



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0042582
(43) 공개일자 2022년04월05일

(51) 국제특허분류(Int. Cl.)
G06Q 30/06 (2012.01) G06Q 30/02 (2012.01)
(52) CPC특허분류
G06Q 30/0631 (2013.01)
G06Q 30/0201 (2013.01)
(21) 출원번호 10-2020-0125566
(22) 출원일자 2020년09월28일
심사청구일자 2020년09월28일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
임일
서울특별시 서대문구 통일로 395, 109동 1304호(홍제동, 홍제 센트럴 아이파크)
서석윤
서울특별시 서대문구 신촌로7길 49-26 FULL HOUSE 302호
(74) 대리인
김인철

전체 청구항 수 : 총 17 항

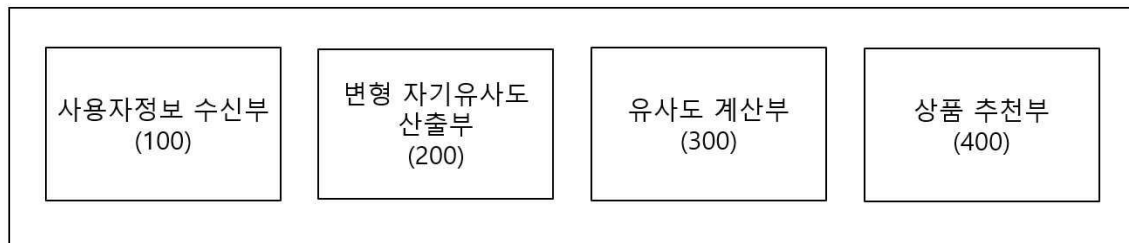
(54) 발명의 명칭 변형 자기유사도를 이용한 반복구매상품 추천시스템 및 추천방법

(57) 요약

본 발명은 사용자 단말기, 데이터베이스 및 웹서버가 네트워크로 연결되고, 컴퓨터에 의해 실행되는 상품추천시스템으로서, 사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신하는 사용자정보 수신부(100); 상기 사용자정보 수신부(100)에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출하는 변형 자기유사도 산출부(200); 상기 변형 자기유사도 산출부(200)에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 산입하여 상품간 유사도의 합을 산출하는 유사도 계산부(300); 및 상기 유사도 계산부(300)에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송하는 상품추천부(400)를 포함하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템이다.

대 표 도 - 도1

반복구매 상품추천시스템(10)



(52) CPC특허분류
G06Q 30/0202 (2013.01)

명세서

청구범위

청구항 1

사용자 단말기, 데이터베이스 및 웹서버가 네트워크로 연결되고, 컴퓨터에 의해 실행되는 상품추천시스템으로서,

사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신하는 사용자정보 수신부;

상기 사용자정보 수신부에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출하는 변형 자기유사도 산출부;

상기 변형 자기유사도 산출부에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 산입하여 상품간 유사도의 합을 산출하는 유사도 계산부; 및

상기 유사도 계산부에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송하는 상품 추천부를 포함하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

청구항 2

청구항 1에 있어서,

상기 사용자정보 수신부가 수신하는 정보는 사용자의 과거 구매상품정보 및 구매회수정보를 포함하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

청구항 3

청구항 1에 있어서,

상기 변형 자기유사도 산출부는 재구매율 산출부, 선호도 산출부 및 가중치 부여부를 포함하며, 다음 식 1로 변형 자기유사도를 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

[식 1]

$$MSS(i, j) = Repurchase\ rate(i) \times Preference(i, j) \times Weight$$

(여기서, MSS는 변형 자기유사도이고, MSS(i, j)는 상품 i에 대한 소비자 j의 변형 자기유사도값이며, Repurchase rate(i)는 상품 i의 재구매율이며, Preference(i, j)는 상품 i에 대한 소비자 j의 선호도이며, weight는 가중치이다)

청구항 4

청구항 3에 있어서,

상기 재구매율 산출부는 상기 재구매율을 다음 식 2로 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

[식 2]

$$Repurchase\ rate(i) = 1 - \frac{Users(i, 1)}{\sum_{k=1}^{max} Users(i, k)}$$

(여기서, Users(i, k)는 상품 i에 대해 k번 구매한 소비자 수이다)

청구항 5

청구항 3에 있어서,

상기 선호도 산출부는 선호도를 다음 식 3으로 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

[식 3]

$$Preference(i, j) = \frac{amount(i, j)}{\sum_{i=1}^n amount(i, j)}$$

(여기서, amount(i, j)는 상품 i에 대한 소비자 j의 구매 수량이며, n은 총 상품의 수이다)

청구항 6

청구항 3에 있어서,

상기 변형 자기유사도 산출부는 구매확률 산출부를 더 구비하여, 구매확률을 다음 식 4로 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

[식 4]

$$P_c(i, j) = \frac{1}{|(R_t - C_t(i, j)) - P_c(i)|}$$

(여기서, $P_c(i, j)$ 는 구매 주기 도래 정도에 따른 소비자 j의 상품 i에 대한 구매확률이고, R_t 는 추천 시스템의 추천 시점이고, $C_t(i, j)$ 는 소비자 j가 상품 i를 마지막으로 구매한 시점이고, $(R_t - C_t(i, j))$ 의 값은 추천 시점으로부터 과거 소비자 j가 상품 i를 마지막으로 구매한 시점간의 거리를 의미한다)

청구항 7

청구항 6에 있어서,

상기 변형 자기유사도 산출부는 식 5와 같이 상기 구매확률을 결합하여 상품구매주기를 가진 상품에 대한 변형 자기유사도를 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

[식 5]

$$MSS(i, j) = Repurchase\ rate(i) \times Preference(i, j) \times P_c(i, j) \times Weight$$

청구항 8

청구항 3에 있어서,

상기 변형 자기유사도 산출부는 반복구매선호성향 산출부를 더 구비하여, 반복구매선호성향을 다음 식 6 및 식 7로 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

[식 6]

$$Stickiness(j) = \left(1 - \frac{\sum_{i=1}^n unique\ item(i, j)}{\sum_{i=1}^n amount(i, j)}\right) \times 100$$

[식 7]

$$unique\ item(i, j) = 1\ if\ amount(i, j) \geq 1, else\ 0$$

(여기서, Stickiness(j)는 소비자 j의 반복 구매 선호 성향이며, amount(i, j)는 상품 i에 대한 소비자 j의 구매 수량이며, n은 총 상품의 수이다. unique item(i, j)는 소비자 j가 상품 i를 구매했을 경우 1, 구매하지 않았을 경우 0의 값을 가진다)

청구항 9

청구항 8에 있어서,

상기 변형 자기유사도 산출부는 다음 식 8과 같이 상기 반복구매선호성향을 변형자기유사도에 결합시켜 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

[식 8]

$$MSS(i,j) = Repurchase\ rate(i) \times Preference(i,j) \times Stickiness(j) \times Weight$$

청구항 10

청구항 1에 있어서,

상기 유사도 계산부는 상기 변형 자기유사도 산출부가 산출한 변형 자기유사도(MSS)를 이용하여, 사용자가 구매한 상품의 개수(m)와 전체 상품의 개수(n)를 $m \times n$ 형태의 유사도 매트릭스로 나타내고, 상품간 유사도의 합을 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

청구항 11

청구항 1에 있어서,

상기 상품 추천부는 상품간 유사도의 합이 높은 순서에서 기 설정된 개수의 상품을 추천하거나, 기 설정된 유사도 이상인 상품을 추천하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템.

청구항 12

사용자 단말기, 데이터베이스 및 웹서버가 네트워크로 연결되고, 컴퓨터에 의해 실행되는 상품추천방법으로서,

사용자정보 수신부는 사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신하는 S100 단계;

변형 자기유사도 산출부가 상기 사용자정보 수신부에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출하는 S200 단계;

유사도 계산부가 상기 변형 자기유사도 산출부에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 삽입하여 상품간 유사도의 합을 산출하는 S300 단계; 및

상품 추천부가 상기 유사도 계산부에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송하는 S400 단계를 포함하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천방법.

청구항 13

청구항 12에 있어서,

S200 단계는

변형 자기유사도 산출부가 재구매율을 산출하는 S210 단계, 선호도를 산출하는 S220 단계 및 가중치를 부여하는 S230 단계를 포함하며, 다음 식 1로 변형 자기유사도를 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천방법.

[식 1]

$$MSS(i,j) = Repurchase\ rate(i) \times Preference(i,j) \times Weight$$

(여기서, MSS는 변형 자기유사도이고, $MSS(i,j)$ 는 상품 i에 대한 소비자 j의 변형 자기유사도값이며, $Repurchase\ rate(i)$ 는 상품 i의 재구매율이며, $Preference(i,j)$ 는 상품 i에 대한 소비자 j의 선호도이며, weight는 가중치이다)

청구항 14

청구항 13에 있어서,

S210 단계에서, 상기 변형 자기유사도 산출부에 구비된 재구매율 산출부가 상기 재구매율을 다음 식 2로 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천방법.

[식 2]

$$Repurchase\ rate(i) = 1 - \frac{Users(i,1)}{\sum_{k=1}^{max} Users(i,k)}$$

(여기서, Users(i,k)는 상품 i에 대해 k번 구매한 소비자 수이다)

청구항 15

청구항 13에 있어서,

S220 단계에서, 상기 변형 자기유사도 산출부에 구비된 선호도 산출부가 선호도를 다음 식 3으로 산출하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천방법.

[식 3]

$$Preference(i,j) = \frac{amount(i,j)}{\sum_{i=1}^n amount(i,j)}$$

(여기서, amount(i,j)는 상품 i에 대한 소비자 j의 구매 수량이며, n은 총 상품의 수이다)

청구항 16

컴퓨터에 의해 실행되며,

사용자정보 수신부는 사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신하는 S100 단계;

변형 자기유사도 산출부가 상기 사용자정보 수신부에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출하는 S200 단계;

유사도 계산부가 상기 변형 자기유사도 산출부에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 산입하여 상품간 유사도의 합을 산출하는 S300 단계; 및

상품 추천부가 상기 유사도 계산부에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송하는 S400 단계를 실행시키기 위한 프로그램을 기록한 컴퓨터 해독 가능한 기록 매체.

청구항 17

하드웨어와 결합되어, 청구항 12 또는 청구항 15 중 어느 한 항에 기재된 변형 자기유사도를 이용한 반복구매상품 추천방법을 컴퓨터에 의해 실행시키기 위하여 컴퓨터가 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

발명의 설명**기술 분야**

[0001] 본 발명은 상품추천시스템 및 추천방법에 관한 것이다. 구체적으로는 변형 자기유사도를 이용한 반복구매상품 추천시스템 및 추천방법에 관한 것이다.

배경 기술

[0002] 빅데이터를 수집, 가공, 분석하는 테크놀로지가 크게 발전함에 따라 전자상거래 기업은 추천 시스템을 통해 고

객에게 적합한 아이템을 추천함으로써, 고객에게 더 관련성 있고 의미 있는 고객 경험을 제공하는 것이 가능해졌다.

[0003] 추천 시스템에는 여러 기법들이 있지만, 그중 다수의 전자상거래 기업이 활발하게 활용하고 있는 대표적인 추천 시스템 기법이 협업 필터링 기법(Collaborative Filtering, CF)이다. 협업 필터링 기법은 과거의 구매 기록을 바탕으로 소비자 간, 또는 아이템 간의 유사도를 측정하여 추천 아이템을 제안하는 기법이다. 이를 통해 소비자의 과거 구매 상품 이력을 바탕으로 소비자가 아직 구매하지 않은 상품들에 대한 선호도를 예측한 후 예상 선호도가 높은 순으로 상품을 추천한다.

[0004] 하지만, 종래의 협업 필터링 추천 기법은 전술한 바와 같이 새로운 상품을 추천하는 것을 기본 구조로 하고 있기 때문에 소비자가 이미 구매한 상품에 대한 추천은 불가능한 한계가 있다.

[0005] 즉, 새로운 상품의 추천이 적합한 음악, 드라마, 영화, 책과 같은 콘텐츠 상품군의 경우 협업 필터링 추천 기법이 높은 성능을 보이지만, 기저귀, 생수, 화장지, 비타민 보충제 등과 같이 반복적 재구매가 관찰되는 상품군의 경우 협업 필터링 추천 기법은 추천 성능이 떨어지게 된다. 본 명세서에서는 소비자가 이미 구매한 상품이라도 다시 추천해주는 추천을 '반복 구매 추천'이라고 정의한다.

[0007] 협업 필터링 알고리즘은 현재 사용자로부터 기존에 구매하여 사용하였던 상품에 관한 평가를 점수로 입력 받고, 이 평가 정보를 가지고 다른 사용자 중 비슷한 평가를 한 사람들을 추출(이하, 이들을 '준거 집단'으로 칭함)한 후, 이들이 좋게 평가한 제품 중에서 현재 사용자가 구입하지 않은 상품을 골라서 추천한다.

[0009] 협업 필터링 추천 기법은 상품의 추천 순위를 정할 때 소비자 간 또는 상품 간 유사도 계산에 기반한다. 협업 필터링 추천 기법에서의 유사도 계산에는 소비자 간 유사도 또는 상품 간 유사도를 활용한다. 소비자 유사도의 경우, 소비자 간 구매 상품 구성이 얼마나 비슷한지를 의미한다. 예를 들어 소비자 유사도를 활용할 경우 소비자 A가 상품 1, 2, 3을 구매했고, 소비자 B가 상품 1, 2, 3, 4를 구매했을 경우 소비자 A는 소비자 B와 유사도가 높다고 판단할 수 있다. 이를 통해 소비자 A에게는 소비자 A와 유사도가 높은 소비자 B가 구매한 상품인면서, 소비자 A가 아직 구매하지 않은 상품인 상품 4를 추천한다. 사용자 간 유사도를 바탕으로 한 협업 필터링 추천 기법의 경우 소비자 A가 이미 구매한 상품에 대한 반복 구매 추천은 이뤄지지 않게 된다.

[0011] 한편, 상품 간 유사도의 경우 상품 간 얼마나 비슷한 구매 이력을 가진 소비자가 구매했는지를 의미한다. 상품 간 유사도를 활용할 경우 상품 1을 구매한 소비자가 소비자 A, B, C, D이고, 상품 2를 구매한 소비자가 소비자 A, B, C 일 경우, 상품 1은 상품 2와 유사도가 높다고 할 수 있다. 이를 통해 상품 1은 구매했지만, 상품 2를 아직 구매하지 않은 소비자에게는 상품 2를 추천하게 된다. 상품 간 유사도를 활용한 협업 필터링 추천 기법에서도 사용자 간 유사도를 활용한 협업 필터링 기법과 마찬가지로 소비자가 이미 구매한 상품에 대한 반복 구매 추천은 이루어지지 않는다.

[0013] 종래의 상품 유사도를 기반으로 한 협업 필터링 추천 기법이 소비자가 과거 구매한 상품까지 추천 리스트에 포함하고자 할 경우, 과거 구매한 상품은 모두 상품 자기 자신에 대한 유사도가 1이 된다. 본 명세서에서는 이러한 상품 자기 자신에 대한 유사도를 '자기 유사도(self similarity)'라고 정의한다.

[0014] 예를 들어 상품 A와 상품 A의 자기 유사도는 1의 값을 갖는다. 다시 말해, 소비자가 구매한 상품들과 유사도가 높은 상품을 뽑는 과정에서 소비자가 한 번 구매한 상품이든, 여러 번 구매한 상품이든 소비자가 이미 구매한 상품은 모두 자기 유사도 값으로 1을 가지게 됨으로써 소비자가 구매한 상품 간 구매 횟수 차이가 반영되지 않는 문제가 발생하게 된다.

[0015] 즉, 종래의 협업 필터링 추천 기법을 통해 반복 구매 추천을 하고자 할 경우, 소비자의 구매 상품 간 선호도 차이를 반영하지 못하게 된다. 이는 소비자가 여러 번 재구매한 상품이 한 번만 구매한 상품보다 다시 재구매할 확률이 높다는 점을 추천 시스템에 반영하지 못하는 한계점이라고 할 수 있다.

[0017] 또한, 종래의 협업 필터링 추천 기법에서 반복 구매 추천을 하고자 할 경우, 상품의 특성상 반복적인 재구매가 관찰되는 상품이나 반복적인 재구매가 전혀 이루어지지 않는 상품을 구분하지 않고, 모두 소비자가 일단 한 번이라도 구매하면 똑같이 상품의 자기유사도가 1이 된다. 이는 상품의 특성상 재구매가 이루어지지 않는 상품의 경우도 재구매가 빈번하게 이루어지는 상품과 차이 없이 소비자에게 추천되는 문제를 발생시키게 된다.

[0019] 또한, 종래의 협업 필터링 추천 기법은 소비자의 구매 내역을 바탕으로 한 유사도 정보 이외의 다른 정보를 추천에 활용하지 못하는 한계가 있다.

[0021] 이미 구매한 상품에 대한 추천 시스템 연구들도 최근 진행되고 있다. 하지만 협업필터링 추천 기법을 활용해 협업 필터링 추천 기법의 장점인 상품 간 유사도 정보를 반복 구매 추천 시스템에 활용하려는 시도는 아직 없었다.

[0023] 또한 기존의 반복적 구매가 이뤄지는 상품에 대한 추천 시스템에 관한 연구들은 어떤 소비자들에게 반복 구매 추천 시스템을 활용하는 것이 좋은지에 대한 측정 및 평가 방법을 제안하지는 못하고 있는 실정이다.

선행기술문헌

특허문헌

[0024] (특허문헌 0001) (문헌 1) 한국공개특허공보 제10-2020-0080024호(2020.07.06)

(특허문헌 0002) (문헌 2) 한국등록특허공보 제10-2050738호 (2019.11.26)

발명의 내용

해결하려는 과제

[0025] 본 발명에 따른 변형 자기유사도를 이용한 반복구매상품 추천시스템 및 추천방법은 다음과 같은 해결과제를 가진다.

[0026] 첫째, 사용자의 반복구매 성향을 고려하여 상품을 추천하고자 한다.

[0027] 둘째, 반복구매추천시스템에 협업 필터링의 장점인 상품 간 유사도를 추천에 활용하는 협업 필터링 추천시스템을 적용하고자 한다.

[0028] 셋째, 반복 구매 추천에 필요한 정보들을 기업 및 분석가가 자유롭게 고안 및 선택하는 것이 가능하게 한다.

[0029] 본 발명의 해결과제는 이상에서 언급한 것들에 한정되지 않으며, 언급되지 아니한 다른 해결과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0030] 본 발명은 사용자 단말기, 데이터베이스 및 웹서버가 네트워크로 연결되고, 컴퓨터에 의해 실행되는 상품추천시스템으로서, 사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신하는 사용자정보 수신부; 상기 사용자정보 수신부에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출하는 변형 자기유사도 산출부; 상기 변형 자기유사도 산출부에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 삽입하여 상품간 유사도의 합을 산출하는 유사도 계산부; 및 상기 유사도 계산부에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송하는 상품 추천부를 포함하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템이다.

[0031] 본 발명에 있어서, 사용자정보 수신부가 수신하는 정보는 사용자의 과거 구매상품정보 및 구매회수정보를 포함할 수 있다.

[0032] 본 발명에 있어서, 변형 자기유사도 산출부는 재구매율 산출부, 선호도 산출부 및 가중치 부여부를 포함하며,

다음 식 1로 변형 자기유사도를 산출할 수 있다.

[식 1]

$$MSS(i, j) = Repurchase\ rate(i) \times Preference(i, j) \times Weight$$

(여기서, MSS는 변형 자기유사도이고, MSS(i, j)는 상품 i에 대한 소비자 j의 변형 자기유사도값이며, Repurchase rate(i)는 상품 i의 재구매율이며, Preference(i, j)는 상품 i에 대한 소비자 j의 선호도이며, weight는 가중치이다)

본 발명에 있어서, 재구매율 산출부는 상기 재구매율을 다음 식 2로 산출할 수 있다.

[식 2]

$$Repurchase\ rate(i) = 1 - \frac{Users(i, 1)}{\sum_{k=1}^{max} Users(i, k)}$$

(여기서, Users(i, k)는 상품 i에 대해 k번 구매한 소비자 수이다)

본 발명에 있어서, 선호도 산출부는 선호도를 다음 식 3으로 산출할 수 있다.

[식 3]

$$Preference(i, j) = \frac{amount(i, j)}{\sum_{i=1}^n amount(i, j)}$$

본 발명에 있어서, 변형 자기유사도 산출부는 구매확률 산출부를 더 구비하여, 구매확률을 다음 식 4로 산출할 수 있다.

[식 4]

$$P_c(i, j) = \frac{1}{|(R_t - C_t(i, j)) - P_c(i)|}$$

본 발명에 있어서, 변형 자기유사도 산출부는 식 5와 같이 상기 구매확률을 결합하여 상품구매주기를 가진 상품에 대한 변형 자기유사도를 산출할 수 있다.

[식 5]

$$MSS(i, j) = Repurchase\ rate(i) \times Preference(i, j) \times P_c(i, j) \times Weight$$

본 발명에 있어서, 변형 자기유사도 산출부는 반복구매선호성향 산출부를 더 구비하여, 반복구매선호성향을 다음 식 6으로 산출할 수 있다.

[식 6]

$$Stickiness(j) = \left(1 - \frac{\sum_{i=1}^n unique\ item(i, j)}{\sum_{i=1}^n amount(i, j)}\right) \times 100$$

[식 7]

$$unique\ item(i, j) = 1\ if\ amount(i, j) \geq 1, else\ 0$$

본 발명에 있어서, 변형 자기유사도 산출부는 다음 식 8과 같이 상기 반복구매선호성향을 변형자기유사도에 곱시켜 산출할 수 있다.

[0055] [식 8]

$$[0056] \quad MSS(i, j) = Repurchase\ rate(i) \times Preference(i, j) \times Stickiness(j) \times Weight$$

[0057] 본 발명에 있어서, 유사도 계산부는 상기 변형 자기유사도 산출부가 산출한 변형 자기유사도(MSS)를 이용하여, 사용자가 구매한 상품의 개수(m)와 전체 상품의 개수(n)를 $m \times n$ 형태의 유사도 매트릭스로 나타내고, 상품간 유사도의 합을 산출할 수 있다.

[0058] 본 발명에 있어서, 상품 추천부는 상품간 유사도의 합이 높은 순서에서 기 설정된 개수의 상품을 추천하거나, 기 설정된 유사도 이상인 상품을 추천할 수 있다.

[0059] 본 발명은 사용자 단말기, 데이터베이스 및 웹서버가 네트워크로 연결되고, 컴퓨터에 의해 실행되는 상품추천방법으로서, 사용자정보 수신부는 사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신하는 S100 단계; 변형 자기유사도 산출부가 상기 사용자정보 수신부에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출하는 S200 단계; 유사도 계산부가 상기 변형 자기유사도 산출부에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 산입하여 상품간 유사도의 합을 산출하는 S300 단계; 및 상품 추천부가 상기 유사도 계산부에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송하는 S400 단계를 포함할 수 있다.

[0060] 본 발명에 있어서, S200 단계는 변형 자기유사도 산출부가 재구매율을 산출하는 S210 단계, 선호도를 산출하는 S220 단계 및 가중치를 부여하는 S230 단계를 포함하며, 식 1로 변형 자기유사도를 산출할 수 있다.

[0061] 본 발명에 있어서, S210 단계에서, 변형 자기유사도 산출부에 구비된 재구매율 산출부가 상기 재구매율을 식 2로 산출할 수 있다.

[0062] 본 발명에 있어서, S220 단계에서, 변형 자기유사도 산출부에 구비된 선호도 산출부가 선호도를 식 3으로 산출할 수 있다.

[0063] 본 발명은 기록매체로서, 컴퓨터에 의해 실행되며, 사용자정보 수신부는 사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신하는 S100 단계; 변형 자기유사도 산출부가 상기 사용자정보 수신부에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출하는 S200 단계; 유사도 계산부가 상기 변형 자기유사도 산출부에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 산입하여 상품간 유사도의 합을 산출하는 S300 단계; 및 상품 추천부가 상기 유사도 계산부에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송하는 S400 단계를 실행시키기 위한 프로그램을 기록한 컴퓨터 해독 가능한 기록매체로 구현될 수 있다.

[0064] 본 발명은 컴퓨터 프로그램으로서, 하드웨어와 결합되어, 본 발명에 따른 변형 자기유사도를 이용한 반복구매상품 추천방법을 컴퓨터에 의해 실행시키기 위하여 컴퓨터가 판독 가능한 기록매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다.

발명의 효과

[0065] 본 발명에 따른 변형 자기유사도를 이용한 반복구매상품 추천시스템 및 추천방법은 다음과 같은 효과를 가진다.

[0066] 첫째, 사용자의 반복구매 성향을 고려하여 상품을 추천하는 효과가 있다.

[0067] 둘째, 반복구매추천시스템에 협업 필터링의 장점인 상품 간 유사도를 추천에 활용하는 협업 필터링 추천시스템을 적용하는 효과가 있다.

[0068] 셋째, 반복 구매 추천에 필요한 정보들을 기업 및 분석가가 자유롭게 고안 및 선택하는 것이 가능하게 하는 효과가 있다.

[0069] 본 발명의 효과는 이상에서 언급된 것들에 한정되지 않으며, 언급되지 아니한 다른 효과들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0070] 도 1은 본 발명에 따른 변형 자기유사도를 이용한 반복구매상품 추천시스템의 구성을 나타낸다.

도 2는 본 발명에 따른 변형 자기유사도 산출부의 구성을 나타낸다.

도 3은 본 발명에 따른 변형 자기유사도를 이용한 반복구매상품 추천방법의 순서도를 나타낸다.

도 4는 본 발명에 따른 컴퓨팅 장치를 나타내는 도면이다.

도 5는 협업 필터링에서 상품 간 유사도 계산 방식의 일 실시예를 나타낸다.

도 6은 소비자의 기 구매 상품을 포함한 유사도 계산 방식의 일 실시예를 나타낸다.

도 7은 자기 유사도를 포함한 상품 유사도 매트릭스의 일 실시예를 나타낸다.

도 8은 변형 자기유사도(Modified Self Similarity: MSS)를 활용한 상품간 유사도 계산 매트릭스의 일 실시예를 나타낸다.

도 9는 여러 추천시스템의 성능을 비교한 것이다.

발명을 실시하기 위한 구체적인 내용

- [0071] 이하, 첨부한 도면을 참조하여, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명의 실시예를 설명한다. 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 용이하게 이해할 수 있는 바와 같이, 후술하는 실시예는 본 발명의 개념과 범위를 벗어나지 않는 한도 내에서 다양한 형태로 변형될 수 있다. 가능한 한 동일하거나 유사한 부분은 도면에서 동일한 도면부호를 사용하여 나타낸다.
- [0072] 본 명세서에서 사용되는 전문용어는 단지 특정 실시예를 언급하기 위한 것이며, 본 발명을 한정하는 것을 의도하지는 않는다. 여기서 사용되는 단수 형태들은 문구들이 이와 명백히 반대의 의미를 나타내지 않는 한 복수 형태들도 포함한다.
- [0073] 본 명세서에서 사용되는 "포함하는"의 의미는 특정 특성, 영역, 정수, 단계, 동작, 요소 및/또는 성분을 구체화하며, 다른 특정 특성, 영역, 정수, 단계, 동작, 요소, 성분 및/또는 군의 존재나 부가를 제외시키는 것은 아니다.
- [0074] 본 명세서에서 사용되는 기술용어 및 과학용어를 포함하는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 일반적으로 이해하는 의미와 동일한 의미를 가진다. 사전에 정의된 용어들은 관련기술문헌과 현재 개시된 내용에 부합하는 의미를 가지는 것으로 추가 해석되고, 정의되지 않는 한 이상적이거나 매우 공식적인 의미로 해석되지 않는다.
- [0076] 본 발명에서는 소비자가 이미 구매한 상품들도 추천 시스템의 추천 리스트에 포함하는 협업 필터링 추천 시스템 및 그 방법을 제공하고자 한다.
- [0077] 이를 위해 본 발명은 첫째로, 반복 구매 추천 시스템에서도 협업 필터링의 장점인 상품 간 유사도를 추천에 활용하는 협업 필터링 추천 시스템 및 그 방법을 제공하고자 한다. 본 발명은 소비자가 여러 번 반복 구매한 상품 일수록 해당 소비자에게 더 높은 확률로 추천될 수 있도록 소비자의 과거 반복 구매 횟수를 추천에 반영하는 협업 필터링 추천 시스템 및 그 방법을 제공하고자 한다. 본 발명은 반복 구매가 자주 발생하는 상품에 추천이 이루어질 확률이 높도록 하고, 반복 구매가 적게 발생하는 상품도 일반적인 추천 알고리즘에 따라 추천하는 협업 필터링 추천 시스템 및 방법을 제공하고자 한다.
- [0078] 둘째로, 본 발명은 반복 구매 추천에 필요한 여러 정보들을 기업 및 분석가가 자유롭게 고안 및 선택하여, 본 발명에서 제시하는 협업 필터링 추천 시스템 알고리즘에 필요에 따라 자유롭게 결합 및 분리해서 활용할 수 있는 협업 필터링 추천 시스템 및 그 방법을 제공하고자 한다.
- [0079] 셋째로, 본 발명은 소비자의 반복 구매 성향을 측정하여 반복 구매 추천 시스템을 활용하기에 적합한 소비자를 평가할 수 있는 추천 시스템 및 그 방법을 제공하고자 한다. 반복 구매가 발생하는 정도는 제품 간에 차이가 있을 뿐 아니라 소비자 간에도 차이가 있다. 이러한 소비자의 반복 구매 성향을 추천에 반영하는 방법을 본 발명에서 제공하고자 한다.
- [0081] 본 발명에서는 종래의 반복 구매 추천 시스템에는 활용되지 않던 협업 필터링 추천 기법을 반복 구매 추천 시스

템에 맞도록 변형한 새로운 반복 협업 필터링 추천 기법을 제시한다. 이러한 방법은 다음과 같은 부분에서 기존 방식과 다른 새로운 방법이다.

- [0082] 첫째, 종래의 협업 필터링 기법이 소비자가 과거에 구매하지 않은 새로운 상품만을 추천한다면, 본 발명에서 제시하는 협업 필터링 기법은 소비자가 이미 구매한 상품에 대해서도 반복구매의 성향에 따라 적절한 추천이 가능해 반복적 재구매가 관찰되는 상품의 추천에 있어 종래의 협업 필터링 기법보다 더 정확한 추천 결과를 얻을 수 있다.
- [0083] 둘째, 종래의 협업 필터링 추천 시스템에서는 과거 소비자의 상품에 대한 구매 유무만을 이진(binary) 데이터로 파악하여 활용한다. 여기서 소비자의 구매 횟수 정보에 대한 정보 손실이 발생한다. 본 발명에서 제시하는 협업 필터링 추천 시스템은 소비자의 상품의 구매 횟수를 정보의 손실 없이 반영할 수 있다.
- [0084] 셋째, 종래의 협업 필터링은 소비자의 상품에 대한 구매 유무만을 이진(binary) 데이터로 파악하여 활용하기 때문에 각 상품이 반복적 재구매가 이뤄지는 상품인지 여부에 대한 정보를 추천에 활용할 수 없다. 본 발명에서 제시하는 협업 필터링 추천 시스템은 상품의 재구매 가능성과 재구매 확률 정보를 기존 협업 필터링 추천 시스템에 결합한 새로운 추천 시스템을 제시한다.
- [0085] 넷째, 종래의 협업 필터링 기법은 분석가가 추천에 도움이 될만한 다른 형태의 정보를 알고 있는 경우에도, 유사도 산출 시 단순히 소비자의 상품 구매 유무 내역만을 참조한다. 하지만 본 발명에서 제시하는 협업 필터링 추천 시스템은 유사도 계산에 분석가가 추천에 유의미하다고 판단하는 여러가지 정보를 결합할 수 있다. 결합할 수 있는 정보의 예로는 소비자의 상품 구매 횟수, 상품의 재구매율, 소비자의 반복 구매 경향 등이 있다.
- [0086] 다섯째, 종래의 반복 구매 추천 시스템 연구들은 어떤 소비자에게 반복 구매 추천 시스템을 활용하는 것이 좋은지에 대한 측정 및 평가 지표를 제시하지는 못했다. 본 발명에서 제시하는 시스템은 소비자의 반복 구매 성향에 대한 측정 및 평가를 통해 반복 구매 추천 시스템의 활용에 적합한 소비자와 제품을 파악할 수 있다.
- [0088] 종래의 협업 필터링 추천 기법은 새로운 상품의 추천을 위해 소비자 간 또는 상품 간 유사도를 계산해서 활용한다. 먼저 소비자 간 유사도를 활용할 경우, 소비자 1이 A, B, C라는 상품을 구매했고, 소비자 1과 구매성향의 유사도가 높은 소비자 2가 A, B, C, D라는 상품을 구매했을 경우 소비자 1에게 D라는 상품을 추천한다. 이와 같이 소비자 간 유사도를 활용할 경우에는 새로운 상품을 추천할 수 있을 뿐, 기존 구매한 상품에 대한 반복 구매 추천은 불가능하다.
- [0090] 한편 상품 간 유사도를 활용할 경우 상품 A를 구매한 소비자가 소비자 1, 2, 3, 4이고 상품 B를 구매한 소비자가 소비자 1, 2, 3라면 상품 A와 상품 B는 높은 유사도를 갖고 있다고 할 수 있다. 이때 소비자 4가 상품 A는 구매했지만 상품 B는 아직 구매하지 않았을 경우, 소비자 4에게 상품 B를 추천하게 된다. 이러한 상품 간 유사도의 계산은 도 5의 형태로 나타난다.
- [0092] 도 5는 협업 필터링에서 상품 간 유사도 계산 방식의 일 실시예를 나타낸다.
- [0093] 예를 들어 전체 아이템이 n 개이고, 소비자가 Item 4, 57, 96, 200을 구매했을 경우, 전체 아이템 n 개에서 소비자가 구매한 아이템 4개를 제외한 나머지 아이템들을 대상으로 소비자가 구매한 아이템과의 유사도를 구한다. 그리고 유사도가 가장 높은 N 개의 아이템을 해당 소비자에게 추천하게 된다.
- [0095] 이와 같이 종래의 협업 필터링에서는 소비자가 구매한 상품을 추천에서 제외하게 된다. 하지만 본 발명에서는 이미 소비자가 구매한 상품까지도 상품 추천이 가능한 반복 구매 추천 시스템을 구축하기 위해 소비자가 이미 구매한 상품 또한 상품 간 유사도 계산에 포함하는 형식을 취한다. 이를 나타내면 도 6과 같다.
- [0097] 도 6은 소비자의 기 구매 상품을 포함한 유사도 계산 방식의 일 실시예를 나타낸다.
- [0098] 즉 소비자가 구매한 상품의 개수를 m 개, 전체 상품의 개수를 n 개라고 했을 때, 종래의 협업 필터링 상품 간 유사도 계산 방식에서는 $m \times (n - m)$ 형태의 상품 유사도 매트릭스가 생성된다면, 본 발명에서는 $m \times n$ 형태의 유

사도 매트릭스를 활용한다. 이는 소비자가 이미 구매한 상품도 추천에 포함할 수 있도록 하기 위함이다. 본 명세서에서는 상품 i 와 상품 j 의 유사도를 S_{ij} 로 표현하기로 한다. 그리고, 상품 i 와 상품 i 간의 유사도는 '자기 유사도(Self Similarity)'로 정의하며, S_{ii} 로 표현하기로 한다.

- [0100] 도 7은 자기 유사도를 포함한 상품 유사도 매트릭스의 일 실시예를 나타낸다.
- [0101] 상품 i 와 상품 j 의 상품간 유사도 S_{ij} 는 본래 상품 i 의 소비자와 상품 j 의 소비자가 얼마나 비슷한 구매 패턴을 보이는지에 대한 정보를 나타내는 값이라고 할 수 있다. 따라서 어떤 상품이든 상품 i 와 상품 i 의 유사도인 자기 유사도는 본래 1의 값을 가지게 된다. 다른 상품 간의 상품 간 유사도(S_{ij})는 1을 넘을 수 없고 통상 1보다 훨씬 작은 값을 가지기에, S_{ii} 의 값에 1을 넣은 뒤 소비자별 구매 상품 유사도 매트릭스에 기반해 상위 N 개의 상품 추천을 할 경우, 과거 구매 상품이 추천 리스트에 포함될 확률이 매우 높아지게 된다(도 7 참조).
- [0103] 자기 유사도에 1의 값이 들어갈 경우, Top N 추천 시 추천 상품 개수를 N , 소비자 k 의 과거 구매 상품 개수를 M_k 라고 할 때, $N > M_k$ 일 경우, $(N - M_k)$ 개의 상품은 기존 상품 기반 협업 필터링 방식과 같이 소비자가 과거 구매한 상품들과 유사도 총합이 높은 새로운 상품이 추천되게 된다. 반대로, $N < M_k$ 인 경우 M_k 개의 상품 중에서 상품 유사도 총합이 가장 높은 N 개의 상품이 추천되게 된다.
- [0105] 이와 같이 상품 i (Item $_i$)과 상품 i (Item $_i$)의 유사도 S_{ii} 를 활용하면 소비자가 기존에 구매했던 상품들을 추천 리스트에 포함할 수 있고, 과거 구매 상품이 적은 소비자에게도 상품 유사도에 기반한 새로운 상품 추천이 기술적으로는 가능해진다.
- [0106] 그러나 앞서 기술했듯, 소비자가 과거에 한 번이라도 구매했던 상품들은 모두 필연적으로 자기 유사도 값으로 1을 가지게 되어 소비자가 과거 구매했던 상품들에 대한 재구매 횟수 정보가 반영되지 않는 문제가 발생한다.
- [0107] 예를 들어 소비자가 과거 10번 구매한 상품이나 1번 구매한 상품이나 모두 자기 유사도 값이 1이 되어 과거 구매 상품에 대한 소비자의 재구매 성향 정보가 반영되지 않는다. 또한 상품의 성격상 재구매가 이뤄지지 않는 상품들도 소비자가 과거 해당 상품을 구매했을 경우 자기 유사도 값으로 1을 갖게 되어, 일반적으로 재구매가 이뤄지지 않는 상품을 소비자에게 추천할 확률이 높아지는 문제가 발생한다.
- [0109] 이하에서는 도면을 참고하여 본 발명을 더욱 구체적으로 설명하고자 한다. 참고로, 도면은 본 발명의 특징을 설명하기 위하여, 일부 과장되게 표현될 수도 있다. 이 경우, 본 명세서의 진 취지에 비추어 해석되는 것이 바람직하다.
- [0111] 도 1은 본 발명에 따른 변형 자기유사도를 이용한 반복구매상품 추천시스템의 구성을 나타낸다.
- [0112] 본 발명은 사용자 단말기, 데이터베이스 및 웹서버가 네트워크로 연결되고, 컴퓨터에 의해 실행되는 상품추천시스템으로서, 사용자정보 수신부(100), 변형 자기유사도 산출부(200), 유사도 계산부(300) 및 상품 추천부(400)를 포함하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천시스템이다.
- [0114] 본 발명에 따른 사용자정보 수신부(100)는 사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신할 수 있다.
- [0115] 사용자 단말기(UE, User Equipment)는 통신네트워크를 경유하여 웹 서버(WS, Web Server)와 데이터를 송수신할 수 있는 단말기를 의미한다. 이동통신 단말기(Mobile Communication Terminal), 개인휴대용 정보단말기(PDA: Personal Digital Assistant), 태블릿, 노트북, 데스크탑 컴퓨터 등 중 어느 하나일 수 있으며, 유무선 데이터 송수신을 위한 프로그램을 저장하기 위한 메모리, 프로그램을 실행하여 연산 및 제어하기 위한 마이크로프로세

서 등을 구비하고 있는 단말기를 의미한다.

- [0116] 컴퓨터를 이용한 네트워크는 유/무선 통신네트워크가 될 수 있으며, LAN, WAN 등을 포함하는 어떠한 네트워크도 가능하나, 인터넷 네트워크로 구현하는 것이 더욱 바람직하다.
- [0117] 웹 서버(WS)는 사용자에게 추천 상황에 따라 추천 정보를 제공하게 되는 쇼핑몰 사이트 또는 포털 사이트(또는 추천 사이트) 등 일 수 있다. 또한, 다른 구현 방법으로서 포털 사이트 등과 링크되어 사용자의 요청에 따라 정보를 제공하는 일반적인 웹 사이트일 수 있다.
- [0118] 데이터베이스 웹 서버 내부에 구비될 수도 있고, 웹 서버와 별개로 구비될 수도 있을 것이다.
- [0120] 한편, 상품추천은 사용자 본인이 직접 요청하는 경우도 있고, 또는 사용자가 접속하고 있는 웹 사이트가 사용자의 명시적 요청이 없더라도 독자적인 판단에 따라 사용자를 대신해서 추천할 수도 있다. 예를 들어, 사용자의 웹 사이트 방문시 추천을 위한 웹 페이지 제공시에는 사용자의 추천 요청이 없어도 웹 사이트에서 자체적으로 추천 목록을 제공할 수가 있다.
- [0122] 본 발명에 따른 사용자정보 수신부(100)가 수신하는 정보는 사용자의 과거 구매상품정보 및 구매회수정보를 포함한다.
- [0124] 본 발명에 따른 변형 자기유사도 산출부(200)는 사용자정보 수신부(100)에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출할 수 있다.
- [0126] 본 발명에 따른 유사도 계산부(300)는 변형 자기유사도 산출부(200)에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 산입하여 상품간 유사도의 합을 산출할 수 있다.
- [0128] 본 발명에 따른 상품 추천부(400)는 유사도 계산부(300)에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송할 수 있다.
- [0130] 이하에서는, 본 발명에 따른 변형 자기유사도 산출부(200)에서 변형 자기유사도를 산출하는 구성을 설명하고자 한다.
- [0131] 본 발명에 따른 변형 자기유사도 산출부(200)는 재구매율 산출부(210), 선호도 산출부(220) 및 가중치 부여부(230)를 포함하며, 다음 식 1로 변형 자기유사도를 산출할 수 있다.
- [0132] [식 1]
- [0133]
$$MSS(i, j) = Repurchase\ rate(i) \times Preference(i, j) \times Weight$$
- [0135] 본 발명의 경우, 소비자가 이미 구매한 상품을 유사도 계산 대상에 포함하면서도, 이미 구매한 상품을 유사도 매트릭스에 추가할 때 발생하는 전술한 문제점들을 해결하기 위해, 이미 구매한 제품의 자기 유사도를 1이 아닌 값으로 변형할 수 있다. 이러한 변형의 목적은 상품 i에 대해서 소비자 j가 재구매할 확률을 계산하기 위한 것이다.
- [0136] 본 발명에서는 이를 변형 자기 유사도(Modified Self Similarity: MSS)로 정의한다(식 1 참조).
- [0138] 식 1에서, MSS는 변형 자기유사도이고, MSS(i, j)는 상품 i에 대한 소비자 j의 변형 자기유사도값이다. 상품 i에 대한 소비자 j의 변형 자기 유사도라는 것은 곧 상품의 자기 유사도값이 소비자에 따라 다르게 정해진다는

것을 의미한다.

[0139] Repurchase rate(i)는 상품 i의 재구매율이며, Preference(i,j)는 상품 i에 대한 소비자 j의 선호도이다. weight 즉 가중치는 해당 함수의 파라미터이다.

[0140] 본 발명에 따른 가중치는 재구매율과 선호도가 계산된 후에 소비자의 구매 데이터를 사용하여 다양한 가중치의 정확도를 비교해서 그중 가장 정확도가 높은 값으로 선택될 수 있다.

[0142] 본 발명에 따른 재구매율 산출부(210)는 재구매율을 다음 식 2로 산출할 수 있다.

[0143] [식 2]

$$Repurchase\ rate(i) = 1 - \frac{Users(i,1)}{\sum_{k=1}^{max} Users(i,k)}$$

[0144]

[0145] 여기서, Repurchase rate(i)는 상품 i의 재구매율로서 재구매 성향을 의미한다. Users(i, k)는 상품 i에 대해 k번 구매한 소비자 수를 의미한다.

[0146] 이와 같이, 상품의 재구매율을 변형 자기 유사도에 활용함으로써 상품 특성상 재구매가 이루어지지 않는 상품에 대한 필터링이 가능해진다. 상품 특성상 재구매가 이루어지지 않는 상품은 서적, 영화 등이 이에 해당될 수 있다. 물론 소비자가 동일한 서적을 재구매할 수도 있을 것이나, 일반적으로 서적은 재구매가 거의 이루어지지 않는 상품에 해당된다.

[0148] 한편, 재구매율이 0인 상품은 변형 자기 유사도가 0이 되어 이미 해당 상품을 구매한 사람에게는 추천이 이루어질 확률이 매우 낮아지게 된다. 또한 재구매율이 높은 상품일수록 변형 자기 유사도 값이 높아짐에 따라 다른 상품들에 비해 더 높은 확률로 추천되며, 재구매율이 낮은 상품일수록 다른 상품들에 비해 더 낮은 확률로 추천되게 된다.

[0150] 본 발명에 따른 선호도 산출부(220)는 선호도를 다음 식 3으로 산출할 수 있다.

[0151] [식 3]

$$Preference(i,j) = \frac{amount(i,j)}{\sum_{i=1}^n amount(i,j)}$$

[0152]

[0153] 상품 i에 대한 소비자 j의 선호도인 Preference(i, j)는 식 3과 같이 계산된다. 식 3에서 amount(i, j)는 상품 i에 대한 소비자 j의 구매 수량을, n은 총 상품의 수를 의미한다. 소비자의 상품 선호도를 활용함으로써 과거 소비자가 더 많이 구매한 상품일수록 더 높은 변형 자기 유사도 값을 갖게 되므로, 더 높은 확률로 추천되게 된다.

[0155] 한편, 두 가지 이상 상품이 소비자의 상품 선호도와 상품의 재구매율이 동일할 경우에는 변형 자기 유사도 값이 같게 되므로, 이 경우에는 종래의 협업 필터링의 상품 간 유사도 값이 높은 상품이 추천되게 된다.

[0157] 본 발명에 따른 변형 자기 유사도(MSS)를 활용한 협업 필터링의 상품 간 유사도 계산 매트릭스를 표현하면 도 8과 같다. 도 8은 변형 자기유사도(MSS)를 활용한 상품간 유사도 계산 매트릭스의 일 실시예를 나타낸다.

[0159] 즉, 본 발명에서 제안하는 변형 자기 유사도를 활용한 협업 필터링 추천 시스템은 변형 자기 유사도를 활용한 유사도 계산을 통해 전체 상품 중 사용자가 구매한 상품들과 유사도가 가장 높은 N개의 상품을 추천하게 된다.

[0161] 본 발명에서는 변형 자기 유사도에 상품의 재구매율과 상품에 대한 소비자의 선호도, 소비자의 반복 구매 성향, 가중치 파라미터만을 사용했지만, 변형 자기 유사도는 이 외에도 기업이나 분석가가 속한 산업 분야에서 재구매에 유의미한 영향을 미친다고 판단되는 변수를 변형 자기 유사도의 구성 요소로 추가하여 활용할 수 있다.

[0163] 이하에서는, 상품의 구매주기가 명확한 상품군이라는 요소를 추가하여 상품을 추천하는 실시예와, 소비자의 반복구매성향이라는 요소를 추가하여 상품을 추천하는 실시예에 대하여 설명하고자 한다.

[0165] 먼저, 상품의 구매주기가 명확한 상품군이라는 요소를 추가하여 상품을 추천하는 실시예를 설명한다.

[0166] 예를 들어, 생수, 기저귀 등 추천 대상 상품군이 주기가 명확한 상품군일 경우, 소비자의 상품 구매 주기에 따른 구매 확률 변수를 정의하여 변형 자기 유사도에 추가하여 활용할 수 있다.

[0167] 이를 위해, 본 발명에 따른 변형 자기유사도 산출부(200)는 구매확률 산출부(240)를 더 구비하여, 구매확률을 다음 식 4로 산출할 수 있다.

[0168] [식 4]

$$P_c(i, j) = \frac{1}{|(R_t - C_t(i, j)) - P_c(i)|}$$

[0171] 본 발명은 식 4를 이용하여, 소비자의 상품 구매 주기를 구할 수 있다.

[0172] 식 4에서 $P_c(i, j)$ 는 구매 주기 도래 정도에 따른 소비자 j 의 상품 i 에 대한 구매 확률을 의미한다. R_t 는 추천 시스템의 추천 시점을, $C_t(i, j)$ 는 소비자 j 가 상품 i 를 마지막으로 구매한 시점을 의미한다. 따라서 $(R_t - C_t(i, j))$ 의 값은 추천 시점으로부터 과거 소비자 j 가 상품 i 를 마지막으로 구매한 시점간의 거리를 의미한다.

[0173] $P_c(i)$ 의 값은 일정 기간 동안 관찰된 상품 i 의 구매 주기 최빈값을 의미한다.

[0174] 예를 들어, $R_t = 2020/07/31$, $C_t(i, j) = 2020/07/01$, $P_c(i) = 30$ 일 경우, 추천 시점에서 소비자는 마지막으로 상품 i 를 구매한지 31일이 지났으며 이는 상품의 구매 주기인 30일에 매우 가깝다. 따라서 소비자 j 가 상품 i 를 구매할 확률인 $P_c(i, j)$ 는 높은 값을 갖게 된다.

[0175] 반면에, $R_t = 2020/07/31$, $C_t(i, j) = 2020/07/30$, $P_c(i) = 30$ 일 경우, 추천 시점에서 소비자 j 가 상품 i 를 마지막으로 구매한지 하루밖에 되지 않았고, 상품의 구매 주기 최빈값은 30일이므로, 소비자 j 가 상품 i 를 재구매할 확률은 낮아지게 된다. 이를 변형 자기 유사도에 결합하면 아래 식 5와 같다.

[0177] 즉 본 발명에 따른 변형 자기유사도 산출부(200)는 식 5와 같이 구매확률을 결합하여 상품구매주기를 가진 상품에 대한 변형 자기유사도를 산출할 수 있다.

[0178] [식 5]

$$MSS(i, j) = Repurchase\ rate(i) \times Preference(i, j) \times P_c(i, j) \times Weight$$

[0181] 다음으로, 소비자의 반복구매성향이라는 요소를 추가하여 상품을 추천하는 실시예를 설명한다.

[0182] 예를 들어, 소비자의 반복 구매 성향을 변형 자기 유사도에 결합하여 활용할 수 있다.

[0183] 이를 위해, 본 발명에 따른 변형 자기유사도 산출부(200)는 반복구매선호성향 산출부(250)를 더 구비하여, 반복

구매선호성향을 다음 식 6 및 식 7로 산출할 수 있다.

[식 6]

$$Stickiness(j) = \left(1 - \frac{\sum_{i=1}^n unique\ item(i,j)}{\sum_{i=1}^n amount(i,j)}\right) \times 100$$

[식 7]

$$unique\ item(i,j) = 1\ if\ amount(i,j) \geq 1, else\ 0$$

여기서, $Stickiness(j)$ 는 소비자 j 의 반복 구매 선호 성향이며, $amount(i,j)$ 는 상품 i 에 대한 소비자 j 의 구매 수량이며, n 은 총 상품의 수이다. $unique\ item(i,j)$ 는 소비자 j 가 상품 i 를 구매했을 경우 1, 구매하지 않았을 경우 0의 값을 가진다.

소비자 j 의 반복구매 선호성향(고착도)인 $Stickiness(j)$ 는 도 6과 같이 산출될 수 있다.

$Stickiness$ 는 각 소비자가 얼마나 반복 구매 성향이 강한지를 나타낸다. 예를 들어 소비자 j 의 $Stickiness$ 가 80이라면 소비자 j 는 전체 구매 상품의 80%가 재구매의 경우에 해당한다.

한편, 변형 자기 유사도에 이와 같이 상품의 추천에 도움을 줄 수 있는 정보들을 변수로 반영하여 변수들을 결합하는 방식에는 여러 가지 방법으로서, 선형적 결합 및 비선형적 결합 등이 있을 수 있다.

본 발명에 따른 변형 자기유사도 산출부(200)는 다음 식 8과 같이 반복구매선호성향을 변형자기유사도에 결합시켜 산출할 수 있다.

[식 8]

$$MSS(i,j) = Repurchase\ rate(i) \times Preference(i,j) \times Stickiness(j) \times Weight$$

다음으로, 유사도 계산부(300)를 설명한다.

본 발명에 따른 유사도 계산부(300)는 변형 자기유사도 산출부(200)에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 삽입하여 상품간 유사도의 합을 산출할 수 있다.

유사도 계산부(300)는 상기 변형 자기유사도 산출부(200)가 산출한 변형 자기유사도(MSS)를 이용하여, 사용자가 구매한 상품의 개수(m)와 전체 상품의 개수(n)를 $m \times n$ 형태의 유사도 매트릭스로 나타내고, 상품간 유사도의 합을 산출할 수 있다(도 8 참조).

다음으로, 상품 추천부(400)를 설명한다.

본 발명에 따른 상품 추천부(400)는 상품간 유사도의 합이 높은 순서에서 기 설정된 개수의 상품을 추천하거나, 기 설정된 유사도 이상인 상품을 추천할 수 있다.

예를 들어, 상품간 유사도의 합이 가장 높은 상품을 추천할 수도 있으며, 상위 3개를 추천할 수도 있을 것이다. 또한, 유사도 값을 기준으로 일정한 유사도값 이상에 해당되는 상품을 추천할 수도 있을 것이다.

상품 추천부(400)에서 추천한 상품은 사용자 단말기에 현출되어, 사용자가 추천된 상품을 참고하여 최종 구매 선택을 할 수 있을 것이다.

[0210] 한편, 본 발명에 따른 추천시스템의 성능을 평가하기 위하여, 다른 추천시스템의 성능비교를 수행하였다.

[0212] 본 발명에 따른 반복구매상품 추천시스템과 종래의 여러 반복구매 추천시스템과의 성능 비교를 위해 피트니스 및 건강 보충제 상품을 판매하는 국내 전자상거래 기업의 판매 내역 데이터셋을 활용한 성능 평가를 진행했다. 성능 비교를 위해 사용한 국내 전자상거래 기업의 판매 데이터셋은 2018년 01월 01일부터 2020년 05월 22일까지 총 268,014건의 소비자별 상품 주문 내역으로 이루어진 데이터셋으로, 해당 기간 동안 총 36439명의 소비자가 총 1718개의 상품을 구매했다.

[0214] 여러 추천 시스템과의 성능 비교를 위해서, 아래의 5가지 추천 시스템의 성능을 비교하였다.

[0215] 1) 베스트 셀러 추천 방식

[0216] 2) 종래의 아이템(상품) 유사도 기반 협업 필터링 추천 방식

[0217] 3) 소비자별 과거 최다 구매 상품을 우선 추천하고, 소비자의 과거 구매 상품이 추천 상품의 개수보다 적을 경우 베스트 셀러를 추천하는 방식(도 9에서는 '최다 구매 상품과 베스트셀러 결합 방식'으로 표현)

[0218] 4) 소비자별 과거 최다 구매 상품을 우선 추천하고, 소비자의 과거 구매 상품이 추천 상품의 개수보다 적을 경우 소비자의 구매 상품과 유사도가 높은 상품을 추천하는 방식(도 9에서는 '최다 구매 상품과 협업 필터링의 결합 방식'으로 표현)

[0219] 5) 본 발명에 따른 변형 자기유사도를 이용한 반복구매상품 추천시스템(도 9에서는 '변형 자기 유사도 기반 추천시스템'으로 표현)

[0221] 성능 평가 지표로는 추천 시스템에서 성능 평가 지표로 가장 보편적으로 활용하는 지표인 F1을 사용했다. F1은 정확도(precision)와 재현율(recall)을 조화 평균한 값으로 계산식은 다음의 식 9와 같다.

[0222] [식 9]

$$F1 = 2 \times \frac{Recall \times Precision}{Recall + Precision}$$

[0223]

[0225] 성능 평가 결과는 도 9와 같다. 성능 비교 평가 결과 본 발명에서 제안하는 '변형 자기 유사도 기반 추천시스템'이 추천 개수에 상관없이 성능이 가장 우수한 것으로 관찰되었다.

[0226] 특히 추천 목록 개수가 적을수록 높은 예측력을 나타냈는데, 추천시스템은 소비자가 좋아할 만한 적은 수의 상품을 추천하는 것이므로, 이는 의미 있는 결과라고 할 수 있다.

[0227] 특히 본 실험에서 활용한 전자상거래 업체의 실험 기간 동안 거래 당 평균 상품 구매 수량이 2.73개인 점을 고려할 때 Top 3 추천에서 가장 성능이 좋고 다른 추천시스템과의 성능 차이가 가장 두드러진다는 것은 긍정적으로 해석할 수 있다.

[0229] 이와 같은 성능대비 평가를 통해, 본 발명에 다음과 같은 장점이 있음을 확인하였다.

[0230] 본 발명에서 제시하는 자기 유사도를 변형해 다수의 변수를 변형 자기 유사도에 결합하고, 이를 종래의 협업 필터링의 아이템 유사도와 함께 활용하여 추천에 반영하는 방법은 아래의 장점을 가진다.

[0231] 첫째, 반복 구매가 이뤄지는 상품에 대한 추천 결과의 정확성을 향상시킬 수 있다. 상기 실험 결과를 통해 알 수 있듯, 변형 자기 유사도를 활용한 추천 방법은 반복 구매가 관찰되는 상품의 추천에 있어 종래의 추천 기법

들보다 소비자에게 더 정확한 추천을 제공한다.

[0232] 둘째, 추천 대상 상품을 확장하여 소비자에게 폭 넓은 선택 기회를 제공한다. 종래의 추천 시스템은 새로운 상품만을 추천하거나, 과거 구매했던 상품만을 추천하는 방식이었다. 본 발명에서는 새로운 상품만을 추천하는 협업 필터링 기법을 소비자가 과거 이미 구매했던 상품까지 추천이 가능하도록 변형하였다. 이에 추천 상품 대상을 확장하는 효과가 있다.

[0233] 셋째, 본 발명의 변형 자기 유사도는 분석가가 추천에 유의미하다고 판단되는 다수의 변수를 결합 및 분리할 수 있도록 고안되어, 분석가의 역량과 선택에 따라 추가적으로 추천시스템의 추천 성능을 제고할 수 있으며, 상황과 업체별 특징에 따라 여러 변수를 결합 및 분리하여 유연하게 활용할 수 있다.

[0235] 한편, 본 발명은 변형 자기유사도를 이용한 반복구매상품 추천방법으로 구현될 수 있다. 추천방법은 추천시스템과 실질적으로 동일한 기술구성을 가지므로, 중복되는 설명은 생략하고 요지 위주로 설명하고자 한다.

[0237] 본 발명은 사용자 단말기, 데이터베이스 및 웹서버가 네트워크로 연결되고, 컴퓨터에 의해 실행되는 상품추천방법으로서, 사용자정보 수신부(100)는 사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신하는 S100 단계; 변형 자기유사도 산출부(200)가 상기 사용자정보 수신부(100)에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출하는 S200 단계; 유사도 계산부(300)가 상기 변형 자기유사도 산출부(200)에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 산입하여 상품간 유사도의 합을 산출하는 S300 단계; 및 상품 추천부(400)가 상기 유사도 계산부(300)에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송하는 S400 단계를 포함하는 것을 특징으로 하는 변형 자기유사도를 이용한 반복구매상품 추천방법이다

[0239] 본 발명에 따른 S200 단계는 변형 자기유사도 산출부(200)가 재구매율을 산출하는 S210 단계, 선호도를 산출하는 S220 단계 및 가중치를 부여하는 S230 단계를 포함하며, 다음 식 1로 변형 자기유사도를 산출할 수 있다.

[0240] [식 1]

$$[0241] \quad MSS(i, j) = Repurchase\ rate(i) \times Preference(i, j) \times Weight$$

[0243] 본 발명에 따른 S210 단계에서, 변형 자기유사도 산출부(200)에 구비된 재구매율 산출부(210)가 상기 재구매율을 다음 식 2로 산출할 수 있다.

[0244] [식 2]

$$[0245] \quad Repurchase\ rate(i) = 1 - \frac{Users(i, 1)}{\sum_{k=1}^{max} Users(i, k)}$$

[0247] 본 발명에 따른 S220 단계에서, 변형 자기유사도 산출부(200)에 구비된 선호도 산출부(220)가 선호도를 다음 식 3으로 산출할 수 있다.

[0248] [식 3]

$$[0249] \quad Preference(i, j) = \frac{amount(i, j)}{\sum_{i=1}^n amount(i, j)}$$

- [0252] 한편, 본 발명은 컴퓨터로 해독 가능한 기록매체로 구현될 수 있다.
- [0253] 본 발명에 따른 기록매체는 컴퓨터에 의해 실행되며, 사용자정보 수신부(100)는 사용자 단말기가 웹 서버에 접속되면, 데이터베이스로부터 사용자 정보를 수신하는 S100 단계; 변형 자기유사도 산출부(200)가 상기 사용자정보 수신부(100)에서 수신된 사용자 정보를 이용하여 변형 자기유사도(MSS)를 산출하는 S200 단계; 유사도 계산부(300)가 상기 변형 자기유사도 산출부(200)에서 산출된 변형 자기유사도를 상품간 유사도 계산 매트릭스에 삽입하여 상품간 유사도의 합을 산출하는 S300 단계; 및 상품 추천부(400)가 상기 유사도 계산부(300)에서 산출된 상품간 유사도의 합을 정렬하고, 기 설정된 기준을 충족하는 상품 정보를 사용자 단말기로 전송하는 S400 단계를 실행시키기 위한 프로그램을 기록한 컴퓨터 해독 가능한 기록매체로 구현될 수 있다.
- [0255] 또한, 본 발명은 컴퓨터 프로그램으로 구현될 수 있다.
- [0256] 본 발명에 따른 컴퓨터 프로그램은 하드웨어와 결합되어, 본 발명에 따른 변형 자기유사도를 이용한 반복구매상품 추천방법을 컴퓨터에 의해 실행시키기 위하여 컴퓨터가 판독 가능한 기록매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다.
- [0258] 도 4는 본 발명의 실시예에 따른 컴퓨팅 장치를 나타내는 도면이다. 도 4의 컴퓨팅 장치(TN100)는 본 명세서에서 기술된 사용자 맞춤형 추천정보 제공장치일 수 있다.
- [0259] 도 4의 실시예에서, 컴퓨팅 장치(TN100)는 적어도 하나의 프로세서(TN110), 송수신 장치(TN120), 및 메모리(TN130)를 포함할 수 있다. 또한, 컴퓨팅 장치(TN100)는 저장 장치(TN140), 입력 인터페이스 장치(TN150), 출력 인터페이스 장치(TN160) 등을 더 포함할 수 있다. 컴퓨팅 장치(TN100)에 포함된 구성 요소들은 버스(bus)(TN170)에 의해 연결되어 서로 통신을 수행할 수 있다.
- [0260] 프로세서(TN110)는 메모리(TN130) 및 저장 장치(TN140) 중에서 적어도 하나에 저장된 프로그램 명령(program command)을 실행할 수 있다. 프로세서(TN110)는 중앙 처리 장치(CPU: central processing unit), 그래픽 처리 장치(GPU: graphics processing unit), 또는 본 발명의 실시예에 따른 방법들이 수행되는 전용의 프로세서를 의미할 수 있다. 프로세서(TN110)는 본 발명의 실시예와 관련하여 기술된 절차, 기능, 및 방법 등을 구현하도록 구성될 수 있다. 프로세서(TN110)는 컴퓨팅 장치(TN100)의 각 구성 요소를 제어할 수 있다.
- [0261] 메모리(TN130) 및 저장 장치(TN140) 각각은 프로세서(TN110)의 동작과 관련된 다양한 정보를 저장할 수 있다. 메모리(TN130) 및 저장 장치(TN140) 각각은 휘발성 저장 매체 및 비휘발성 저장 매체 중에서 적어도 하나로 구성될 수 있다. 예를 들어, 메모리(TN130)는 읽기 전용 메모리(ROM: read only memory) 및 랜덤 액세스 메모리(RAM: random access memory) 중에서 적어도 하나로 구성될 수 있다.
- [0262] 송수신 장치(TN120)는 유선 신호 또는 무선 신호를 송신 또는 수신할 수 있다. 송수신 장치(TN120)는 네트워크에 연결되어 통신을 수행할 수 있다.
- [0264] 한편, 앞서 설명된 본 발명의 실시예에 따른 방법들은 다양한 컴퓨터수단을 통하여 판독 가능한 프로그램 형태로 구현되어 컴퓨터로 판독 가능한 기록매체에 기록될 수 있다. 여기서, 기록매체는 프로그램 명령, 데이터 파일, 데이터구조 등을 단독으로 또는 조합하여 포함할 수 있다. 기록매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 예컨대 기록매체는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CDRom, DVD와 같은 광 기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치를 포함한다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어를 포함할 수 있다. 이러한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0266] 본 명세서에서 설명되는 실시예와 첨부된 도면은 본 발명에 포함되는 기술적 사상의 일부를 예시적으로 설명하

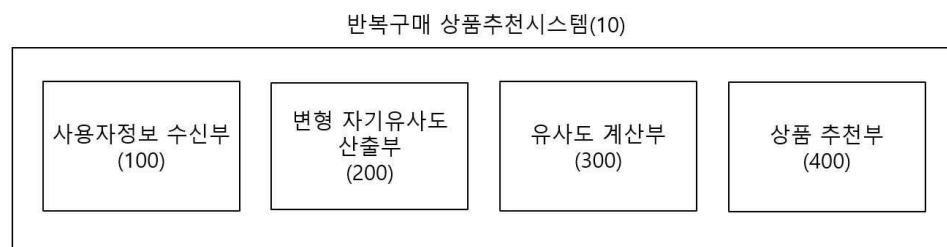
는 것에 불과하다. 따라서, 본 명세서에 개시된 실시예들은 본 발명의 기술적 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이므로, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아님은 자명하다. 본 발명의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형예와 구체적인 실시 예는 모두 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

10 : 반복구매상품 추천시스템
100 : 사용자정보 수신부
200 : 변형 자기유사도 산출부
210 : 재구매율 산출부
220 : 선호도 산출부
230 : 가중치 부여부
240 : 구매확률 산출부
250 : 반복구매선호성향 산출부
300 : 유사도 계산부
400 : 상품 추천부

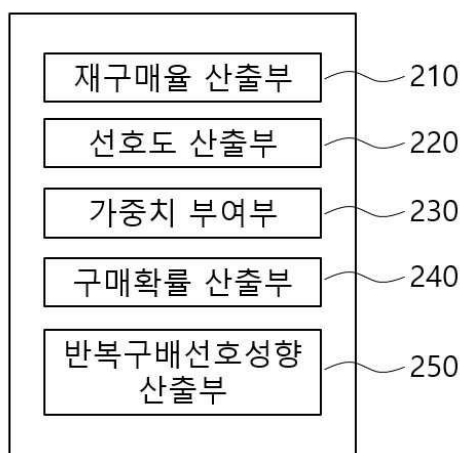
도면

도면1

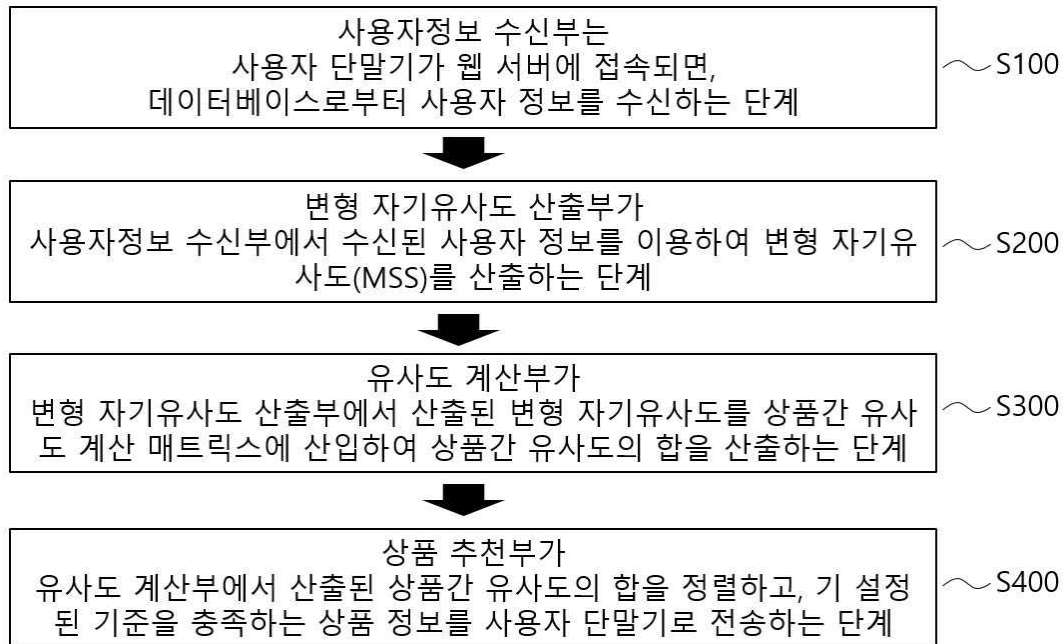


도면2

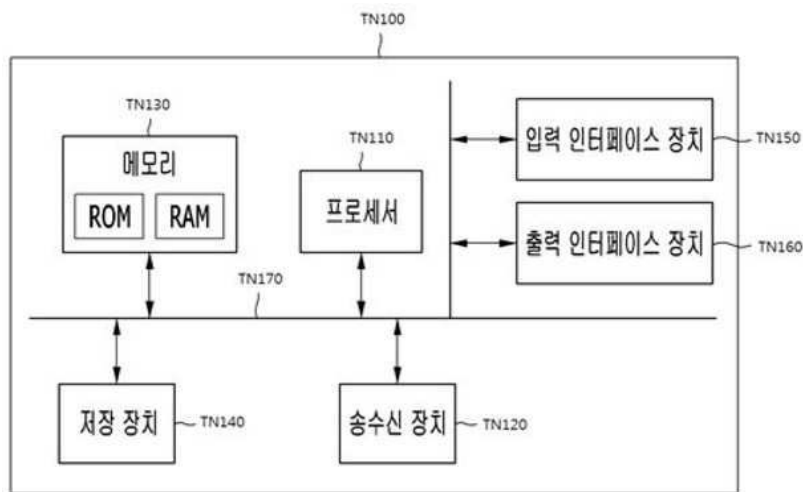
변형자기유사도산출부(200)



도면3



도면4



도면5

(n - 4) 개						
	Item 1	Item 2	Item 3	Item 4	...	Item n
사용자가 구매한 아이템 (4개)	Item 4					
	Item 57					
	Item 96					
	Item 200					

도면6

		전체 Item (n개)					
사용자가 구매 Item (4개)		Item 1	Item 2	Item 3	Item 4	...	Item n
	Item 4						
	Item 57						
	Item 96						
	Item 200						

도면7

		전체 Item (n개)						
사용자 구매 Item (2개)		Item 1	...	Item 4	...	Item 57	...	Item n
	Item 4			1				
	Item 57					1		
	유사도합	0.xx	0.xx	1.xx	0.xx	1.xx	0.xx	0.xx

도면8

		전체 Item (n개)				
사용자 j의 구매 Item		Item 1	...	Item i	...	Item n
	...					
	Item i			MSS (i, j)		
	...					
	Item m					
	유사도 합					

도면9

	Top 3 추천	Top 5 추천	Top 7 추천	Top 10 추천
베스트셀러 추천	0.1066	0.1155	0.1104	0.1030
아이템 유사도 기반 협업 필터링	0.0554	0.0607	0.0609	0.0590
최다 구매 상품과 베스 트셀러 결합 방식	0.2194	0.2098	0.1969	0.1776
최다 구매 상품 추천과 협업 필터링 결합 방식	0.0665	0.0748	0.0780	0.0795
변형 자기 유사도 기반 추천시스템	0.2294	0.2180	0.2034	0.1840