(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년12월28일
(11) 등록번호 10-1913591
(24) 등록일자 2018년10월25일

(51) 국제특허분류(Int. Cl.)
G06Q 30/02 (2012.01) *G06Q 40/06* (2012.01)
(52) CPC특허분류
G06Q 30/0255 (2013.01)
G06F 17/27 (2013.01)
(21) 출원번호 10-2017-0108158
(22) 출원일자 2017년08월25일
심사청구일자 2017년08월25일
(56) 선행기술조사문헌
KR1020170072644 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
권오병
서울특별시 강남구 삼성로 150, 107동 1001호(대치동, 한보미도맨션)
이재웅
서울특별시 강남구 논현로159길 57(신사동)
(74) 대리인
서재승

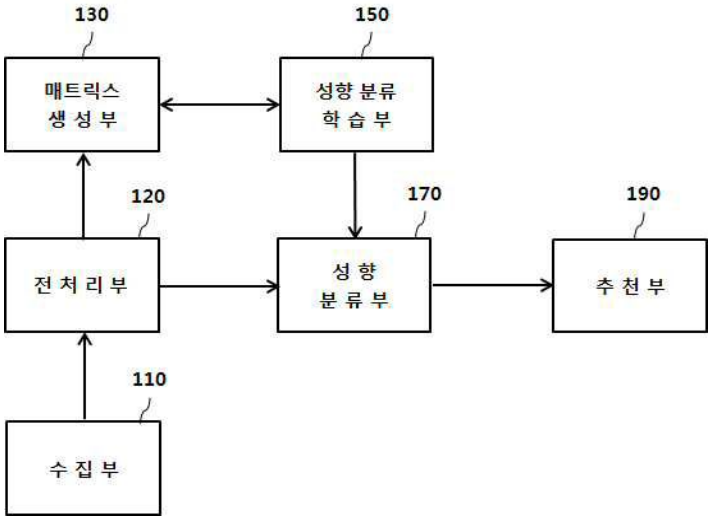
전체 청구항 수 : 총 10 항 심사관 : 이충근

(54) 발명의 명칭 **비정형 데이터를 정형화하여 금융상품을 추천하는 방법**

(57) 요 약

본 발명은 비정형 데이터를 정형화하여 금융 소비자에 금융 상품을 추천하는 방법에 관한 것으로, 보다 구체적으로 금융소비자가 작성한 문서나 금융소비자의 상담 내용에 포함된 텍스트 등과 같은 비정형의 상담 데이터로부터 금융소비자의 투자성향을 분류하기 위한 성향 분류 알고리즘을 생성하며, 금융소비자의 신규 투자성향 텍스트가 수집되는 경우 성향 분류 알고리즘을 통해 신규 투자성향 텍스트로부터 금융소비자가 가지는 투자성향을 자동 분류할 수 있는 금융상품의 추천 방법에 관한 것이다.

대 표 도 - 도2



(52) CPC특허분류
G06Q 40/06 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 2017-0-01122
부처명 미래창조과학부
연구관리전문기관 정보통신기술진흥센터
연구사업명 ICT유망기술개발지원(ICT융합핵심기술개발지원)
연구과제명 지능형 잇츠마이스토리 서비스 제공을 위한 언어 사용패턴 분석 기반 개인 프로파일링 및 개인 맞춤형 챗봇 기술 개발
기 여 율 1/1
주관기관 경희대 산학협력단
연구기간 2017.05.01 ~ 2018.12.31
공지예외적용 : 있음

명 세 서

청구범위

청구항 1

금융 소비자의 투자성향에 기초한 금융 상품의 추천 방법에 있어서,

투자성향 텍스트에서 명사를 추출하고, 추출한 명사별 가중치로 이루어진 텍스트-명사 매트릭스(Document-Term Matrix)를 생성하는 단계;

상기 텍스트-명사 매트릭스로부터 학습 데이터와 테스트 데이터를 구분하여 생성하고, 상기 학습 데이터와 상기 테스트 데이터로부터 투자성향을 분류하는데 이용되는 성향 분류 알고리즘을 생성하는 단계; 및

신규 투자성향 텍스트를 전처리하여 추출한 명사의 가중치를 상기 성향 분류 알고리즘에 적용하고 상기 신규 투자성향 텍스트가 속한 투자성향을 판단하는 단계를 포함하며,

상기 성향 분류 알고리즘은

포트폴리오에 기반한 기본요인에 대한 성향을 분류하는데 이용되는 기본요인 분류 알고리즘; 및

귀인이론에 기반한 내적요인에 대한 성향을 분류하는데 이용되는 내적요인 분류 알고리즘과 귀인이론에 기반한 외적요인에 대한 성향을 분류하는데 이용되는 외적요인 분류 알고리즘인 것을 특징으로 하는 금융상품의 추천 방법.

청구항 2

제 1 항에 있어서, 상기 금융 상품의 추천 방법은

상기 신규 투자성향 텍스트가 속한 투자성향에 기초하여 상기 투자성향에 매핑된 금융상품을 상기 금융 소비자에 추천하는 단계를 더 포함하는 것을 특징으로 하는 금융상품의 추천 방법.

청구항 3

제 2 항에 있어서, 상기 텍스트-명사 매트릭스를 생성하는 단계는

투자성향 텍스트를 형태소 분석기에 입력하여 상기 투자성향 텍스트를 구성하는 명사를 추출하는 단계;

추출한 명사에 대한 가중치를 계산하는 단계; 및

투자성향과 추출한 명사에 대한 가중치로 이루어진 텍스트-명사 매트릭스를 생성하는 단계를 포함하는 특징으로 하는 금융상품의 추천 방법.

청구항 4

삭제

청구항 5

제 3 항에 있어서, 상기 텍스트-명사 매트릭스를 생성하는 단계는

상기 투자성향 텍스트를 구성하는 명사 중 상기 투자성향 텍스트를 수집하는데 사용된 키워드를 제외하고 명사를 추출하는 것을 특징으로 하는 금융상품의 추천 방법.

청구항 6

제 3 항에 있어서, 상기 추출한 명사에 대한 가중치는

단위 수의 투자성향 텍스트 중 상기 명사가 존재하는 투자성향 텍스트의 수의 비율과 상기 투자성향 텍스트에 상기 추출한 명사가 나타난 빈도로부터 계산되는 것을 특징으로 하는 금융상품의 추천 방법.

청구항 7

제 6 항에 있어서, 상기 신규 투자성향 텍스트가 속하는 투자성향을 판단하는 단계

상기 신규 투자성향 텍스트를 형태소 분석기에 입력하여 상기 신규 투자성향 텍스트를 구성하는 명사를 추출하는 단계;

상기 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 계산하는 단계; 및

상기 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 상기 성향 분류 알고리즘에 적용하여 상기 신규 투자성향 텍스트가 속하는 투자성향을 판단하는 단계를 포함하는 것을 특징으로 하는 금융상품의 추천 방법.

청구항 8

제 3 항에 있어서, 상기 신규 투자성향 텍스트가 속하는 투자성향을 판단하는 단계는

상기 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 상기 기본요인 분류 알고리즘에 적용하여 상기 신규 투자성향 텍스트가 속하는 기본요인을 판단하는 단계;

상기 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 상기 내적요인 분류 알고리즘에 적용하여 상기 신규 투자성향 텍스트가 속하는 내적요인을 판단하는 단계;

상기 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 상기 외적요인 분류 알고리즘에 적용하여 상기 신규 투자성향 텍스트가 속하는 외적요인을 판단하는 단계; 및

상기 신규 투자성향 텍스트가 속하는 기본요인, 내적요인 및 외적요인의 조합으로부터 상기 신규 투자성향 텍스트가 속하는 투자성향을 판단하는 단계를 포함하는 것을 특징으로 하는 금융상품의 추천 방법.

청구항 9

제 1 항 또는 제 8 항에 있어서, 상기 기본요인은 수익성과 안정성인 것을 특징으로 하는 금융상품의 추천 방법.

청구항 10

제 1 항 또는 제 8 항에 있어서, 상기 내적요인은 경험성과 정보성인 것을 특징으로 하는 금융상품의 추천 방법.

청구항 11

제 1 항 또는 제 8 항에 있어서, 상기 외적요인은 반응성과 신뢰성인 것을 특징으로 하는 금융상품의 추천 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 비정형 데이터를 정형화하여 금융 소비자에 금융 상품을 추천하는 방법에 관한 것으로, 보다 구체적으로 금융소비자가 작성한 문서나 금융소비자의 상담 내용에 포함된 텍스트 등과 같은 비정형의 상담 데이터로부터 금융소비자의 투자성향을 분류하기 위한 성향 분류 알고리즘을 생성하며, 금융소비자의 신규 투자성향 텍스트가 수집되는 경우 성향 분류 알고리즘을 통해 신규 투자성향 텍스트로부터 금융소비자가 가지는 투자성향을 자동 분류할 수 있으며, 분류한 투자성향에 따라 금융소비자에 투자성향에 해당하는 금융상품을 추천할 수 있는 금융상품의 추천 방법에 관한 것이다.

배 경 기 술

[0002] 새로운 디지털 비즈니스 전략으로서 핀테크(Fintech)가 새롭고 가치 있는 재무금융서비스로 등장하고 있다. 핀테크는 전자결재부터 자산관리까지, 또한 개인간 금융거래부터 크라우드펀딩까지 아우르는 금융 산업을 재편할 수 있는 차세대 재무적 기술이며, 전통적인 재무금융서비스로부터 제공되는 재품 및 서비스를 혁신하는 재무금융서비스와 지능정보기술을 동시에 가진 기술 중심의 스타트업으로 정의되기도 한다.

[0003] 재무금융서비스 고객들에 대한 데이터가 축적되면서 은행들은 새로운 고객 경험을 제공하고 데이터 분석에 기반한 새로운 가치를 창출하려고 한다. 이러한 새로운 가치 중 하나는 새로운 투자 고객의 개발이다. 전통적이고 고비용의 고객관리 방법을 보완하기 위해 소셜미디어(Social Media) 기술을 활용하여 고객에게 가치가 있는 정보를 소비자들 스스로 창출하고 유통하게 함으로써 롱테일(long-tail)에 위치한 고객을 유지하거나 새로운 고객 개발 비용을 효율화한 것처럼, 핀테크, 특히 에이전트 기반의 디지털 어드바이저(digital advisor) 또는 로보 어드바이저(robo-advisor)도 기존의 고비용적 고객 투자 상담 외에 롱테일 군에 속하는 고객들에게 비용 효율적으로 고객의 투자성향을 파악하고 투자 상품을 추천하는 노력을 경감시킬 수 있다.

[0004] 이를 위해 고객들이 자발적으로 남긴 데이터를 활용할 수 있을 것이다. 인터넷 상에 공유된 각종 투자 상품관련 오피니언이나 고객 상담 내용 중 VOC(Voice Of Customer)로 인식될 수 있는 자료들이 그 일부이다. 그러나 이들은 정형화되거나 투자의향을 파악할 목적으로 사전에 준비된 유형으로 준비되어 있지 않다.

[0005] 더욱이 핀테크에서는 비대면 의사소통에 대응해야 한다. 따라서 직접 접촉하여 투자자의 성향을 파악하는 것이 거의 불가능하고 투자자가 남긴 음성 정보 혹은 텍스트 정보를 통해 간접적으로 투자자의 성향을 추론해야 한다.

[0006] 금융상품의 복잡한 구조와 전문성, 시간성, 특수성 등의 금융상품 고유특성과 수익성에 대하여 금융소비자는 금융상품의 탐색과 이해 및 평가 등에서 어려움을 겪는다. 이러한 금융소비자의 텍스트로부터 추출할 수 있는 많은 정보 중에 유용한 정보의 하나가 금융소비자의 투자성향이다.

[0007] 금융소비자가 작성한 문서의 텍스트로부터 금융소비자의 투자성향을 파악함으로써, 금융상품을 선택하려는 고객의 요구사항을 패턴으로 분석하여 고객에게 가장 적절한 금융상품을 실시간으로 설계 및 제공할 수 있는 자동화된 시스템에 대한 요구가 있다.

[0008] 본원발명의 배경기술로 한국공개특허 제10-2012-0109943호가 있다.

발명의 내용

해결하려는 과제

[0009] 본 발명이 이루고자 하는 목적은 금융소비자가 작성한 문서나 금융소비자의 상담 내용에 포함된 텍스트 등과 같은 비정형의 상담 데이터로부터 금융소비자의 투자성향을 판단하고 판단한 투자성향에 따라 금융상품을 추천할 수 있는 방법을 제공하는 것이다.

[0010] 본 발명이 이루고자 하는 다른 목적은 종래 다른 금융소비자가 작성한 투자성향 텍스트로부터 투자성향을 분류하기 위한 성향 분류 알고리즘을 생성하며, 신규 투자성향 텍스트가 수집되는 경우 성향 분류 알고리즘을 통해 신규 투자성향 텍스트가 가지는 투자성향을 자동 분류할 수 있는 금융상품의 추천 방법을 제공하는 것이다.

과제의 해결 수단

[0011] 본 발명의 목적을 달성하기 위하여 본 발명에 따른 금융 상품의 추천 방법은 투자성향 텍스트에서 명사를 추출하고 추출한 명사별 가중치로 이루어진 텍스트-명사 매트릭스(Document-Term Matrix)를 생성하는 단계와, 텍스트-명사 매트릭스로부터 학습 데이터와 테스트 데이터를 구분하여 생성하고 학습 데이터와 테스트 데이터로부터 투자성향을 분류하는데 이용되는 성향 분류 알고리즘을 생성하는 단계와, 신규 투자성향 텍스트를 전처리하여 추출한 명사의 가중치를 성향 분류 알고리즘에 적용하고 신규 투자성향 텍스트가 속한 투자성향을 판단하는 단계를 포함하는 것을 특징으로 한다.

[0012] 바람직하게, 본 발명에 따른 금융 상품의 추천 방법은 신규 투자성향 텍스트가 속한 투자성향에 기초하여 투자성향에 매핑된 금융상품을 금융 소비자에 추천하는 단계를 더 포함하는 것을 특징으로 한다.

[0013] 바람직하게, 본 발명의 일 실시예에서 텍스트-명사 매트릭스를 생성하는 단계는 투자성향 텍스트를 형태소 분석기에 입력하여 상기 투자성향 텍스트를 구성하는 명사를 추출하는 단계와, 추출한 명사에 대한 가중치를 계산하는 단계와, 투자성향과 추출한 명사에 대한 가중치로 이루어진 텍스트-명사 매트릭스를 생성하는 단계를 포함하는 특징으로 한다.

[0014] 여기서 성향 분류 알고리즘은 포트폴리오에 기반한 기본요인에 대한 성향을 분류하는데 이용되는 기본요인 분류 알고리즘과, 귀인이론에 기반한 내적요인에 대한 성향을 분류하는데 이용되는 내적요인 분류 알고리즘과 귀인이

론에 기반한 외적요인에 대한 성향을 분류하는데 이용되는 외적요인 분류 알고리즘인 것을 특징으로 한다.

[0015] 바람직하게, 본 발명의 일 실시예에서 텍스트-명사 매트릭스를 생성하는 단계는 투자성향 텍스트를 구성하는 명사 중 투자성향 텍스트를 수집하는데 사용된 키워드를 제외하고 명사를 추출하는 것을 특징으로 한다.

[0016] 여기서 추출한 명사에 대한 가중치는 단위 수의 투자성향 텍스트 중 명사가 존재하는 투자성향 텍스트의 수의 비율과 투자성향 텍스트에 추출한 명사가 나타난 빈도로부터 계산되는 것을 특징으로 한다.

[0017] 바람직하게, 본 발명의 일 실시예에서 신규 투자성향 텍스트가 속하는 투자성향을 판단하는 단계 신규 투자성향 텍스트를 형태소 분석기에 입력하여 신규 투자성향 텍스트를 구성하는 명사를 추출하는 단계와, 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 계산하는 단계와, 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 성향 분류 알고리즘에 적용하여 신규 투자성향 텍스트가 속하는 투자성향을 판단하는 단계를 포함하는 것을 특징으로 한다.

[0018] 바람직하게, 본 발명의 일 실시예에서 신규 투자성향 텍스트가 속하는 투자성향을 판단하는 단계는 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 기본요인 분류 알고리즘에 적용하여 신규 투자성향 텍스트가 속하는 기본요인을 판단하는 단계와, 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 내적요인 분류 알고리즘에 적용하여 신규 투자성향 텍스트가 속하는 내적요인을 판단하는 단계와, 신규 투자성향 텍스트로부터 추출한 명사에 대한 가중치를 외적요인 분류 알고리즘에 적용하여 신규 투자성향 텍스트가 속하는 외적요인을 판단하는 단계와, 신규 투자성향 텍스트가 속하는 기본요인, 내적요인 및 외적요인의 조합으로부터 신규 투자성향 텍스트가 속하는 투자성향을 판단하는 단계를 포함하는 것을 특징으로 한다.

[0019] 여기서 기본요인은 수익성과 안정성인 것을 특징으로 하며, 내적요인은 경험성과 정보성인 것을 특징으로 하며, 외적요인은 반응성과 신뢰성인 것을 특징으로 한다.

발명의 효과

[0020] 본 발명에 따른 금융상품의 추천 방법은 다음과 같은 효과를 가진다.

[0021] 본 발명에 따른 금융상품의 추천 방법은 금융소비자가 작성한 문서나 금융소비자의 상담 내용에 포함된 텍스트 등과 같은 비정형의 상담 데이터로부터 금융소비자의 투자성향을 판단할 수 있는 정형 데이터, 즉 텍스트-명사 매트릭스를 생성함으로써, 금융소비자의 비정형 상담 데이터로부터 금융소비자에 맞춤형 금융상품을 추천할 수 있다.

[0022] 한편 본 발명에 따른 금융상품의 추천 방법은 종래 다른 금융소비자가 작성한 투자성향 텍스트로부터 투자성향을 분류하기 위한 성향 분류 알고리즘을 생성하며 신규 투자성향 텍스트가 수집되는 경우 성향 분류 알고리즘을 통해 신규 투자성향 텍스트가 가지는 투자성향을 자동 분류함으로써, 전문적이고 복잡한 금융상품을 금융소비자의 투자성향에 맞추어 추천할 수 있다.

도면의 간단한 설명

[0023] 도 1은 본 발명에 따른 금융상품의 추천 시스템을 설명하기 위한 기능 블록도이다.

도 2는 본 발명의 일 실시예에 따른 추천 서버를 설명하기 위한 기능 블록도이다.

도 3은 본 발명에 따른 전처리부의 일 예를 설명하기 위한 기능 블록도이다.

도 4는 본 발명의 일 실시예에 따른 금융상품의 추천 방법을 설명하기 위한 흐름도이다.

도 5는 텍스트-명사 매트릭스를 생성하는 단계의 일 예를 설명하기 위한 흐름도이다.

도 6은 데이터베이스부에서 수집한 투자성향 텍스트를 형태소 분석하여 추출한 명사의 일 예를 도시하고 있다.

도 7은 다수의 투자성향 텍스트로부터 생성되는 텍스트-명사 매트릭스의 일 예를 도시하고 있다.

발명을 실시하기 위한 구체적인 내용

[0024] 본 발명에서 사용되는 기술적 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아님을 유의해야 한다. 또한, 본 발명에서 사용되는 기술적 용어는 본 발명에서 특별히 다른 의미로 정의되지 않는 한, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 의미로 해석되어야 하며, 과도하게 포괄적인 의미로 해석되거나, 과도하게 축소된 의미로 해석되지 않아야 한다.

또한, 본 발명에서 사용되는 기술적인 용어가 본 발명의 사상을 정확하게 표현하지 못하는 잘못된 기술적 용어일 때에는, 당업자가 올바르게 이해할 수 있는 기술적 용어로 대체되어 이해되어야 할 것이다.

[0025] 또한, 본 발명에서 사용되는 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한 복수의 표현을 포함한다. 본 발명에서, "구성된다" 또는 "포함한다" 등의 용어는 발명에 기재된 여러 구성 요소들, 또는 여러 단계를 반드시 모두 포함하는 것으로 해석되지 않아야 하며, 그 중 일부 구성 요소들 또는 일부 단계들은 포함되지 않을 수도 있고, 또는 추가적인 구성 요소 또는 단계들을 더 포함할 수 있는 것으로 해석되어야 한다.

[0026] 또한, 첨부된 도면은 본 발명의 사상을 쉽게 이해할 수 있도록 하기 위한 것일 뿐, 첨부된 도면에 의해 본 발명의 사상이 제한되는 것으로 해석되어서는 아니 됨을 유의해야 한다.

[0027] 이하 첨부한 도면을 참고로 본 발명에 따른 인문 사상의 분류 방법에 대해 보다 구체적으로 살펴본다.

[0028] 도 1은 본 발명에 따른 금융상품의 추천 시스템을 설명하기 위한 기능 블록도이다.

[0029] 도 1을 참고로 보다 구체적으로 살펴보면, 네트워크(30)에는 데이터베이스부(10), 추천 서버(50), 상담자 단말기(70) 및 사용자 단말기(90)가 접속되어 있다.

[0030] 데이터베이스부(10)에는 금융소비자가 금융기관과 상담하며 남긴 상담 내용, 금융소비자가 소정 금융상품과 관련하여 남긴 사용자의 의견, 금융소비자가 투자하고자 하는 금융상품에 대한 의견 등의 다양한 종류의 투자성향 텍스트가 저장되어 있다. 여기서 상담 내용은 금융소비자와 금융기관의 상담자 사이에 통화 내역 또는 상담 내역 등이 텍스트로 생성되어 저장되어 있다. 한편, 투자성향 텍스트는 금융소비자의 투자성향을 판단할 수 있는 다양한 종류의 텍스트가 사용될 수 있는데 투자성향 텍스트는 대면 상담 과정 또는 비대면 상담 과정에서 생성된 상담 텍스트 또는 온라인상에 남긴 다양한 종류의 텍스트가 사용될 수 있다. 바람직하게, 데이터베이스부(10)에 저장되어 있는 투자성향 텍스트는 투자성향에 대한 기본요인, 내적요인 또는 외적요인 등의 투자성향요인으로 구분되어 저장될 수 있다.

[0031] 추천 서버(50)는 네트워크(30)를 통해 데이터베이스부(10)로부터 투자성향 텍스트를 수집하며, 수집한 투자성향 텍스트로부터 투자성향을 분류하기 위한 성향 분류 알고리즘을 학습 생성하는데, 데이터베이스부(10)로부터 신규 투자성향 텍스트를 수집하거나 또는 네트워크(30)를 통해 다른 소스로 신규 투자성향 텍스트를 수집하는 경우 성향 분류 알고리즘에 기초하여 신규 투자성향 텍스트에 나타난 금융소비자의 투자성향을 분류하거나 분류한 투자성향에 상응하는 금융상품을 검색하여 금융소비자 또는 상담자에 검색한 금융상품을 추천 금융상품으로 제공한다.

[0032] 즉, 추천 서버(50)는 네트워크(30)를 통해 추천 금융상품에 대한 정보를 금융기관에 종사하는 상담자가 사용하는 상담자 단말기(70)로 제공하거나, 금융소비자가 사용하는 사용자 단말기(90)로 제공한다.

[0033] 도 2는 본 발명의 일 실시예에 따른 추천 서버를 설명하기 위한 기능 블록도이다.

[0034] 도 2를 참고로 보다 구체적으로 살펴보면, 수집부(110)는 입력된 키워드에 기초하여 데이터베이스부 또는 다른 소스를 통해 투자성향 텍스트를 수집한다. 바람직하게, 수집부(110)는 네트워크를 통해 데이터를 송수신하기 위한 통신 모듈을 구비하며, 수집부(110)는 네트워크를 통해 데이터베이스부로부터 또는 다른 소스로부터 입력된 키워드에 해당하는 투자성향 텍스트를 수집한다.

[0035] 전처리부(120)는 투자성향 텍스트를 형태소 분석기에 입력하여 투자성향 텍스트에서 명사를 추출하고 추출한 명사별 가중치를 계산하며, 매트릭스 생성부(130)는 이미 알고 있는 투자성향요인과 각 투자성향 텍스트에서 추출한 명사의 가중치로 이루어진 텍스트-명사 매트릭스(Document-term matrix)를 생성한다.

[0036] 성향 분류 학습부(150)는 매트릭스 생성부(130)에서 생성된 텍스트-명사 매트릭스로부터 트레이닝 데이터와 테스트 데이터를 추출하고 추출한 트레이닝 데이터와 테스트 데이터로부터 금융소비자의 투자성향을 판단하는데 이용되는 성향 분류 알고리즘을 생성한다. 여기서 성향 분류 학습부(150)에서 사용되는 텍스트-명사 매트릭스는 투자성향에 대한 투자성향요인을 이미 알고 있는 텍스트와 각 테스트에서 추출한 명사의 가중치로 이루어지며, 트레이닝 데이터와 테스트 데이터는 텍스트-명사 매트릭스에서 임의적으로 구분하여 추출할 수 있다.

[0037] 한편, 수집부(110)를 통해 신규 투자성향 텍스트가 수집되는 경우, 전처리부(120)는 신규 투자성향 텍스트를 형태소 분석기에 입력하여 신규 투자성향 텍스트에서 명사를 추출하고 추출한 명사의 가중치를 계산한다. 여기서 신규 투자성향 텍스트는 금융상품을 추천하고자 하는 금융소비자가 작성한 투자성향 텍스트를 의미한다.

[0038] 신규 투자성향 텍스트의 명사별 가중치는 성향 분류부(170)로 제공되는데, 성향 분류부(170)는 신규 투자성향

텍스트의 명사별 가중치를 생성한 성향 분류 알고리즘에 적용하여 신규 투자성향 텍스트가 속하는 투자성향을 판단한다. 수집부(110)는 적어도 1개 이상의 신규 투자성향 텍스트를 수집할 수 있으며, 성향 분류부(170)는 적어도 1개 이상의 신규 투자성향 텍스트를 성향 분류 알고리즘에 적용하여 각 신규 투자성향 텍스트가 속하는 투자성향을 판단할 수 있다.

[0039] 바람직하게, 다수의 신규 투자성향 텍스트를 성향 분류 알고리즘에 적용하여 각 신규 투자성향 텍스트가 서로 상이한 투자성향을 가지는 것으로 판단되는 경우, 다수의 신규 투자성향 텍스트에서 가장 많이 동일한 투자성향으로 판단되는 투자성향을 금융소비자의 투자성향으로 판단할 수 있다.

[0040] 한편, 추천부(190)는 성향 분류부(170)에서 판단한 신규 투자성향 텍스트가 속하는 투자성향에 기초하여 투자성향에 상응하는 금융상품을 금융상품 데이터베이스부에서 검색하고, 검색한 금융상품을 금융소비자의 추천 금융상품으로 상담자 단말기 또는 사용자 단말기로 제공한다.

[0041] 바람직하게, 성향 분류 학습부(150)는 기본요인, 내적요인 및 외적요인의 투자성향요인별로 성향 분류 알고리즘을 생성할 수 있는데, 성향 분류부(170)에서는 기본요인, 외적요인 및 내적요인으로 구분하여 투자성향을 판단할 수 있으며, 추천부(190)는 기본요인, 외적요인 및 내적요인으로 이루어진 투자성향패턴에 상응하는 금융상품을 금융상품 데이터베이스부에서 검색할 수 있다.

[0042] 도 3은 본 발명에 따른 전처리부의 일 예를 설명하기 위한 기능 블록도이다.

[0043] 도 3을 참고로 보다 구체적으로 살펴보면, 형태소 분석부(121)는 수집한 투자성향 텍스트로부터 명사를 추출하는데, 여기서 형태소는 뜻을 가진 가장 작은 말의 단위로 형태소 분석 방식은 규칙기반 형태 분석 방식과 음절 단위 형태 분석 방식 등 다양한 형태소 분석 알고리즘이 사용되고 있으며, 이에 대한 자세한 설명은 생략한다.

[0044] 형태소 분석부(121)를 통해 투자성향 텍스트로부터 추출한 명사에서 투자성향 텍스트를 수집하는데 사용된 키워드를 제거한다. 추출한 명사에서 투자성향 텍스트를 수집하는데 사용된 키워드를 제거함으로써, 수집한 투자성향 텍스트로부터 생성되는 성향 분류 알고리즘이 추출한 명사에 과적합(overfitting)되는 것을 방지할 수 있다.

[0045] 가중치 계산부(125)는 추출한 명사가 투자성향 텍스트에서 가지는 가중치를 계산하는데, 추출한 명사의 가중치는 단위 수의 투자성향 텍스트 중 해당 명사가 존재하는 투자성향 텍스트 수의 비율, 해당 투자성향 텍스트에서 추출한 명사가 나타난 빈도로부터 계산될 수 있다.

[0046] 도 4는 본 발명의 일 실시예에 따른 금융상품의 추천 방법을 설명하기 위한 흐름도이다.

[0047] 도 4를 참고로 보다 구체적으로 살펴보면, 데이터베이스부에 저장되어 있는 투자성향 텍스트를 수집하고 수집한 투자성향 텍스트에서 명사를 추출하여 추출한 명사의 가중치로 이루어진 텍스트-명사 매트릭스를 생성한다(S110).

[0048] 생성한 텍스트-명사 매트릭스로부터 트레이닝 데이터와 테스트 데이터를 분류하여 추출하고, 추출한 트레이닝 데이터와 테스트 데이터로부터 성향 분류 알고리즘을 생성한다(S130). 본 발명이 적용되는 분야에 따라 성향 분류 알고리즘을 생성하기 위한 다양한 종류의 학습 알고리즘이 사용될 수 있는데, 결정 트리, Deep Neural Network, k-NN classifier, Multinomial logistic regression, Naive Bayes, Random Forest, SVM 등의 학습 알고리즘을 사용할 수 있으며, 이에 대한 자세한 설명은 생략한다.

[0049] 바람직하게, 금융소비자의 투자성향을 판단하는데 포트폴리오 이론에 기반한 기본요인, 귀인이론에 기반한 내적요인과 외적요인 등이 투자성향요인으로 사용될 수 있다. 기본요인에 대한 투자성향을 판단하는데 사용할 수 있는 성향 분류 알고리즘, 내적요인에 대한 투자성향을 판단하는데 사용할 수 있는 성향 분류 알고리즘 및 외적요인에 대한 투자성향을 판단하는데 사용할 수 있는 성향 분류 알고리즘을 각각 생성할 수 있다.

[0050] 포트폴리오 이론에 기반한 기본요인은 투자의 기본으로 생각되는 투자성향요인으로, 투자에 있어 가장 기본적이고 중요한 함수는 투자의 기대수익과 위험이 될 것이다. 즉, 투자에 따른 안정성과 수익성을 기본요인으로 사용할 수 있는데, 일반적으로 금융투자상품의 선택 목적은 이익극대화 및 손실최소화에 있다. 수익과 위험은 현대 투자론의 이론인 평균-분산 포트폴리오 분석이라는 투자행동의 기본적인 규범적 모델에서 지적하고 있듯이 투자 대안을 선택하는데 있어 핵심적인 요인이다. 투자를 성공적으로 하기 위해서는 투자대상의 기대수익률과 위험을 동시에 고려하여 최적의 투자대안을 선택하는 것이 필요하다.

[0051] 한편, 귀인이론에 따른 내적요인은 경험성과 정보성으로,

[0052] 경험성(experientiality)은 과거에 겪었던 투자경험이 현재 또는 미래의 투자의 의사결정에 영향을 미칠 수 있

다는 특성이다. 예를 들면, A사에 판매했던 펀드를 가입했던 투자자가 큰 손실을 경험했을 경우 향우 펀드라는 금융상품을 불신하여 펀드가입을 꺼려할 수도 있고, 이러한 금융 소비자는 투자성향이 보다 위험회피형 투자자로 변모할 수 있다. 경험성은 객관화하기 어려운 척도이고, 주관적인 내적 척도로 구분할 수 있는 근거가 된다.

[0053] 정보성(informativeness)은 금융회사가 금융소비자라고 볼 수 있는 투자자들이 이용가능한 정보를 제공하는 정도로 정의한다. 투자자로 통칭되는 금융소비자는 넓게 금융거래에서 금융업자와 금융계약을 체결하는 상대방으로 정의되고 있는데, 금융시장에서 정보의 비대칭성이 존재하고 거래상대방과 충분한 교섭력을 가지지 못하기 때문에 정보성은 또 다른 측면에서 금융소비자의 보호와 관련이 있다.

[0054] 금융상품은 상당히 복잡하여 설명이 난해한 특성도 정보의 비대칭적인 성향을 부각시킨다. 다시 말하면, 전문성으로 요약할 수 있을 것이다. 금융상품에 대한 기초지식이 있는 경우는 정보에 우위에 있다면, 투자의 실행이 용이하지만, 금융에 대한 기초지식이 없다면, 사실상 감(feeling)이나 경험에 의존한 투자를 실행할 수밖에 없다.

[0055] 귀인이론에 따른 외적요인이라는 것은 외부적인 요소가 투자결정에 미칠 수 있는 요인을 말하는 것으로, 귀인이론에 따른 다양한 외적 요인 가운데 거시경제변수 등과 같은 거시적인 경제요인을 중요시하며 반응하는 반응성과 자산을 운용하는 운용회사나 판매회사에 대한 일반적인 평판에 대한 신뢰성을 외적요인으로 사용할 수 있다.

[0056] 거시 경제 변수의 움직임은 투자를 실행하는데 있어 중요한 요인임을 시사한다. 반응성은 이러한 불확실성 하에 나타나는 통제 불가능한 거시경제변화에 따른 금융소비자의 반응을 의미한다.

[0057] 신뢰성은 금융상품을 판매하거나 운용하는 회사의 명성이나 이미지, 무형자산 등을 투자자가 신뢰할 만한지에 대한 객관적인 것을 의미한다. 기업의 명성이 높은 경우, 기업의 경영활동에서 실질적으로 나타나는 주가상승과 매출증가 등의 경제적인 이득뿐만 아니라 경제적인 수치로 환산하기는 어렵지만 기업의 생존에 장기적인 이득을 가져다준다. 예를 들어, 금융 소비자는 평소 높은 명성을 가진 기업의 부정적인 신문보도를 접할 경우, 그렇지 않은 기업들의 부정적인 신문보도를 접할 경우보다 덜 부정적인 태도를 형성하는 것으로 나타났다.

[0058] 신규 투자성향 텍스트가 수집되는 경우(S150), 신규 투자성향 텍스트를 전처리하여 신규 투자성향 텍스트에서 명사를 추출하고 추출한 명사의 가중치를 계산한다(S160). 추출한 명사의 가중치를 생성한 성향 분류 알고리즘에 적용하여 신규 투자성향 텍스트가 속하는 투자성향을 판단한다(S170). 여기서 신규 투자성향 텍스트를 기본요인을 판단하는데 사용되는 성향 분류 알고리즘에 적용하여 금융소비자의 기본요인에 대한 투자성향이 안정성을 추구하는지 아니면 수익성을 추구하는지 판단할 수 있다. 또한, 신규 투자성향 텍스트를 외적요인을 판단하는데 사용되는 성향 분류 알고리즘에 적용하여 금융소비자의 내적요인에 대한 투자성향이 경험성을 추구하는지 아니면 정보성을 추구하는지 판단할 수 있다. 또한, 신규 투자성향 텍스트를 외적요인을 판단하는데 사용되는 성향 분류 알고리즘에 적용하여 금융소비자의 외적요인에 대한 투자성향이 반응성을 추구하는지 아니면 신뢰성을 추구하는지 판단할 수 있다.

[0059] 이와 같이 각 투자성향요인에 대한 판단 결과에 기초하여, 기본요인, 외적요인 및 내적요인으로 이루어진 투자성향 패턴을 생성할 수 있으며, 이러한 투자성향 패턴에 매핑되어 있는 금융상품을 금융상품 데이터베이스부에서 검색하고, 검색한 금융상품을 금융소비자에 추천할 금융상품으로 사용자 단말기 또는 상담자 단말기로 제공한다(S190).

[0060] 도 5는 텍스트-명사 매트릭스를 생성하는 단계의 일 예를 설명하기 위한 흐름도이다.

[0061] 도 5를 참고로 보다 구체적으로 살펴보면, 데이터베이스부에서 수집한 투자성향 텍스트를 형태소 분석하여 투자성향 텍스트를 구성하는 명사를 추출한다(S111). 도 6은 데이터베이스부에서 수집한 투자성향 텍스트를 형태소 분석하여 추출한 명사의 일 예를 도시하고 있다. 데이터베이스부에서 수집한 투자성향 텍스트의 경우 해당 투자성향 텍스트의 투자성향요인을 이미 알고 있으며, 투자성향 텍스트를 구성하는 명사를 추출한다.

[0062] 추출한 명사에서 해당 투자성향 텍스트를 수집하는데 이용한 키워드를 제거한다(S113). 추출한 명사에서 해당 투자성향 텍스트를 수집하는데 이용한 키워드를 제거하는 것은 투자성향 텍스트에서 추출한 명사를 이용하여 성향 분류 알고리즘을 생성시 과적합의 문제가 발생하는 것을 미연에 방지하기 위한 것으로, 예를 들어 '안정성'를 키워드로 투자성향 텍스트를 수집하는 경우 수집한 투자성향 텍스트에서 추출한 명사 중 키워드로 사용한 '안정성'를 제거하여 성향 분류 알고리즘에서 과적합 문제가 발생하는 것을 방지한다.

[0063] 추출한 각 명사가 해당 투자성향 텍스트에서 가지는 가중치를 계산한다(S115). 여기서 가중치는 동일한 투자성

향요인으로 분류된 단위 수의 투자성향 텍스트 중 해당 명사가 존재하는 투자성향 텍스트의 수의 비율과 해당 투자성향 텍스트에 해당 명사가 나타난 빈도로부터 계산되는데, 바람직하게 각 명사(i)의 가중치(w_i)는 아래의 수학식 (1)과 같이 계산될 수 있다.

[0064] [수학식 1]

[0065]
$$w_i = \left(\frac{n_i}{N}\right)^{-1} \times f_i$$

[0066] 여기서 w_i는 명사 i의 가중치이며, N은 단위 수의 투자성향 텍스트, n_i은 단위 수의 투자성향 텍스트에서 해당 명사(i)가 존재하는 문언의 수, f_i는 해당 투자성향 텍스트에서 명사(i)가 나타난 빈도를 의미한다.

[0067] 이와 같이 단위 수의 투자성향 텍스트 중 해당 명사가 존재하는 투자성향 텍스트의 수의 비율에 대한 역수로부터 명사의 가중치를 계산함으로써, 추출한 명사가 해당 투자성향 텍스트가 속하는 투자성향요인과 관련하여 중요한 명사인지를 판단할 수 있다.

[0068] 투자성향 텍스트를 구성하는 각 명사의 가중치와 해당 투자성향 텍스트의 투자성향요인으로 이루어진 텍스트-명사 매트릭스를 생성한다(S117).

[0069] 도 7은 다수의 투자성향 텍스트로부터 생성되는 텍스트-명사 매트릭스의 일 예로 텍스트-명사 매트릭스에는 수집한 투자성향 텍스트의 각 명사에 대한 가중치와 각 투자성향 텍스트가 속하는 투자성향요인이 매트릭스 형태로 생성된다.

[0070] 한편, 상술한 본 발명의 실시 예들은 컴퓨터에서 실행될 수 있는 프로그램으로 작성 가능하고, 컴퓨터로 읽을 수 있는 기록 매체를 이용하여 상기 프로그램을 동작시키는 범용 디지털 컴퓨터에서 구현될 수 있다.

[0071] 상기 컴퓨터로 읽을 수 있는 기록 매체는 마그네틱 저장 매체(예를 들어, 롬, 플로피 디스크, 하드디스크 등), 광학적 판독 매체(예를 들면, 시디롬, 디브이디 등) 및 캐리어 웨이브(예를 들면, 인터넷을 통한 전송)와 같은 저장 매체를 포함한다.

[0072] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

[0073]
10: 데이터베이스부 30: 네트워크

50: 추천 서버 70: 상담자 단말기

90: 사용자 단말기 110: 수집부

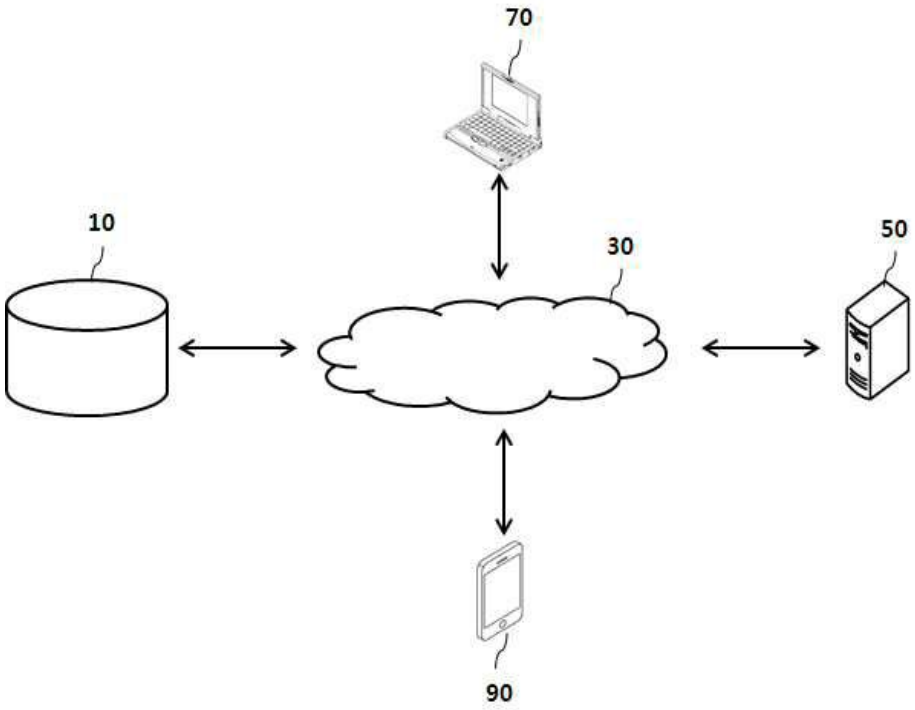
120: 전처리부 130: 매트릭스 생성부

150: 사상 분류 학습부 170: 성향 분류부

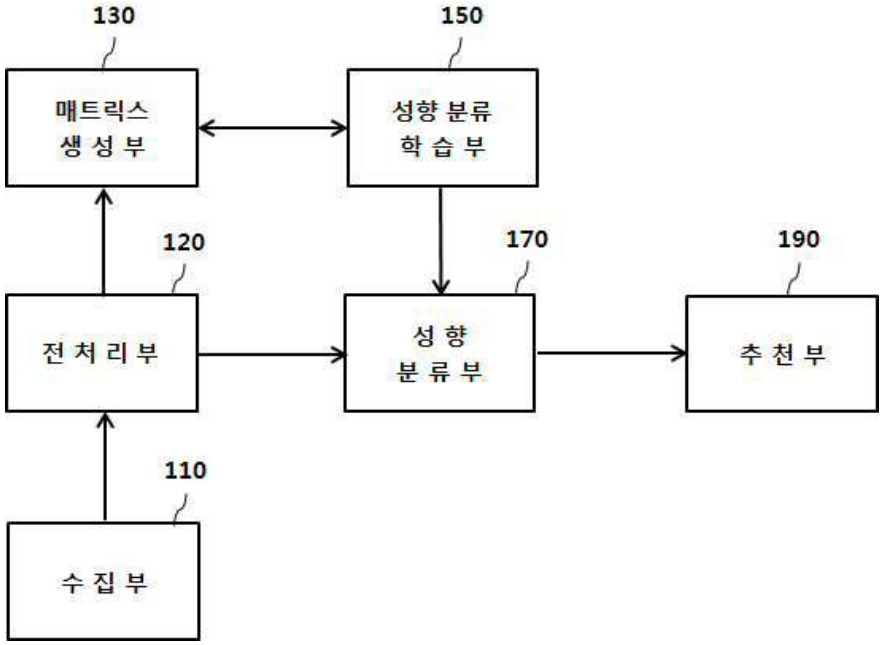
190: 시계열 판단부

도면

도면1


70
10
30
50
90

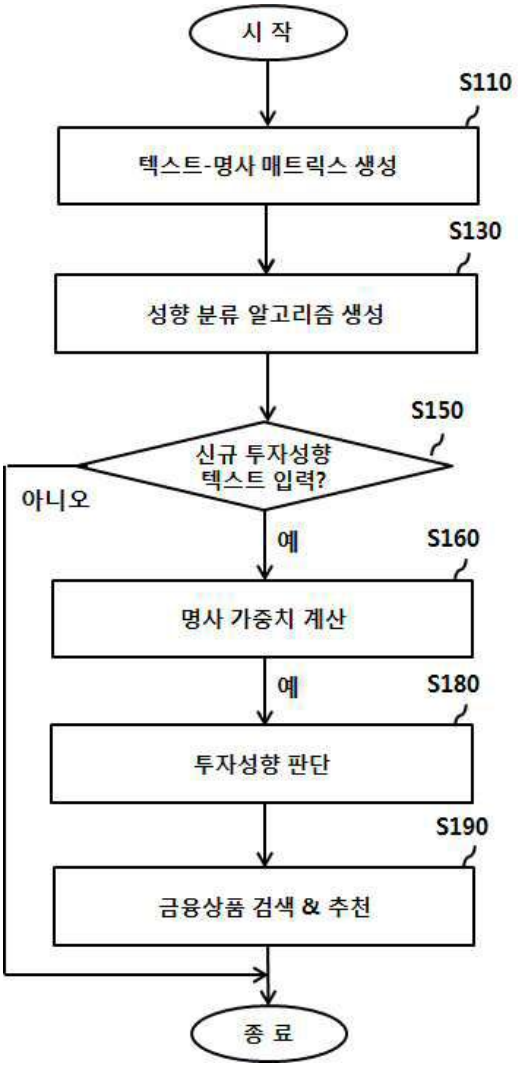
도면2


130
150
매트릭스
생성부
성향 분류
학습부
120
170
190
전처리부
성향
분류부
추천부
110
수집부

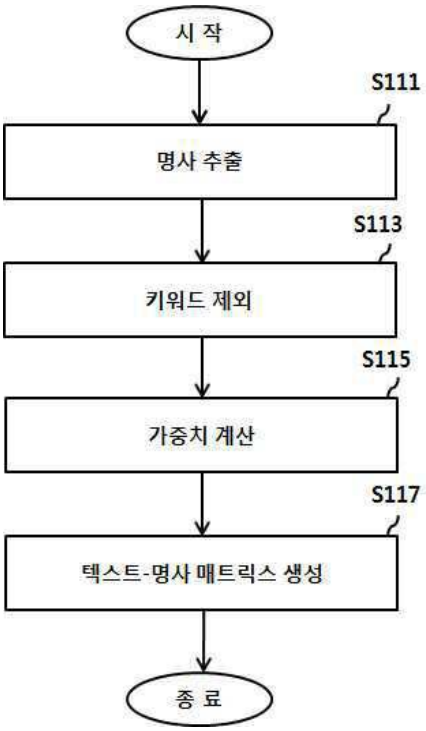
도면3


121
123
125
형태소 분석부
키워드 제거부
가중치 계산부

도면4


시작
S110
텍스트-명사 매트릭스 생성
S130
성향 분류 알고리즘 생성
S150
신규 투자성향 텍스트 입력?
아니오
예
S160
명사 가중치 계산
예
S180
투자성향 판단
S190
금융상품 검색 & 추천
종료

도면5



도면6

철학자	추출 명사
수익성	투자처, 주식, 대상, 금액, 증권, 이상, 단기, 수익, 위험, 기회, 리스크, 최대, 메리트, 개인, 이익 등
	⋮ ⋮
안정성	기초, 은행, 회사, 믿음, 요건, 안정, 수익, 안전, 채권, 우량주, 블루, 성향, 대표, 연간 등
	⋮ ⋮
정보성	운용, 분야, 형태, 필요, 설명, 관심, 이유, 연금, 수수료, 연말, 환매, 경고, 이자 등
	⋮ ⋮
경험	펀드, 주식, 국외, 코스닥, 개인연금, 국내, 코스피, 가치, 성장, 경기, 수익, 연금 등
신뢰성	장기, 언론, 회사, 신탁, 평가, 성장, 기준, 기관, 계열, 전문, 수익, 기업, 투자 등
반응성	안정, 현시점, 국내, 채권, 은행, 우량, 금리, 변동성, 침체, 글로벌, 보도 등

도면7

텍스트	우량주	투자	가치	원금	리스크	
1	0.017359	0	0.003626	0.001526	0	안정성
2	0	0.003849	0.003131	0.000988	0.002661	안정성
⋮	⋮	⋮	⋮	⋮	⋮	⋮
41	0.021479	0.027086	0.029158	0.004908	0.003305	정보성
42	0.050914	0.012589	0.019199	0	0	정보성
⋮	⋮	⋮	⋮	⋮	⋮	⋮
105	0.010902	0.048524	0.019734	0.01661	0.005592	반응성
106	0	0.031681	0.025769	0.004067	0	반응성
⋮	⋮	⋮	⋮	⋮	⋮	⋮

【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1의 15번째 줄

【변경전】

외적요인 분류 알고리즘인 것을 특징으로 하는 금융상품의 추천 방법. 것을 특징 으로 하는 금융상품의 추천 방법.

【변경후】

외적요인 분류 알고리즘인 것을 특징으로 하는 금융상품의 추천 방법.