(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년07월10일
(11) 등록번호 10-1877161
(24) 등록일자 2018년07월04일

(51) 국제특허분류(Int. Cl.)
G06Q 30/02 (2012.01) *G06F 17/18* (2006.01)
G06F 17/27 (2006.01) *G06N 3/04* (2006.01)
(52) CPC특허분류
G06Q 30/0254 (2013.01)
G06F 17/18 (2013.01)
(21) 출원번호 10-2017-0002890
(22) 출원일자 2017년01월09일
심사청구일자 2017년01월09일
(56) 선행기술조사문헌
KR1020040099092 A*
KR1020120047079 A*
KR1020130118597 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
(72) 발명자
유환조
경상북도 포항시 남구 지곡로 155, 9동 102호
오진오
전라남도 해남군 산이면 상공길 46-9
(뒷면에 계속)
(74) 대리인
특허법인이상

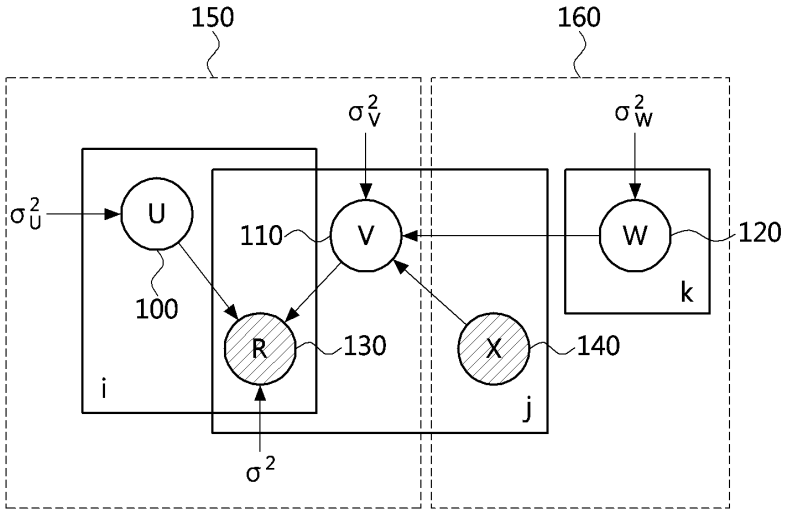
전체 청구항 수 : 총 12 항 심사관 : 이충근

(54) 발명의 명칭 **문서 문맥정보를 고려하는 상황기반 추천 방법 및 장치**

(57) 요 약

프로세서에 의해 수행되는 상황기반 추천 방법이 개시된다. 사용자 학습변수와 품목 문서정보 관측변수 및 문맥 학습변수가 반영된 품목 학습변수로부터 사용자-품목 평점정보 관측변수를 획득하는 단계; 사용자-품목 평점정보 관측변수를 통해 산출된 평점 예측도를 기반으로 추천 서비스를 제공하는 단계를 포함한다. 따라서, 품목의 문서 문맥정보를 고려하여 보다 신뢰할 수 있는 추천 서비스가 제공될 수 있다.

대 표 도 - 도1



(52) CPC특허분류
G06F 17/27 (2013.01)
G06N 3/0472 (2013.01)
G06Q 30/0255 (2013.01)
G06Q 30/0269 (2013.01)
(72) 발명자
박찬영
경기도 성남시 분당구 예원로6번길 52, 101호

김동현
경상남도 진주시 진주대로 789, 103동 1202호

이 발명을 지원한 국가연구개발사업
과제고유번호 B0101160307000 1002
부처명 미래창조과학부
연구관리전문기관 정보통신기술진흥센터
연구사업명 산업원천기술개발사업(정보통신)
연구과제명 인간 수준의 평생 기계학습 SW 기초 연구 (기계학습연구센터)
기 여 율 1/3
주관기관 포항공과대학교 산학협력단
연구기간 2016.03.01 ~ 2017.02.28
이 발명을 지원한 국가연구개발사업
과제고유번호 2016913993
부처명 미래창조과학부
연구관리전문기관 서울대학교 산학협력단
연구사업명 차세대정보컴퓨팅기술개발사업
연구과제명 비정형 빅 데이터 분석 및 개인화 기술 개발
기 여 율 1/3
주관기관 포항공과대학교 산학협력단
연구기간 2016.07.01 ~ 2017.06.30
이 발명을 지원한 국가연구개발사업
과제고유번호 10049079
부처명 산업통상자원부
연구관리전문기관 경희대학교산학협력단
연구사업명 핵심기반기술개발사업
연구과제명 퍼스널 빅데이터를 활용한 마이닝마인즈 핵심기술개발
기 여 율 1/3
주관기관 포항공과대학교 산학협력단
연구기간 2016.06.01 ~ 2017.05.31
공지예외적용 : 있음

명 세 서

청구범위

청구항 1

프로세서에 의해 수행되는 상황기반 추천 방법으로서,

상품의 정보가 포함된 문서의 정보를 지시하는 문서정보 관측변수 및 상기 문서의 문맥정보 파악을 위해 사용되는 문맥 학습변수가 반영된 품목 학습변수를 획득하는 단계;

상품을 구매한 사용자의 정보를 지시하는 사용자 학습변수와 상기 품목 학습변수로부터 상기 사용자와 품목 간의 평점정보 관계를 지시하는 사용자-품목 평점정보 관측변수를 획득하는 단계;

상기 사용자-품목 평점정보 관측변수를 통해 각 상품 별 예상 평가 점수를 지시하는 평점 예측도를 산출하는 단계; 및

상기 산출된 평점 예측도를 기반으로 추천 서비스를 제공하는 단계를 포함하고,

상기 품목 학습변수를 획득하는 단계는,

상기 문서정보 관측변수 및 상기 문맥 학습변수를 포함하는 인공신경망이 원본 문서를 숫자 행렬로 변형하는 단계;

상기 숫자 행렬에서 문맥특징을 추출하는 단계;

상기 문맥특징 중 대표 데이터를 획득하는 단계; 및

상기 대표 데이터를 이용해 품목 학습변수를 획득하는 단계를 포함하는, 상황기반 추천 방법.

청구항 2

청구항 1에 있어서,

상기 품목 학습변수는 상기 문서정보 관측변수 및 상기 문맥 학습변수를 기초로 일 방향 함수를 지시하는 피드포워드(feed-forward) 기법을 통해 획득되는, 상황기반 추천 방법.

청구항 3

청구항 1에 있어서,

상기 품목 학습변수는 상기 문맥 학습변수로부터 아웃풋의 차이를 비교하는 오차역전파법(back propagation algorithm)을 통해 미리 정해진 기준보다 낮은 오류값을 갖도록 설정되는, 상황기반 추천 방법.

청구항 4

청구항 1에 있어서,

상기 추천 서비스는 상기 산출된 평점 예측도가 미리 설정된 임계값을 초과하면 제공되는, 상황기반 추천 방법.

청구항 5

청구항 1에 있어서,

상기 사용자 학습변수 및 상기 품목 학습변수 각각은 최대 사후 확률을 통해 획득된 손실함수(optimization function)를 최소화 하는 좌표감소법(coordinate descent)을 이용하여 획득되는, 상황기반 추천 방법.

청구항 6

삭제

청구항 7

청구항 1에 있어서,

상기 평점 예측도는 상기 사용자 학습변수와 상기 품목 학습변수에 기초하여 생성된 확률 그래프 모델을 통해 산출되는, 상황기반 추천 방법.

청구항 8

상황기반 추천 장치로서,

프로세서(processor); 및

상기 프로세서를 통해 실행되는 적어도 하나의 명령이 저장된 메모리(memory)를 포함하고,

상기 적어도 하나의 명령은,

상품의 정보가 포함된 문서의 정보를 지시하는 문서정보 관측변수 및 상기 문서의 문맥정보 파악을 위해 사용되는 문맥 학습변수가 반영된 품목 학습변수를 획득하고,

상품을 구매한 사용자의 정보를 지시하는 사용자 학습변수와 상기 품목 학습변수로부터 상기 사용자와 품목 간의 평점정보 관계를 지시하는 사용자-품목 평점정보 관측변수를 획득하고,

상기 사용자-품목 평점정보 관측변수를 통해 각 상품 별 예상 평가 점수를 지시하는 평점 예측도를 산출하고, 그리고

상기 산출된 평점 예측도를 기반으로 추천 서비스를 제공하도록 실행되고,

상기 품목 학습변수를 획득하는 경우,

상기 적어도 하나의 명령은,

상기 문서정보 관측변수 및 상기 문맥 학습변수를 포함하는 인공신경망이 원본 문서를 숫자 행렬로 변형하고,

상기 숫자 행렬에서 문맥특징을 추출하고,

상기 문맥특징 중 대표 데이터를 획득하고, 그리고

상기 대표 데이터를 이용해 품목 학습변수를 획득하도록 실행되는, 상황기반 추천 장치.

청구항 9

청구항 8에 있어서,

상기 품목 학습변수는 상기 문서정보 관측변수 및 상기 문맥 학습변수를 기초로 일 방향 함수를 지시하는 피드 포워드(feed-forward) 기법을 통해 획득되는, 상황기반 추천 장치.

청구항 10

청구항 8에 있어서,

상기 품목 학습변수는 상기 문맥 학습변수로부터 아웃풋의 차이를 비교하는 오차역전파법(back propagation algorithm)을 통해 미리 정해진 기준보다 낮은 오류값을 갖도록 설정되는, 상황기반 추천 장치.

청구항 11

청구항 8에 있어서,

상기 추천 서비스는 상기 산출된 평점 예측도가 미리 설정된 임계값을 초과하면 제공되는, 상황기반 추천 장치.

청구항 12

청구항 8에 있어서,

상기 사용자 학습변수 및 상기 품목 학습변수 각각은 최대 사후 확률을 통해 획득된 손실함수(optimization function)를 최소화 하는 좌표감소법(coordinate descent)을 이용하여 획득되는, 상황기반 추천 장치.

청구항 13

삭제

청구항 14

청구항 8에 있어서,

상기 평점 예측도는 상기 사용자 학습변수와 상기 품목 학습변수에 기초하여 생성된 확률 그래프 모델을 통해 산출되는, 상황기반 추천 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 상황기반 추천 방법 및 장치에 관한 것으로, 더욱 상세하게는 문서 문맥정보를 고려하는 상황기반 추천 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 모바일 디바이스(Mobile Device)의 다양성과 성능의 발전은 모바일 디바이스 사용자 수를 급증하게 하였다. 이러한 모바일 클라우드 컴퓨팅 환경에서는 사용자의 정보를 이용하여 사용자에게 적절한 자원과 서비스를 제공할 수 있는 상황기반 추천의 중요성이 더욱 부각되고 있다.

[0003] 상황기반 추천은 수많은 품목 중 사용자가 구매할 가능성이 높은 품목, 즉 사용자가 높은 평점을 줄 가능성이 높은 품목을 사용자에게 빈번하게 노출시키는 것을 의미한다.

[0004] 상황기반 추천은 사용자가 수많은 품목 중에서 사용자에게 적합한 품목을 검색하기 위한 시간을 절약하도록 도울 수 있다. 상황기반 추천은 주로 전자 상거래 분야에서 적용될 수 있다.

[0005] 상황기반 추천을 제공하기 위해서는 서비스를 제공하는 시스템들이 사용자가 처해있는 상황에 대한 인식과 함께 사용자의 현재 상황, 행동, 선호도 등과 같은 사용자 프로파일(User Profile)을 분석한다. 상황기반 추천은 대체로 평점데이터를 이용하여 사용자와 품목 간의 유사도를 도출한다.

[0006] 품목에 의한 평점 데이터의 희소성은 점점 증가하고 있고, 사용자는 점점 증가하고 있다. 이러한 증가는 상황기반 추천의 평점 예측 성능을 저해하는 주 요소가 되고 있다.

[0007] 희소성 문제를 해결하기 위해 평점 데이터 이외에 사용자와 품목과 관련 있는 부가적인 데이터를 이용하는 방법론들 중 품목 문서정보를 이용하는 방법론이 주목을 받았으나, 평점 예측이 정확하지 않은 문제가 있었다.

발명의 내용

해결하려는 과제

[0008] 상기와 같은 문제점을 해결하기 위한 본 발명의 목적은 상황기반 추천에서 평점 데이터가 희소할 때 나선형 인공신경망을 통해 문서의 문맥정보를 효과적으로 추출하고, 이를 통해 사용자에게 알맞은 품목을 추천하는 방법을 제공하는데 있다.

과제의 해결 수단

[0009] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 프로세서에 의해 수행되는 상황기반 추천 방법은, 상품의 정보가 포함된 문서의 정보를 지시하는 문서정보 관측변수 및 상기 문서의 문맥정보 파악을 위해 사용되는 문맥 학습변수가 반영된 품목 학습변수를 획득하는 단계; 상품을 구매한 사용자의 정보를 지시하는 사용자 학습변수와 상기 품목 학습변수로부터 상기 사용자와 품목 간의 평점정보 관계를 지시하는 사용자-품목 평점정보 관측변수를 획득하는 단계; 상기 사용자-품목 평점정보 관측변수를 통해 각 상품 별 예상 평가 점수를 지시하는 평점 예측도를 산출하는 단계; 및 상기 산출된 평점 예측도를 기반으로 추천 서비스를 제공하는 단계를 포함할 수 있다.

[0010] 여기서, 상기 품목 학습변수는 상기 문서정보 관측변수 및 상기 문맥 학습변수를 기초로 일 방향 함수를 지시하

는 피드포워드(feed-forward) 기법을 통해 획득될 수 있다.

[0011] 여기서, 상기 품목 학습변수는 상기 문맥 학습변수로부터 아웃풋의 차이를 비교하는 오차역전파법(back propagation algorithm)을 통해 미리 정해진 기준보다 낮은 오류값을 갖도록 설정될 수 있다.

[0012] 여기서, 상기 추천 서비스는 상기 산출된 평점 예측도가 미리 설정된 임계값을 초과하면 제공될 수 있다.

[0013] 여기서, 상기 사용자 학습변수 및 상기 품목 학습변수 각각은 최대 사후 확률을 통해 획득된 손실함수(optimization function)를 최소화 하는 좌표감소법(coordinate descent)을 이용하여 획득될 수 있다.

[0014] 여기서, 상기 품목 학습변수를 획득하는 단계는, 상기 문서정보 관측변수 및 상기 문맥 학습변수를 포함하는 인공신경망이 원본 문서를 숫자 행렬로 변형하는 단계; 상기 밀집한 숫자 행렬에서 문맥특징을 추출하는 단계; 상기 문맥특징 중 대표 데이터를 획득하는 단계; 및 상기 대표 데이터를 이용해 품목 학습변수를 획득하는 단계를 포함할 수 있다.

[0015] 여기서, 상기 평점 예측도는 상기 사용자 학습변수와 상기 품목 학습변수에 기초하여 생성된 확률 그래프 모델을 통해 산출될 수 있다.

[0016] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 상황기반 추천 장치는, 프로세서(processor); 및 상기 프로세서를 통해 실행되는 적어도 하나의 명령이 저장된 메모리(memory)를 포함하고, 상기 적어도 하나의 명령은, 상품의 정보가 포함된 문서의 정보를 지시하는 문서정보 관측변수 및 상기 문서의 문맥정보 파악을 위해 사용되는 문맥 학습변수가 반영된 품목 학습변수를 획득하고, 상품을 구매한 사용자의 정보를 지시하는 사용자 학습변수와 상기 품목 학습변수로부터 상기 사용자와 품목 간의 평점정보 관계를 지시하는 사용자-품목 평점정보 관측변수를 획득하고, 상기 사용자-품목 평점정보 관측변수를 통해 각 상품 별 예상 평가 점수를 지시하는 평점 예측도를 산출하고, 그리고 상기 산출된 평점 예측도를 기반으로 추천 서비스를 제공하도록 실행 가능하다.

[0017] 여기서, 상기 품목 학습변수는 상기 문서정보 관측변수 및 상기 문맥 학습변수를 기초로 일 방향 함수를 지시하는 피드포워드(feed-forward) 기법을 통해 획득되도록 실행 가능할 수 있다.

[0018] 여기서, 상기 품목 학습변수는 상기 문맥 학습변수로부터 아웃풋의 차이를 비교하는 오차역전파법(back propagation algorithm)을 통해 미리 정해진 기준보다 낮은 오류값을 갖도록 설정되도록 실행 가능할 수 있다.

[0019] 여기서, 상기 추천 서비스는 상기 산출된 평점 예측도가 미리 설정된 임계값을 초과하면 제공되도록 실행 가능할 수 있다.

[0020] 여기서, 상기 사용자 학습변수 및 상기 품목 학습변수 각각은 최대 사후 확률을 통해 획득된 손실함수(optimization function)를 최소화 하는 좌표감소법(coordinate descent)을 이용하여 획득되도록 실행 가능할 수 있다.

[0021] 여기서, 상기 품목 학습변수를 획득하는 경우, 상기 적어도 하나의 명령은, 상기 문서정보 관측변수 및 상기 문맥 학습변수를 포함하는 인공신경망이 원본 문서를 숫자 행렬로 변형하고, 상기 밀집한 숫자 행렬에서 문맥특징을 추출하고, 상기 문맥특징 중 대표 데이터를 획득하고, 그리고 상기 대표 데이터를 이용해 품목 학습변수를 획득하도록 실행되도록 실행 가능할 수 있다.

[0022] 여기서, 상기 평점 예측도는 상기 사용자 학습변수와 상기 품목 학습변수에 기초하여 생성된 확률 그래프 모델을 통해 산출되도록 실행 가능할 수 있다.

발명의 효과

[0023] 본 발명에 의하면, 평점 데이터가 희소할 때 나선형 인공신경망을 통해 문서의 문맥정보를 효과적으로 추출할 수 있으며, 이를 통해 문서 의미 해석에 특화된 상황기반 추천을 제공할 수 있다. 또한, 문서 의미 해석에 특화된 상황기반 추천을 제공함으로써 사용자에게 더 나은 수준의 신뢰할만한 추천 서비스를 제공할 수 있다.

도면의 간단한 설명

[0024] 도 1은 행렬분해기법이 통합된 나선형 인공신경망의 그래픽 모델을 도시한 개념도이다.

도 2는 상황기반 추천 장치의 일 실시예를 도시한 블록도이다.

도 3은 상황기반 추천 장치의 동작 방법을 도시한 흐름도이다.

도 4는 행렬분해기법이 통합된 나선형 인공신경망(ConvMF)의 아키텍처를 도시한 개념도이다.

도 5는 각 데이터 베이스에 포함된 품목 평점 수의 왜곡도를 도시한 그래프이다.

도 6은 품목 평점정보를 포함하는 문서의 개수를 누적적으로 도시한 그래프이다.

도 7은 아마존 데이터세트에서 임베딩 단어의 차원 크기 효과를 도시한 그래프이다.

도 8은 ConvMF에 대비한 ConvMF+의 개선도를 도시한 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0025] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0026] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0027] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.

[0028] 본 출원에서 사용한 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, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 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.

[0029] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가진 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0030] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다. 본 발명을 설명함에 있어 전체적인 이해를 용이하게 하기 위하여 도면상의 동일한 구성요소에 대해서는 동일한 참조부호를 사용하고 동일한 구성요소에 대해서 중복된 설명은 생략한다.

[0032] 도 1은 행렬분해기법이 통합된 나선형 인공신경망의 그래픽 모델을 도시한 개념도이다.

[0033] 도 1을 참조하면, 나선형 인공신경망(convolutional neural network; CNN)을 확률행렬분해기법(probabilistic matrix factorization; PMF)에 통합하는 나선형 행렬분해기법(convolutional matrix factorization; ConvMF)의 확률모델(또는 상황기반 추천 장치)가 확인될 수 있다.

[0034] 상황기반 추천 장치는 사용자정보(user; i)를 포함하는 사용자 잠재 모델(또는 사용자 학습변수)(U)(100), 품목정보(items; j)를 포함하는 품목 잠재 모델(또는 품목 학습변수)(V)(110), 문맥의 가중치를 반영하는 공유 가중치(또는 문맥 학습변수)(W)(120), 사용자-품목 간의 평점에 관한 정보를 포함하는 평점정보 관측변수(또는 평점정보)(R)(130), 문서정보를 포함하는 품목 문서정보 세트(또는 문서정보 관측변

수)(X)(140)를 포함할 수 있다.

[0035] 왼쪽 점선은 확률행렬분해기법(150)을 의미할 수 있다. 오른쪽 점선은 나선형 인공신경망(160)의 동작 과정을 의미할 수 있다.

[0036] 상황기반 추천 장치는 다음과 같이 구성될 수 있다.

[0038] 도 2는 상황기반 추천 장치의 일 실시예를 도시한 블록도이다.

[0039] 도 2를 참조하면, 상황기반 추천 장치(200)는 적어도 하나의 프로세서(210) 및 메모리(220)를 포함할 수 있다. 또한, 상황기반 추천 장치(200)는 입력 인터페이스 장치(240), 출력 인터페이스 장치(150), 저장 장치(160) 등을 더 포함할 수 있다. 상황기반 추천 장치(200)에 포함된 각각의 구성들은 버스(bus)(270)에 의해 연결되어 서로 통신을 수행할 수 있다.

[0040] 프로세서(210)는 메모리(220) 및/또는 저장 장치(260)에 저장된 프로그램 명령을 실행할 수 있다. 프로세서(210)는 중앙 처리 장치(central processing unit, CPU), 그래픽 처리 장치(graphics processing unit, GPU), 본 발명에 따른 방법들이 수행되는 전용의 프로세서를 의미할 수 있다. 메모리(220)와 저장 장치(260)는 휘발성 저장 매체 및/또는 비휘발성 저장 매체로 구성될 수 있다. 예를 들어, 메모리(220)는 읽기 전용 메모리(read only memory, ROM) 및/또는 랜덤 액세스 메모리(random access memory, RAM)로 구성될 수 있다.

[0041] 상황기반 추천 장치의 동작방법은 다음과 같이 표현될 수 있다.

[0043] 도 3은 상황기반 추천 장치의 동작 방법을 도시한 흐름도이다.

[0044] 도 3을 참조하면, 상황기반 추천 장치는 원본 문서로부터 문서정보 관측변수를 획득할 수 있다(S300). 상황기반 추천 장치는 원본 문서로부터 문서의 문맥 정보를 통해 문서의 공유 가중치(또는 문맥 학습변수)를 획득할 수 있다(S310).

[0045] 상황기반 추천 장치는 나선형 인공신경망(160)을 통해 문맥 학습변수(120)와 문서정보 관측 변수(140)로부터 품목 학습변수(110)를 획득할 수 있다(S320).

[0046] 품목 학습변수(110)는 문서정보 관측변수(140) 및 문맥 학습변수(120)를 기초로 일 방향 함수를 지시하는 피드포워드(feed-forward) 기법을 통해 획득될 수 있다. 상황기반 추천 장치는 사용자 정보로부터 사용자 학습변수를 획득할 수 있다(S330).

[0047] 상황기반 추천 장치는 확률행렬분해기법(150) 내에서 피드포워드 기법을 통해 획득된 품목 학습변수(110)와 사용자 학습변수(100)로부터 평점정보 관측변수(130)를 획득할 수 있다(S340).

[0048] 상황기반 추천 장치는 N명의 사용자와, M개의 품목 정보를 포함할 수 있다. 평점정보 관측변수(130)가 $R \in \mathbb{R}^{N \times M}$ 이라고 가정할 수 있다.

[0050] 다시 도 1을 참조하면, 상황기반 추천 장치는 사용자 잠재 모델(100)($U \in \mathbb{R}^{k \times N}$)과 품목 잠재 모델(110)($V \in \mathbb{R}^{k \times M}$)들의 결과물($U^T V$)을 평점 행렬인 R로 재구성할 수 있다.

[0051] 확률론적 관점에서 관찰된 평점에 대한 조건부 분포는 아래와 같은 수학식으로 표현될 수 있다.

수학식 1

[0053]
$$p(R|U,V,\sigma^2) = \prod_{i}^{N} \prod_{j}^{M} N(r_{ij}|u_i^T v_j, \sigma^2)^{I_{ij}}$$

[0055] 여기서 $N(x|\mu,\sigma^2)$ 은 평균(μ) 및 분산(σ^2)을 가지는 가우시안(Gaussian) 정규 분포의 확률 밀도 함수일 수 있다. I_{ij} 는 지시자일 수 있다.

[0056] 사용자 잠재 모델(100)에 대한 생성 모델로서, 제로 평균 구형 가우시안은 분산(σ_U^2)을 가지는 상황기반 추천 장치의 사용자 잠재 모델(100)에 배치될 수 있으며, 사용자 잠재 모델(100)은 다음과 같이 표현될 수 있다.

수학식 2

$$p(U|\sigma_U^2) = \prod_i^N N(u_i|0,\sigma_U^2 I)$$

[0058]

[0060] 기존 PMF의 잠재 모델에 대한 확률 모델과는 달리 상황기반 추천 장치는 품목 잠재 모델(110)을 아래와 같은 세 가지 변수를 이용하여 생성할 수 있다.

[0061] 1) CNN에서의 공유 가중치(120) W , 2) 품목 j 의 문서를 나타내는 X_j, 3) 가우스 노이즈(noise)로서 엡실론(epsilon) 변수

[0062] 위 변수들은 평점에 대한 품목 잠재 모델(110)을 더욱 최적화할 수 있다. 따라서, 최종 품목 잠재 모델(110)은 다음 수학식으로 표현될 수 있다.

수학식 3

$$v_j = cnn(W, X_j) + \epsilon_j$$
$$\epsilon_j \sim N(0, \sigma_V^2 I)$$

[0064]

[0066] 상황기반 추천 장치는 W 의 각각의 공유 가중치(120) w_k 에 대해, 제로 평균 구형 가우스를 배치할 수 있으며, 이는 다음과 같은 수학식으로 표현될 수 있다.

수학식 4

$$p(W|\sigma_W^2) = \prod_k N(w_k|0,\sigma_W^2)$$

[0068]

[0070] 따라서, 품목 잠재 모델(110)에 대한 조건부 분포는 다음과 같은 수학식으로 표현될 수 있다.

수학식 5

$$p(V|W,X,\sigma_V^2) = \prod_j^M N(v_j|cnn(W,X_j),\sigma_V^2 I)$$

[0072]

[0074] 여기서 X 는 품목 문서정보 세트일 수 있다. 상황기반 추천 장치는 CNN 모델에서 얻은 문서 잠재 벡터를 가우시안 분포 평균으로 사용할 수 있다. 상황기반 추천 장치는 품목의 가우시안 노이즈를 PMF와 CNN의 관계 설정에 있어서, 중요한 역할을 하는 가우시안 분포의 분산으로 사용할 수 있다.

[0075] CNN과 PMF는 문서정보 및 평점을 완전하게 분석하는데 있어서 도움이 될 수 있다. 상황기반 추천 장치의 동작 방법을 도시한 흐름도는 다음과 같을 수 있다.

[0076] 행렬분해기법이 통합된 나선형 인공신공망의 아키텍처는 다음과 같이 표현될 수 있다.

[0078] 도 4는 행렬분해기법이 통합된 나선형 인공신경망(ConvMF)의 아키텍처를 도시한 개념도이다.

[0079] 도 4를 참조하면, 상황 기반 추천 장치에 포함되는 4개의 계층으로 구성된 CNN 아키텍처가 확인될 수 있다. 4개의 계층은 임베딩 계층(embedding layer)(400), 콘볼루션 계층(convolution layer)(410), 풀링 계층(pooling layer)(420), 출력 계층(output layer)(430)을 포함할 수 있다.

[0080] 상황기반 추천 장치는 품목 잠재 모델(110)을 작성하는데 사용되는 원본 문서(440)에서 문서 잠재 벡터를 CNN 아키텍처를 통해 생성할 수 있다. 엡실론 변수는 CNN 아키텍처를 통해 사용될 수 있다.

[0081] 임베딩 계층(400)은 원본 문서(440)를 다음 컨볼루션 계층의 문서를 나타내는 밀집된 숫자 행렬로 변환시킬 수 있다. 구체적으로, 원본 문서(440)를 l 단어의 시퀀스로 간주할 때, 원본 문서(440)에 기재된 단어의 단어 벡터를 연결함으로써, 문서가 행렬로 표현될 수 있다.

[0082] 단어 벡터는 사전 학습된 단어 임베딩 모델을 사용하여 초기화 되거나, 무작위로 초기화 될 수 있다. 단어 벡터는 최적화 과정을 통해 추가로 학습될 수 있다.

[0083] 그러면 문서 행렬 $D \in \mathbb{R}^{p \times l}$ 는 다음과 같이 될 수 있다.

수학식 6

$$D = \begin{bmatrix} | & | & | & | & | \\ \cdots & w_{i-1} & w_i & w_{i+1} & \cdots \\ | & | & | & | & | \end{bmatrix}$$

[0085]

[0087] 여기서, l 은 원본 문서(440)의 길이일 수 있고, p 는 각 단어 w_i 에 대한 임베딩 차원의 크기일 수 있다.

[0088] 콘볼루션 계층(410)은 문맥특징을 추출할 수 있다. 문맥특징인 $c_i^j \in \mathbb{R}$ 는 j 번째 공유 가중치(120) $W_c^j \in \mathbb{R}^{p \times ws}$ 에 의해 추출될 수 있다. 이 공유 가중치(120)의 윈도우 사이즈 WS는 주변 단어의 수

를 결정할 수 있다. 문맥특징은 다음과 같은 수학식으로 표현될 수 있다.

수학식 7

[0090] $$c_i^j = f(W_c^j * D_{(:,i:(i+ws-1))} + b_c^j)$$

[0092] 여기서 $*$ 은 콘볼루션 연산자이고, $b_c^j \in \mathbb{R}$ 는 W_c^j 에 대한 바이어스이고, f 는 비선형 활성화 함수일 수 있다. 수정된 선형 유닛(rectified linear unit; ReLU)을 통해, 느린 최적화 컨버전스(convergence)를 야기하는 소멸 기울기의 문제가 방지될 수 있다. 또한 수정된 선형 유닛을 통해 극한의 지역 최소값이 유도될 수 있다.

[0093] W_c^j 를 포함하는 원본 문서(440)의 문맥특징 벡터 $c^j \in \mathbb{R}^{l-ws+1}$ 는 다음과 같은 수학식으로 표현될 수 있다.

수학식 8

[0095] $$c^j = [\, c_1^j,\ c_2^j,\ \ldots,\ c_i^j,\ \ldots,\ c_{l-ws+1}^j]$$

[0097] 하나의 공유 가중치(120)는 문맥특징의 한 종류를 포착할 수 있다. 따라서, 여러 유형의 문맥특징을 포착하기 위해 다중 공유 가중치(120)를 사용할 수 있다. 문맥특징은 W_c 의 숫자 n_c 만큼의 문맥특징 벡터를 생성하는데 이용될 수 있다. 즉, W_c^j 에서 $j = 1, 2, ..., n_c$ 일 수 있다.

[0098] 풀링 계층(420)은 콘볼루션 계층(410)에서 대표적인 특징을 추출하고, 고정 길이 특성 벡터(fixed-length feature vector)를 구성하는 풀링 연산을 통해 가변 길이의 원본 문서(440)를 처리할 수 있다.

[0099] 콘볼루션 계층(410) 이후에 문서는 문맥특징 벡터로서 표현될 수 있다. 여기서 각각의 문맥특징 벡터는 가변 길이(즉, $l - ws + 1$ 문맥 특징)를 가질 수 있다. 대표 값은 성능 향상에 도움이 되지 않는 정보(예를 들어, c_i)를 포함할 수 있다. 문맥특징 벡터의 길이가 가변이기 때문에 다음과 같은 계층을 구성하기 어려울 수 있다.

[0100] 따라서, 최대 풀링(max-pooling)은 다음과 같이 각 문맥 특징 벡터로부터 최대 문맥 특징만을 추출함으로써, 문서의 표현을 n_c 고정 길이 벡터로 줄일 수 있다. 문서의 표현을 n_c 고정 길이 벡터로 표현한 수학식은 다음과 같을 수 있다.

수학식 9

[0102] $$d_f = [\max(c^1), \max(c^2), \ldots, \max(c^j), \ldots, \max(c^{n_c})]$$

[0104] 여기서 c^j 는 j 번째 공유 가중치(120)(W_c^j)에 의해 추출된 길이 ($l - ws + 1$) 의 문맥특징 벡

터일 수 있다.

[0106] 출력 계층(430)은 풀링 계층(420)에서 얻은 상위 레벨의 특징들을 특정 작업에 맞게 변환할 수 있다. 추천 작업을 위해 사용자 잠재 모델(100) 및 품목 잠재 모델(110)의 K차원 공간에 d_f가 투영될 수 있다. 추천 작업은 기존의 비선형 투영 통해 문서 잠재 벡터를 생성하는 것을 의미할 수 있다.

수학식 10

$$s = \tanh(W_{f2}\{\tanh(W_{f1}d_f + b_{f1})\} + b_{f2})$$

[0108]

[0110] 위 수학식에서 $W_{f1} \in \mathbb{R}^{f \times n_c}, W_{f2} \in \mathbb{R}^{k \times f}$는 투영 행렬일 수 있으며, $b_{f1} \in \mathbb{R}^f, b_{f2} \in \mathbb{R}^k$는 W_{f1}, W_{f2}(단, $s \in \mathbb{R}^k$)를 위한 바이어스 벡터일 수 있다.

[0112] 위 프로세스를 통해, CNN 아키텍처는 원본 문서를 입력으로 받고, 각 문서의 잠재 벡터를 출력으로 반환하는 함수가 될 수 있다.

수학식 11

$$s_j = cnn(W, X_j)$$

[0114]

[0116] 위 수학식에서 W는 혼란을 방지하기 위한 모든 공유 가중치(120) 및 바이어스 변수를 나타낼 수 있다. X_j는 품목 j의 원본 문서(440)를 나타내며, s_j는 품목 j의 잠재 벡터를 나타낼 수 있다.

[0118] 다시 도 1을 참조하면, 사용자 학습변수, 품목 학습변수, 문맥 학습변수가 최적화 되는 과정이 확인될 수 있다. 최적화 과정에서 사용되는 최대 사후 확률(maximum a posteriori; MAP) 추정 기법은 다음과 같이 표현될 수 있다.

수학식 12

$$\max_{U,V,W} p(U,V,W|R,X,\sigma^2,\sigma_U^2,\sigma_V^2,\sigma_W^2)$$
$$= \max_{U,V,W} [p(R|U,V,\sigma^2)p(U|\sigma_U^2)p(V|W,X,\sigma_V^2)p(W|\sigma_W^2)]$$

[0120]

[0122] 수학식 12의 정리를 통해 다음과 같은 수학식이 표현될 수 있다.

수학식 13

$$\mathcal{L}(U,V,W) = \sum_{i}^{N}\sum_{j}^{M} \frac{I_{ij}}{2}(r_{ij} - u_i^T v_j)_2 + \frac{\lambda_U}{2}\sum_{i}^{N}\|u_i\|_2 + \frac{\lambda_V}{2}\sum_{j}^{M}\|v_j - cnn(W, X_j)\|_2 + \frac{\lambda_W}{2}\sum_{k}^{|w_k|}\|w_k\|_2,$$

[0124]

[0126] 수학식 13에서 λ_U 는 σ^2/σ_U^2 이고, λ_V 는 σ^2/σ_V^2 이고, λ_W는 σ^2/σ_W^2일 수 있다. 상황기반 추천 장치은 학습변수들의 최적화를 위해 나머지 변수들을 고정하면서 반복적으로 학습변수들을 최적화하는 좌표감소법(coordinate descent)을 사용할 수 있다.

[0127] 수학식 13은 W 와 V (또는 U)를 일정하다고 가정한 U (또는 V)에 대해 2차 함수일 수 있다. 이 경우, 상황기반 추천 장치는 최적화된 U (또는 V)을 다음과 같은 u_i (또는 v_j)에 대한 최적화 함수 $\mathcal{L}$ 에 대한 미분을 통해 닫힌 형태 내에서 분석적으로 계산할 수 있다.

수학식 14

$$u_i \leftarrow (VI_iV^T + \lambda_U I_K)^{-1} VR_i$$
$$v_j \leftarrow (UI_jU^T + \lambda_V I_K)^{-1}(UR_j + \lambda_V cnn(W, X_j))$$

[0129]

[0131] 수학식 14에서 I_i 은 I_{ij} 에서 직교 요소로서 $j = 1, \ldots, M$ 인 직교 행렬일 수 있다. R_i는 사용자 i 에 대해 $(r_{ij})_{j=1}^{M}$ 인 벡터일 수 있다. 품목 j 에 대해서, I_j 및 R_j는 각각 I_i 및 R_i로서 유사하게 정의될 수 있다.

[0132] 수학식 14를 통해 품목 잠재 모델(110) v_j 를 생성할 때, CNN의 문서 잠재 벡터의 영향이 확인될 수 있다. 여기서 λ_V 는 균형 매개변수일 수 있다.

[0133] W 는 최대 풀링 계층 및 비선형 활성화 함수와 같이 CNN 아키텍처의 특징과 밀접하게 관련되어 있기 때문에 U 와 V 처럼 분석 솔루션으로 최적화될 수 없다.

[0134] 그럼에도 불구하고, $\mathcal{L}$ 은 U 와 V 가 일시적으로 일정할 때, 다음과 같이 L_2 정규화 항을 갖는 제곱오차 함수로 해석될 수 있다.

수학식 15

$$\mathcal{E}(W) = \frac{\lambda_V}{2} \sum_{j}^{M} \|(v_j - cnn(W, X_j))\|^2 + \frac{\lambda_W}{2} \sum_{k}^{|w_k|} \|w_k\|^2 + \text{constant}$$

[0136]

[0138] 상황기반 추천 장치는 W 의 최적화를 위해, 오차역전파법(back propagation)을 사용할 수 있다. 상황기반 추천 장치는 전체 최적화 프로세스(U , V 및 W 가 번갈아가면서 갱신 되는 프로세스)가 수렴될 때까지 오차역전파법을 반복할 수 있다.

[0139] 상황기반 추천 장치는 최적화된 U , V 및 W 를 통하여 품목에 대한 사용자의 비공개된 평가를 예측할 수 있다. 비공개된 평가의 예측은 다음과 같은 수학식으로 표현될 수 있다.

수학식 16

$$r_{ij} \approx \mathbb{E}[r_{ij}|u_i^T v_j, \sigma^2] = u_i^T v_j = u_i^T (cnn(W, X_j) + \epsilon_j)$$

[0141]

수학식 17

$$v_j = cnn(W, X_j) + \epsilon_j$$

[0143]

[0145] 상황기반 추천 장치를 이용한 실시예를 얻기 위해 시간 복잡도가 분석될 수 있다. 모든 사용자 잠재 모델(100) 및 품목 잠재 모델(110)은 $O(k^2 n_R + k^3 N + k^3 M)$ 으로 갱신 될 수 있다.

[0146] 상황기반 추천 장치는 각 기간에서 W 의 갱신 동안 문서 잠재 벡터를 미리 계산할 수 있다. 상황기반 추천 장치는 CNN의 모든 공유 가중치(120) 및 바이어스 변수 각각을 $O(n_c \cdot p \cdot l \cdot M)$ 에서 갱신할 수 있다. n_R 은 관측 된 평가의 숫자일 수 있다.

[0147] 결과적으로, 기간 당 전체 시간 복잡도는 $O(k^2 n_R + k^3 N + k^3 M + n_c \cdot p \cdot l \cdot M)$ 일 수 있고, 상황기반 추천 장치는 이 최적화 프로세스를 주어진 데이터의 크기에 따라 선형적으로 조정할 수 있다.

[0148] 상황기반 추천 장치는 등급 예측 측면에서 본 발명의 효과를 입증하기 위해 무비렌즈(movie-lens) 데이터세트 2가지와 아마존 데이터세트 한가지, 즉, 총 세 가지 실제 데이터세트가 이용할 수 있다.

[0149] 무비렌즈 데이터세트에는 100만개의 평점정보(130)로 구성된 무비렌즈-1m 데이터세트와 1000만개의 평점정보(130)로 구성된 무비렌즈-10m 데이터세트가 존재할 수 있다.

[0150] 세 가지 데이터세트들은 품목에 대한 사용자의 명시적 등급이 1에서 5까지의 범위로 구성되어 있을 수 있다. 아마존 데이터세트에는 품목 설명 문서와 같은 품목 리뷰(review)가 포함될 수 있다.

[0151] 무비렌즈 데이터세트에는 품목 설명 문서가 포함되어 있지 않을 수 있다. 따라서 www.imdb.com(이하 'IMDB'라 한다.)에서 해당 품목에 대한 문서(즉, 플롯 요약)를 얻을 수 있다.

[0152] 모든 데이터세트에 대한 설명 문서는 다음과 같이 사전 처리될 수 있다. 상황기반 추천 장치는 1) 원본 문서의 최대 길이를 300으로 설정하고, 2) 중지 단어를 제거하고, 3) 각 단어에 대해 tf-idf를 계산하고, 4) 문서 빈도가 0.5보다 높은 코퍼스(corpus) 특정 중지 단어를 제거하고, 5) 탑 8000개의 상이한 단어를 어휘로서 선택하고, 6) 원본 문서에서 어휘가 아닌 단어는 삭제할 수 있다.

[0153] 결과적으로, 문서 당 평균 단어 수는 무비렌즈-1m 데이터세트가 97.09개, 무비렌즈-10m 데이터세트가 92.05개, 아마존 인스턴트 비디오(이하 '아마존 데이터세트'라 한다)(amazon instant video; AIV)가 91.50개일 수 있다.

[0154] 각 데이터세트에 설명 문건이 없는 항목은 삭제될 수 있고, 아마존 데이터세트의 경우 등급이 3 미만인 사용자는 삭제될 수 있다. 아마존 데이터세트는 무비렌즈 데이터세트와 비교할 때 자료가 상대적으로 희소할 수 있다. 서로 다른 특징을 갖는 세 개의 데이터세트는 다음과 같은 표 1로 나타낼 수 있다.

표 1

[0156]

데이터세트	# 사용자정보	# 품목정보	# 평가정보	밀집도
ML-1m	6,040	3,544	993,482	4.641%
ML-10m	69,878	10,073	9,945,875	1.413%
AIV	29,757	15,149	135,188	0.030%

[0158] 밀집도(density)는 사용자정보(user;i)와 품목정보(items;j)의 곱으로 평가정보(ratings)를 나눈 값일 수 있다. 예를 들어, 무비렌즈-1m 데이터세트의 밀집도는 993,482/(6,040*3,544)=4.641%일 수 있다.

[0159] 무비렌즈-10m 데이터세트의 밀집도는 9,945,875/(69,878*10,073)=1.413%일 수 있다. 아마존 데이터세트의 밀집도는 135,188/(29,757*15,149)=0.030%일 수 있다.

[0160] ConvMF의 두 버전(ConvMF, ConvMF+)을 다음과 같은 기준으로 비교할 수 있다. PMF은 평점정보(130)만 사용하는 평가 예측 모델일 수 있다.

[0161] CTR(Collaborative Topic Regression) 및 CDL(Collaborative Deep Learning) 은 문서정보와 평점정보(130)만 사용하는 평가 예측 모델일 수 있다.

[0162] ConvMF는 본 발명에서 제안된 모델로서, 문서정보, 평점정보(130), 문맥정보를 모두 사용하는 평가 예측 모델일 수 있다. ConvMF는 무작위로 형성된 단어 데이터세트를 사용한 평가 예측 모델일 수 있다.

[0163] ConvMF+(Convolutional Matrix Factorization with pre-trained word embedding model)은 본 발명에서 제안된 또다른 모델로서, 문서정보, 평점정보(130), 문맥정보를 모두 사용하는 평가 예측 모델일 수 있다. ConvMF+는 사전 학습된 단어 임베딩 모델을 이용한 ConvMF를 사용한 평가 예측 모델일 수 있다. ConvMF+는 단어의 의미를 반영한 초기값을 설정할 수 있다.

[0165] 상황기반 추천 장치는 U, V의 잠재 크기를 50으로 설정하고 U, V를 0에서 1까지 임의로 초기화할 수 있다. 아래 표 2에서는 검색에 의해 발견된 각 모델의 공통 매개변수(λ_U, λ_V)의 최상의 수행값이 나타날 수 있다.

표 2

[0167]

모델	ML-1m		ML-10m		AIV	
	λ_U	λ_V	λ_U	λ_V	λ_U	λ_V
PMF	0.01	10000	10	100	0.1	0.1
CTR	100	1	10	100	10	0.1
CDL	10	100	100	10	0.1	100
ConvMF	100	10	10	100	1	100
ConvMF+	100	10	10	100	1	100

[0169] 명시적인 데이터세트를 이용하기 때문에, r_{ij} 가 관찰되면 CTR과 CDL의 정밀도 매개 변수는 1로 설정되고 그렇지 않으면 0으로 설정될 수 있다. 구현 세부 사항은 다음과 같을 수 있다.

[0170] 1) 상황기반 추천 장치는 문서의 최대 길이를 300으로 설정할 수 있다, 2-1) 상황기반 추천 장치는 ConvMF에서 단어 잠재 벡터를 무작위로 차원 크기 200을 기준으로 초기화할 수 있다.

[0171] 2-2) 상황기반 추천 장치는 ConvMF+에서 단어 잼재 벡터를 차원 크기가 200인 미리 학습된 단어 삽입 모델을 기준으로 초기화 할 수 있다. ConvMF+의 단어 잠재 벡터는 최적화 과정을 통해 획득될 수 있다.

[0172] 3) 상황기반 추천 장치는 콘볼루션 계층에서 공유 가중치(120)에 따라 다양한 윈도우 크기(3, 4, 5)를 사용하여 다양한 길이의 주변 단어를 고려할 수 있다. 본 실시예에서는 윈도우 크기 당 100개의 공유 가중치(120)가 사용될 수 있다.

[0173] 4) 상황기반 추천 장치는 CNN의 공유 가중치(120)와 관련된 L_2 정규 표현식 대신 CNN이 과적합이 되지 않도록 드롭 아웃을 사용할 수 있다. 드롭 아웃의 비율은 0.2로 설정될 수 있다.

[0175] 상황기반 추천 장치는 실제 데이터세트에서 각 모델의 전반적인 성능을 평가하기 위해, 각 데이터세트를 학습세트(80%), 유효성 검사세트(10%), 테스트세트(10%)로 무작위로 분리할 수 있다.

[0176] 학습세트에는 PMF가 모든 사용자 및 품목을 처리 할 수 있도록 모든 사용자 및 품목에 대한 적어도 하나의 평가가 포함되어 있을 수 있다. 상황기반 추천 장치는 평가 척도로는 RMSE(root mean squared error)를 사용할 수 있다. RMSE는 다음과 같은 수학식으로 표현될 수 있다.

수학식 18

[0178]

$$\text{RMSE} = \sqrt{\frac{\sum_{i,j}^{N,M}(r_{ij} - \hat{r}_{ij})^2}{\#\text{ of ratings}}}$$

표 3

모델	데이터세트		
	ML-1m	ML-10m	AIV
PMF	0.8971	0.8311	1.4118
CTR	0.8969	0.8275	1.5496
CDL	0.8879	0.8186	1.3594
ConvMF	**0.8531**	**0.7958**	**1.1337**
ConvMF+	**0.8549**	**0.7930**	**1.1279**
개선도	3.92%	2.79%	16.60%

[0180]

[0182] 표 3을 통해 무비렌즈 데이터세트 및 아마존 데이터세트의 정량 결과에 관한 정보가 확인 될 수 있다. 구체적으로 표 3을 통해 ConvMF, ConvMF+, PMF, CTR, CDL의 전반적인 평가 예측 오류가 확인될 수 있다.

[0183] 표 3은 Improve는 최고의 경쟁 업체(PMF, CTR, CDL 중 하나)와 비교할 때 ConvMF의 상대적 개선도를 나타내는 지표일 수 있다. 표 3을 통해 세가지 모델(PMF, CTR, CDL)에 비해 ConvMF 및 ConvMF+는 모든 데이터세트에서 개선된 점이 있는 것이 확인될 수 있다.

[0184] 상대적으로 밀도가 높은 데이터세트인 무비렌즈-1m 데이터세트의 경우, 최고의 경쟁 업체인 CDL에 비해 ConvMF의 성능 개선의 정도가 3.92%인 것이 확인될 수 있다. 무비렌즈-10m 데이터세트의 경우, 최고의 경쟁 업체인 CDL에 비해 ConvMF의 성능 개선의 정도가 2.79%인 것이 확인될 수 있다.

[0185] PMF는 CTR, CDL과 비교할 때, 무비렌즈 데이터세트-10m에서 성능의 차이가 미미한 것이 확인될 수 있다. 이를 통해 사용자 잠재 모델(100)과 품목 잠재 모델(110)이 충분한 환경에서 문맥정보를 고려하지 못하는 문서 분석은 정확한 잠재 모델을 생성하는데 도움이 되지 않는 것이 확인될 수 있다.

[0186] PMF와 ConvMF를 비교할 때, 사용자 잠재 모델(100)과 품목 잠재 모델(110)이 충분한 환경에서도 문맥정보를 고려하는 문서 분석은 잠재 모델을 보다 정확하게 조정하는데 도움이 되는 점이 확인될 수 있다.

[0187] 매우 희소한 데이터세트를 포함하는 아마존 데이터세트에서 ConvMF는 최고의 경쟁 업체인 CDL과 비교할 때, 16.60%의 성능 개선이 있는 것이 확인될 수 있다.

[0188] 무비렌즈 데이터세트에서보다 아마존 데이터세트에서 ConvMF의 성능 개선의 정도가 큰 것은 ConvMF는 희박하고 왜곡된 데이터로부터도 문서를 효과적으로 분석하여 정확한 품목 잠재 모델(110)을 구성하는 것을 나타낼 수 있다.

표 4

모델	전체 데이터세트에 대한 학습 세트의 비율(밀집도)						
	20% (0.93%)	30% (1.39%)	40% (1.86%)	50% (2.32%)	60% (2.78%)	70% (3.25%)	80% (3.71%)
PMF	1.0168	0.9711	0.9497	0.9354	0.9197	0.9083	0.8971
CTR	1.0124	0.9685	0.9481	0.9337	0.9194	0.9089	0.8969
CDL	1.0044	0.9639	0.9377	0.9211	0.9068	0.8970	0.8879
ConvMF	**0.9745**	**0.9330**	**0.9063**	**0.8897**	**0.8726**	**0.8676**	**0.8531**
개선도	2.98%	3.20%	3.36%	3.41%	3.77%	3.27%	3.92%

[0190]

[0192] 상황기반 추천 장치는 무비렌즈-1m 데이터세트를 임의로 샘플링하고, 각 경쟁 업체와 비교된 ConvMF의 성능을 희소성의 차이가 있는 7가지 추가 데이터세트에서 실험을 할 수 있다. 실험 결과 값은 표 4에서 확인될 수 있다.

[0193] 표 4에서 ConvMF는 3가지 다른 경쟁 업체보다 성능 개선의 정도가 큰 것이 확인될 수 있다. 구체적으로 데이터 밀도가 0.93%에서 3.71%로 증가하면서, ConvMF는 최고의 경쟁 업체(CDL) 대비 2.98%에서 3.92%의 지속적인 성능 개선이 확인될 수 있다.

[0194]

표 5

[0195]

W_c^{11}에 의해 포착된 문장	최대 (c^{11})	W_c^{86} 에 의해 포착된 문장	최대 (c^{86})
people **trust** the man	*0.0704*	betray his **trust** finally	*0.1009*
W_c^{11}를 위한 테스트 문장	$c^{11}_{테스트}$	W_c^{86}를 위한 테스트 문장	$c^{86}_{테스트}$
people **believe** the man	*0.0391*	betray his **believe** finally	*0.0682*
people **faith** the man	*0.0374*	betray his **faith** finally	*0.0693*
people **tomas** the man	*0.0054*	betray his **tomas** finally	*0.0480*

[0197] ConvMF의 성능은 공유 가중치(120) W_c 에 의해 추출된 문맥 특징에 의해 영향을 받을 수 있다. 각 문맥특징은 문구와 관련이 있을 수 있다. 구체적으로, 표 5에 'people trust the man' 기재된 'trust'의 의미는 '신뢰하다'라는 동사로서 사용될 수 있다. 이 때의 공유 가중치(120)는 0.0704 일 수 있다.

[0198] 세 가지 문장 'people believe the man' 'people faith the man' 'people tomas the man'을 테스트 해본 결과 'believe'의 공유 가중치(120)는 0.0391일 수 있고, 'faith'의 공유 가중치(120)는 0.0371일 수 있고, 'tomas'의 공유 가중치(120)는 0.0054일 수 있다.

[0199] 'believe'는 문장에서 동사로 사용되었고, '신뢰하다'와 가장 유사한 '믿다'라는 의미로서 'trust'의 공유 가중치(120)인 0.0704와 가장 유사한 공유 가중치(120) 값 0.0391을 가질 수 있다.

[0200] 'faith'는 문장에서 명사로 사용되었고, '신뢰하다'와 유사한 '신념'이라는 의미로서 'trust'의 공유 가중치(120)인 0.0704와 유사한 공유 가중치(120) 값 0.0374을 가질 수 있다.

[0201] 'tomas'는 문장에서 명사로 사용되었고, '신뢰하다'와 차이가 있는 '토마스'라는 의미로서 'trust'의 공유 가중치(120)인 0.0704와 가장 상이한 공유 가중치(120) 값 0.0054를 가질 수 있다.

[0202] 위와 같은 비교 실시예는 ConvMF가 용어 "신뢰"의 미묘한 문맥상의 차이점을 구별한다는 것을 의미할 수 있다.

[0204] 도 5는 각 데이터 베이스에 포함된 품목 평점 수의 왜곡도를 도시한 그래프이다.

[0205] 도 5를 참조하면, 그래프 세로축은 평점정보(130) 개수를 최대 평점정보(130) 개수로 나눈 값을 의미할 수 있다. 그래프 가로축은 1을 기준으로 결정된 데이터 개수의 누진적 합을 의미할 수 있다.

[0206] 즉, 도 5는 데이터를 평점정보(130) 개수로 정규화한 그래프일 수 있다. 무비렌즈-1m 데이터세트의 경우 밀집도의 변화와 관계 없이 일정한 정규화 그래프가 생성될 수 있다.

[0207] 무비렌즈-1m(sparse) 데이터세트는 10-20만개의 평점정보(130)로 구성될 수 있고, 무비렌즈-1m(dense) 데이터세트는 100만개의 평점정보(130)로 구성될 수 있고, 무비렌즈-10m 데이터세트는 1000만개의 평점정보(130)로 구성될 수 있다. 아마존 데이터세트는 10만개의 평가로 구성될 수 있다.

[0208] 아마존 데이터세트(AIV)의 경우 자료가 희소하기 때문에 데이터의 왜곡도가 클 수 있다. 무비렌즈-10m은 무비렌

즈-1m과 비교할 때 품목대비 상대적으로 많은 평점정보(130)를 포함하기 때문에 왜곡도가 상대적으로 클 수 있다.

[0209] 무비렌즈-1m(sparse) 데이터세트와 무비렌즈-1m(dense) 데이터세트는 데이터 희소성에 차이가 있더라도 데이터의 왜곡도의 차이가 크지 않을 수 있다. 도 5에서 무비렌즈-1m(sparse) 데이터세트와 무비렌즈-1m(dense) 데이터세트의 그래프가 거의 일치하는 것이 확인될 수 있다.

[0211] 도 6은 품목 평점정보를 포함하는 문서의 개수를 누적적으로 도시한 그래프이다.

[0212] 도 6을 참조하면, 그래프 세로축은 가로축 평정점보의 개수보다 낮은 개수의 평점정보(130)를 포함하는 품목의 비율을 1을 기준으로 정규화한 값을 의미할 수 있다. 그래프 가로축은 평가정보 개수를 의미할 수 있다.

[0213] 구체적으로, 아마존 데이터세트가 그래프에서 지시하는 (2, 0.5)는 아마존 데이터세트에서 평점정보(130)의 개수가 2보다 낮은(즉, 평점정보(130)의 개수가 1인) 품목의 비율이 50%인 것을 의미할 수 있다.

[0214] 아마존 데이터세트의 경우, 평점정보(130)의 개수가 3개 미만인 데이터가 전체의 80%일 수 있다. 따라서, 아마존 데이터세트는 평점정보(130)의 희소성 때문에 정보의 왜곡도가 상대적으로 클 수 있다.

[0215] 무비렌즈-1m(sparse) 데이터세트, 무비렌즈-1m(dense) 데이터세트, 무비렌즈-10m 데이터세트의 경우 평점정보(130)의 개수가 3개 미만인 데이터가 전체의 10% 미만일 수 있다. 따라서, 무비렌즈 데이터세트들의 경우 정보의 왜곡도가 상대적으로 작을 수 있다.

[0217] 도 7은 아마존 데이터세트에서 임베딩 단어의 차원 크기 효과를 도시한 그래프이다.

[0218] 도 7을 참조하면, 세로축은 제곱근 평균오차(RMSE)를 의미할 수 있다. 가로축은 단어 임베딩 차원 크기를 의미할 수 있다. ConvMF의 경우 랜덤하게 획득된 데이터세트를 포함하고 있기 때문에, 단어 임베딩 차원 크기에 따른 성능의 차이가 크지 않을 수 있다.

[0219] 그러나, ConvMF+는 사전 학습된 단어 임베딩 모델을 이용한 ConvMF를 사용한 평가 예측 모델로서, 단어의 의미를 반영한 초기값을 설정하기 때문에 단어가 많아질수록 많은 정보가 포함되어 성능이 좋아질 수 있다.

[0220] 그래프에서 ConvMF의 경우 단어 임베딩 차원 크기에 따른 제곱근 평균오차의 차이가 크지 않은 것이 확인될 수 있다. 그래프에서 ConvMF+의 경우 임베팅 차원 크기가 100에서 300으로 증가할수록 제곱근 평균오차가 약 1.155에서 약 1.125으로 줄어드는 것이 확인될 수 있다.

[0221] 무비렌즈-1m 데이터세트, 무비렌즈-10m 데이터세트의 경우 평점정보(130)가 아마존 데이터세트보다 상대적으로 풍부하지만, 아마존 데이터세트는 평점정보(130)의 희소성 때문에 정보의 왜곡도가 상대적으로 클 수 있다. 따라서, 단어의 의미를 반영한 초기값 설정으로 인한 성능 차이가 클 수 있다.

[0223] 도 8은 ConvMF에 대비한 ConvMF+의 개선도를 도시한 그래프이다.

[0224] 도 8을 참조하면, 아마존 데이터세트, 무비렌즈-1m 데이터세트, 무비렌즈-1m 데이터세트 세 가지 경우에 문서정보를 고정시키고, 평점 매개변수(λ_V)를 가변하면서 관찰된 그래프가 확인될 수 있다.

[0225] 그래프 세로축은 상대적 개선도를 의미할 수 있다. 그래프 가로축은 무비렌즈 데이터세트의 경우와, 아마존 데이터세트의 경우에 품목의 수를 의미할 수 있다.

[0226] 무비렌즈-1m 데이터세트, 무비렌즈-10m 데이터세트의 경우 평점정보(130)가 아마존 데이터세트보다 상대적으로 풍부하기 때문에, 평점정보(130)가 많아지더라도 성능 개선의 효과가 크지 않을 수 있다.

[0227] 아마존 데이터세트의 경우, 초기 단계에서는 평점정보(130)가 많아질수록 성능이 개선되는 것이 확인될 수 있다. 다만, 문서정보와 비교할 때, 평점정보(130)가 너무 많아지면 성능 개선이 오히려 떨어지는 것을 확인 할 수 있다.

[0228] 즉, 그래프에서 문서정보와 평점정보(130)가 균형을 이룰 때 성능 개선의 효과가 가장 큰 것이 확인될 수 있다.

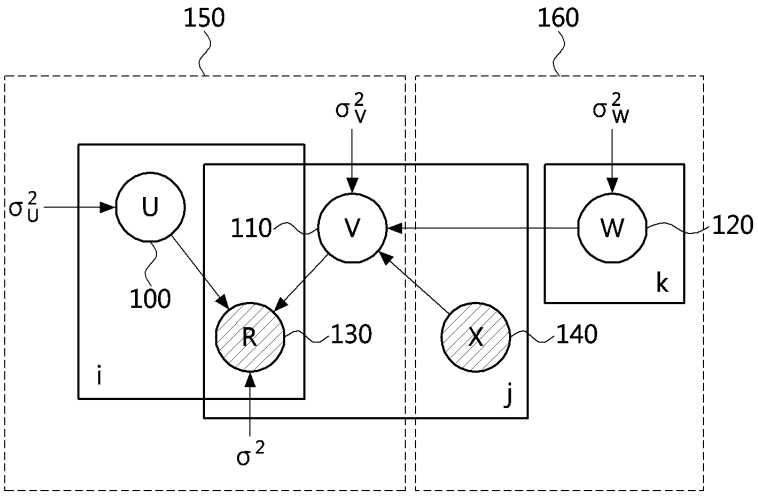
[0230] 본 발명에 따른 방법들은 다양한 컴퓨터 수단을 통해 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은 본 발명을 위해 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다.

[0231] 컴퓨터 판독 가능 매체의 예에는 롬, 램, 플래시 메모리(flash memory) 등과 같이 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러(compiler)에 1의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터(interpreter) 등을 사용해서 컴퓨터에 의해 실행될 수 있는 고급 언어 코드를 포함한다. 상술한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 적어도 하나의 소프트웨어 모듈로 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

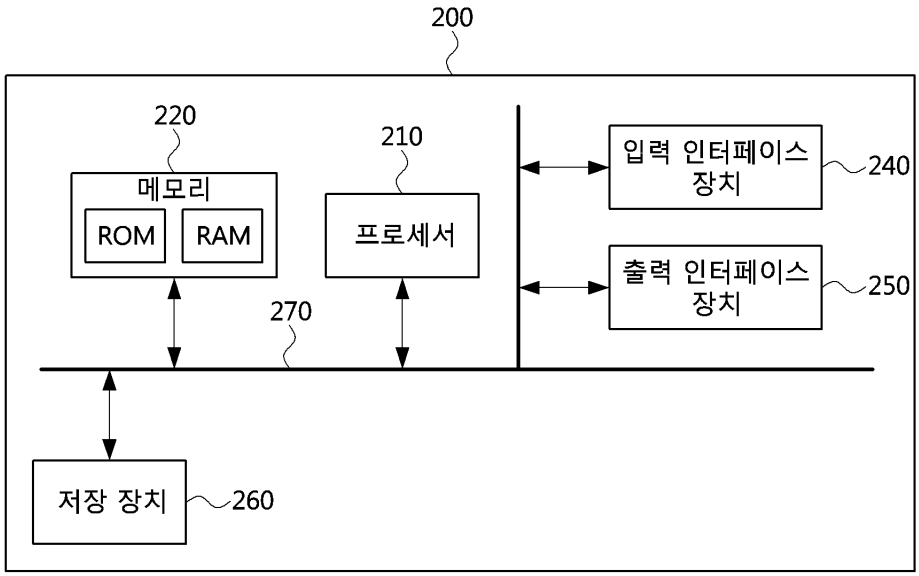
[0232] 이상 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

도면

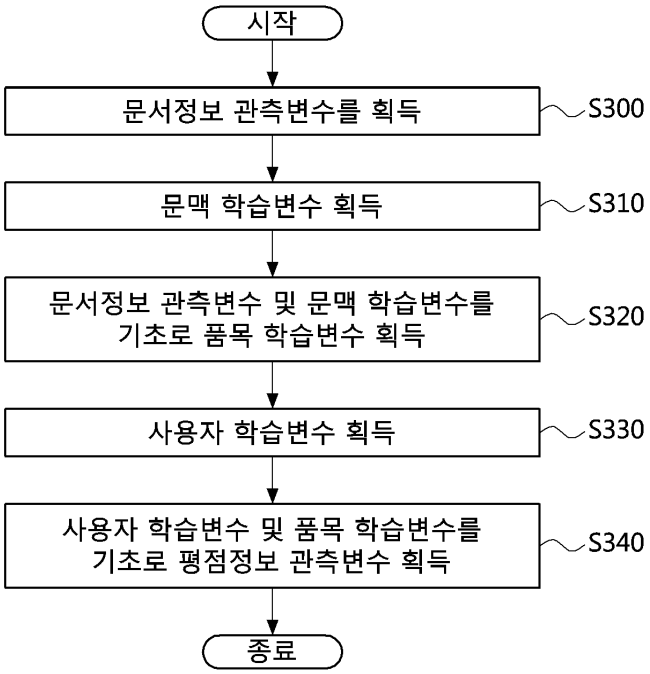
도면1



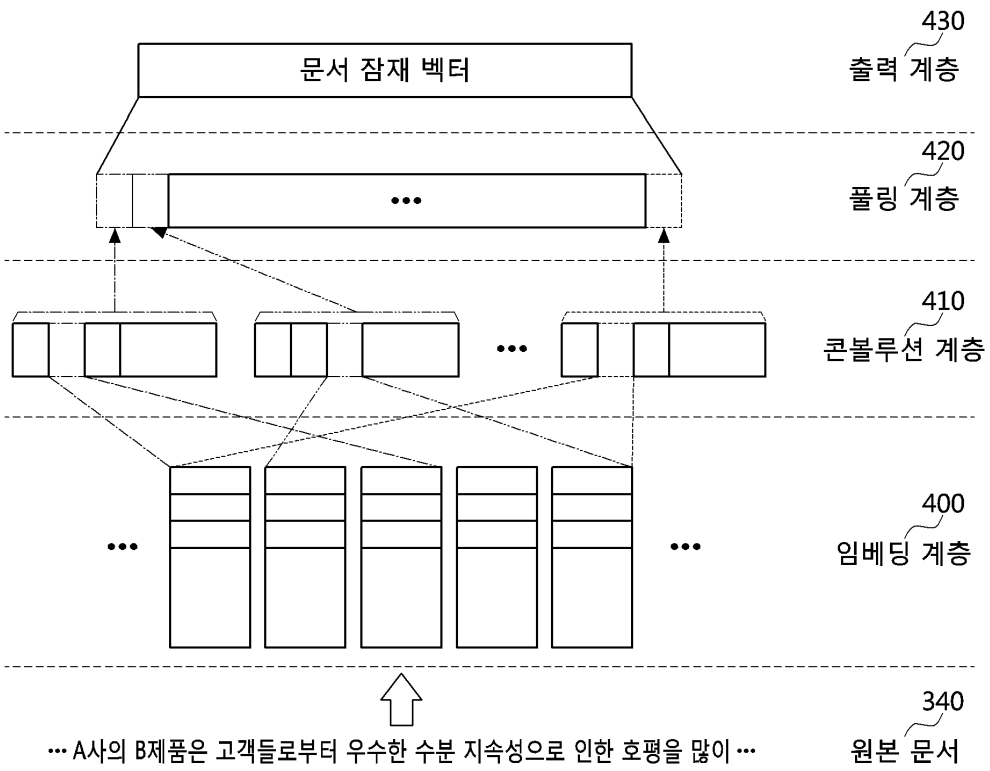
도면2



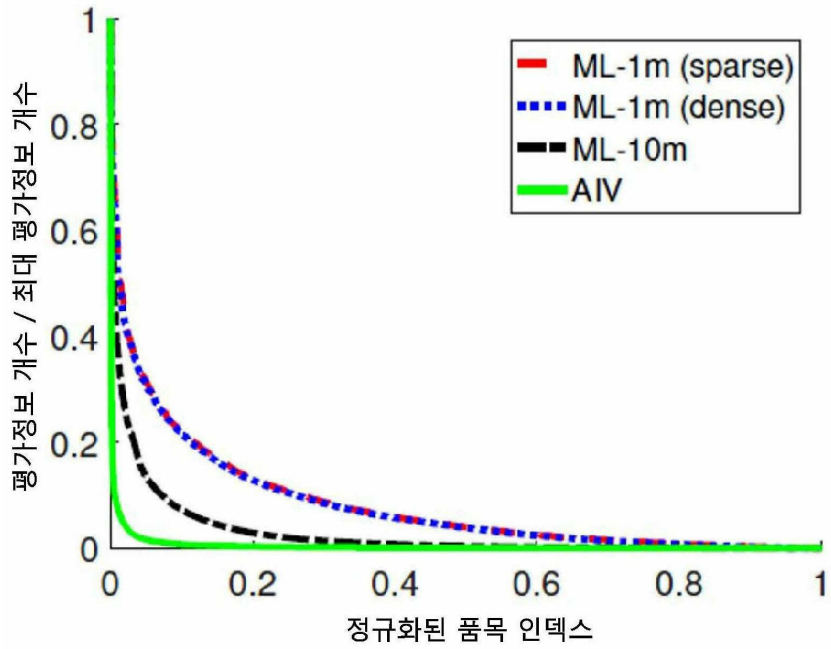
도면3


시작
문서정보 관측변수를 획득
S300
문맥 학습변수 획득
S310
문서정보 관측변수 및 문맥 학습변수를 기초로 품목 학습변수 획득
S320
사용자 학습변수 획득
S330
사용자 학습변수 및 품목 학습변수를 기초로 평점정보 관측변수 획득
S340
종료

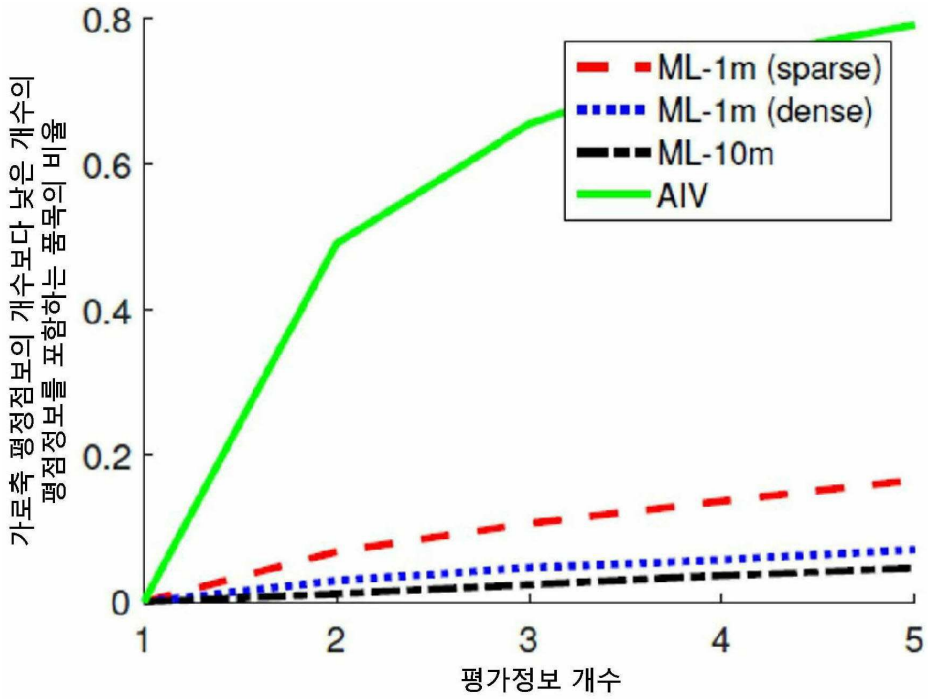
도면4


430
출력 계층
문서 잠재 벡터
420
풀링 계층
410
콘볼루션 계층
400
임베딩 계층
340
원본 문서
… A사의 B제품은 고객들로부터 우수한 수분 지속성으로 인한 호평을 많이 …

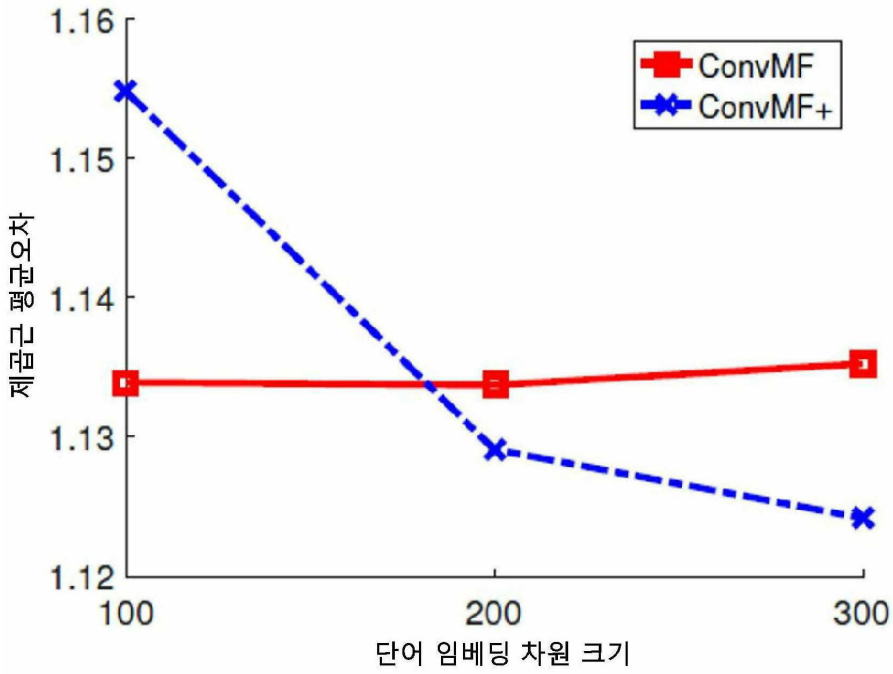
도면5


ML-1m (sparse)
ML-1m (dense)
ML-10m
AIV
평가정보 개수 / 최대 평가정보 개수
정규화된 품목 인덱스
0
0.2
0.4
0.6
0.8
1

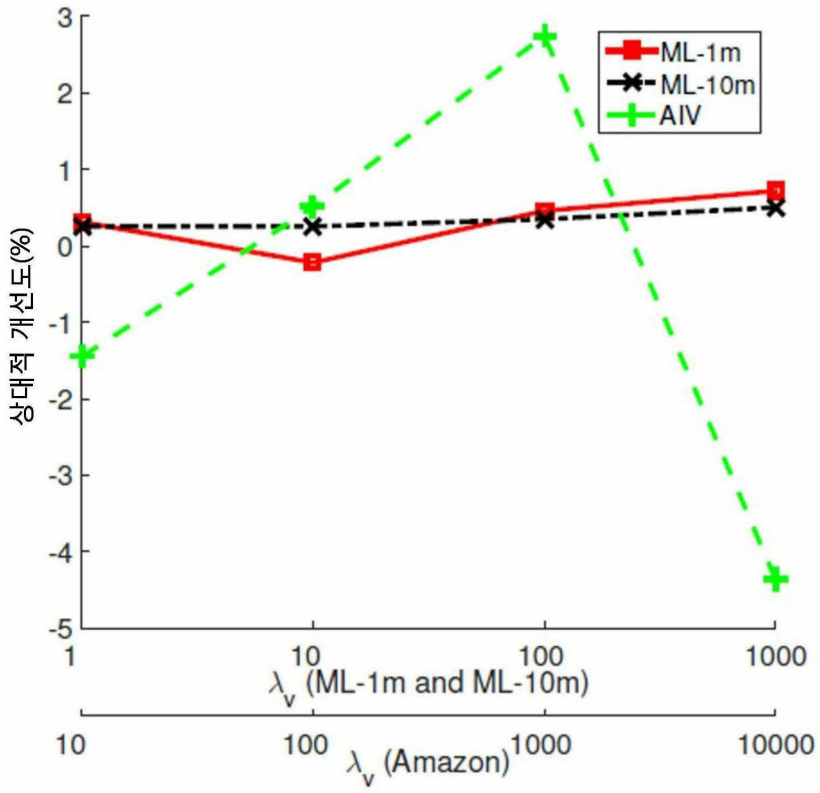
도면6


ML-1m (sparse)
ML-1m (dense)
ML-10m
AIV
가로축 평정점보의 개수보다 낮은 개수의 평점정보를 포함하는 품목의 비율
평가정보 개수
0
0.2
0.4
0.6
0.8
1
2
3
4
5

도면7


1.16
1.15
1.14
1.13
1.12
제곱근 평균오차
100
200
300
단어 임베딩 차원 크기
ConvMF
ConvMF+

도면8


3
2
1
0
-1
-2
-3
-4
-5
상대적 개선도(%)
ML-1m
ML-10m
AIV
1
10
100
1000
λv (ML-1m and ML-10m)
10
100
1000
10000
λv (Amazon)