail.net)*, Young-Ho Lee(leeyh@gachon.ac.kr)**

요약

차세대 IT융합기술의 발전에 따라 개인화 건강관리 서비스를 위한 인프라스트럭처가 구축되면서, 사용자의 선호도 기반 서비스의 중요성이 부각되고 있다. 만성질환자가 증가함에 따라 건강관리 서비스는 특정 질병의 치료 및 관리에서 서비스 대상자에 대한 식이영양 관리로 진화하고 있다. 본 논문에서는 건강관리를 위한 혼합 필터링을 이용한 개인화 식이영양 콘텐츠 추천을 제안하였다. 제안한 방법에서는 유사한 콘텐츠만을 제공하는 추천 특수화 경향을 개선하기 위해 협력적 필터링과 이미지 기반 필터링을 결합한 혼합 필터링을 사용한다. 이를 웹 어플리케이션으로 구축하여 논리적 타당성과 유효성을 검증하기 위해 실험적인 적용을 시도한다. 따라서 식이영양 콘텐츠를 추천하여 건강관리에 대한 만족도와 서비스의 질을 향상시켰다. 연구결과를 활용하면 시장성 증대와 고부가 가치를 창출할 수 있을 것으로 기대하며 다양한 응용분야에 활용이 가능하다.

■ 중심어 : | 선호도 | 식이영양 | 추천 | 시뮬레이션 | 협력적 필터링 | 만성질환자 |

Abstract

With the development of next IT convergence technology and the construction of infrastructure for personalized healthcare services, the importance of services based on user's preference is being spotlighted. Healthcare service have been progressed as treatment and management for specific diseases and dietary nutrition managements to customers according to the increase in chronic patients. In this paper, we proposed the personalized dietary nutrition contents recommendation using the hybrid filtering for managing health. The proposed method uses the hybrid filtering through combining the collaborative filtering and the image filtering in order to reinforce the special trend that recommendation provides similar contents. We developed the Web application for this purpose, and experimented with it to verify the logical validity and effectiveness. Accordingly, the satisfaction and the quality of services will be improved the healthcare by recommending the dietary nutrition contents. This evaluation found that the difference of satisfaction by service was statistically meaningful and showed high satisfaction.

■ keyword : | Preference | Nutrition Contents | Recommendation | Simulation | Collaborative Filtering | Chronic Patient |

"본 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임 (No. 2011-0008934)"

접수번호 : #110504-001

접수일자 : 2011년 05월 04일

심사완료일 : 2011년 06월 16일

교신저자 : 이영호, e-mail : leeyh@gachon.ac.kr

I. 서론

최근 IT융합기술의 발전으로 지능화된 환경과 개인이 모바일 기기를 이용한 상호 통신이 가능해지면서 건강 관리 서비스에 대한 연구가 부각되기 시작했으며 사회, 경제, 문화면에서 관심의 초점이 되고 있다. 기존의 경제 환경에 의존하던 기업들이 새로운 환경에 적응하기 위해 인터넷이라는 흐름에 합류하기 시작하였고, 인터넷 기반으로 하는 원투원 마케팅 사업에 과감한 투자를 아끼지 않고 있다[1][4][8]. 특히, 전자상거래의 발달은 새로운 시장의 창출 및 고객에 대한 서비스 강화를 위하여, 고객들의 선호도와 구매 내역을 분석하여 차별화된 정보를 자동으로 필터링하여 추천하는 개인화 추천 시스템을 출현하게 하였다[2][3]. 개인화 추천은 사용자에게 자신의 선호도에 부합하는 상품이나 정보를 제공함으로써 이를 얻기 위한 시간과 비용을 줄여주고, 손쉽게 상품이나 정보에 접근하도록 하는 장점을 갖는다. 또한 개인화를 통해 자신과 비슷한 유형을 갖는 사람들과 교류할 수 있는 기회를 가질 수도 있다. 이러한 개인화 콘텐츠 추천 시스템은 정보 검색이나 정보 필터링의 일부분으로 구축된다. 추천 시스템은 사용자가 어떤 정보에 관심이 있는지, 어떤 정보가 유용한지를 예측하는 일에 초점을 맞춘다. 추천 시스템의 예측은 사용자의 프로파일을 활용하여 개인화된다. 이러한 개인화 아이템 추천 시스템에서 사용되는 추천 기법은 내용 기반 필터링과 협력적 필터링이 주류를 이루고 있다. 본 논문에서는 협력적 필터링과 이미지 필터링을 결합한 혼합필터링을 이용하여 개인화 식이영양 콘텐츠 추천을 제안하였다. 만성질환자를 위한 건강관리 서비스를 위해서는 의료 서비스와 의학적 가이드라인에 기반을 두어 자기관리를 통해 건강증진을 위해 노력을 해야 한다[10]. 최근 고혈압과 당뇨병자 대상의 건강관리 서비스가 소개되고 있으며, 다양한 디바이스와 IT융합기술을 활용한 솔루션에 많은 투자와 연구가 진행되고 있다.

본 논문의 구성은 2장에서는 관련기술에 대해서 기술하고, 3장에서는 제안하는 건강관리를 위한 혼합 필터링을 이용한 식이영양 콘텐츠 추천에 대해서 기술한다. 4장에서는 성능 평가를 기술하고 5장에서는 결론과 향후

연구에 대해서 기술한다.

II. 관련 기술

본 논문에서 제안한 방법을 기술하기에 앞서 콘텐츠 추천 시스템에 대해서 기술한다.

1. 콘텐츠 추천 시스템

전자상거래 웹 콘텐츠 사이트는 구매 사용자들의 만족을 극대화시키기 위해서 사용자 별로 개인화 서비스를 제공하고 있다. 다양한 정보원으로부터 생산되는 수많은 정보를 하나씩 처리하는 것은 많은 시간과 노력을 필요로 한다. 최근에 정보처리와 IT융합기술의 발달로 인해 이러한 현상은 더욱 심화되고 있다. 이러한 문제를 해결하기 위해 학계에서 구체적인 연구가 진행되어 왔으며, 정보 필터링을 중심으로 개인화 콘텐츠 추천은 IT융합 분야의 핫이슈가 되고 있다. 추천이 적용된 분야로는 유즈넷 뉴스, 웹 페이지, 비디오, 음악, 패션, 헬스케어, 식단, 소방방재, 기상청, 도서 분야 등이며, 다양한 분야에서 추천 방법론이 적용되고 있다[4][7].

콘텐츠 추천 시스템에서 사용되는 방법론은 다양한 기준에 의해 분류할 수 있다. 콘텐츠 추천 시스템은 지금까지 광범위하게 연구되어 왔으며, 개인화 추천을 하는데 기초가 되는 선호도 데이터의 출처와 그 데이터가 입력되어 어떻게 이용되는지를 기준에 따라 내용 기반 필터링, 사례 기반 필터링, 규칙 기반 필터링, 협력적 필터링의 범주로 분류된다. 가장 대표적인 방법론인 협력적 필터링은 사용자와 유사한 선호도를 가진 사용자 군집의 선호도에 따라 새로운 사용자가 관심을 가질 것으로 생각되는 아이템을 추천해 주는 기법이다[8]. 선호도를 반영하기 위해서 각 사용자는 정보에 따라 선호도를 제공하여, 이로부터 상관관계를 계산하여 유사 선호도 사용자 군집을 형성하고 이들의 선호도를 종합하여 정보를 추천한다[6]. 협력적 필터링은 정보의 내용을 직접 분석할 필요없이 사용자들의 관계만을 이용하며, 정보 추천의 범위를 넘어 뜻하지 않은 것들을 추천될 수 있다. 또한 정보의 내용뿐만 아니라 우수성에 따라 아이템을 추

천할 수 있다[5]. Ringo 및 Grouplens[4]는 피어슨 상관 계수를 통해 사용자들간의 정보에 대한 선호도를 비교하고 상호관계를 계산하여 유사한 사용자를 얻는다. 유사한 사용자들의 선호도에 기반하여 정보를 추천한다. 또한 선호도를 SVD로 전환하여 다중퍼셉트론의 입력으로 전환하여 정보에 대한 선호 여부를 결정하는 분류 시스템이 있다[6]. 내용 기반 필터링은 아이템의 특징을 기술하는 속성 정보와 사용자의 선호도를 가지고 있는 프로파일을 비교하여 사용자에게 필요한 정보를 추천하는 방법이다.

III. 건강관리를 위한 혼합 필터링을 이용한 식이영양 콘텐츠 추천

1. 개인화 식이영양 콘텐츠 추천 시스템 구성도

건강관리를 위한 혼합 필터링을 이용한 개인화 식이영양 콘텐츠 추천 시스템의 구조는 [그림 1]과 같으며 세부 단계는 [그림 2]와 같다. 사용자는 웹 브라우저를 통해서 웹서버와 연결되며, 웹서버는 사용자의 요구에 적절한 추천을 제공하기 위하여 시스템과 통신한다. 사용자가 추천을 요구하면 시스템은 데이터베이스에 저장되어 있는 사용자의 평가 데이터베이스, 상관관계 데이터베이스, 내용 정보 데이터베이스, 이미지 데이터베이스를 참고하여 대상 사용자에게 유용한 추천을 계산하게 된다. 이렇게 계산된 결과는 웹서버에게 전달되고 웹서버는 HTML 형태로 사용자에게 추천 결과를 보여준다.

그림 1. 개인화 식이영양 콘텐츠 추천 시스템의 구조

그림 2. 개인화 식이영양 콘텐츠 추천의 세부 단계

본 논문에서 제안한 건강관리를 위한 혼합 필터링을 이용한 개인화 식이영양 콘텐츠 추천 과정을 구체적으로 살펴보면 [그림 2]와 같으며 세부 단계의 작업 과정은 다음과 같다. 개인화 식이영양 콘텐츠 추천 시스템은 크게 웹 로봇 에이전트, 관리 모듈, 추천 엔진 모듈, 사용자 인터페이스 모듈의 네가지 모듈로 나누어진다.

2. 식이영양 콘텐츠 추천 시스템의 모듈 구성

개인화 식이영양 콘텐츠 추천 시스템은 웹 로봇 에이전트[9]에 의한 식이영양 콘텐츠를 수집한다. 역문서 빈도를 이용하여 가중치를 부여하고 식이영양 콘텐츠의 키워드를 추출한다. 식이영양 콘텐츠에서 새로운 키워드가 추출되면 내용 정보 데이터베이스의 역화일을 갱신하지 않고 새로운 역화일로 저장하며 또한 콘텐츠 식별자도 저장한다. 식이영양 콘텐츠의 특성상 소멸과 생성이 계속적으로 반복되기 때문에 수집 작업과 저장 작업을 주기적으로 수행해야 한다. 또한 수집된 콘텐츠의 키워드는 빈도에 의해 계산되어지기 때문에 기존의 역화일에서 적절하게 추출된 키워드는 새로 수집한 역화일에서 키워드가 아닐 수 있다. 그러나 사용자는 개인화 식이영양 콘텐츠 추천 시스템이 언제 새롭게 키워드를 수집하였는지 알 수 없다. 그러므로 이미지 필터링에서 첫번째 입력한 질의를 다음번 검색할 때에도 그대로 이용하여, 원하는

정보를 얻지 못하였을 경우에는 원하는 정보를 얻을 때까지 계속하여 개인화 식이영양 콘텐츠 추천 시스템에 질의를 보내게 된다. 이렇게 사용자가 질의를 반복하여 입력하는 문제를 해결하기 위해서 본 논문에서는 개인화 식이영양 콘텐츠 추천 결과에 대해 적합성 피드백을 이용하여 사용자에게 최적화된 추천을 한다. 적합성 피드백은 평가 데이터를 수정하게 하여 추천의 횟수가 증가함에 따라 추천 엔진은 정확한 추천이 가능해진다. 개인화 식이영양 콘텐츠 추천의 관리 모듈은 추천 시스템의 스케줄링, 위크플러우, 사용자의 정보 입력, 평가한 선호도를 내용 정보 데이터베이스에 저장하거나 추천 엔진에 전달하는 관리 작업을 담당한다.

개인화 식이영양 콘텐츠 추천의 핵심 모듈인 추천 엔진 모듈은 이미지 필터링, 협력적 필터링, 혼합 필터링에서 개인화 식이영양 콘텐츠 추천의 세부부분으로 구성된다.

이미지 필터링에서 추천은 사용자의 선호도와 식이영양 콘텐츠의 이미지 색 정보 사이의 정적 상관관계를 이용하여 추천한다. 이미지 필터링 방법론은 기존연구인 [7]에 자세한 내용을 기술하였다. 협력적 필터링에서 추천은 {사용자-콘텐츠} 행렬에서 사용자간의 연관 관계를 유지하면서 차원수를 감소시키기 위해서 사용자를 군집한다. 협력적 필터링에서 희박성 문제를 해결하기 위해서 {사용자-콘텐츠} 행렬에서 결측치 값을 내용 기반 필터링의 예측 값을 이용하여 {사용자-콘텐츠} 병합 행렬을 구성한다. {사용자-콘텐츠} 병합 행렬에서 분류된 식이영양 콘텐츠에 Naïve Bayes 분류자를 사용하여 클래스별로 추정치를 다르게 부여한다. 이는 식이영양 콘텐츠에 사용자의 선호도에 가중치를 달리 부여하여 선호도 정도만을 이용하는 것이 아닌 통계적인 값에 의해 추정치를 부여한다[8].

혼합 필터링에서 개인화 식이영양 콘텐츠 추천은 이미지 필터링과 협력적 필터링을 조합하여 사용자의 선호도를 예측하는 것이다. 사용자에게 제공하는 추천 리스트 중에서 비선호하는 식이영양 콘텐츠가 있을 경우 정보검색 분야의 적합성 피드백을 사용함으로써 추천 리스트를 수정한다. 제안한 개인화 식이영양 콘텐츠 추천 시스템은 혼합 필터링에 의한 추천을 사용자에게 제공한다. 혼합 필터링 방법에 의한 추천은 이미지 필터링에 의한 추천

과 협력적 필터링을 조합한 혼합 필터링에 의한 추천 시스템이다.

추천대상: 이미지 필터링에 의한 콘텐츠 추천

순위	제목	순위	제목
1	김치찌개	6	시금치와 당근을 곁들인 닭볶음탕
2	미나리국	7	조청이득 꿀꿀이
3	나스알라데	8	생선살 완자 찜밥
4	알록달록 시금치 달걀 수제비	9	참깨 오렌지 쇠고기
5	다시마전	10	산호두부

2. 혼합 필터링에 의한 개인화 식이영양 콘텐츠 추천 리스트

순위	제목	순위	제목
1	김치찌개	6	돼지고기 찜
2	미나리국	7	조청이득 꿀꿀이
3	나스알라데	8	호두 시금치 볶음
4	알록달록 시금치 달걀 수제비	9	송편구이 10개
5	두부 샌드위치	10	누룽지 찜

그림 4. 이미지 필터링에 의해 추천된 리스트

이미지 필터링에 의한 추천은 사용자의 선호도 차이와 색 히스토그램 거리의 정적 상관관계를 이용한 필터링이다. 이미지 필터링은 이미지 데이터간의 비교를 하기 위해서 색 히스토그램 간의 유클리드 거리를 계산한 후, 거리의 범위에 따른 군집을 한다. 이미지 군집 속에서 평균 선호도의 가중치가 부여된 합으로 추천을 위한 예측을 하여 이미지 추천 리스트를 구성한다. [그림 4]는 사용자가 입력한 '김치찌개'이라는 질의어에 대해서 이미지 필터링으로 추천된 식이영양 콘텐츠 리스트이다.

협력적 필터링에 의한 추천은 사용자가 식이영양 콘텐츠에 대해 평가한 선호도를 대상으로 {사용자-콘텐츠} 행렬을 구성한다. [그림 5]는 협력적 필터링에 의한 추천의 초기 화면에서 사용자가 식이영양 콘텐츠에 대하여 선호도를 평가하는 예를 보인다. [그림 6]의 협력적 필터링을 위한 선호도 평가에서는 6단계로 나누어 식이영양 콘텐츠에 대한 평가를 하도록 한다. 사용자들이 식이영양 콘텐츠에 평가를 할 때 분류된 클래스에 따른 메뉴타입별, 메뉴분류별, 연령별, 열량별로 나누어 평가를 함으로써 전체 데이터베이스를 검색하거나 무작위 식이영양 콘텐츠를 선택하는 것을 피할 수 있다. 이는 사용자가 선택한 식이영양 콘텐츠에 대해서 효율적인 평가가 이루어지도록 한다. [그림 6]은 협력적 필터링만으로 사용자에게 식이영양 콘텐츠가 추천되는 예를 보인다.

그림 5. 협력적 필터링을 위한 선호도 평가

그림 6. 협력적 필터링에 의해 추천된 식이영양 콘텐츠

3. 혼합 필터링을 이용한 추천 기법

이미지 필터링과 협력적 필터링을 조합한 혼합 필터링에서 사용자의 선호도를 예측하는 방법에 대해서 기술한다. 혼합 필터링에서 콘텐츠 추천을 위하여 연관 사용자 유사도 가중치에 중요도 가중치를 적용하여 개인화 식이영양 콘텐츠를 추천한다[1]. 중요도 가중치는 새로운 사용자가 개인화 콘텐츠 추천 시스템에서 선호도를 평가한 콘텐츠의 개수가 매우 적을 경우에 유사도 가중치가 매우 높게 나오는 경우가 발생하므로 예측의 정확도가 떨어진다.

따라서 새로운 사용자가 평가한 콘텐츠 개수의 제한을 정하고 제한된 개수에 미치지 못하는 사용자 유사도 가중치에 중요도 가중치를 부여한다. 본 논문에서는 중요도 가중치를 $sw_{a,u}$ 로 표현하고 임계치를 60으로 설정한다. 혼합 필터링에서 개인화 식이영양 콘텐츠에 대한 선호도의 예측은 (식 4)와 같이 정의한다[8].

$$P_{a,k} = \bar{v}_a + \frac{(IM_{a,k} - \bar{v}_a) + \sum_{i=1}^m sw_{a,u} w(a,i) (v_{u,j} - \bar{v}_u)}{\sum_{i=1}^m sw_{a,u} w(a,i)} \quad (4)$$

(식 4)에서 $P_{a,k}$ 는 선호도가 수정 및 보완된 {사용자-콘텐츠} 행렬에서 새로운 사용자가 개인화 식이영양 콘텐츠 k 에 대한 예측한 선호도이고, $\bar{v}_a$ 는 새로운 사용자의 선호도 평균값이다. $IM_{a,k}$ 은 사용자 a 가 콘텐츠 k 에 대해서 이미지 필터링에서 예측한 선호도이다. $sw_{a,k}$ 는 $n/60$ 의 중요도 가중치를 부여한다. 공통으로 선호도를 평가한 콘텐츠 개수 제한이 60이고 60보다 작은 공통으로 평가한 식이영양 콘텐츠 개수 n 을 가지는 사용자간의 유사도 가중치에는 $n/60$ 의 중요도 가중치를 곱한다. n 이 60보다 클 경우에는 중요도 가중치를 1로 준다. $w(a,i)$ 는 사용자가 평가한 식이영양 콘텐츠의 선호도를 대상으로 사용자 유사도 가중치이다. m 은 새로운 사용자와 다른 사용자간의 연관 사용자 유사도 가중치가 0이 아닌 사용자 수이다[8].

혼합 필터링에 의한 개인화 식이영양 콘텐츠 추천 시스템은 [그림 7]과 같이 이미지 필터링과 협력적 필터링을 조합하여 사용한다. 여기서 왼쪽 프레임은 이미지 필터링을 위한 내용이며 오른쪽 프레임은 협력적 필터링을 위한 내용이다. 협력적 필터링을 위한 데이터베이스가 구축되지 않았을 경우 이미지 필터링만을 사용하여 추천이 이루어지며, 협력적 필터링을 위한 데이터베이스가 구축된 경우에 이미지 필터링과 협력적 필터링을 혼합하여 추천이 이루어진다.

그림 7. 혼합 필터링에 의한 추천된 개인화 식이영양 콘텐츠

[그림 7]은 이미지 필터링과 협력적 필터링을 조합한 혼합 필터링에서 슈퍼맨이라는 질의어에 대해 추천된 식이영양 콘텐츠를 나타낸다. [그림 7]에서 추천된 콘텐츠는 [그림 4]의 이미지 필터링에 의한 추천과는 달리 관련 없는 콘텐츠가 추천됨을 보인다. 이는 협력적 필터링에 의해 비슷한 흥미를 갖는 연관 사용자들이 선호하는 콘텐츠를임을 의미한다.

V. 성능평가

1. 실험환경

본 논문에서 사용하는 실험 대상은 식이영양 분야의 콘텐츠이다. 식이영양 분야는 건강메뉴에서 한식형, 중식형, 일식형, 양식형, 동남아시아식형을 기반으로 주식형, 부식형, 후식형으로 분류한다. 실험 대상으로 사용하기 위한 500개의 서로 다른 식이영양 콘텐츠를 추출하기 위해서 웹 로봇 에이전트를 이용하였다. 수집한 결과, 각 집합에는 내용이 없거나 한국어의 기본 문법 체계를 갖추지 않는 문장으로 구성된 콘텐츠가 존재하기도 하고, 식이영양과 관련없는 콘텐츠가 수집되기도 하였다. 식이영양과 관련되지 않은 콘텐츠로부터 추출한 단어는 식이영양 콘텐츠의 특징을 정확하게 표현한 단어이나 본 논문에서는 추출한 식이영양에 관련된 내용 정보 데이

터베이스에 포함되지 않는다. 따라서 이러한 단어는 오류로 처리되는 문제가 발생한다. 이러한 이유는 웹 로봇 에이전트가 식이영양에 관련된 콘텐츠의 링크를 기반으로 무작위로 콘텐츠를 수집하기 때문이다[5][8]. 수집한 500개의 식이영양 콘텐츠 중에서 250개의 콘텐츠는 학습을 위해 사용하며, 250개의 콘텐츠는 실험 데이터로서 사용하였다. 250개의 훈련 콘텐츠는 내용 기반 데이터베이스에 속한 사용자가 분류한 콘텐츠를 기반으로 클래스로 분류한다. 제안한 혼합 필터링을 이용한 개인화 식이영양 콘텐츠 추천의 타당성을 검증하기 위해 개인화 식이영양 콘텐츠 추천 방법을 개발하고, 개발된 시스템에서 식이영양 콘텐츠 추천 과정을 통해 성능을 평가한다.

실험 환경으로는 Sun Blade 2000, 2.8GHz, 4G RAM의 컴퓨터 사양에서 Windows Server XP 환경에 MS Visual Studio C++ 2008과 Jakarta-Tomcat-5.0.24 환경에서 JSP를 사용하여 제안한 알고리즘을 구현하였고 시뮬레이션을 하였다. 내용 정보 데이터베이스 구축을 위한 DBMS는 MS SQL Server 2008을 활용하였다. 인터넷 상에서 설문을 받는 설문 조사와 클라이언트 프로그램은 JSP로 구현하였고, 평가한 선호도는 MS SQL Server 2008의 데이터베이스에 저장된다. 저장된 데이터베이스를 기반으로 제안한 방법론에 의해 추천 결과를 제공한다.

이미지 필터링과 협력적 필터링을 조합한 혼합 필터링에 의한 개인화 식이영양 콘텐츠 추천 시스템에서 성능을 평가하기 위해서 데이터베이스는 369명의 사용자와 500개의 서로 다른 식이영양 콘텐츠로 구성하였다. 369명의 사용자 중에서 269명의 사용자는 250개의 식이영양 콘텐츠에 대해 선호도를 평가한 사용자이며, 500개의 콘텐츠는 프로파일 학습에 사용된 콘텐츠와 동일하다. 269명의 사용자는 식이영양 콘텐츠에 대해 총 64,412번 평가를 하였다. 식이영양 데이터에서부터 적어도 100번 이상의 선호도를 평가하였고, 나이 및 성별을 제공한 321명을 추출하여 모든 식이영양 콘텐츠를 평가한 사용자 중에서 최종적으로 150명의 사용자를 선별하였다.

웹기반 설문조사는 클래스별로 식이영양 콘텐츠를 분류하고 각각의 클래스에는 50개의 식이영양 콘텐츠로

구성된다. [그림 8]은 식이영양 콘텐츠에 대한 사용자의 6단계의 선호도를 표시하기 위한 설문조사 화면을 나타낸다. ‘주식형’이라는 클래스에서 설문하기 위한 식이영양 콘텐츠를 나타낸다.

그림 8. 선호도를 평가하기 위한 설문조사 화면

2. 제안한 방법론의 성능 평가

본 논문에서는 성능 평가를 위해 제안한 이미지 필터링과 협력적 필터링을 혼합한 추천 방법, 피어슨 상관 계수를 이용한 협력적 필터링을 이용한 추천 방법, 내용 기반 필터링만을 이용한 추천 방법을 군집된 사용자의 수를 변화시키면서 MAE 성능을 비교하였다. MAE는 예측의 정확도를 측정하기 위해서 실제로 사용자가 평가한 값과 예측된 값의 차이에 대한 절대값의 평균을 나타낸다. MAE는 절대적으로 알고리즘이 얼마나 정확하게 예측을 했는지를 알 수 있다[10].

제안한 혼합필터링을 이용한 추천 방법은 H_Col_IM, 기존의 협력적 필터링을 이용한 추천 방법은 P_Corr, 이미지 필터링을 이용한 추천 방법은 IM이라고 표기하였다. [표 1]은 사용자 수의 변화에 따른 H_Col_IM, P_Corr, IM의 MAE 성능평가를 나타낸다.

사용자들의 수가 많아짐에 따라 H_Col_IM, IM의 MAE는 낮아짐을 보인다. 반면, P_Corr를 이용한 방법의 성능은 사용자 수가 많아짐에 따라 다소 낮아지기는 하나 큰 차이를 없음을 나타낸다. IM은 P_Corr를 이용한 방법보다 MAE의 정확도가 높음을 나타낸다. 결론적

으로, 사용자 수를 변화시키에 따른 MAE의 결과를 보면, 본 논문에서 제안한 H_Col_IM 방법이 예측의 정확도가 높음을 보였으며, 특히 사용자의 수가 많아짐에 따른 예측 정확도는 다른 방법보다 더욱 우수함을 보인다.

표 1. 사용자 수의 변화에 따른 MAE 성능평가

사용자 수	MAE		
	H_Col_IM	P_Corr	IM
20	0.191	0.235	0.226
40	0.182	0.227	0.214
60	0.184	0.207	0.204
80	0.183	0.207	0.200
100	0.182	0.204	0.195
120	0.177	0.203	0.193
180	0.174	0.202	0.191
240	0.171	0.201	0.185
300	0.164	0.201	0.184
330	0.153	0.200	0.183
360	0.152	0.200	0.181
평균	0.173	0.207	0.196

VI. 결론

기존의 추천 시스템에서 가장 많이 사용하는 협력적 필터링만으로 사용자의 관심이나 선호도에 맞는 식이영양 콘텐츠에 대한 예측은 잘 수행할 수 있으나, 특정 식이영양 콘텐츠의 내용에 대한 사용자의 정보 요구를 해결하는 것은 효율적이지 못하다. 본 논문에서는 보다 정확한 예측을 하기 위해서 이미지 필터링과 협력적 필터링을 조합한 혼합 필터링을 이용한 개인화 식이영양 콘텐츠 추천 시스템을 제안하고 구현하였다. 이는 이미지 필터링이 협력적 필터링을 이용한 식이영양 콘텐츠 예측에 보정하는 역할을 하기 때문에 정확한 추천이 가능하다. 혼합 필터링을 이용한 개인화 식이영양 콘텐츠 추천 시스템의 성능을 평가하기 위해 MAE를 이용하여 사용자의 수를 변화시키면서 비교하였다. 비교 결과, 제안된 방법은 사용자의 수가 많아짐에 따라 추천에 있어서 정확도가 높아짐을 보였다. 본 연구의 결과를 일반화시키기 위해서는 실험에 사용된 데이터의 규모를 확대시킨 광범위한 연구가 필요하며, 이러한 연구 결과를 이용하여 실제 전자상거래 시스템을 구축함으로써 사용자에게 보다 다양한 서비스를 제공할 수 있을 것이다.

향후 본 연구의 결과를 토대로 다양한 분야에 추천을 적용하여 제안한 혼합 필터링을 이용한 방법론에 대한 유효성 검증이 필요하다. 선호도를 분석하여 추천을 하기 위한 사용하였던 정보검색 분야의 방법론에 대해서 구체적인 연구가 필요하다. 웹 솔루션 업체와 구체적인 시스템 출시를 통하여 시장성 증대와 고부가가치를 창출할 수 있을 것으로 기대함으로써 다양한 응용분야에 활용이 가능하다.

참 고 문 헌

- [1] M. Jalali, N. Mustapha, Md. N. Sulaiman, A. Mamat, "WebPUM: A Web-based Recommendation System to Predict User Future Movements," J. of Expert Systems with Applications, Vol.37, Issue9, pp.6201-6212, 2010.
- [2] S. U. Jang, J. Y. Cho, K. S. Jeong, and G. S. Cho, "Exploring Possibilities of ECG Electrodes for Bio-monitoring Smartwear with Cu Sputtered Fabrics," J. of Human-Computer Interaction, Vol.4551, pp.1130-1137, 2007.
- [3] H. N. Kim, A. T. Jia, I. A. Haa, and G. S. Joa, "Collaborative Filtering based on Collaborative Tagging for Enhancing the Quality of Recommendation," J. of Electronic Commerce Research and Applications, Vol.9, Issue1, pp.73-83, 2010.
- [4] P. Melville, R. J. Mooney, and R. Nagarajan, "Content-Boosted Collaborative Filtering for Improved Recommendations," Proc. of National Conference on Artificial Intelligence, pp.187-192, 2002.
- [5] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating Collaborative Filtering Recommender Systems," J. of ACM Trans. on Information Systems, Vol.22, No.1, pp.5-53, 2004.

- [6] 정경용, 조선문, "내용 기반 필터링을 위한 프로파일 학습에 의한 선호도 발견", 한국콘텐츠학회논문지, 제8권, 제2호, pp.1-8, 2008.
- [7] 정경용, "이미지 기반 필터링을 이용한 개인화 아이템 추천", 한국콘텐츠학회논문지, 제8권, 제3호, pp.1-7, 2008.
- [8] 정경용, "혼합 필터링과 연관 이웃 마이닝을 이용한 개인화 아이템 추천 기법", 인하대학교 컴퓨터정보공학과 박사학위논문, 2005.
- [9] K. Y. Chung, "Sensibility Ergonomics Fashion Recommendation System using Weather WebBot," Proc. of the International Conference on Information Science and Applications 2011, pp.712-717, IEEE Computer Society, 2011.
- [10] 박동균, 김종훈, 김재권, 정은영, 이용호, "멀티플랫폼 환경의 만성 질환자 건강관리를 위한 유헬스 서비스 모델", 한국콘텐츠학회논문지, 제11권, 제8호, pp.23-32, 2011.

저 자 소 개

정 경 용(Kyung-Yong Chung)

정회원

- 2000년 2월 : 인하대학교 전자계산공학과(공학사)
- 2002년 2월 : 인하대학교 컴퓨터정보공학과(공학석사)
- 2005년 8월 : 인하대학교 컴퓨터정보공학과(공학박사)

- 2005년 8월 : 한국소프트웨어진흥원 KSI 책임
- 2005년 9월 ~ 2006년 2월 : 한세대학교 IT학부 교수
- 2006년 3월 ~ 현재 : 상지대학교 컴퓨터정보공학부 교수

<관심분야> : 지능시스템, 데이터마이닝, 상황인식, 웨어러블 컴퓨팅, HCI, 바이오센서, IT융합

이 영 호(Young-Ho Lee)

정회원

- 1996년 2월 : 한국외국어대학교
응용전산학과(이학석사)
- 2005년 8월 : 아주대학교 의과대
학 의료정보학과(이학박사)
- 1999년 ~ 2002년 : IBM Korea
BI & CRM EM

- 2002년 ~ 현재 : 가천의과학대학교 의료공학부 교수
 - 2007년 ~ 현재 : ISO/TC215 전문위원
 - 2005년 ~ 현재 : 가천의과학대학교 u-헬스케어연구
소 책임연구원
 - 2008년 ~ 현재 : 수송물류분야 단체표준 전문위원
 - 2011년 ~ 현재 : 스마트의료정보표준포럼사무총장
- <관심분야> : 메디컬인포매틱스, 유헬스케어