



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년05월13일
(11) 등록번호 10-2252188
(24) 등록일자 2021년05월10일

(51) 국제특허분류(Int. Cl.)
G06Q 30/06 (2012.01) G06F 16/951 (2019.01)
G06F 40/20 (2020.01)
(52) CPC특허분류
G06Q 30/0631 (2013.01)
G06F 16/951 (2019.01)
(21) 출원번호 10-2019-0058596
(22) 출원일자 2019년05월20일
심사청구일자 2019년05월20일
(65) 공개번호 10-2020-0133439
(43) 공개일자 2020년11월30일
(56) 선행기술조사문헌
KR101230744 B1*
KR1020060099377 A*
KR1020180052489 A*
KR1020180080492 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
배재대학교 산학협력단
대전광역시 서구 배재로 155-40 (도마동)
(72) 발명자
정희경
대전광역시 서구 둔산로 155, 112동 1303호(둔산동, 크로바아파트)
황도연
대전광역시 대덕구 송촌로25번길 12, 202호 (송촌동)
(74) 대리인
유병욱, 한승범

전체 청구항 수 : 총 12 항

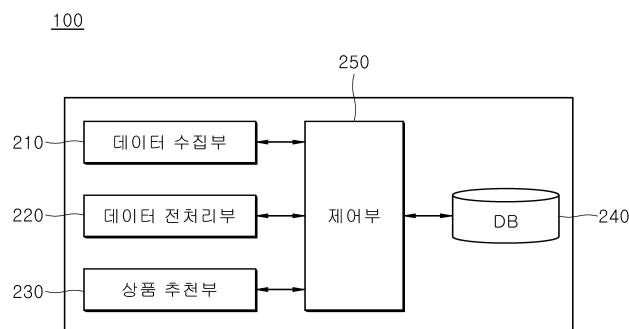
심사관 : 권태현

(54) 발명의 명칭 사용자 구매 기준을 반영한 상품 추천 시스템 및 방법

(57) 요약

본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 시스템은 상품의 검색 시 사용자에게 의해 선택된 상기 상품에 대한 구매 기준을 기반으로, 상기 사용자에게 의해 검색된 상품에 관한 상품 정보 및 상품 리뷰 데이터를 온라인 쇼핑몰 시스템으로부터 크롤링(crawling)하여 수집하는 데이터 수집부; 상기 상품 리뷰 데이터를 대상으로 형태소 분석 및 감정 분석을 수행하여, 상기 상품 리뷰 데이터에 내포되어 있는 긍정 평가와 부정 평가를 수치화하여 상기 구매 기준에 대한 상기 상품의 감성 점수를 산출하는 데이터 전처리부; 및 상기 감성 점수에 기초하여 상기 상품 정보에 따른 상품 리스트로부터 상기 사용자에게 추천할 상품을 추출하여 상기 사용자에게 맞춤형 추천 정보로서 제공하는 상품 추천부를 포함한다.

대표도



(52) CPC특허분류
G06F 40/268 (2020.01)

명세서

청구범위

청구항 1

상품의 검색 시 일정 기간 동안의 구매 기준을 분석하여 각 구매 기준별 사용자 선택 횟수를 산출하고, 상기 산출된 각 구매 기준별 사용자 선택 횟수에 근거하여 각 구매 기준별로 우선 순위를 부여함으로써 최우선 순위의 구매 기준을 사용자에게 추천하고, 상기 추천된 구매 기준이 상기 사용자에게 의해 선택되는 경우, 상기 선택된 구매 기준을 기반으로, 상기 사용자에게 의해 검색된 상품에 관한 상품 정보 및 상품 리뷰 데이터를 온라인 쇼핑 플랫폼 시스템으로부터 크롤링(crawling)하여 수집하는 데이터 수집부;

상기 상품 리뷰 데이터를 대상으로 형태소 분석 및 감정 분석을 수행하여, 상기 상품 리뷰 데이터에 내포되어 있는 긍정 평가와 부정 평가를 수치화하여 상기 구매 기준에 대한 상기 상품의 감성 점수를 산출하는 데이터 전처리부; 및

상기 감성 점수에 기초하여 상기 상품 정보에 따른 상품 리스트로부터 상기 사용자에게 추천할 상품을 추출하여 상기 사용자에게 맞춤형 추천 정보로서 제공하는 상품 추천부를 포함하고,

상기 구매 기준은 상기 상품의 내구성, 디자인, 가성비, 및 인플루언서(influencer)를 포함하고,

상기 데이터 수집부는 상기 산출된 각 구매 기준별 사용자 선택 횟수에 근거하여 각 구매 기준별로 가중치를 부여함으로써 가중치가 가장 높은 구매 기준을 상기 사용자에게 추천하되, 가중치가 가장 높은 구매 기준(A)의 가중치와 다른 구매 기준(B)의 가중치가 설정치 이하인 경우에는 구매 기준(A)와 구매 기준(B)를 모두 상기 사용자에게 추천함으로써 상기 사용자에게 다른 사용자들이 선호하는 구매 기준을 직접 선택하도록 구매 기준 선호 목록(A, B)을 제공하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 시스템.

청구항 2

제1항에 있어서,

상기 데이터 전처리부는

상기 구매 기준과 연관되는 키워드를 선정하고, 상기 크롤링한 상품 리뷰 데이터 중 상기 선정된 키워드가 포함된 상품 리뷰 데이터를 대상으로 상기 형태소 분석을 수행하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 시스템.

청구항 3

제1항에 있어서,

상기 데이터 전처리부는

상기 형태소 분석을 통해, 특징과 구매 기준을 나타내는 명사, 및 그 특징에 대한 의견과 감정을 나타내는 동사와 형용사를 추출하고, 상기 추출된 명사 및 동사와 형용사에 기초하여 상기 감정 분석을 수행하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 시스템.

청구항 4

제1항에 있어서,

상기 데이터 전처리부는

상기 감정 분석의 결과를 바탕으로, 상기 상품에 대한 전체 상품 리뷰 중 긍정적 평가의 개수 및 부정적 평가의

개수를 세어 긍정 평가 비율 및 부정 평가 비율을 산출하고, 상기 긍정 평가 비율 및 상기 부정 평가 비율에 기초하여 긍정 점수 및 부정 점수를 포함하는 상기 감성 점수를 산출하며, 상기 긍정 점수 및 상기 부정 점수를 합한 값을 상기 전체 상품 리뷰의 개수로 나누어 상기 감성 점수의 평균값을 산출하며,

상기 상품 추천부는

상기 감성 점수의 평균값이 높은 상위 N(상기 N은 자연수)개의 상품을 상기 사용자에게 추천할 상품으로서 추출하고, 상기 추출된 상품을 오름차순으로 정렬하여 상기 사용자에게 맞춤형 추천 정보로서 제공하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 시스템.

청구항 5

제1항에 있어서,

상기 데이터 전처리부는

상기 구매 기준이 상기 사용자에게 의해 복수개 선택된 경우, 상기 복수개의 선택 시 미리 설정된 우선순위에 따라 각 구매 기준별로 가중치를 부여하고, 상기 각 구매 기준별로 부여된 가중치를 반영하여 상기 상품의 감성 점수를 산출하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 시스템.

청구항 6

제1항에 있어서,

상기 데이터 전처리부는

상기 상품 리뷰 데이터를 크롤링 할 때, 추후의 형태소 분석에 유용하도록 마침표를 기준으로 한 문장씩 문장 단위로 분할하여 읽어와 텍스트 파일 형태로 저장하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 시스템.

청구항 7

제1항에 있어서,

상기 데이터 수집부는

상기 수집된 데이터를 데이터베이스(DB)에 저장하되, 상기 상품의 URL, 상품명 및 ID를 포함하는 상기 상품 정보는 상기 데이터베이스의 제1 테이블에 저장하고, 상기 상품 리뷰 데이터는 상기 데이터베이스의 제2 테이블에 저장하며, 상기 구매 기준 및 상기 구매 기준과 관련된 연관어가 포함된 상품 리뷰 데이터는 상기 데이터베이스의 제3 테이블에 저장하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 시스템.

청구항 8

제1항에 있어서,

상기 데이터 전처리부는

상기 상품 리뷰 데이터에 대해 노이즈 필터링(noise filtering)을 수행하여, 상기 상품 리뷰 데이터에 포함된 특수문자와 의미 없는 불용어, 및 상기 상품 리뷰 데이터의 문장 안에 긍정 혹은 부정을 판단할 수 있는 형용사가 포함되어 있지 않은 문장을 노이즈로 판단하여 상기 형태소 분석의 대상에서 제외시키는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 시스템.

청구항 9

사용자 구매 기준을 반영한 상품 추천 시스템을 이용한 상품 추천 방법에 있어서,

상기 상품 추천 시스템의 데이터 수집부가 상품의 검색 시 일정 기간 동안의 구매 기준을 분석하여 각 구매 기준별 사용자 선택 횟수를 산출하고, 상기 산출된 각 구매 기준별 사용자 선택 횟수에 근거하여 각 구매 기준별로 우선 순위를 부여함으로써 최우선 순위의 구매 기준을 사용자에게 추천하고, 상기 추천된 구매 기준이 상기 사용자에게 의해 선택되는 경우, 상기 선택된 구매 기준을 기반으로, 상기 사용자에게 의해 검색된 상품에 관한 상품 정보 및 상품 리뷰 데이터를 온라인 쇼핑몰 시스템으로부터 크롤링(crawling)하여 수집하는 단계;

상기 상품 추천 시스템의 데이터 전처리부가 상기 상품 리뷰 데이터를 대상으로 형태소 분석 및 감정 분석을 수행하여, 상기 상품 리뷰 데이터에 내포되어 있는 긍정 평가와 부정 평가를 수치화하여 상기 구매 기준에 대한 상기 상품의 감성 점수를 산출하는 단계; 및

상기 상품 추천 시스템의 상품 추천부가 상기 감성 점수에 기초하여 상기 상품 정보에 따른 상품 리스트로부터 상기 사용자에게 추천할 상품을 추출하여 상기 사용자에게 맞춤형 추천 정보로서 제공하는 단계를 포함하고,

상기 구매 기준은 상기 상품의 내구성, 디자인, 가성비, 및 인플루언서(influencer)를 포함하고,

상기 데이터 수집부는 상기 산출된 각 구매 기준별 사용자 선택 횟수에 근거하여 각 구매 기준별로 가중치를 부여함으로써 가중치가 가장 높은 구매 기준을 상기 사용자에게 추천하되, 가중치가 가장 높은 구매 기준(A)의 가중치와 다른 구매 기준(B)의 가중치가 설정치 이하인 경우에는 구매 기준(A)와 구매 기준(B)를 모두 상기 사용자에게 추천함으로써 상기 사용자에게 다른 사용자들이 선호하는 구매 기준을 직접 선택하도록 구매 기준 선호 목록(A, B)을 제공하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 방법.

청구항 10

제9항에 있어서,

상기 형태소 분석을 수행하는 단계는

상기 구매 기준과 연관되는 키워드를 선정하는 단계; 및

상기 크롤링한 상품 리뷰 데이터 중 상기 선정된 키워드가 포함된 상품 리뷰 데이터를 대상으로 상기 형태소 분석을 수행하는 단계

를 포함하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 방법.

청구항 11

제9항에 있어서,

상기 구매 기준이 상기 사용자에게 의해 복수개 선택된 경우,

상기 상품의 감성 점수를 산출하는 단계는

상기 복수개의 선택 시 미리 설정된 우선순위에 따라 각 구매 기준별로 가중치를 부여하는 단계; 및

상기 각 구매 기준별로 부여된 가중치를 반영하여 상기 상품의 감성 점수를 산출하는 단계

를 포함하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 방법.

청구항 12

제9항에 있어서,

상기 데이터 전처리부가 상기 상품 리뷰 데이터에 대해 노이즈 필터링(noise filtering)을 수행하여, 상기 상품 리뷰 데이터에 포함된 특수문자와 의미 없는 불용어, 및 상기 상품 리뷰 데이터의 문장 안에 긍정 혹은 부정을 판단할 수 있는 형용사가 포함되어 있지 않은 문장을 노이즈로 판단하여 상기 형태소 분석의 대상에서 제외시키

는 단계

를 더 포함하는 것을 특징으로 하는 사용자 구매 기준을 반영한 상품 추천 방법.

발명의 설명

기술 분야

[0001] 본 발명의 실시예들은 상품 추천 시스템 및 방법에 관한 것으로, 더욱 상세하게는 상품 구매 시 사용자가 선택한 구매 기준을 반영하여 사용자에게 추천할 상품을 제공할 수 있는 사용자 구매 기준을 반영한 상품 추천 시스템 및 방법에 관한 것이다.

배경 기술

[0003] 최근 상품에 대한 사용자의 의견을 공유할 수 있도록 하는 사용자 리뷰가 여러 구매자들에 의해 상품 리뷰가 작성되고, SNS, 블로그 등 다양한 경로를 통해서 공유되어 활성화 되고 있다. 이러한 상품 리뷰에는 해당 상품에 관해 사용자가 갖고 있는 선호에 대한 보다 상세하고 신뢰할 수 있는 정보를 담고 있어 추천 시스템에 유용한 활용이 가능하여 실제로 많은 연구가 진행되고 있다.

[0004] 또한, 개방적인 웹의 특성에 의해 소비자들이 온라인상에서 작성한 상품 리뷰 데이터의 양이 급격하게 증가하였고, 자연스럽게 상품 리뷰에 담긴 방대한 양의 의견 정보를 분석하는 여러 방법이 제안되었다.

[0005] 그러나, 기존의 상품 추천 시스템은 객관적이지 못한 정보를 바탕으로 하거나, 사용자의 선호도를 고려하지 않아 만족도가 낮은 편이다. 사용자가 상품을 구매할 때 중요하게 여기는 기준을 고려하여 반영한다면 추천 상품에 대한 만족도를 높일 수 있다.

[0006] 관련 선행기술로는 대한민국 등록특허공보 제10-1621938호(발명의 명칭: 사용자 등록 게시정보를 이용한 온라인 쇼핑정보 추천시스템, 등록일자: 2016.05.11)가 있다.

발명의 내용

해결하려는 과제

[0008] 본 발명의 일 실시예는 상품 구매 시 사용자가 선택한 구매 기준을 반영하여 사용자에게 추천할 상품을 제공할 수 있는 사용자 구매 기준을 반영한 상품 추천 시스템 및 방법을 제공한다.

[0010] 본 발명이 해결하고자 하는 과제는 이상에서 언급한 과제(들)로 제한되지 않으며, 언급되지 않은 또 다른 과제(들)는 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0012] 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 시스템은 상품의 검색 시 사용자에게 의해 선택된 상기 상품에 대한 구매 기준을 기반으로, 상기 사용자에게 의해 검색된 상품에 관한 상품 정보 및 상품 리뷰 데이터를 온라인 쇼핑몰 시스템으로부터 크롤링(crawling)하여 수집하는 데이터 수집부; 상기 상품 리뷰 데이터를 대상으로 형태소 분석 및 감정 분석을 수행하여, 상기 상품 리뷰 데이터에 내포되어 있는 긍정 평가와 부정 평가를 수치화하여 상기 구매 기준에 대한 상기 상품의 감성 점수를 산출하는 데이터 전처리부; 및 상기 감성 점수에 기초하여 상기 상품 정보에 따른 상품 리스트로부터 상기 사용자에게 추천할 상품을 추출하여 상기 사용자에게 맞춤형 추천 정보로서 제공하는 상품 추천부를 포함한다.

[0013] 상기 데이터 전처리부는 상기 구매 기준과 연관되는 키워드를 선정하고, 상기 크롤링한 상품 리뷰 데이터 중 상기 선정된 키워드가 포함된 상품 리뷰 데이터를 대상으로 상기 형태소 분석을 수행할 수 있다.

[0014] 상기 데이터 전처리부는 상기 형태소 분석을 통해, 특징과 구매 기준을 나타내는 명사, 및 그 특징에 대한 의견과 감정을 나타내는 동사와 형용사를 추출하고, 상기 추출된 명사 및 동사와 형용사에 기초하여 상기 감정 분석을 수행할 수 있다.

[0015] 상기 데이터 전처리부는 상기 감정 분석의 결과를 바탕으로, 상기 상품에 대한 전체 상품 리뷰 중 긍정적 평가의 개수 및 부정적 평가의 개수를 세어 긍정 평가 비율 및 부정 평가 비율을 산출하고, 상기 긍정 평가 비율 및 상기 부정 평가 비율에 기초하여 긍정 점수 및 부정 점수를 포함하는 상기 감성 점수를 산출하며, 상기 긍정 점

수 및 상기 부정 점수를 합한 값을 상기 전체 상품 리뷰의 개수로 나누어 상기 감성 점수의 평균값을 산출하며, 상기 상품 추천부는 상기 감성 점수의 평균값이 높은 상위 N(상기 N은 자연수)개의 상품을 상기 사용자에게 추천할 상품으로서 추출하고, 상기 추출된 상품을 오름차순으로 정렬하여 상기 사용자에게 맞춤형 추천 정보로서 제공할 수 있다.

- [0016] 상기 데이터 전처리부는 상기 구매 기준이 상기 사용자에게 의해 복수개 선택된 경우, 상기 복수개의 선택 시 미리 설정된 우선순위에 따라 각 구매 기준별로 가중치를 부여하고, 상기 각 구매 기준별로 부여된 가중치를 반영하여 상기 상품의 감성 점수를 산출할 수 있다.
- [0017] 상기 데이터 전처리부는 상기 상품 리뷰 데이터를 크롤링 할 때, 추후의 형태소 분석에 유용하도록 마침표를 기준으로 한 문장씩 문장 단위로 분할하여 읽어와 텍스트 파일 형태로 저장할 수 있다.
- [0018] 상기 데이터 수집부는 상기 수집된 데이터를 데이터베이스(DB)에 저장하되, 상기 상품의 URL, 상품명 및 ID를 포함하는 상기 상품 정보는 상기 데이터베이스의 제1 테이블에 저장하고, 상기 상품 리뷰 데이터는 상기 데이터베이스의 제2 테이블에 저장하며, 상기 구매 기준 및 상기 구매 기준과 관련된 연관어가 포함된 상품 리뷰 데이터는 상기 데이터베이스의 제3 테이블에 저장할 수 있다.
- [0019] 상기 데이터 전처리부는 상기 상품 리뷰 데이터에 대해 노이즈 필터링(noise filtering)을 수행하여, 상기 상품 리뷰 데이터에 포함된 특수문자와 의미 없는 불용어, 및 상기 상품 리뷰 데이터의 문장 안에 긍정 혹은 부정을 판단할 수 있는 형용사가 포함되어 있지 않은 문장을 노이즈로 판단하여 상기 형태소 분석의 대상에서 제외시킬 수 있다.
- [0020] 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 방법은 상기 상품 추천 시스템의 데이터 수집부가 상품의 검색 시 사용자에게 의해 선택된 상기 상품에 대한 구매 기준을 기반으로, 상기 사용자에게 의해 검색된 상품에 관한 상품 정보 및 상품 리뷰 데이터를 온라인 쇼핑몰 시스템으로부터 크롤링(crawling)하여 수집하는 단계; 상기 상품 추천 시스템의 데이터 전처리부가 상기 상품 리뷰 데이터를 대상으로 형태소 분석 및 감정 분석을 수행하여, 상기 상품 리뷰 데이터에 내포되어 있는 긍정 평가와 부정 평가를 수치화하여 상기 구매 기준에 대한 상기 상품의 감성 점수를 산출하는 단계; 및 상기 상품 추천 시스템의 상품 추천부가 상기 감성 점수에 기초하여 상기 상품 정보에 따른 상품 리스트로부터 상기 사용자에게 추천할 상품을 추출하여 상기 사용자에게 맞춤형 추천 정보로서 제공하는 단계를 포함한다.
- [0021] 상기 형태소 분석을 수행하는 단계는 상기 구매 기준과 연관되는 키워드를 선정하는 단계; 및 상기 크롤링한 상품 리뷰 데이터 중 상기 선정된 키워드가 포함된 상품 리뷰 데이터를 대상으로 상기 형태소 분석을 수행하는 단계를 포함할 수 있다.
- [0022] 상기 구매 기준이 상기 사용자에게 의해 복수개 선택된 경우, 상기 상품의 감성 점수를 산출하는 단계는 상기 복수개의 선택 시 미리 설정된 우선순위에 따라 각 구매 기준별로 가중치를 부여하는 단계; 및 상기 각 구매 기준별로 부여된 가중치를 반영하여 상기 상품의 감성 점수를 산출하는 단계를 포함할 수 있다.
- [0023] 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 방법은 상기 데이터 전처리부가 상기 상품 리뷰 데이터에 대해 노이즈 필터링(noise filtering)을 수행하여, 상기 상품 리뷰 데이터에 포함된 특수문자와 의미 없는 불용어, 및 상기 상품 리뷰 데이터의 문장 안에 긍정 혹은 부정을 판단할 수 있는 형용사가 포함되어 있지 않은 문장을 노이즈로 판단하여 상기 형태소 분석의 대상에서 제외시키는 단계를 더 포함할 수 있다.
- [0025] 기타 실시예들의 구체적인 사항들은 상세한 설명 및 첨부 도면들에 포함되어 있다.

발명의 효과

- [0027] 본 발명의 일 실시예에 따르면, 상품 구매 시 사용자가 선택한 구매 기준을 반영하여 사용자에게 추천할 상품을 제공할 수 있다.
- [0028] 본 발명의 일 실시예에 따르면, 상품 리뷰 데이터를 형태소 분석 및 감정 분석하여 산출한 감성 점수의 평균값을 이용하여, 상품 리뷰가 많은 상품이 상위권에 랭크되지 않도록 하는 작업을 통해, 상품 리뷰의 개수와 상관없이 동일한 점수가 산출되도록 하며, 이를 통해 상품 리뷰의 개수가 많은 상품이 구매 평가와 상관없이 상위권에 랭크되는 기존의 문제점을 해결할 수 있다.
- [0029] 본 발명의 일 실시예에 따르면, 구매 기준별 가중치를 달리함으로써 사용자의 구매 기준에 최적화된 상품을 상위권에 랭크되도록 할 수 있다.

도면의 간단한 설명

- [0031] 도 1은 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 시스템의 네트워크 구성을 도시한 도면이다.
 도 2는 도 1의 상품 추천 시스템의 상세 구성을 설명하기 위해 도시한 블록도이다.
 도 3은 도 2의 데이터베이스(DB)의 상세 구성을 설명하기 위해 도시한 블록도이다.
 도 4 및 도 5는 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 방법을 설명하기 위해 도시한 흐름도이다.
 도 6은 Twitter 클래스에서 제공하는 형태소 태그와 그에 대한 설명을 나타낸 도면이다.
 도 7은 '가격'을 1순위로, '배송'을 2순위로 선택한 경우로, 구매 기준과 연관된 키워드가 포함된 사용자 리뷰 데이터를 한 문장씩 형태소 분석을 한 예시이다.
 도 8은 각 구매 기준에 대한 긍정/부정 점수를 합산한 결과 상위 10위를 정리하여 나타낸 도면이다.
 도 9는 감성(긍정/부정) 점수를 바탕으로 순위를 매긴 상위 15개 상품의 구매 기준 비중을 나타낸 도면이다.
 도 10은 상위 15개 상품의 구매 기준 비중을 보여주는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0032] 본 발명의 이점 및/또는 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나, 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성요소를 지칭한다.
- [0033] 또한, 이하 실시되는 본 발명의 바람직한 실시예는 본 발명을 이루는 기술적 구성요소를 효율적으로 설명하기 위해 각각의 시스템 기능구성에 기 구비되어 있거나, 또는 본 발명이 속하는 기술분야에서 통상적으로 구비되는 시스템 기능 구성은 가능한 생략하고, 본 발명을 위해 추가적으로 구비되어야 하는 기능 구성을 위주로 설명한다. 만약 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자라면, 하기에 도시하지 않고 생략된 기능 구성 중에서 종래에 기 사용되고 있는 구성요소의 기능을 용이하게 이해할 수 있을 것이며, 또한 상기와 같이 생략된 구성 요소와 본 발명을 위해 추가된 구성 요소 사이의 관계도 명백하게 이해할 수 있을 것이다.
- [0034] 또한, 이하의 설명에 있어서, 신호 또는 정보의 "전송", "통신", "송신", "수신" 기타 이와 유사한 의미의 용어는 일 구성요소에서 다른 구성요소로 신호 또는 정보가 직접 전달되는 것뿐만이 아니라 다른 구성요소를 거쳐 전달되는 것도 포함한다. 특히 신호 또는 정보를 일 구성요소로 "전송" 또는 "송신"한다는 것은 그 신호 또는 정보의 최종 목적지를 지시하는 것이고 직접적인 목적지를 의미하는 것이 아니다. 이는 신호 또는 정보의 "수신"에 있어서도 동일하다.
- [0036] 이하에서는 첨부된 도면을 참조하여 본 발명의 실시예들을 상세히 설명하기로 한다.
- [0037] 도 1은 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 시스템의 네트워크 구성을 도시한 도면이고, 도 2는 도 1의 상품 추천 시스템의 상세 구성을 설명하기 위해 도시한 블록도이며, 도 3은 도 2의 데이터베이스(DB)의 상세 구성을 설명하기 위해 도시한 블록도이다.
- [0038] 도 1 내지 도 3을 참조하면, 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 시스템(100)은 데이터 수집부(210), 데이터 전처리부(220), 상품 추천부(230), 데이터베이스(DB)(240), 및 제어부(250)를 포함할 수 있다.
- [0039] 상기 데이터 수집부(210)는 사용자에 의해 검색된 상품에 관한 상품 정보 및 상품 리뷰 데이터를 온라인 쇼핑몰 시스템(102)으로부터 크롤링(crawling)하여 수집할 수 있다. 이때, 상기 데이터 수집부(210)는 상품의 검색 시 사용자에게 의해 선택된 상품에 대한 구매 기준을 기반으로, 상기 상품 정보 및 상품 리뷰 데이터를 수집할 수 있다.

- [0040] 구체적으로, 사용자는 사용자 단말기(101)를 통해 온라인 쇼핑물 시스템(102)에 접속하여 구매하고자 하는 상품을 검색할 수 있다. 이때, 상기 사용자는 상기 검색된 상품에 대한 구매 기준을 하나 선택할 수 있으며, 이와 달리 여러 개의 구매 기준을 선택할 수도 있다. 이처럼 상기 사용자에 의해 상기 상품에 대한 구매 기준이 선택되면, 상기 데이터 수집부(210)는 상기 상품에 대한 구매 기준을 활용하여 상기 온라인 쇼핑물 시스템(102)과의 연동을 통해 상기 검색된 상품에 관한 상품 정보와 상품 리뷰 데이터를 상기 온라인 쇼핑물 시스템(102)으로부터 크롤링하여 수집할 수 있다.
- [0041] 여기서, 상기 구매 기준은 상기 사용자가 상품을 선택할 때 우선시 하는 기준으로서, 상기 상품의 내구성, 디자인, 가성비 등을 포함할 수 있다. 예를 들어, 상기 사용자가 상기 상품의 검색 시 가성비를 구매 기준으로 선택한 경우, 상기 데이터 수집부(210)는 상기 상품과 관련된 리뷰 중에서 가격과 연관된 문장을 포함하는 리뷰를 선별하고, 상기 선별된 리뷰를 상기 상품 리뷰 데이터로서 수집할 수 있다.
- [0042] 또한, 상기 구매 기준은 인플루언서(influencer)를 더 포함할 수 있다. 여기서, 인플루언서(influencer)는 포털 사이트에서 영향력이 큰 블로그를 운영하는 ‘파워블로거’와 수십만 명의 팔로워 수를 가진 소셜네트워크서비스(SNS) 사용자, 1인 방송 진행자들을 통칭한 용어다. 인플루언서는 콘텐츠를 생산하는 크리에이터형 인플루언서와 패션뷰티 분야의 모델형 인플루언서로 구분할 수 있다. 한편, 영향력을 행사하는 사람을 활용한 마케팅 방법이 인플루언서 마케팅인데, 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 시스템(100)은 인플루언서 마케팅에도 적용될 수 있다.
- [0043] 예컨대, 상기 사용자가 상기 상품의 검색 시 인플루언서를 구매 기준으로 선택한 경우, 상기 데이터 수집부(210)는 상기 사용자가 검색한 상품 중 상기 인플루언서와 관련된 상품, 다시 말해 연예인이나 셀럽, 유튜브 스타 등과 같은 인플루언서가 사용하거나 추천한 상품을 선별하고, 상기 선별된 상품과 관련된 리뷰를 상기 상품 리뷰 데이터로서 수집할 수 있다. 이때, 상기 데이터 수집부(210)는 데이터 마이닝(data mining)을 통해 상기 사용자가 검색한 상품 중 상기 인플루언서와 관련된 상품을 선별할 수 있다.
- [0044] 이와 같이 상기 상품의 검색 시 상기 사용자가 상기 상품에 대한 구매 기준을 직접 선택할 수도 있지만, 다른 실시예로서 상기 상품의 검색 시 상기 상품에 대한 구매 기준의 선택 과정 없이 자동으로 추천된 구매 기준이 적용될 수도 있다. 예를 들면, 상기 데이터 수집부(210)는 일정 기간 동안(예: 1년)의 구매 기준을 분석하여 각 구매 기준별 사용자 선택 횟수를 산출하고, 상기 산출된 각 구매 기준별 사용자 선택 횟수에 근거하여 각 구매 기준별로 우선 순위를 부여함으로써 최우선 순위의 구매 기준을 상기 사용자에게 추천하여 상기 사용자의 선택에 도움을 주거나 최우선 순위의 구매 기준을 사용자가 선택한 구매 기준으로서 바로 적용할 수도 있다.
- [0045] 다른 예를 들면, 상기 데이터 수집부(210)는 일정 기간 동안(예: 1년)의 구매 기준을 분석하여 각 구매 기준별 사용자 선택 횟수를 산출하고, 상기 산출된 각 구매 기준별 사용자 선택 횟수에 근거하여 각 구매 기준별로 가중치를 부여함으로써 가중치가 가장 높은 구매 기준을 상기 사용자에게 추천하되, 가중치가 가장 높은 구매 기준(A)의 가중치와 다른 구매 기준(B)의 가중치가 설정치 이하인 경우에는 구매 기준(A)과 구매 기준(B)를 모두 상기 사용자에게 추천함으로써 상기 사용자에게 다른 사용자들이 선호하는 구매 기준을 직접 선택하도록 구매 기준 선호 목록(A, B)을 제공할 수도 있다.
- [0046] 또 다른 예를 들면, 상기 데이터 수집부(210)는 성별, 연령별 조건으로 일정 기간 동안(예: 1년)의 구매 기준을 분석하여 성별, 연령별 조건에 따른 각 구매 기준별 사용자 선택 횟수를 산출하고, 상기 산출된 각 구매 기준별 사용자 선택 횟수에 근거하여 각 구매 기준별로 가중치를 부여할 수도 있다.
- [0047] 상기 데이터 수집부(210)는 상기 수집된 데이터(상품 정보, 상품 리뷰 데이터)를 데이터베이스(DB)(240)에 저장할 수 있다. 이를 위해, 상기 데이터베이스(240)는 도 3에 도시된 바와 같이 제1 테이블(310), 제2 테이블(320) 및 제3 테이블(330)을 포함할 수 있다.
- [0048] 다시 말해, 상기 데이터 수집부(210)는 상기 상품의 URL, 상품명 및 ID를 포함하는 상품 정보는 상기 데이터베이스(240)의 제1 테이블(310)에 저장하고, 상기 상품 리뷰 데이터는 상기 데이터베이스(240)의 제2 테이블(320)에 저장할 수 있다. 그리고, 상기 데이터 수집부(210)는 상기 구매 기준 및 상기 구매 기준과 관련된 연관어가 포함된 상품 리뷰 데이터는 상기 데이터베이스(240)의 제3 테이블(330)에 저장할 수 있다.
- [0049] 상기 데이터 전처리부(220)는 상기 상품 리뷰 데이터를 크롤링 할 때, 추후의 형태소 분석에 유용하도록 마침표를 기준으로 한 문장씩 문장 단위로 분할하여 읽어와 텍스트 파일 형태로 저장할 수 있다.
- [0050] 상기 데이터 전처리부(220)는 상기 상품 리뷰 데이터에 대해 노이즈 필터링(noise filtering)을 수행하여, 상기

상품 리뷰 데이터에 포함된 특수문자와 의미 없는 불용어를 노이즈로 판단하여 상기 형태소 분석의 대상에서 제외시킬 수 있다. 또한, 상기 데이터 전처리부(220)는 상기 상품 리뷰 데이터의 문장 안에 긍정 혹은 부정을 판단할 수 있는 형용사가 포함되어 있지 않은 문장을 노이즈로 판단하여 상기 형태소 분석의 대상에서 제외시킬 수 있다.

- [0051] 이를 위해, 상기 데이터 전처리부(220)는 상기 상품 리뷰 데이터에 포함된 특수문자와 의미 없는 불용어를 제거하여 노이즈 필터링을 수행할 수 있다. 또한, 상기 데이터 전처리부(220)는 상기 상품 리뷰 데이터의 문장 안에 긍정 혹은 부정을 판단할 수 있는 형용사가 포함되어 있지 않은 문장을 제거하여 노이즈 필터링을 수행할 수 있다.
- [0052] 상기 데이터 전처리부(220)는 상기 상품 리뷰 데이터를 대상으로 형태소 분석 및 감정 분석을 수행하여, 상기 상품 리뷰 데이터에 내포되어 있는 긍정 평가와 부정 평가를 수치화하여 상기 구매 기준에 대한 상기 상품의 감성 점수를 산출할 수 있다.
- [0053] 구체적으로, 상기 데이터 전처리부(220)는 상기 구매 기준과 연관되는 키워드를 선정하고, 상기 크롤링한 상품 리뷰 데이터 중 상기 선정된 키워드가 포함된 상품 리뷰 데이터를 대상으로 형태소 분석을 수행할 수 있다. 상기 데이터 전처리부(220)는 상기 형태소 분석을 통해, 특징과 구매 기준을 나타내는 명사, 및 그 특징에 대한 의견과 감정을 나타내는 동사와 형용사를 추출하고, 상기 추출된 명사 및 동사와 형용사에 기초하여 상기 상품 리뷰 데이터에 대한 감정 분석을 수행할 수 있다.
- [0054] 상기 데이터 전처리부(220)는 상기 감정 분석의 결과를 바탕으로, 상기 상품에 대한 전체 상품 리뷰 중 긍정적 평가의 개수 및 부정적 평가의 개수를 세어 긍정 평가 비율 및 부정 평가 비율을 산출할 수 있다. 상기 데이터 전처리부(220)는 상기 긍정 평가 비율 및 상기 부정 평가 비율에 기초하여 상기 긍정 평가에 해당하는 긍정 점수 및 상기 부정 평가에 해당하는 부정 점수를 포함하는 감성 점수를 산출할 수 있다. 상기 데이터 전처리부(220)는 상기 긍정 점수 및 상기 부정 점수를 합한 값을 상기 전체 상품 리뷰의 개수로 나누어 상기 감성 점수의 평균값을 산출할 수 있다.
- [0055] 구체적으로, 상기 데이터 전처리부(220)는 각 상품별로 나누어 상기 상품 리뷰 데이터를 형태소 분석하고, 상기 형태소 분석한 결과를 바탕으로 상기 각 상품에 대한 전체 상품 리뷰 중 긍정적 평가의 개수 및 부정적 평가의 개수를 카운트(count)할 수 있다. 상기 데이터 전처리부(220)는 상기 각 상품에 대한 전체 상품 리뷰의 개수 대비 긍정적 평가의 개수를 토대로 긍정 평가 비율을 산출하고, 상기 각 상품에 대한 전체 상품 리뷰의 개수 대비 부정적 평가의 개수를 토대로 부정 평가 비율을 산출할 수 있다. 상기 데이터 전처리부(220)는 상기 산출된 긍정 평가 비율에 기초하여 상기 긍정 평가에 해당하는 긍정 점수를 산출하고, 상기 산출된 부정 평가 비율에 기초하여 상기 부정 평가에 해당하는 부정 점수를 산출함으로써 상기 감성 점수를 도출해낼 수 있다. 상기 데이터 전처리부(220)는 상기 감성 점수, 즉 상기 긍정 점수 및 부정 점수를 합한 값을 상기 전체 상품 리뷰의 개수로 나누는 나눗셈 연산을 수행하여 상기 상품에 대한 감성 점수의 평균값을 산출할 수 있다.
- [0056] 이처럼 본 발명의 일 실시예에서는 상기 상품에 대한 감성 점수의 평균값을 이용함으로써 상품 리뷰가 많은 상품이 상위권에 랭크되지 않도록 하는 작업을 통해 상품 리뷰의 개수와 상관없이 동일한 점수가 산출되도록 하며, 이는 상품 리뷰의 개수가 많은 상품이 구매 평가와 상관없이 상위권에 랭크되는 기존의 문제점을 해결할 수 있다.
- [0057] 한편, 상기 사용자는 상기 상품의 검색 시 복수개의 구매 기준을 선택할 수도 있다. 이때, 상기 사용자는 상기 복수개의 구매 기준 각각에 우선순위를 미리 설정할 수도 있다. 이와 같이 복수개의 구매 기준이 선택된 경우, 상기 데이터 전처리부(220)는 미리 설정된 우선순위에 따라 각 구매 기준별로 가중치를 부여하고, 상기 각 구매 기준별로 부여된 가중치를 반영하여 상기 상품의 감성 점수를 산출할 수도 있다.
- [0058] 상기 상품 추천부(230)는 상기 데이터 전처리부(220)에 의해 산출된 감성 점수에 기초하여, 상기 상품 정보에 따른 상품 리스트로부터 상기 사용자에게 추천할 상품을 추출하고, 상기 추출된 추천 상품을 상기 사용자에게 맞춤형 추천 정보로서 제공할 수 있다.
- [0059] 이때, 상기 상품 추천부(230)는 상기 감성 점수의 평균값이 높은 상위 N(상기 N은 자연수)개, 예를 들면 10개의 상품을 상기 사용자에게 추천할 상품으로서 추출하고, 상기 추출된 상품을 오름차순으로 정렬하여 상기 사용자에게 맞춤형 추천 정보로서 제공할 수 있다.
- [0060] 상기 데이터베이스(240)는 상기 데이터 수집부(210)에 의해 수집된 데이터를 저장할 수 있다. 특히, 상기 데이터베이스(240)는 앞서 설명한 바와 같이 제1 내지 제3 테이블(310, 320, 330)에 상기 수집된 데이터를 분산하여

저장할 수 있다.

- [0061] 구체적으로, 상기 제1 테이블(310)에는 상기 상품의 URL, 상품명 및 ID를 포함하는 상품 정보가 저장되고, 상기 제2 테이블(320)에는 상기 상품 리뷰 데이터가 저장될 수 있다. 그리고, 상기 제3 테이블(330)에는 상기 사용자가 상품의 검색 시 선택한 구매 기준 및 상기 구매 기준과 관련된 연관어가 포함된 상품 리뷰 데이터가 저장될 수 있다.
- [0062] 상기 제어부(250)는 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 시스템(100), 즉 상기 데이터 수집부(210), 상기 데이터 전처리부(220), 상기 상품 추천부(230), 상기 데이터베이스(240) 등의 동작을 전반적으로 제어할 수 있다.
- [0064] 이상에서 설명된 장치는 하드웨어 구성 요소, 소프트웨어 구성 요소, 및/또는 하드웨어 구성 요소 및 소프트웨어 구성 요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성 요소는, 예를 들어, 프로세서, 컨트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.
- [0065] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0067] 도 4 및 도 5는 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 방법을 설명하기 위해 도시한 흐름도이다.
- [0068] 여기서 설명하는 상품 추천 방법은 본 발명의 하나의 실시예에 불과하며, 그 이외에 필요에 따라 다양한 단계들이 추가될 수 있고, 하기의 단계들도 순서를 변경하여 실시될 수 있으므로, 본 발명이 하기에 설명하는 각 단계 및 그 순서에 한정되는 것은 아니다.
- [0069] 먼저 도 1, 도 2 및 도 4를 참조하면, 단계(410)에서 사용자는 사용자 단말기(101)를 통해 온라인 쇼핑몰 시스템(102)에 접속하여 구매하고자 하는 상품을 검색할 수 있다.
- [0070] 다음으로, 단계(420)에서 상기 사용자는 상기 상품의 검색 시 상기 상품에 대한 구매 기준을 선택할 수 있다. 여기서, 상기 구매 기준은 상기 사용자가 상품을 선택할 때 우선시 하는 기준으로서, 상기 상품의 내구성, 디자인, 가성비 등을 포함할 수 있다. 또한, 상기 사용자는 상기 상품에 대해 복수개의 구매 기준을 선택할 수도 있으며, 이때에는 각 구매 기준에 대해 우선순위를 미리 설정할 수 있다.
- [0071] 다음으로, 단계(430)에서 상기 상품 추천 시스템(100)의 데이터 수집부(210)는 상기 사용자에게 의해 검색된 상품에 관한 상품 정보 및 상품 리뷰 데이터를 수집할 수 있다.
- [0072] 다음으로, 단계(440)에서 상기 상품 추천 시스템(100)의 데이터 수집부(210)는 상기 수집된 데이터(상품 정보, 상품 리뷰 데이터 등)를 데이터베이스(240)에 저장할 수 있다.
- [0073] 다음으로, 단계(450)에서 상기 상품 추천 시스템(100)의 데이터 전처리부(220)는 상기 상품 리뷰 데이터를 대상으로 형태소 분석 및 감정 분석을 수행하여, 상기 선택된 구매 기준에 대한 상품의 감성 점수를 산출할 수 있다. 이에 대해 도 5를 참조하여 구체적으로 설명하면 다음과 같다.

- [0074] 즉, 단계(510)에서 상기 데이터 전처리부(220)는 형태소 분석을 통해, 특징과 구매 기준을 나타내는 명사, 및 그 특징에 대한 의견과 감정을 나타내는 동사와 형용사를 추출할 수 있다. 이후, 단계(520)에서 상기 데이터 전처리부(220)는 상기 추출된 명사 및 동사와 형용사에 기초하여 상기 상품 리뷰 데이터에 대한 감정 분석을 수행할 수 있다. 이후, 단계(530)에서 상기 데이터 전처리부(220)는 상기 감정 분석의 결과를 바탕으로, 상기 상품에 대한 전체 상품 리뷰 중 긍정적 평가의 개수 및 부정적 평가의 개수를 세어 긍정 평가 비율 및 부정 평가 비율을 산출할 수 있다. 이후, 단계(540)에서 상기 데이터 전처리부(220)는 상기 긍정 평가 비율 및 상기 부정 평가 비율에 기초하여 상기 긍정 평가에 해당하는 긍정 점수 및 상기 부정 평가에 해당하는 부정 점수를 포함하는 감성 점수를 산출할 수 있다. 이후, 단계(550)에서 상기 데이터 전처리부(220)는 상기 긍정 점수 및 상기 부정 점수를 합한 값을 상기 전체 상품 리뷰의 개수로 나누어 상기 감성 점수의 평균값을 산출할 수 있다.
- [0075] 다시 도 4를 참조하면, 단계(460)에서 상기 상품 추천 시스템(100)의 상품 추천부(230)는 상기 데이터 전처리부(220)에 의해 산출된 감성 점수에 기초하여 상품 리스트로부터 상기 사용자에게 추천할 상품을 추출할 수 있다. 이때, 상기 상품 추천부(230)는 상기 평가 점수의 평균값이 높은 상위 N(상기 N은 자연수)개의 상품을 상기 사용자에게 추천할 상품으로서 추출할 수 있다.
- [0076] 다음으로, 단계(470)에서 상기 상품 추천 시스템(100)의 상품 추천부(230)는 상기 추출된 상품을 오름차순으로 정렬하여 상기 사용자 단말기(101)를 통해 상기 사용자에게 맞춤형 추천 정보로서 제공할 수 있다.
- [0078] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 관독 가능 매체에 기록될 수 있다. 상기 컴퓨터 관독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 관독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CDROM, DVD와 같은 광기록 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 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0080] 이하에서는 본 발명의 일 실시예에 따른 사용자 구매 기준을 반영한 상품 추천 시스템을 구현하는 일례를 상세히 설명하기로 한다.
- [0082] 시스템 구현
- [0084] 본 실험에서 사용한 컴퓨터는 Windows 운영체제 기반의 CPU(Intel i5-4690), RAM-8G를 사용하였고, 구현은 Pycharm을 사용하였다. 시스템 구현을 위해 사용자가 가격과 배송을 구매 기준으로 선택한 경우를 예시로 들었다. 즉, 사용자가 다른 디자인이나 브랜드 인지도 등의 구매 기준보다 가격이 저렴하고 배송이 빠른 상품을 구매하길 원하는 경우를 예로 들었다.
- [0086] 1. 데이터 수집(Crawling)
- [0088] 실험을 위하여 Python을 이용하여 국내 쇼핑몰의 카테고리 중 계절가전의 '전기히터'를 수집 대상으로 선정하고, 전기 히터 상품의 정보와 상품에 대해 사용자가 작성한 리뷰를 수집한다. 상품 리뷰의 양질이 이번 실험을 좌우하기 때문에 비교적 솔직한 사용 후기가 많이 작성되어 있는 가전용품으로 정하였고, 국내에서 약 7~8조 이상의 거래액을 기록한 인터넷 사이트로부터 수집했다. 거래액이 많은 만큼 소비자들이 많이 이용하고, 많이 이용하는 만큼 많은 리뷰 데이터가 있을 것이라고 생각했다. '계절가전' 카테고리 항목에서 가장 많은 리뷰가 존재하는 '전기히터' 상품 항목의 상품명, 상품 URL과 상품 리뷰를 BeautifulSoup 웹크롤러를 이용하여 수집한다. 분석 데이터가 될 수 있도록 리뷰의 개수가 20개 초과인 상품만을 선별하여 추출한다.
- [0089] 크롤링 하고자 하는 온라인 쇼핑몰 시스템에서 제공하는 상품 리뷰는 '더보기' 버튼을 누르는 행위와 같이, 일정 수의 리뷰가 작성되면 페이지를 넘겨 볼 수 있는 웹 브라우저 상의 액션이 이루어져야 모든 리뷰의 목록을 볼 수가 있다. 따라서 이러한 액션들을 크롤링 할 때 자동으로 수행할 수 있는 웹 페이지 테스트 툴인 Selenium을 이용한다. 이 툴은 초기에 브라우저를 조정하여 웹을 테스트하는 용도로 쓰이며, 웹 페이지 내에 태그명이나 HTML 태그 내에 있는 텍스트를 자바스크립트를 포함하여 크롤링 하는 기능들이 있어 손쉽게 데이터 수집이 가능하여 크롤링을 할 때 널리 쓰이는 툴이기도 하다.

- [0090] Selenium 패키지는 Python 환경에서 상품의 URL을 통해 웹 브라우저를 실행시키고 HTML 태그 이름을 이용하여 상품의 이름, 상품의 분류, 리뷰 등을 수집한다. 상품 정보와 상품 리뷰를 크롤링할 때 상품 리뷰의 수가 20개 초과인 상품만을 DB에 저장한다.
- [0092] 2. 전처리(Pre-processing)
- [0094] 위와 같이 상품 정보와 상품 리뷰를 크롤링 할 때 정성적 정보인 상품 리뷰를 정량적인 정보로 수치화해야 하므로 적절한 수준의 상품 리뷰가 존재해야 한다. 일정 수준 이하의 상품 리뷰를 가지고 가중치를 계산한다면, 해당 상품을 평가할 때 적절하지 못하기 때문이다.
- [0095] 따라서 본 실시예에서는 상품 리뷰의 수가 20개 초과인 상품만을 DB에 저장한다. 상품 특징의 점수를 계산하기 위해서는 각 특징이 속한 문장의 점수를 구하는 것이 선행되어야 한다. 그러기 위해서는 문장을 분리하는 작업이 필요하다. 각 리뷰마다 적게는 1개, 많게는 10개의 문장으로 이루어지는 것이 보편적이다. 따라서 상품 리뷰를 크롤링할 때는 마침표를 기준으로 한 문장씩을 읽어와 텍스트 파일로 저장하여 이후 형태소 분석에 용이하도록 했다. 또한 특수 문자나 'ㅋㅋㅋㅋ'와 같은 의미 없는 불용어(stopword)를 제거하는 전처리 작업을 한다.
- [0096] 본 실시예에서는 구매 기준을 표현하는 명사(noun)뿐만 아니라 '가격' 등과 같은 구매 기준의 키워드를 설명하는 형용사(adjective)를 함께 고려한다. 예를 들면, 사용자가 가격이 저렴한 상품을 원하기 때문에 '가격'이라는 키워드와 함께 '저렴-' 혹은 '싸-'와 같은 형용사를 키워드에 추가하여 사용자 리뷰 데이터로부터 추출한다. 명사와 함께 형용사를 고려함으로써 더 정확하게 사용자가 원하는 구매 기준에 부합하는 상품을 추출할 수 있고, 또한 분석에 필요한 정보만을 고려함으로써 불필요한 데이터를 필터링하는 기능도 할 수 있다.
- [0098] 3. 형태소 분석
- [0100] Python 프로그램의 한글을 처리하는 패키지인 'KoNLP'를 설치하여 Twitter 클래스를 기반으로 하여 수행한다. Twitter 클래스는 트위터에서 만든 스칼라로 작성된 오픈 소스 한국어 토큰나이저로, 보다 간결하고 이해하기 쉽게 결과를 출력 한다. 도 6은 Twitter 클래스에서 제공하는 형태소 태그와 그에 대한 설명을 나타낸 도면이다. Twitter 클래스는 도 6에 포함시키지 않은 트위터 해쉬태그(Hashtag), 트위터 아이디(ScreenName), 이메일 주소(Email) 등을 포함한 약 20개의 태그를 가지고 있다.
- [0101] 이전 단계에서 저장한 텍스트 파일을 읽어와 리스트로 저장하여 한 문장씩 형태소 분석을 수행한다. 이 과정에서 리뷰의 문장 중 긍정 혹은 부정을 판단할 수 있는 형용사가 포함되어 있지 않은 문장들은 노이즈로 판단하여 분석 대상에서 제거한다.
- [0102] 도 7은 '가격'을 1순위로, '배송'을 2순위로 선택한 경우로, 구매 기준과 연관된 키워드가 포함된 사용자 리뷰 데이터를 한 문장씩 형태소 분석을 한 예시이다.
- [0104] 4. 감성(긍정/부정) 점수 부여
- [0106] 상품 추천을 위해서는 앞서 추출한 상품 리뷰 중 사용자가 선택한 구매 기준이 긍정적으로 평가한 상품 리뷰가 많은지, 부정적으로 평가한 상품 리뷰가 많은지 판별한다. 각각 1순위와 2순위의 구매 기준이 포함된 문장에서 그 구매 기준에 대한 평가가 있는 '형용사' 혹은 '동사'를 가지고 감성(긍정/부정) 점수를 매긴다. 예를 들어, "배송이 정말 빠르네요."라는 리뷰 문장이 있다면, '배송'이라는 구매 기준에 대해 '빠르네요.'라는 긍정적인 표현을 했으므로 +1점을 매긴다. 이렇게 긍정적인 평가이면 +1점, 부정적인 평가이면 -1점으로 하여 합산하는 방식으로 한다.
- [0107] 전체 리뷰 중 '가격'이라는 키워드가 포함된 리뷰를 $PR=\{pr_1, pr_2, \dots, pr_n\}$ 이라고 하고, 그 리뷰를 긍정/부정 점수를 매긴 것을 PS(PriceScore)라고 했을 때, 구매 기준 '가격'에 대한 최종 리뷰 점수(Keyword Review)는 $KR_1=PS/RP$ 으로 구할 수 있다. 마찬가지로 두 번째 키워드인 '배송'의 경우에는 키워드가 포함된 리뷰(Delivery Review)는 $DR=\{dr_1, dr_2, \dots, dr_n\}$ 이고, 그 리뷰의 긍정/부정 점수를 매긴 것을 DS(DeliveryScore)이면, '배송' 키워드의 최종 리뷰 점수는 $KR_2=DS/DR$ 으로 계산할 수 있다.
- [0108] 그러나 중립적인 의견은 판별하는 데 어려움이 있어 노이즈로 판단하여 제외시켰고, 전체적인 상품 리뷰가 긍정적이지만 정렬 기준 키워드인 '가격' 혹은 '배송'에 대한 좋은 평가가 없는 상품 리뷰도 제외 목록에 포함시켰다.
- [0109] 긍정과 부정 점수를 합산한 점수에 각 구매 기준의 가중치를 곱하여 합산하고 이를 백분율로 나타내면 각 구매

기준에 대한 비중을 구할 수 있다. 이 비율은 높을수록 상위권에 랭크될 확률이 높아진다는 것을 알 수 있다. 같은 방식으로 구매 기준 2순위도 산출하여 두 개의 점수를 합산한 점수가 최종 점수가 되고, 이 점수를 가지고 순위를 매기게 된다. 이를 식으로 나타내면 아래 수학적 식 1과 같다.

[0110] [수학적 식 1]

$$ReviewSentimentScore = \frac{(KR_1 \times W_1) + (KR_2 \times W_2)}{100}$$

[0111]

[0112] 그 결과로, 전체 상품 리스트 중 가장 좋은 평가를 받은 상품을 사용자에게 추천한다. 도 8은 각 구매 기준에 대한 긍정/부정 점수를 합산한 결과 상위 10위를 정리하여 나타낸 도면이다.

[0114] 본 실시예에서 제안하는 상품 추천 시스템은 Beautifulsoup 웹 크롤러를 사용하여 데이터를 수집하고 Python 환경에서 상품 리뷰 데이터의 형태소 분석과 감성 분석을 통해 사용자에게 적합한 추천 정보를 제공한다. 또한, 사용자 맞춤형 추천 정보를 도출하기 위해 상품 구매 의사 결정에 영향을 미치는 구매 기준을 직접 선택하고 이를 알고리즘에 반영함으로써 기존의 추천 시스템과 차별화하였다.

[0115] 앞서 산출한 감성 점수를 전체 상품 리뷰의 개수로 나누어 비중을 구할 수 있다. 이것은 그 상품에서 해당 구매 기준이 차지하는 비중을 알 수 있다. 그리고 가중치가 상품 목록의 생성에 큰 영향을 끼치는지도 알 수 있다.

[0116] 도 9는 감성(긍정/부정) 점수를 바탕으로 순위를 매긴 상위 15개 상품의 구매 기준 비중을 나타낸 도면이다. 11위 상품과 같이 상위 10위의 상품보다 더 높은 비중을 가지고 있는 데도 불구하고 그보다 낮은 순위에 랭크된 상품이 있다. 이는 1순위와 2순위의 가중치에 따라 순위가 결정된다는 것을 보여준다.

[0117] 도 10은 상위 15개 상품의 구매 기준 비중을 보여주는 도면이다. 도 9에서 높은 비율에도 불구하고 11위에 랭크된 상품과 낮은 비율임에도 불구하고 상위 5위에 랭크된 상품의 경우를 비교해 보면 2순위였던 '배송'의 감성 점수 비율보다 1순위인 '가격'의 감성 점수 비율이 월등히 높은 것을 알 수 있다. 2순위의 가중치를 1순위의 가중치의 절반으로 반영했기 때문에 전체 리뷰의 비율이 더 크지만 상위권에 랭크되지 않은 결과가 도출된 것이다.

[0118] 이와 같이 사용자가 선택한 구매 기준의 우선순위에 따라 추천 순위가 바뀌는 것을 확인할 수 있다. 나아가 이 결과는 사용자의 선호도와 구매 성향을 반영하여 각각의 개인 사용자별 맞춤형 상품 추천을 할 수 있다는 의미가 될 수 있다. 따라서 구매 기준별 가중치를 달리함으로써 사용자의 구매 기준에 최적화된 상품을 상위권에 랭크되도록 할 수 있다.

[0120] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0121] 그러므로, 다른 구현들, 다른 실시예들 및 청구범위와 균등한 것들도 후술하는 청구범위의 범위에 속한다.

부호의 설명

- [0123] 100: 상품 추천 시스템
- 101: 사용자 단말기
- 102: 온라인 쇼핑몰 시스템
- 210: 데이터 수집부
- 220: 상품 평가부
- 230: 상품 추천부
- 240: DB
- 250: 제어부

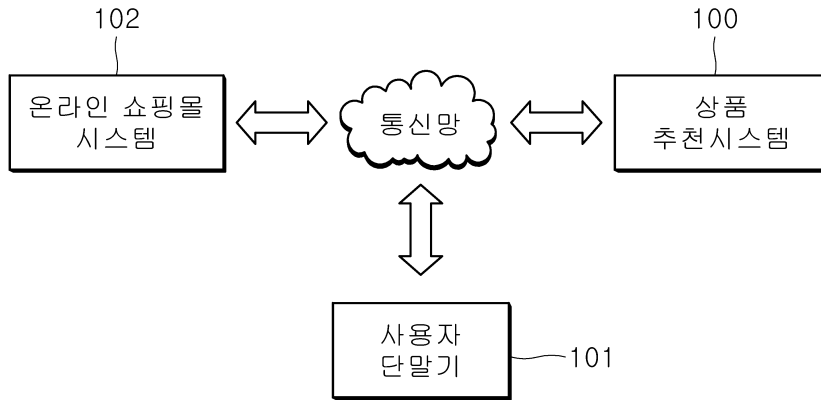
310: 제1 테이블

320: 제2 테이블

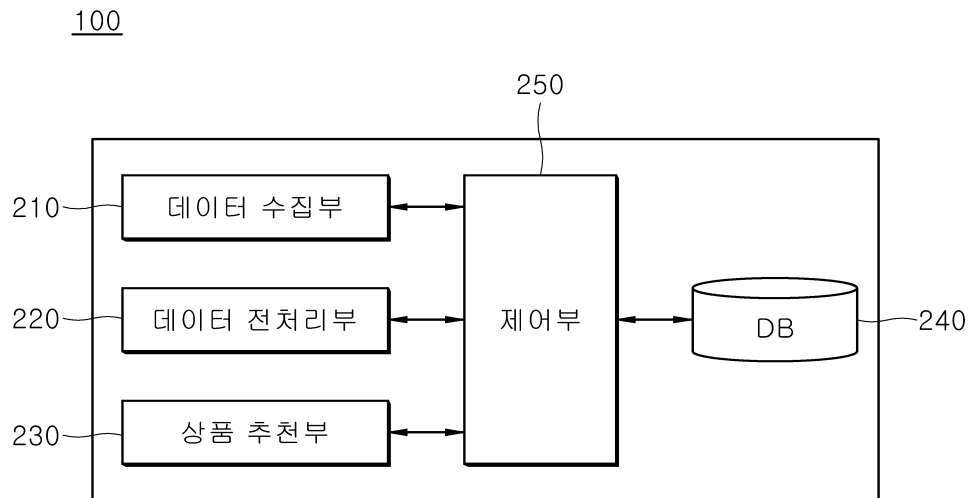
330: 제3 테이블

도면

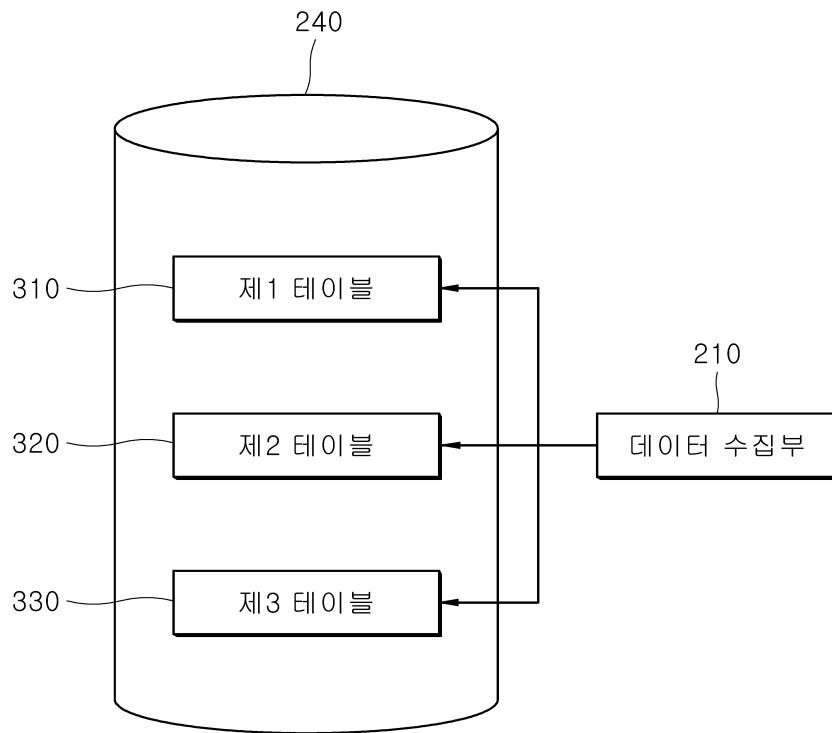
도면1



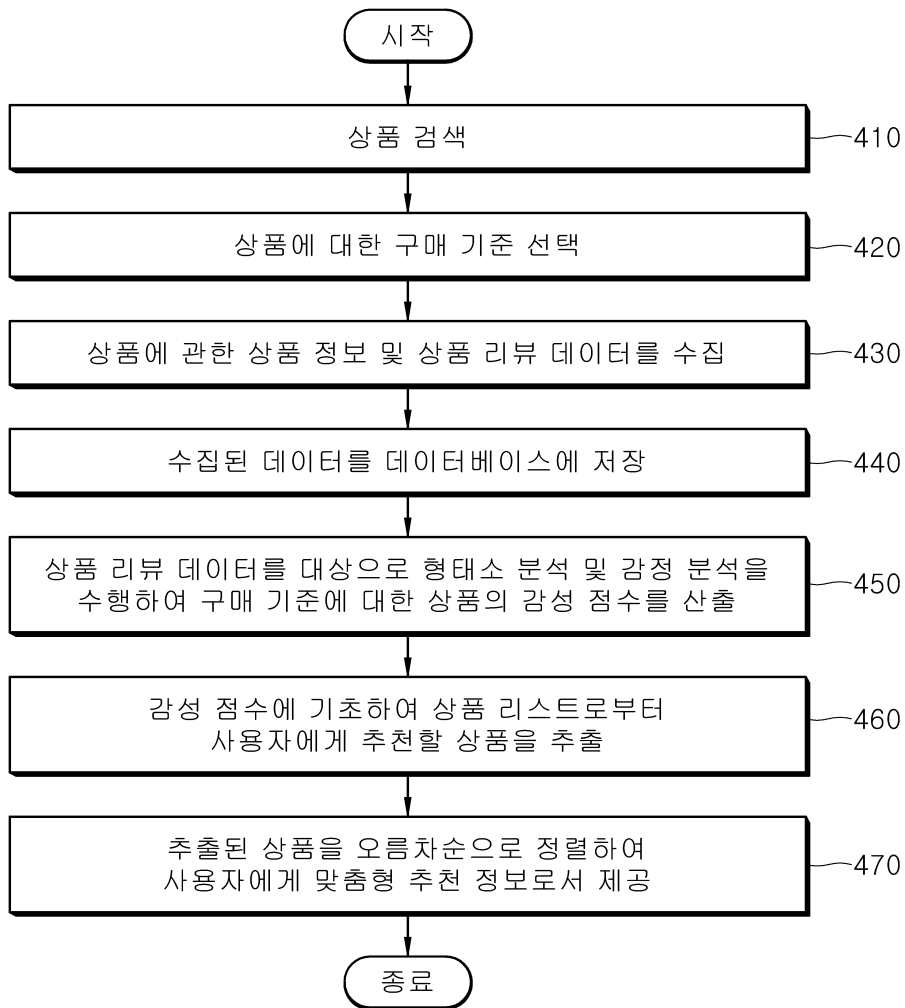
도면2



도면3

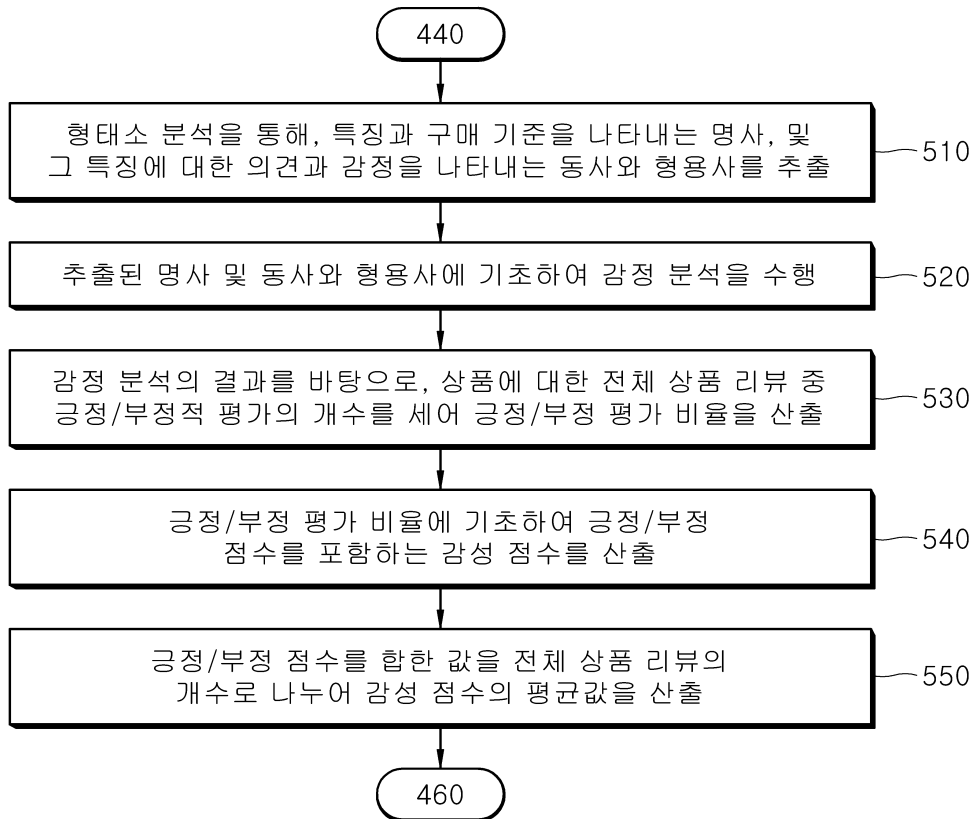


도면4



도면5

450



도면6

Twitter Korean Text	
Tag	Description
Noun	명사
Verb	동사
Adjective	형용사
Determiner	관형사(ex.새, 헌, 참,이, 그, 저)
Adverb	부사(ex.잘, 매우, 빨리, 반드시)
Conjunction	접속사
Exclamation	감탄사
Josa	조사(ex.의, 에서, 예)
PreEomi	선어말어미(ex. 었)
Eomi	어미(ex. 다, 요, 여)
Suffix	접미사
Punctuation	구두점
Foreign	외국어, 한자 및 기타기호
Alpha	알파벳
Unknown	미등록어

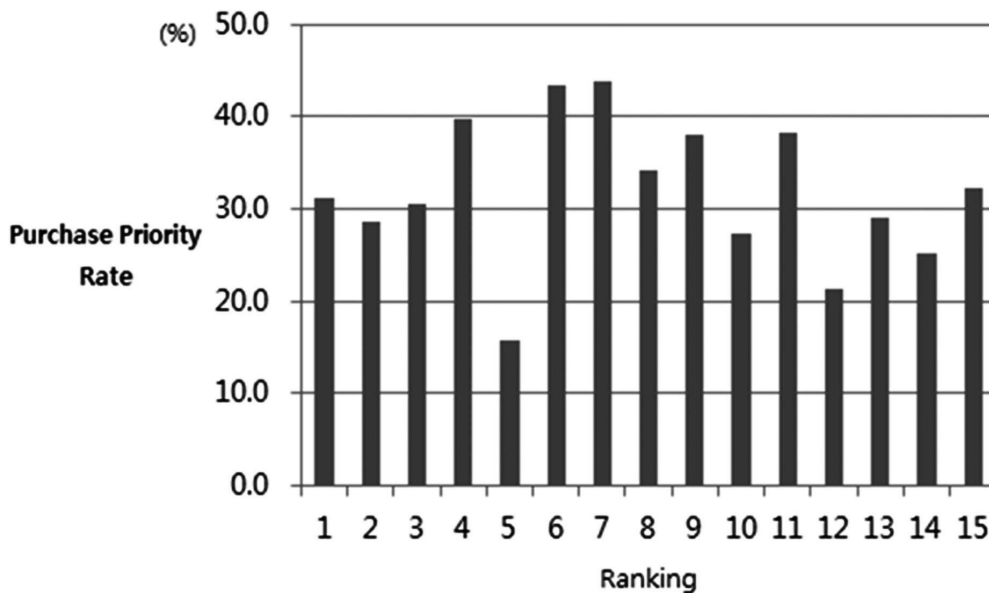
도면7

```
[('배송', 'Noun'), ('이', 'Josa'), ('완전', 'Noun'), ('빨랐어요', 'Adjective'),
('실물', 'Noun'), ('이', 'Josa'), ('생각', 'Noun'), ('보다', 'Josa'), ('꼬꼬매',
'Noun'), ('서', 'Josa'), ('제대로', 'Noun'), ('작동', 'Noun'), ('하러나',
'Verb'), ('했는데', 'Verb'), ('들자마자', 'Verb'), ('후끈하니', 'Adjective'),
('좋네요', 'Adjective'), ('잘쓸게요', 'Verb'), ('\n', 'Foreign')]
-----
[('배송', 'Noun'), ('포장', 'Noun'), ('잘', 'VerbPrefix'), ('왔어요', 'Verb'),
('생각', 'Noun'), ('보다', 'Josa'), ('작고', 'Noun'), ('제법', 'Noun'), ('열기',
'Noun'), ('가', 'Josa'), ('있네요', 'Adjective'), ('사무실', 'Noun'), ('자리',
'Noun'), ('에', 'Josa'), ('놓고', 'Verb'), ('사용', 'Noun'), ('하려고', 'Verb'),
('합니다', 'Verb'), ('\n', 'Foreign')]
-----
[('작', 'Noun'), ('지만', 'Josa'), ('따뜻해요', 'Adjective'), ('타이머', 'Noun'),
('가', 'Josa'), ('없고', 'Adjective'), ('조절', 'Noun'), ('이', 'Josa'),
('안되는건', 'Adjective'), ('아쉽지만', 'Adjective'), ('싼', 'Verb'), ('가격',
'Noun'), ('에서는', 'Josa'), ('좋아요', 'Adjective'), ('\n', 'Foreign')]
```

도면8

Ranking	Product Name	Price Sentiment Score Rate	Delivery Sentiment Score Rate	Total Score
1	2018 NEW 신일산업 SEH-G800 전기히터 미니히터	76.0	18	94.0
2	대용 2단 멀티히터 탁상용 전기난로/발난로 /난로/히터	72.7	20.45	93.2
3	스위스 밀리터리 서클레이터형 히터 (DWH-PTC4021M)	81.8	9.05	90.9
4	전기히터/전기난로/전기스토브 대용 히트 CZ-HT7007C	59.2	29.6	88.8
5	바름 전기히터 카본 난로 스토브 가정용 사무실 히터	66.6	22.2	88.8
6	스탠드히터 전기히터 전기난로 세라믹히터 난로 1490H	76.9	11.5	88.4
7	(hit) 신일산업 SEH-G800/전기히터/석영관 히터	55.5	25	80.5
8	대용 3단 석영관 전기히터/전기난로 CZ-6500KH(고급형)	62.5	15	77.5
9	미니히터 미니난로 전기히터 발히터 WINDPIA-1580HS X2	50	22.9	72.9
10	블랙앤데커 전기컨벡터 전기히터 온풍기 BXSH1801-A	40.0	32.5	72.5

도면9



도면10

