



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0057209
(43) 공개일자 2020년05월26일

(51) 국제특허분류(Int. Cl.) G06Q 30/02 (2012.01) G06F 16/00 (2019.01) G06Q 30/06 (2012.01)	(71) 출원인 이윤열 서울 영등포구 여의나루로 7 10동 409호 (여의도, 광장아파트)
(52) CPC특허분류 G06Q 30/0271 (2013.01) G06F 16/337 (2019.01)	(72) 발명자 이윤열 서울 영등포구 여의나루로 7 10동 409호 (여의도, 광장아파트)
(21) 출원번호 10-2018-0141268	(74) 대리인 김애라
(22) 출원일자 2018년11월16일 심사청구일자 없음	

전체 청구항 수 : 총 3 항

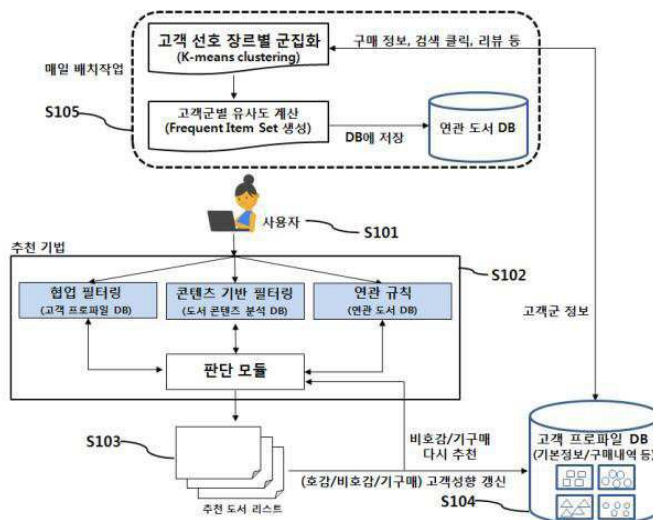
(54) 발명의 명칭 K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법 및 그 장치

(57) 요약

본 발명은 K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법에 관한 것으로, 도서 취향이 비슷한 사람들을 모으고 이들의 선호 장르를 기반으로 정교한 필터링 모듈을 통하여 도서를 추천하는 방법 및 그 시스템에 관한 것이다.

본 발명에 따르면 사용자가 자신의 온라인 구매 또는 반품 이력을 관리할 수 있으므로 자신의 취향에 맞는지 미리 확인해 볼 수 있어 사용자의 구매만족도를 높이고 도서 상품 반품율을 현저하게 낮출 수 있다.

대표도 - 도3



(52) CPC특허분류

G06F 16/35 (2019.01)

G06Q 30/0255 (2013.01)

G06Q 30/0631 (2013.01)

명세서

청구범위

청구항 1

네트워크를 통해서 사용자에게 맞춤형 도서를 추천해주는 시스템에 있어서, 사용자 클라이언트와 연결되어 맞춤형 도서 추천 서비스를 운영하는 서버를 포함하며 상기 서버는 연관 도서 DB와 도서 콘텐츠 분석 DB를 비교하여 해당 고객에게 최적화된 도서를 추천하는 판단 모듈 및 고객 구매내역 및 고객 성향 정보를 저장하는 고객 프로파일 DB; 및 상기 고객 프로파일 DB를 기반으로 협업 필터링을 통한 추천 목록을 포함하여 이루어지는 필터링 모듈;을 포함하는 것을 특징으로 하는 맞춤형 도서 추천 시스템

청구항 2

K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법에 있어서, 네트워크를 통해서 추천 시스템을 PC나 스마트폰 상에서 실행시키는 단계;

상기 서버내에 필터링 모듈에서는 사용자의 선호 장르에 따라 군집화된 그룹을 기본 단위로 하여 협업 필터링으로 생성된 추천 목록과 도서 콘텐츠 분석 DB, 연관 도서 DB를 혼합한 추천 기법으로 사용자에게 최적화된 맞춤형 도서를 판단하는 단계;

판단 모듈이 검색한 결과를 상기 사용자에게 추천 리스트로 디스플레이 하는 단계;

및 매일 변화하는 고객의 관심도를 반영하고, 실시간 추천을 위해 매일 고객을 K-means clustering을 통해 재그룹화하고, 매일 고객군 별로 나눈 후, 생성된 연관 규칙에 대한 결과 정보(Frequent Item Set)를 실시간 추천을 하기 위해서는 연관 도서 DB에 저장하는 단계를 포함하는 것을 특징으로 하는 K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법

청구항 3

제 2 항에 있어서, 상기 도서를 추천하는 방법은 스마트폰 뿐만 아니라 영상입력 수단과 통신기능이 구비된 개인 휴대용 단말기 상에서 구현될 수 있으며, 나아가 도서 뿐만 아니라 영화, 쇼핑물 상품 등 다양한 부가상품에도 적용될 수 있는 것을 특징으로 하는 K-means Clustering을 활용한 맞춤형 도서를 추천 하는 방법

발명의 설명

기술 분야

[0001] 본 발명은 K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법 및 그 시스템에 관한 것으로, 특히 도서 취향이 비슷한 사람들을 모으고 이들의 선호 장르를 기반으로 정교한 필터링 모듈을 통하여 도서를 추천하는 방법 및 그 시스템에 관한 것이다.

배경 기술

[0002] 정보기술의 발전에 따라 이용자가 접할 수 있는 정보의 양은 기하급수적으로 늘어났다. 다양한 정보를 전달하는 주요 매체인 많은 수의 책, 음악, 영화, 뉴스, 광고, 제품 등과 그에 따른 정보의 범람은 이미 개인이 처리할 수 있는 한계를 넘어섰다. 이러한 정보과잉 속에서 자신이 원하는 정보를 단시간 내에 검색하는 것은 그리 쉬운 일이 아니다.

[0003] 최근에는 필터링(여러가지 항목 중 적당한 항목을 선택하는 기술) 기법을 응용한 추천 시스템이 등장하였으며, 이 시스템을 이용하여 이용자가 원하는 정보를 예측하고 추천함으로써 정보과잉의 문제를 경감시키려는 많은 노력이 진행되고 있다. 온라인 쇼핑 시 구매자가 관심 있을 만한 물건을 보여주거나, 뉴스 기사를 읽을 때 관련 기사의 목록을 제공해 주거나, 사용자가 알 만한 온라인 친구들을 알려주는 등 광범위한 곳에서 추천 시스템이 사용되고 있다.

[0004] 그러나 현재 기존 오프라인 서점이나 온라인 서점에서는 신간, 베스트셀러 위주로 고객들에게 도서를 추천하여,

개인별 취향을 고려하지 않고 사용자가 직접 일일이 책 내용을 보고 구매하는 번거로움이 있다.

- [0005] 또한 현재 추천 시스템에서 성공적으로 사용되는 협업 필터링은 초기 평가 문제, 희박성 문제, 확장성 문제 등 근본적으로 해결해야 할 문제점을 가지고 있다.
- [0006] 초기 평가 문제는 기존에 없던 새로운 항목이 추가되는 경우는 추천이 곤란해지는 문제이다. 초기 평가 문제의 극단적인 예로서, 협업 필터링 시스템이 처음 시작될 때 모든 사용자는 모든 항목에 대해서 초기 평가 문제로부터 영향을 받는다.
- [0007] 희박성 문제는 협업 필터링을 적용하기 위해서는 고객 및 구매 데이터가 충분히 축적되어야 한다. 데이터가 충분하지 않다면 항목을 추천을 하는데 있어서 정확도가 떨어지게 되어 희박성의 문제가 제기된다. 확장성 문제는 협업 필터링 분야에서 사용되는 알고리즘은 사용자와 항목 수에 비례해서 계산 시간이 비례한다. 사용자와 항목 수가 수백만이나 되는 환경 하에서는 이러한 계산 시간이 치명적일 수 있다.
- [0008] 대한민국 특허출원 제10-2009-0054825호(2009.06.19.출원)는 시청자와 유사한 시청 취향을 갖는 시청자 그룹의 시청 프로그램으로부터 협업필터링을 이용한 프로그램을 추출하고 취향 선호도가 높은 순서로 정렬하여 취향 선호도가 높은 몇 개의 프로그램을 자동으로 추천하는 IP TV 프로그램 자동 추천 방법을 개시하고 있다. 그러나 도 1에 도시된 바와 같이 종래의 방법은 추천 대상 시청자와 시청 취향이 비슷한 유사 시청자 그룹의 프로그램 시청 히스토리 데이터와 추천 대상 시청자의 프로그램 시청 히스토리 데이터를 협업 필터링하여 후보 추천 방송 프로그램 콘텐츠를 선별한 후 순위 정렬 모델을 이용하여 우선순위와 추천 순위가 높은 순서대로 IP TV 단말에 방송 프로그램 콘텐츠 추천한다는 점에서, 여전히 문제를 안고 있다.

발명의 내용

해결하려는 과제

- [0009] 본 발명에서는 현재 추천 시스템에서 주로 사용되는 협업 필터링의 초기 평가 문제, 희박성 문제, 확장성 문제 등 근본적인 문제점에 대한 해결을 그 목적으로 한다.

과제의 해결 수단

- [0010] 따라서, 본 발명은 상기와 같은 종래의 협업 필터링만을 사용하여 콘텐츠를 추천하는 방법이 가지는 문제점을 해결하기 위하여 안출된 것으로서, 본 발명에 따르면 새로운 사용자에 대해서도 선호 장르를 지니고 있는 그룹을 형성하여 추천해 주게 되어, 사용자의 관심사에 맞춰 추천받을 수 있게 되어 초기부터 원하는 도서를 추천받을 수 있다.
- [0011] 나아가, 본 발명에 따르면 고객 프로파일 DB를 효과적으로 활용하여 유사 고객끼리 클러스터를 구축한 후, 각 클러스터 내 고객들을 대상으로 협업 필터링을 수행함으로써 추천시간이 단축되고 정확도가 개선되며, 연관 규칙과 콘텐츠 기반 필터링과의 통합된 추천 방법을 이용함으로써 두 추천 방법의 장점들이 결합됨으로써 정확도와 coverage가 향상된다(데이터 희박성 문제 해결).
- [0012] 본 발명은 시간적 효율성 문제 해결을 위해 협업 필터링 추천시 사용자 사이의 유사도를 비교할 때나 선호도 예측 시 유사한 사용자의 선호도 리스트 검색 및 계산할 때의 시간 단축이 필요하기 때문에, 매일 배치 처리에 의한 고객 그룹화를 제안하고, 고객 그룹군 내에서 협업 필터링을 수행하였다. 또한, 연관 규칙을 이용한 추천에서는 상품들의 조합을 생성하는 데에 있어서의 시간 단축이 필요하므로, 매일 배치 처리에 의한 고객군별 연관 도서 DB 구축으로 실시간 추천이 가능하여, 대형 도서 쇼핑몰에도 적용할 수 있다.

발명의 효과

- [0013] 상기에서 살펴본 바와 같이, 본 발명에 따른 K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법 및 그 시스템은 기존 오프라인 서점이나 온라인 서점에서는 신간, 베스트셀러 위주로 추천함으로써, 개인별 취향을 고려하지 않고 사용자가 직접 일일이 책 내용을 보고 구매하는 번거로운 과정이 제거되며, 사용자가 온라인 서점에서 구매한 책 구매 또는 반품 이력을 관리하고 자신의 취향에 맞는지 미리 확인해 볼 수 있도록 함으로써 사용자의 구매만족도를 높이고 도서 상품 반품율을 현저하게 낮출 수 있다.
- [0014] 또한, 본 발명에 따르면 개인 프로파일을 활용하여 고객들을 K-means Clustering 기법을 통해 유사한 선호 장르를 가지고 있는 고객들을 적절히 군집화하고, 군집화된 그룹을 기본 단위로 하여 협업 필터링을 수행하면서,

더불어 추천의 정확도를 높이기 위해 콘텐츠 기반 필터링과 연관 규칙을 혼합한 추천 기법으로 현재 추천 시스템에서 주로 사용되는 협업 필터링의 초기 평가 문제, 희박성 문제, 확장성 문제 등 근본적인 문제점을 해결된다.

도면의 간단한 설명

[0015]

도 1은 종래의 필터링을 이용한 시청자 맞춤형 IP TV 프로그램 자동 추천 방법의 실시예를 설명하기 위한 흐름도

도 2는 본 발명의 선호적인 실시예에 따른 K-means Clustering을 활용한 맞춤형 도서를 추천하는 시스템의 전체 구성도를 도시한 모습

도 3은 본 발명에 따른 K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법을 설명하기 위한 흐름도

도 4는 고객 프로파일 DB 항목의 실시예를 설명하기 위한 도면

도 5는 고객 성향 갱신의 실시예를 설명하기 위한 도면

도 6은 선호 장르별 클러스터링 유형의 실시예를 설명하기 위한 도면

도 7은 K-means clustering을 통한 고객 세분화의 실시예를 설명하기 위한 도면

도 8은 Frequent Item Set 생성을 통한 연관 도서 DB 갱신의 실시예를 설명하기 위한 도면

도 9는 군집화 기반 협업 필터링의 실시예를 설명하기 위한 도면

도 10은 고객의 선호 장르 변경의 실시예를 설명하기 위한 도면

도 11은 군집화 기반 협업 필터링 수행에 대한 실시예를 설명하기 위한 도면

도 12는 연관 규칙 생성 과정을 설명하기 위한 예제 데이터 도면

도 13은 연관 규칙 생성을 위해서 최소 지지도 이상의 지지도를 가지는 항목집합들에 대한 실시예를 설명하기 위한 도면

도 14는 최종 Frequent Item Set의 실시예를 설명하기 위한 도면

도 15는 생성된 연관 규칙의 실시예를 설명하기 위한 도면

도 16은 콘텐츠 기반 필터링의 실시예를 설명하기 위한 도면

발명을 실시하기 위한 구체적인 내용

[0016]

도 2는 본 발명의 선호적인 실시예에 따른 K-means Clustering을 활용한 맞춤형 도서를 추천하는 시스템의 전체 구성도를 도시한 모습이다.

[0017]

상기한 목적을 달성하고 상술한 종래의 문제점을 해결하기 위하여, 본 발명의 선호적인 일 실시예에 따른 네트워크를 통해서 사용자에게 맞춤형 도서를 추천해주는 시스템은, 시스템을 운영 및 제어하는 서버(200); 상기 서버(200)와 연결되는 클라이언트(300); 및 상기 서버(200)내에 고객 프로파일 DB(140)를 기반으로 연관 도서 DB(110)와 도서 콘텐츠 분석 DB(120)를 비교하여 해당 고객에게 최적화된 도서를 추천하는 판단 모듈(130) 및 고객 구매내역 및 고객 성향 정보를 저장하는 고객 프로파일 DB(140)와 상기 고객 프로파일 DB(140)를 기반으로 협업 필터링을 통한 추천 목록(150)을 포함하여 이루어지는 필터링 모듈(100);로 이루어지는 것을 특징으로 한다.

[0018]

상기한 목적을 달성하고 상술한 종래의 문제점을 해결하기 위하여 본 발명의 선호적인 일 실시예에 따른 K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법은, 네트워크를 통해서 추천 시스템을 PC나 스마트폰 상에서 실행시키는 단계와; 상기 서버내에 필터링 모듈에서는 사용자의 선호 장르에 따라 군집화된 그룹을 기본 단위로 하여 협업 필터링으로 생성된 추천 목록과 도서 콘텐츠 분석 DB, 연관 도서 DB를 혼합한 추천 기법으로 사용자에게 최적화된 맞춤형 도서를 판단 하는 단계와; 판단 모듈이 검색한 결과를 상기 사용자에게 추천 리스트로 디스플레이 하는 단계와; 매일 변화하는 고객의 관심도를 반영하고, 실시간 추천을 위해 매일 고객을 K-means clustering을 통해 재그룹화하고, 매일 고객군 별로 나눈 후, 생성된 연관 규칙에 대한 결과 정보(Frequent Item Set)를 실시간 추천을 하기 위해서는 연관 도서 DB에 저장하는 단계로 이루어지는 것을 특징

으로 한다.

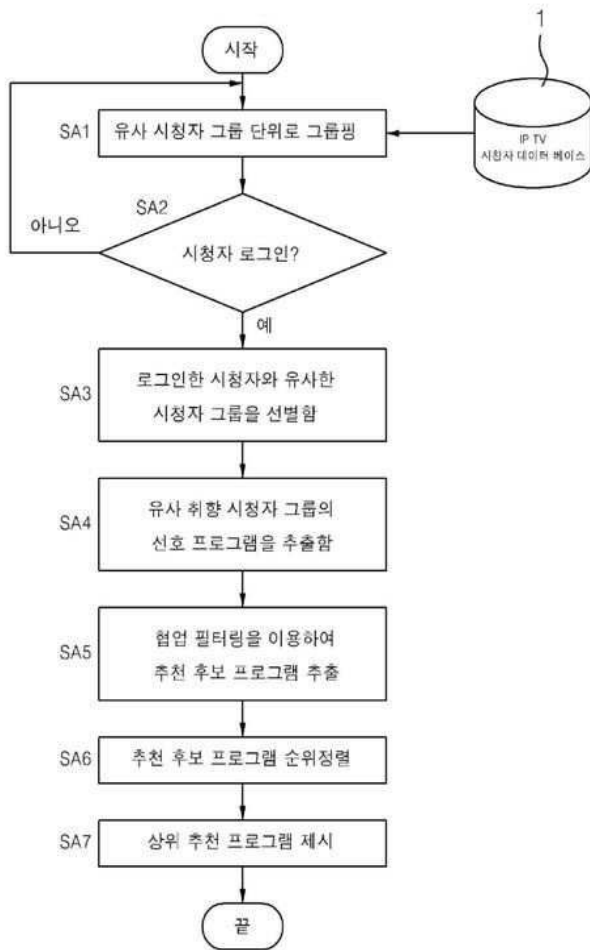
- [0019] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있으며, 특정 실시예들은 도면에 예시하고 상세한 설명에서 구체적으로 설명한다. 그러나 이는 본 발명을 특정한 실시 형태에 대해서 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 구성요소에 대해서는 유사한 도면부호를 사용하였다.
- [0020] 한편 본 출원에서 사용하는 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, “포함하다” 또는 “가지다” 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이며, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0021] 이하에서는 본 발명의 선호적인 실시예와 첨부된 도면을 참고로 하여 더욱 상세하게 설명한다.
- [0022] 도 3는 본 발명에 따른 K-means Clustering을 활용한 맞춤형 도서를 추천 하는 방법을 도시하고 있다. 상기 도면에 도시된 바와 같이 본 발명의 선호적인 실시예에 따른 K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법은 다음과 같은 단계로 이루어진다.
- [0023] 먼저, 본 발명에 따른 스마트폰이나 PC에서 도서를 추천받고자 하는 사용자는 네트워크를 통해서 본 발명에 따른 추천 시스템을 PC나 스마트폰 상에서 실행시킨다(S101). 다음으로 상기 서버내에 필터링 모듈에서는 사용자의 선호 장르에 따라 군집화된 그룹을 기본 단위로 하여 협업 필터링으로 생성된 추천 목록과 도서 콘텐츠 분석 DB, 연관 도서 DB를 혼합한 추천 기법으로 사용자에게 최적화된 맞춤형 도서를 판단한다(S102). 판단 모듈이 검색한 결과를 상기 사용자에게 추천 리스트로 디스플레이 한다(S103). 추천된 도서에 대해 사용자의 반응(읽고 싶어요/취향 아니에요/이미 읽었어요)을 고객 프로파일 DB에도 지속적으로 업데이트함으로써, 다음 추천 판단에 사용한다. 고객이 회원 가입을 하거나 도서를 구매하면 고객 데이터와 도서 구매 데이터가 고객 프로파일 DB에 저장된다(S104). 매일 변화하는 고객의 관심도를 반영하고, 실시간 추천을 위해 매일 고객을 K-means clustering을 통해 재그룹화하고, 매일 고객군 별로 나눈 후, 생성된 연관 규칙에 대한 결과 정보(Frequent Item Set)를 실시간 추천을 하기 위해서는 연관 도서 DB에 저장한다(S105).
- [0024] 본 발명은 고객 프로파일의 정보를 이용함으로써 새로운 사용자에게 대해서도 선호 장르를 지니고 있는 그룹을 형성하여 추천해 주게 되어, 사용자의 관심사에 맞춰 추천 받을 수 있게 되어 초기부터 원하는 도서를 추천 받을 수 있다.
- [0025] 도 4는 고객 프로파일 DB 항목의 실시예를 설명하기 위한 도면이다. 상기 도면에서 도시된 바와 같이, 고객 프로파일에 사용되는 정보는 고객번호, 생년, 성별, 고객 성향, 구매 내역 등을 포함한다. ISBN은 국제표준 도서번호로 도서 서지 정보를 검색하는 식별자이며, 장르는 현재 서점 및 도서관 등에서 일반적으로 사용하는 대표적인 도서 분류이다.
- [0026] 도 5는 고객 성향 갱신의 실시예를 설명하기 위한 도면이다. 상기 도면에서 도시된 바와 같이, 고객의 성향을 분석하는 첫 단계로 선호 장르를 사용한다. 선호 장르 분석을 위한 피드백 정보로는 구매내역, 고객의 리뷰(호감, 기구매), 검색 클릭이 있다. 고객의 리뷰는 고객이 추천 도서에 대해 호감(읽고 싶어요)을 클릭한 경우와 기구매(이미 읽었어요)를 클릭한 경우에 장르별 선호도 점수에 반영한다. 고객이 특정 도서에 대한 검색 클릭시에도 장르별 선호도 점수에 반영한다. 고객이 도서를 구매할 경우 고객의 구매이력을 피드백으로 하여 장르별 선호도를 계산하고, 그 결과를 취합하여 선호 장르를 분석한다. 추천된 도서에 대해 고객의 반응(읽고 싶어요)인 경우에는 프로파일 DB의 장르별 선호도에 반영하고, 고객의 반응(취향 아니에요)인 경우에는 프로파일 DB의 비호감 항목에 반영하고, 고객의 반응(이미 읽었어요)인 경우에는 프로파일 DB의 기구매 항목에 반영하고, 도서를 다시 추천해 준다.
- [0027] 도 6은 선호 장르별 클러스터링 유형의 실시예를 설명하기 위한 도면이다. 상기 도면에서 도시된 바와 같이, 선호 장르를 기반으로 K-means clustering을 수행한다. 선호 장르는 구매 내역, 고객의 리뷰, 검색 클릭을 반영하여 실시간 업데이트 한다. 본 발명에서는 K-means clustering을 활용하여 선호 장르에 따라 고객군을 5개의 유형으로 세분화 한다. 선호 장르는 각 장르별 선호 정도를 나타 내는 정보이며, 장르는 현재 서점 및 도서관에서 도서를 분류하는 기준을 참조한다.
- [0028] 도 7은 K-means clustering을 통한 고객 세분화의 실시예를 설명하기 위한 도면이며, 상기 도면에서 도시된 바

와 같이, 매일 변화하는 고객의 관심도를 반영하고, 실시간 추천을 위해 매일 고객을 K-means clustering을 통해 재그룹화한다.

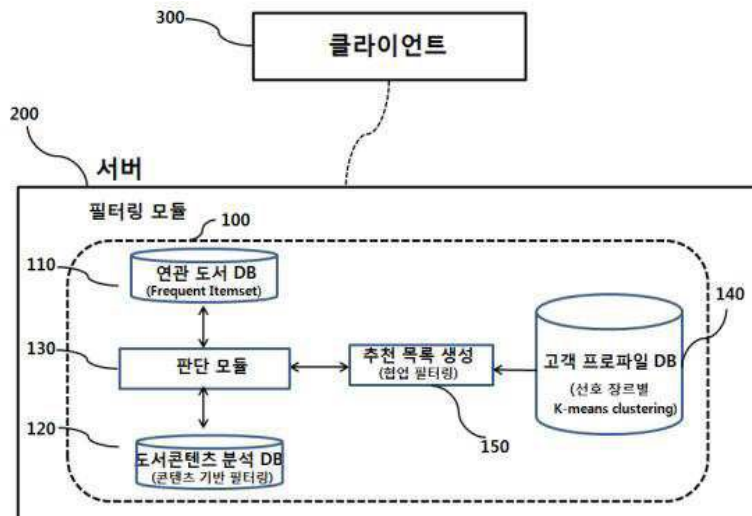
- [0029] 도 8은 Frequent Item Set 생성을 통한 연관 도서 DB 갱신의 실시예를 설명하기 위한 도면이며, 상기 도면에서 도시된 바와 같이, 매일 고객군 별로 나눈 후, 생성된 연관 규칙에 대한 결과 정보(Frequent Item Set)를 실시간 추천을 하기 위해서는 연관 도서 DB에 저장한다.
- [0030] 도 9는 군집화 기반 협업 필터링의 실시예를 설명하기 위한 도면이며, 상기 도면에서 도시된 바와 같이, 유사한 선호 장르를 가지는 고객들을 k-means clustering 기법을 통해 적절히 군집화 하고, 군집화된 그룹을 기본 단위로 하여 협업 필터링을 수행한다. 협업 필터링은 고객군 내 사용자 유사도를 계산하는 단계, 유사도를 가지고 예측값 계산 및 추천 목록 생성 단계로 구성된다.
- [0031] 도 10은 고객의 선호 장르 변경의 실시예를 설명하기 위한 도면이며, 상기 도면에서 도시된 바와 같이, 피드백 정보가 반영된 선호 장르를 도출하기 위해 구매내역, 고객의 리뷰(호감, 기구매), 검색 클릭 정보를 취합한다. 각 경우에 따라 적용되는 가중치 값을 다르게 하여 선호 장르를 계산한다. 예를 들면, 선호 장르 계산시 가장 정량적인 신뢰도가 높은 구매내역은 가중치 5, 반품시 가중치 -5, 고객의 리뷰(기구매) 5, 고객의 리뷰(호감) 2, 검색 클릭은 1로 산정한다. 실시예에서 보여 주듯이, 고객번호 9번의 고객이 잡지를 2권 구매하고, 잡지 1권에 대해서 “읽고 싶어요” 평가를 남기고, 잡지 관련 도서를 2권 검색 했다면, 잡지 장르에 대한 선호 장르 점수는 14점이다. 9번 고객은 선호 장르가 “소설”에서 “잡지”로 바뀐다.
- [0032] 도 11은 군집화 기반 협업 필터링 수행에 대한 실시예를 설명하기 위한 도면 이며, 상기 도면에서 도시된 바와 같이, 군집3(소설/인문 장르를 선호하는 그룹)에 속하는 고객번호 1번을 추천 대상으로 하여 협업 필터링 수행을 위한 예제 데이터이다. 우선, 군집3 고객 중에서 1번 고객과 취향이 비슷한 사용자를 찾는다. 취향은 콘텐츠에 대한 평가로 나타날 것이므로, 각 사용자의 평가의 유사도(유사성)을 계산해 봅니다. 계산해보니, 고객번호 2, 3, 4가 유사도가 가장 높게 나온다. 이 사용자들을 취향이 비슷한 사용자 그룹인 Neighbor로 분류한다. 그 다음 단계는 고객번호 2, 3, 4가 가장 좋게 평가한 콘텐츠를 찾는 것이다. 추천 대상자가 아직 구매하지 않은 콘텐츠 중 하나인 '열한 계단'의 평가의 평균이 가장 높게 나온다. 따라서, 추천 대상자가 이 콘텐츠를 좋아할 것이라고 예상할 수 있다. 따라서 '열한 계단'을 추천 대상자(고객번호 1번)한테 추천 콘텐츠로 제시한다.
- [0033] 도 12는 연관 규칙 생성 과정을 설명하기 위한 예제 데이터 도면이며, 상기 도면에서 도시된 바와 같이, 주어진 예제 데이터를 기준으로 최소 지지도를 50%, 최소 신뢰도를 50%라고 가정하여 연관 규칙 생성 과정을 알아보기 위한 예제 데이터이다.
- [0034] 도 13은 연관 규칙 생성을 위해서 최소 지지도 이상의 지지도를 가지는 항목집합들에 대한 실시예를 설명하기 위한 도면이며, 상기 도면에서 도시된 바와 같이, 도 12의 데이터를 기준으로 계산하면 최소 지지도를 50%로 주었기 때문에 50% 이하의 지지도를 가지는 아이템을 모두 제거한 항목들에 대한 지지도이다.
- [0035] 도 14는 최종 Frequent Item Set의 실시예를 설명하기 위한 도면이며, 상기 도면에서 도시된 바와 같이, 도 13의 데이터를 기반으로 계산한 최종 Frequent Item Set 데이터이다.
- [0036] 도 15는 생성된 연관 규칙의 실시예를 설명하기 위한 도면이며, 상기 도면 에서 도시된 바와 같이, 도14의 최종 Frequent Item Set을 이용해 신뢰도가 50% 이상인 항목과 연관규칙을 추출한다.
- [0037] 도 16은 콘텐츠 기반 필터링의 실시예를 설명하기 위한 도면이며, 상기 도면에서 도시된 바와 같이, 본 발명은 초기 평가, 희박성 문제를 보완하고, 시간, 베스트셀러 등 트렌드를 반영한 추천을 위해 콘텐츠 기반 필터링을 사용한다. 콘텐츠 기반 필터링을 사용하여 고객의 선호 장르에 대한 시간과 베스트셀러를 추천해 주는 사례이다.
- [0038] 본 명세서에서는 본 발명을 설명하기 위한 실시예로서 K-means Clustering을 활용한 맞춤형 도서를 추천하는 방법을 중심으로 설명이 진행될 것이나, 이러한 새로운 추천방법은 스마트폰 뿐만 아니라 영상입력수단과 통신 기능이 구비된 개인 휴대용 단말기 상에서 구현될 수 있으며, 나아가 도서 뿐만 아니라 영화, 쇼핑물 물건 등 다양한 부가상품에도 적용될 수 있음은 본 발명이 속하는 분야의 당업자들에게 자명한 사실이라 할 것이다.
- [0039] 상기에서 기술된 구성과 효과를 가진 본 발명은 다양한 방법으로 변형이 가능하며, 상기에서 기술된 내용은 본 발명의 범위를 한정하고자 하는 것은 아니다. 또한 본 발명의 사상과 범위에서 벗어나지 않는 한도 내에서 다양한 변형과 수정이 가능하며, 본 발명이 속한 분야의 당업자에게 자명한 변형은 다음의 특허 청구범위 범위 내에 포함되어진다.

도면

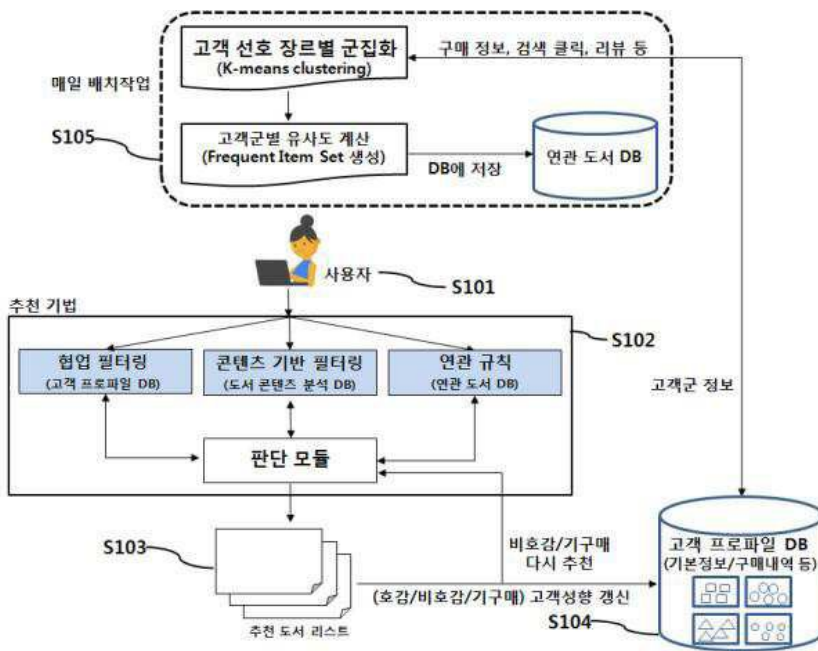
도면1



도면2



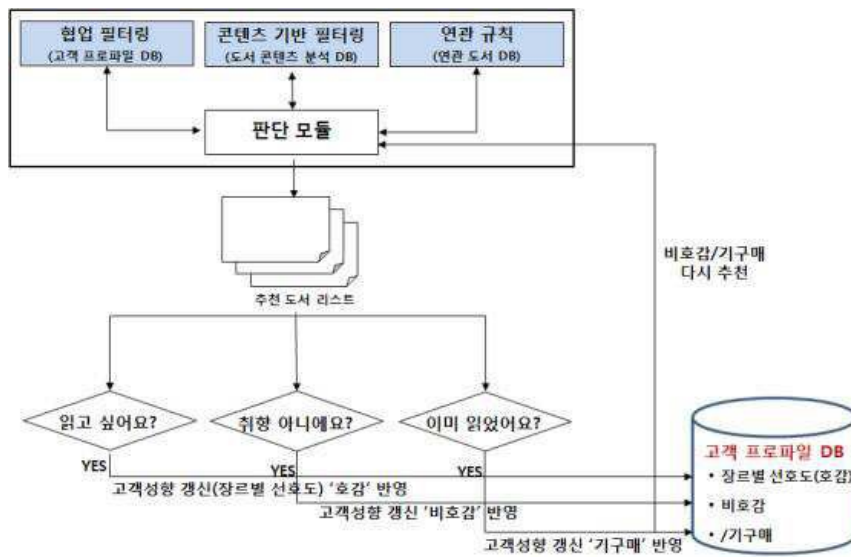
도면3



도면4

No.	이름	내용
1	고객번호	고객에게 부여 된 번호
2	생년	고객이 태어난 해
3	성별	남자, 여자
4	고객 성향	선호 장르 각 도서 장르의 선호도를 취합한 결과
5		장르별 선호도 장르별 선호도 점수
6		비호감 등록일자, ISBN
7		기구매 등록일자, ISBN
8	구매내역(반품 포함)	구매(반품)번호, 구매(반품)날짜, ISBN, 장르

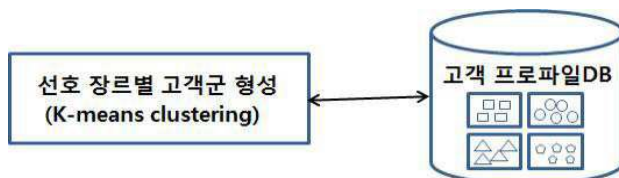
도면5



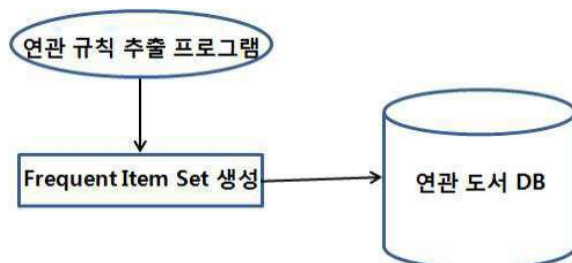
도면6

유형	선호 장르
군집1	컴퓨터/인터넷/자연/과학/공학
군집2	외국어/수험서/청소년 문제집
군집3	소설/문학/역사/인문/시/에세이/만화/종교
군집4	여행/자기 계발/예술/대중문화/잡지/가정/생활/요리
군집5	경영/경제/사회/법률

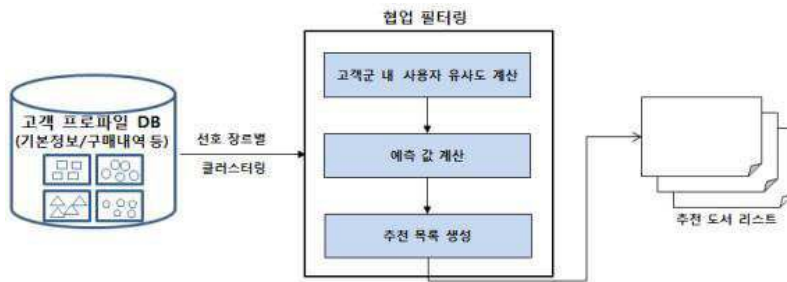
도면7



도면8



도면9



도면10

	A	E	F	G	H
1	고객번호	소셜	잡지	자연/과학	자기 계발
9	8	4	0	0	0
10	9	10	0	0	0
11	10	0	0	0	23
12	11	5	10	0	0
13	12	0	0	0	0
14	13	0	0	0	0
15	14	40	0	0	10

↓

	A	E	F	G	H
1	고객번호	소셜	잡지	자연/과학	자기 계발
9	8	4	0	0	0
10	9	10	14	0	0
11	10	0	0	0	23
12	11	5	10	0	0
13	12	0	0	0	0
14	13	0	0	0	0
15	14	40	0	0	10

도면11

고객번호	안생의 발견	열한 계단	채식주의자	도깨비	남한산성	유사도
1	5	-	1	2	-	-
2	5	5	1	3	-	4.2
3	4	5	2	-	3	4.5
4	5	4	1	1	2	4.1
5	1	2	5	5	4	2.3

도면12

고객번호	도서 구매 내역
1	미움 받을 용기, 구해줘, 자존감 수업
2	미움 받을 용기, 자존감 수업
3	미움 받을 용기, 나미야 잡화점의 기적
4	구해줘, 채식주의자, 완벽한 공부법

도면13

항목집합	지지도
미움 받을 용기	75%
구해줘	50%
자존감 수업	50%
미움 받을 용기, 자존감 수업	50%

도면14

항목집합	지지도
미움 받을 용기, 자존감 수업	50%

도면15

항목집합	신뢰도
미움 받을 용기 => 자존감 수업	2/3 = 66.6%
자존감 수업 => 미움 받을 용기	2/2 = 100%

도면16

고객번호	선호 장르	콘텐츠 기반 필터링 추천 도서	
1	소설	신간	리버 보이, 홀리데이 로맨스
		베스트셀러	돌이킬 수 없는 약속, 인생 우화