

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2024-0022284
(43) 공개일자 2024년02월20일

(51) 국제특허분류(Int. Cl.)

G06Q 40/04 (2023.01) G06N 20/00 (2019.01)

G06Q 10/06 (2012.01) G06Q 10/10 (2023.01)

G06Q 40/06 (2023.01)

(52) CPC특허분류

G06Q 40/04 (2023.01)

G06N 20/00 (2021.08)

(21) 출원번호 10-2022-0100772

(22) 출원일자 2022년08월11일

심사청구일자 2022년08월11일

(71) 출원인

한국과학기술원

대전광역시 유성구 대학로 291(구성동)

울산과학기술원

울산광역시 울주군 언양읍 유니스트길 50

(72) 발명자

김우창

대전광역시 유성구 대학로 291(구성동, 한국과학기술원)

정문기

대전광역시 유성구 대학로 291(구성동, 한국과학기술원)

(뒷면에 계속)

(74) 대리인

양성보

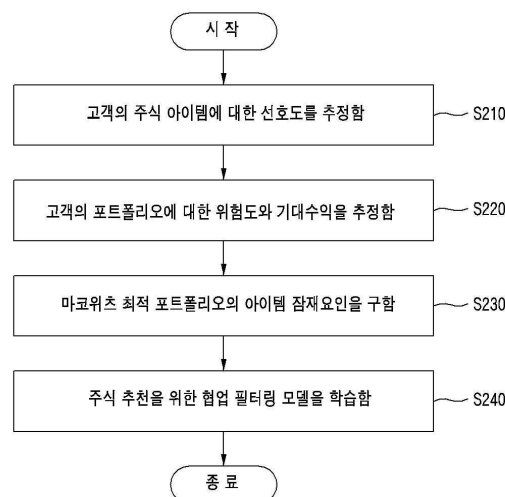
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 주식 추천을 위한 협업 필터링 추천 시스템

(57) 요약

주식 추천을 위한 협업 필터링 추천 시스템이 개시된다. 주식 추천 모델 학습 방법은, 고객의 포트폴리오를 기초로 상기 고객의 주식 아이템에 대한 선호도를 추정하는 단계; 상기 포트폴리오에 대한 위험도와 기대수익을 추정하는 단계; 및 상기 선호도 및 상기 위험도와 상기 기대수익을 기초로 주식 추천을 위한 협업 필터링 모델을 학습하는 단계를 포함한다.

대표도 - 도2



(52) CPC특허분류

G06Q 10/0635 (2023.01)

G06Q 10/06375 (2013.01)

G06Q 10/103 (2013.01)

G06Q 40/06 (2023.01)

(72) 발명자

최화용

대전광역시 유성구 대학로 291(구성동, 한국과학기술원)

이용재

울산광역시 울주군 언양읍 유니스트길 50

이 발명을 지원한 국가연구개발사업

과제고유번호 1345353385

과제번호 2022R1I1A406916311

부처명 교육부

과제관리(전문)기관명 한국연구재단

연구사업명 학제간융합연구지원사업(이공분야)

연구과제명 가계 금융활동 분석을 위한 인공지능 모델 및 맞춤형 생애주기 자산관리 최적화 모

형 개발

기 여 율 1/1

과제수행기관명 울산과학기술원

연구기간 2022.06.01 ~ 2023.02.28

명세서

청구범위

청구항 1

컴퓨터 장치에서 수행되는 주식 추천 모델 학습 방법에 있어서,

상기 컴퓨터 장치는 메모리에 포함된 컴퓨터 판독가능한 명령들을 실행하도록 구성된 적어도 하나의 프로세서를 포함하고,

상기 주식 추천 모델 학습 방법은,

상기 적어도 하나의 프로세서의 의해, 고객의 포트폴리오를 기초로 상기 고객의 주식 아이템에 대한 선호도를 추정하는 단계;

상기 적어도 하나의 프로세서의 의해, 상기 포트폴리오에 대한 위험도와 기대수익을 추정하는 단계; 및

상기 적어도 하나의 프로세서의 의해, 상기 선호도 및 상기 위험도와 상기 기대수익을 기초로 주식 추천을 위한 협업 필터링 모델을 학습하는 단계

를 포함하는 주식 추천 모델 학습 방법.

청구항 2

제1항에 있어서,

상기 위험도와 상기 기대수익을 추정하는 단계는,

상기 고객의 보유 주식에 대한 평균 수익률 벡터와 공분산 행렬을 기초로 상기 위험도와 상기 기대수익을 추정하는 것

을 특징으로 하는 주식 추천 모델 학습 방법.

청구항 3

제1항에 있어서,

상기 학습하는 단계는,

상기 위험도를 감소시키고 상기 기대수익을 증가시키는 방향으로 상기 협업 필터링 모델을 학습하는 단계

를 포함하는 주식 추천 모델 학습 방법.

청구항 4

제1항에 있어서,

상기 학습하는 단계는,

상기 고객의 위험 회피성향을 결정하는 벡터와 투자 성과에 대한 학습 비중을 결정하는 변수를 기초로 상기 협업 필터링 모델을 학습하는 단계

를 포함하는 주식 추천 모델 학습 방법.

청구항 5

제1항에 있어서,

상기 주식 추천 모델 학습 방법은,

상기 적어도 하나의 프로세서의 의해, 복수 개의 위험 레벨에서의 최적 포트폴리오를 이용하여 아이템 잠재요인을 구하는 단계

를 더 포함하는 주식 추천 모델 학습 방법.

청구항 6

제5항에 있어서,

상기 아이템 잠재요인을 구하는 단계는,

상기 최적 포트폴리오를 이용하여 아이템간 PMI(Pointwise Mutual Information) 행렬을 만드는 단계; 및

상기 PMI 행렬을 분해하여 아이템에 대한 잠재요인을 구하는 단계

를 포함하는 주식 추천 모델 학습 방법.

청구항 7

제5항에 있어서,

상기 학습하는 단계는,

상기 아이템 잠재요인을 이용하여 하나의 아이템에 대해 포트폴리오에 공존하는 아이템을 찾을 확률을 높이는 예측을 학습하는 단계

를 포함하는 주식 추천 모델 학습 방법.

청구항 8

제5항에 있어서,

상기 학습하는 단계는,

아이템 잠재요인 행렬의 행 벡터와 선호도에 대한 학습 비중을 결정하는 변수를 기초로 상기 협업 필터링 모델을 학습하는 단계

를 포함하는 주식 추천 모델 학습 방법.

청구항 9

주식 추천 모델 학습 방법을 컴퓨터에 실행시키기 위해 컴퓨터 판독가능한 기록 매체에 저장되는 컴퓨터 프로그램에 있어서,

상기 주식 추천 모델 학습 방법은,

고객의 포트폴리오를 기초로 상기 고객의 주식 아이템에 대한 선호도를 추정하는 단계;

상기 포트폴리오에 대한 위험도와 기대수익을 추정하는 단계; 및

상기 선호도 및 상기 위험도와 상기 기대수익을 기초로 주식 추천을 위한 협업 필터링 모델을 학습하는 단계

를 포함하는, 컴퓨터 프로그램.

청구항 10

컴퓨터 장치에 있어서,

메모리에 포함된 컴퓨터 판독가능한 명령들을 실행하도록 구성된 적어도 하나의 프로세서

를 포함하고,

상기 적어도 하나의 프로세서는,

고객의 포트폴리오를 기초로 상기 고객의 주식 아이템에 대한 선호도를 추정하는 과정;

상기 포트폴리오에 대한 위험도와 기대수익을 추정하는 과정; 및

상기 선호도 및 상기 위험도와 상기 기대수익을 기초로 주식 추천을 위한 협업 필터링 모델을 학습하는 과정

을 처리하는 컴퓨터 장치.

청구항 11

제10항에 있어서,

상기 적어도 하나의 프로세서는,

복수 개의 위험 레벨에서의 최적 포트폴리오를 이용하여 아이템 잠재요인을 구하고,

상기 아이템 잠재요인을 이용하여 하나의 아이템에 대해 포트폴리오에 공존하는 아이템을 찾을 확률을 높이는 예측을 학습하는 것

을 특징으로 하는 컴퓨터 장치.

발명의 설명

기술 분야

[0001] 아래의 설명은 주식 추천 기술에 관한 것이다.

배경 기술

[0002] 인터넷의 발전으로 많은 기업들의 서비스가 인터넷을 통해 제공되고 플랫폼 기업들이 성장하면서 많은 수에 고객들을 유치하는 것이 기업 활동에서 가장 중요한 일이 되고 있다. 이러한 산업 구조의 변화에 발맞춰 추천 시스템에 관한 연구는 빠른 속도로 발전하고 있다.

[0003] 추천 시스템 중 협업 필터링 방법은 사용자들의 과거 아이템 선호도를 이용하여 사용자가 선호하는 아이템과 유사한 아이템을 추천하거나 비슷한 성향을 가지는 다른 사용자가 선호하는 아이템을 추천하는 것으로, 소셜 네트워크 서비스, 전자 상거래, 오디오/비디오 스트리밍 서비스 등의 다양한 분야에서 활용되고 있다.

[0004] 금융산업 또한 새로운 산업 구조에 맞게 발전하며 복잡한 절차를 통해 가입하던 금융 서비스 또한 다른 e-커머스(commerce) 서비스와 다르지 않게 인터넷을 통해 구매할 수 있는 상품이 되었다. 따라서 고객들의 선호도에 맞춘 상품 추천을 제공하여 많은 수의 이용자를 확보하는 것이 금융기업들에게도 가장 중요한 일이라고 할 수 있다.

[0005] 또한, 고객들에게도 주식 추천 시스템은 중요하다. 많은 사람들이 MTS(모바일 트레이딩 시스템)를 이용하여 주식 거래를 하지만 대부분 기업에 대한 정량적인 분석이나 안정적인 투자 포트폴리오를 구성하는 것에 어려움이 있다.

발명의 내용

해결하려는 과제

[0006] 고객이 원하는 특성을 가진 종목을 추천하면서도 고객의 포트폴리오 관점에서 위험대비 기대수익률을 증가시킬 수 있는 방향으로 주식 추천을 제공할 수 있다.

과제의 해결 수단

[0007] 컴퓨터 장치에서 수행되는 주식 추천 모델 학습 방법에 있어서, 상기 컴퓨터 장치는 메모리에 포함된 컴퓨터 관독가능한 명령들을 실행하도록 구성된 적어도 하나의 프로세서를 포함하고, 상기 주식 추천 모델 학습 방법은, 상기 적어도 하나의 프로세서의 의해, 고객의 포트폴리오를 기초로 상기 고객의 주식 아이템에 대한 선호도를 추정하는 단계; 상기 적어도 하나의 프로세서의 의해, 상기 포트폴리오에 대한 위험도와 기대수익을 추정하는 단계; 및 상기 적어도 하나의 프로세서의 의해, 상기 선호도 및 상기 위험도와 상기 기대수익을 기초로 주식 추천을 위한 협업 필터링 모델을 학습하는 단계를 포함하는 주식 추천 모델 학습 방법을 제공한다.

[0008] 일 측면에 따르면, 상기 위험도와 상기 기대수익을 추정하는 단계는, 상기 고객의 보유 주식에 대한 평균 수익률 백터와 공분산 행렬을 기초로 상기 위험도와 상기 기대수익을 추정할 수 있다.

[0009] 다른 측면에 따르면, 상기 학습하는 단계는, 상기 위험도를 감소시키고 상기 기대수익을 증가시키는 방향으로

상기 협업 필터링 모델을 학습하는 단계를 포함할 수 있다.

- [0010] 또 다른 측면에 따르면, 상기 학습하는 단계는, 상기 고객의 위험 회피성향을 결정하는 벡터와 투자 성과에 대한 학습 비중을 결정하는 변수를 기초로 상기 협업 필터링 모델을 학습하는 단계를 포함할 수 있다.
- [0011] 또 다른 측면에 따르면, 상기 주식 추천 모델 학습 방법은, 상기 적어도 하나의 프로세서의 의해, 복수 개의 위험 레벨에서의 최적 포트폴리오를 이용하여 아이템 잠재요인을 구하는 단계를 더 포함할 수 있다.
- [0012] 또 다른 측면에 따르면, 상기 아이템 잠재요인을 구하는 단계는, 상기 최적 포트폴리오를 이용하여 아이터간 PMI(Pointwise Mutual Information) 행렬을 만드는 단계; 및 상기 PMI 행렬을 분해하여 아이터에 대한 잠재요인을 구하는 단계를 포함할 수 있다.
- [0013] 또 다른 측면에 따르면, 상기 학습하는 단계는, 상기 아이터 잠재요인을 이용하여 하나의 아이터에 대해 포트폴리오에 공존하는 아이터를 찾을 확률을 높이는 예측을 학습하는 단계를 포함할 수 있다.
- [0014] 또 다른 측면에 따르면, 상기 학습하는 단계는, 아이터 잠재요인 행렬의 행 벡터와 선호도에 대한 학습 비중을 결정하는 변수를 기초로 상기 협업 필터링 모델을 학습하는 단계를 포함할 수 있다.
- [0015] 주식 추천 모델 학습 방법을 컴퓨터에 실행시키기 위해 컴퓨터 판독가능한 기록 매체에 저장되는 컴퓨터 프로그램에 있어서, 상기 주식 추천 모델 학습 방법은, 고객의 포트폴리오를 기초로 상기 고객의 주식 아이터에 대한 선호도를 추정하는 단계; 상기 포트폴리오에 대한 위험도와 기대수익을 추정하는 단계; 및 상기 적어도 하나의 프로세서의 의해, 상기 선호도 및 상기 위험도와 상기 기대수익을 기초로 주식 추천을 위한 협업 필터링 모델을 학습하는 단계를 포함하는, 컴퓨터 프로그램을 제공한다.
- [0016] 컴퓨터 장치에 있어서, 메모리에 포함된 컴퓨터 판독가능한 명령들을 실행하도록 구성된 적어도 하나의 프로세서를 포함하고, 상기 적어도 하나의 프로세서는, 고객의 포트폴리오를 기초로 상기 고객의 주식 아이터에 대한 선호도를 추정하는 과정; 상기 포트폴리오에 대한 위험도와 기대수익을 추정하는 과정; 및 상기 선호도 및 상기 위험도와 상기 기대수익을 기초로 주식 추천을 위한 협업 필터링 모델을 학습하는 과정을 처리하는 컴퓨터 장치를 제공한다.

발명의 효과

- [0017] 본 발명의 실시예들에 따르면, 고객이 원하는 특성을 가진 종목을 추천하면서도, 고객의 포트폴리오 관점에서 위험대비 기대수익률을 증가시킬 수 있는 방향으로 주식 추천을 제공할 수 있다. 주식 추천을 위해 이용자의 포트폴리오에서 유추할 수 있는 이용자 정보를 반영할 수 있고, 그와 동시에 포트폴리오의 다각화가 가능하다. 이를 통해 주식 추천 시스템의 성능을 평가할 때 추천 성능뿐 아니라 사후 투자 성과를 평가할 수 있다.

도면의 간단한 설명

- [0018] 도 1은 본 발명의 일실시예에 있어서 컴퓨터 장치의 내부 구성의 일례를 설명하기 위한 블록도이다.
- 도 2는 본 발명의 일실시예에 있어서 주식 추천을 위한 협업 필터링 추천 모델의 학습 방법의 일례를 도시한 흐름도이다.
- 도 3과 도 4는 본 발명의 일실시예에 있어서 협업 필터링 추천에 따른 투자 성과와 추천 성능에 대한 평가 결과를 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

- [0019] 이하, 본 발명의 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0020] 본 발명의 실시예들은 주식 추천 기술에 관한 것이다.
- [0021] 본 명세서에서 구체적으로 개시되는 것들을 포함하는 실시예들은 고객이 원하는 특성을 가진 종목을 추천하면서도, 고객의 포트폴리오 관점에서 위험대비 기대수익률을 증가시킬 수 있는 방향으로 주식 추천을 제공할 수 있고, 이를 통해 추천 성능, 서비스 다각화, 효율성, 안정성 등의 측면에 있어서 상당한 장점들을 달성할 수 있다.
- [0022] 금융업계에서 가장 주목받는 서비스는 투자운용일 것이다. 현재까지 주식 추천 시스템에 대한 연구는 크게 세 가지로 분류할 수 있다.

- [0023] 첫째, 주가가 상승할 종목을 예측하는 것에 중점을 둔 연구이다. 하지만 이러한 방식은 사용자 정보를 활용하지 않는다는 점에서 고객의 선호도를 고려한 추천과 고객의 현재 보유한 종목들을 고려하여 안정적인 포트폴리오를 구성하는 관점 모두에서 적절하지 않다고 할 수 있다.
- [0024] 둘째, 기존 협업 필터링 방법론을 주식 추천에 적용하는 방법이다. 현재 보유한 주식과 비슷한 주식이나 비슷한 투자자가 보유한 주식을 매수하는 것은 기존 추천 시스템이 목표하는 대로 투자자의 선호도에 맞는 투자 추천을 할 가능성이 높다. 그러나 이는 투자자의 포트폴리오가 특정 유형의 위험(또는 부문)에 더 많이 노출되도록 하여 투자위험을 더욱 증가시킬 수 있다.
- [0025] 마지막으로, 몇몇 연구들은 기존 추천 시스템에 포트폴리오 다각화를 고려한 연구를 진행하고 있다. 먼저 협업 필터링으로 종목들을 우선 선정한 이후 선택된 종목 군 안에서 포트폴리오 다각화를 고려하는 연구가 있다. 이 방법론은 고객의 선호도와 포트폴리오 다각화를 모두 고려하는 것 같지만, 두 가지 목표가 동시에 고려되는 것이 아니라, 선호도에 기반한 선택이 우선되기 때문에 포트폴리오의 안정성 효과를 기대하기 어렵다.
- [0026] 주식 추천 시스템은 일반적인 추천 시스템과는 다른 관점을 필요로 한다. 가격 불확실성이 높은 주식투자 추천을 위해서는 몇 가지 필수적인 사항이 있다.
- [0027] 첫째, 주식 추천에 투자자의 시각이 반영되어야 한다. 높은 수익률을 예측한 추천이나 기존 포트폴리오 이론을 바탕으로 한 추천은 현재 보유자산과 거의 관련이 없기 때문에 투자자들은 이러한 접근법을 받아들이지 않을 수 있다. 따라서 주식 추천은 이용자의 포트폴리오에서 유추할 수 있는 이용자 정보를 반영해야 한다.
- [0028] 둘째, 주식 추천 시스템은 포트폴리오를 다각화하는 추천을 해야 한다. 투자 전문가들은 분산투자를 통해 투자 수익률의 위험성을 줄인다. 그러나, 협업 필터링을 직접 적용하면 이용자가 현재 보유한 것과 유사한 종목을 추천하기 때문에 이용자의 포트폴리오가 일부 위험요소나 산업부문에 더 집중될 것이다. 따라서, 주식 추천은 이용자의 선호도를 고려하면서도 포트폴리오 다각화를 고려해야 한다.
- [0029] 마지막으로, 주식 추천 시스템의 성능을 평가할 때 추천 성능뿐 아니라 사후 투자 성과를 평가해야 한다. 대부분의 온라인 서비스에서 잘못된 추천이 사용자에게 심각한 피해를 주지는 않지만, 주가 불확실성이 높은 주식시장에서의 잘못된 투자는 막대한 금전적 손해로 직접 이어질 수 있다.
- [0030] 본 발명에서는 세 가지 핵심 요구사항을 모두 고려하여 사용자의 현재 포트폴리오를 기반으로 추천이 이루어지도록 하는 동시에 포트폴리오 다각화 효과를 높이는 협업 필터링 기반 추천 시스템을 제공한다.
- [0031] 본 실시예에서는 주식 추천을 위한 효율적 평균-분산 협업 필터링(Mean-Variance Efficient Collaborative Filtering, 이하, 'MVECF' 이라 칭함) 모델을 제안한다. MVECF 모델은 협업 필터링 알고리즘에서 정규화 및 공동 인수분해(co-factorization) 기술을 사용하여 주가의 불확실성을 처리함으로써 위험-수익 균형의 파레토 최적성을 개선하도록 설계될 수 있다. 다시 말해, MVECF 모델은 투자자의 보유 데이터를 이용하여 선호도를 고려한 추천 시스템을 학습하는 목적식에 포트폴리오 투자 성과를 높이기 위한 항을 추가하여 투자 성과와 추천 성능 사이에 균형을 맞추도록 한다. 또한, 다양한 데이터를 이용하여 주식 추천 모델의 추천 성능과 투자 성능을 평가한다.
- [0032] 도 1은 본 발명의 일실시예에 따른 컴퓨터 장치의 예를 도시한 블록도이다. 예를 들어, 본 발명의 실시예들에 따른 주식 추천을 위한 협업 필터링 추천 시스템은 도 1을 통해 도시된 컴퓨터 장치(100)에 의해 구현될 수 있다.
- [0033] 도 1에 도시된 바와 같이 컴퓨터 장치(100)는 본 발명의 실시예들에 따른 협업 필터링 추천 방법을 실행하기 위한 구성요소로서, 메모리(110), 프로세서(120), 통신 인터페이스(130) 그리고 입출력 인터페이스(140)를 포함할 수 있다.
- [0034] 메모리(110)는 컴퓨터에서 판독 가능한 기록매체로서, RAM(random access memory), ROM(read only memory) 및 디스크 드라이브와 같은 비소멸성 대용량 기록장치(permanent mass storage device)를 포함할 수 있다. 여기서 ROM과 디스크 드라이브와 같은 비소멸성 대용량 기록장치는 메모리(110)와는 구분되는 별도의 영구 저장 장치로서 컴퓨터 장치(100)에 포함될 수도 있다. 또한, 메모리(110)에는 운영체제와 적어도 하나의 프로그램 코드가 저장될 수 있다. 이러한 소프트웨어 구성요소들은 메모리(110)와는 별도의 컴퓨터에서 판독 가능한 기록매체로부터 메모리(110)로 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록매체는 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록매체를 포함할 수 있다. 다른 실시예에서 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록매체가 아닌 통신 인터페이스(130)를 통해 메

모리(110)에 로딩될 수도 있다. 예를 들어, 소프트웨어 구성요소들은 네트워크(160)를 통해 수신되는 파일들에 의해 설치되는 컴퓨터 프로그램에 기반하여 컴퓨터 장치(100)의 메모리(110)에 로딩될 수 있다.

- [0035] 프로세서(120)는 기본적인 산술, 로직 및 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(110) 또는 통신 인터페이스(130)에 의해 프로세서(120)로 제공될 수 있다. 예를 들어 프로세서(120)는 메모리(110)와 같은 기록 장치에 저장된 프로그램 코드에 따라 수신되는 명령을 실행하도록 구성될 수 있다.
- [0036] 통신 인터페이스(130)는 네트워크(160)를 통해 컴퓨터 장치(100)가 다른 장치와 서로 통신하기 위한 기능을 제공할 수 있다. 일례로, 컴퓨터 장치(100)의 프로세서(120)가 메모리(110)와 같은 기록 장치에 저장된 프로그램 코드에 따라 생성한 요청이나 명령, 데이터, 파일 등이 통신 인터페이스(130)의 제어에 따라 네트워크(160)를 통해 다른 장치들로 전달될 수 있다. 역으로, 다른 장치로부터의 신호나 명령, 데이터, 파일 등이 네트워크(160)를 거쳐 컴퓨터 장치(100)의 통신 인터페이스(130)를 통해 컴퓨터 장치(100)로 수신될 수 있다. 통신 인터페이스(130)를 통해 수신된 신호나 명령, 데이터 등은 프로세서(120)나 메모리(110)로 전달될 수 있고, 파일 등은 컴퓨터 장치(100)가 더 포함할 수 있는 저장 매체(상술한 영구 저장 장치)로 저장될 수 있다.
- [0037] 통신 방식은 제한되지 않으며, 네트워크(160)가 포함할 수 있는 통신망(일례로, 이동통신망, 유선 인터넷, 무선 인터넷, 방송망)을 활용하는 통신 방식뿐만 아니라 기기들 간의 근거리 유선/무선 통신 역시 포함될 수 있다. 예를 들어, 네트워크(160)는, PAN(personal area network), LAN(local area network), CAN(campus area network), MAN(metropolitan area network), WAN(wide area network), BBN(broadband network), 인터넷 등의 네트워크 중 하나 이상의 임의의 네트워크를 포함할 수 있다. 또한, 네트워크(160)는 버스 네트워크, 스타 네트워크, 링 네트워크, 메쉬 네트워크, 스타-버스 네트워크, 트리 또는 계층적(hierarchical) 네트워크 등을 포함하는 네트워크 토폴로지 중 임의의 하나 이상을 포함할 수 있으나, 이에 제한되지 않는다.
- [0038] 입출력 인터페이스(140)는 입출력 장치(150)와의 인터페이스를 위한 수단일 수 있다. 예를 들어, 입력 장치는 마이크, 키보드, 카메라 또는 마우스 등의 장치를, 그리고 출력 장치는 디스플레이, 스피커와 같은 장치를 포함할 수 있다. 다른 예로 입출력 인터페이스(140)는 터치스크린과 같이 입력과 출력을 위한 기능이 하나로 통합된 장치와의 인터페이스를 위한 수단일 수도 있다. 입출력 장치(150)는 컴퓨터 장치(100)와 하나의 장치로 구성될 수도 있다.
- [0039] 또한, 다른 실시예들에서 컴퓨터 장치(100)는 도 1의 구성요소들보다 더 적은 혹은 더 많은 구성요소들을 포함할 수도 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다. 예를 들어, 컴퓨터 장치(100)는 상술한 입출력 장치(150) 중 적어도 일부를 포함하도록 구현되거나 또는 트랜시버(transceiver), 카메라, 각종 센서, 데이터베이스 등과 같은 다른 구성요소들을 더 포함할 수도 있다.
- [0040] 이하에서는 주식 추천을 위한 협업 필터링 추천 시스템의 구체적인 실시예를 설명하기로 한다.
- [0041] 본 실시예에 따른 컴퓨터 장치(100)에는 컴퓨터로 구현된 협업 필터링 추천 시스템이 구성될 수 있다. 컴퓨터 장치(100)의 프로세서(120)는 이하의 협업 필터링 추천 방법을 수행하기 위한 구성요소로 구현될 수 있다. 실시예에 따라 프로세서(120)의 구성요소들은 선택적으로 프로세서(120)에 포함되거나 제외될 수도 있다. 또한, 실시예에 따라 프로세서(120)의 구성요소들은 프로세서(120)의 기능의 표현을 위해 분리 또는 병합될 수도 있다.
- [0042] 이러한 프로세서(120) 및 프로세서(120)의 구성요소들은 이하의 협업 필터링 추천 방법이 포함하는 단계들을 수행하도록 컴퓨터 장치(100)를 제어할 수 있다. 예를 들어, 프로세서(120) 및 프로세서(120)의 구성요소들은 메모리(110)가 포함하는 운영체제의 코드와 적어도 하나의 프로그램의 코드에 따른 명령(instruction)을 실행하도록 구현될 수 있다.
- [0043] 여기서, 프로세서(120)의 구성요소들은 컴퓨터 장치(100)에 저장된 프로그램 코드가 제공하는 명령에 따라 프로세서(120)에 의해 수행되는 서로 다른 기능들(different functions)의 표현들일 수 있다.
- [0044] 프로세서(120)는 컴퓨터 장치(100)의 제어와 관련된 명령이 로딩된 메모리(110)로부터 필요한 명령을 읽어들이 수 있다. 이 경우, 상기 읽어들이 명령은 프로세서(120)가 이후 설명될 단계들을 실행하도록 제어하기 위한 명령을 포함할 수 있다.
- [0045] 이후 설명될 단계들은 도시된 순서와 다른 순서로 수행될 수 있으며, 단계들 중 일부가 생략되거나 추가의 과정이 더 포함될 수 있다.
- [0046] 도 2는 본 발명의 일실시예에 있어서 주식 추천을 위한 협업 필터링 추천 모델의 학습 방법의 일례를 도시한 흐

름도이다.

[0047] 추천 시스템 중 가장 많이 사용되는 방식은 모델 기반 협업 필터링이다. m 사용자와 n 개 아이템이 있을 때, 모델 기반 협업 필터링은 고객들의 아이템 사용기록으로 고객-아이템 선호행렬 $Y \in \mathbb{R}^{m \times n}$ 를 만들고 이를 분해하여 고객 잠재요인 행렬 $P \in \mathbb{R}^{m \times l}$ 와 아이템 잠재요인 행렬 $Q \in \mathbb{R}^{n \times l}$ 를 만든다.

[0048] 이를 위한 식은 수학적 식 1과 같다.

[0049] [수학적 식 1]

$$\min_{p_u, q_i} \sum_{u,i} c_{ui} |y_{ui} - p_u q_i^T|^2 + \lambda_p \sum_u \|p_u\|^2 + \lambda_q \sum_i \|q_i\|^2$$

[0051] 여기서, l 은 잠재 인자의 수, $y_{ui} \in \mathbb{R}$ 은 행렬 Y 의 성분이고 $p_u \in \mathbb{R}^l$ 과 $q_i \in \mathbb{R}^l$ 은 각각 P 와 Q 의 행 벡터이다. $c_{ui} \in \mathbb{R}$ 은 관찰 값 y_{ui} 에 대한 신뢰 수준을 나타내는 매개 변수이다. p_u 와 q_i 를 이용하여 $p_u q_i^T$ 로 사용자 u 의 아이템 i 에 대한 선호도를 추정한다.

[0052] 본 발명에서는 두 가지 발전된 모델을 통하여 주식 추천을 투자 성과가 보다 높아질 수 있도록 유도한다. 첫 번째 방법은 추정된 고객의 보유 포트폴리오 위험을 감소시키고 포트폴리오 기대수익을 증가시키는 항을 학습에 추가하는 것이다. 두 번째 방법은 주식 아이템에 대해 아이템 잠재요인들이 마코위츠(Markowitz) 최적 포트폴리오들이 만드는 아이템 잠재요인이 되도록 강제하는 방법이다.

[0053] 도 2를 참조하면, 단계(S210)에서 프로세서(120)는 고객의 보유 포트폴리오를 기초로 해당 고객의 아이템에 대한 선호도를 추정하여 선호도 기반의 모델 학습을 진행할 수 있다.

[0054] 기본 협업 필터링 기반 추천 모델의 학습 목적식은 수학적 식 1과 같고, 고객의 아이템 선호도는 수학적 식 1에서 $\min_{p_u, q_i} \sum_{u,i} c_{ui} |y_{ui} - p_u q_i^T|^2$ 항을 통하여 추정할 수 있다.

[0055] 단계(S220)에서 프로세서(120)는 고객의 보유 포트폴리오에 대한 위험도와 기대수익을 추정하여 투자 성과 기반의 모델 학습을 진행할 수 있다.

[0056] 고객 u 의 추정된 보유 벡터는 $\hat{y} = p_u Q^T$ 이므로 포트폴리오 위험을 σ_u 라 하고 기대수익을 μ_u 라 할 때, 추천 시스템을 학습하는 수식은 수학적 식 2와 같다.

[0057] [수학적 식 2]

$$\min_{p_u, q_i} \sum_{u,i} c_{ui} |y_{ui} - \hat{y}_{ui}|^2 + \lambda \left(\sum_u \|p_u\|^2 + \sum_i \|q_i\|^2 \right) + \sum_u \lambda_{MV} \left(\frac{Y}{2} \sigma_u^2 - \mu_u \right)$$

[0059] 여기서, Y 는 고객의 위험 회피성향을 결정하는 벡터이고, λ_{MV} 는 투자 성과 부분의 학습 비중을 결정하는 변수이다. 현재 주식의 평균 수익률 벡터와 공분산 행렬이 μ, Σ 일 때 고객 u 의 포트폴리오 위험과 기대수익을 추정하는 식은 수학적 식 3, 수학적 식 4와 같다.

[0060] [수학적 식 3]

$$\sigma_u^2 = \hat{y} \Sigma \hat{y}^T = p_u Q^T \Sigma Q p_u^T$$

[0062] [수학적 식 4]

$$\mu_u = \hat{y} \mu^T = p_u Q^T \mu^T$$

[0064] 단계(S230)에서 프로세서(120)는 마코위츠 최적 포트폴리오의 아이템 잠재요인을 구할 수 있다.

[0065] 먼저, 다양한 샘플 수의 다양한 위험 레벨에서의 마코위츠 최적 포트폴리오들을 구하고 이를 이용하여 아이템간

점별 상호정보량(PMI: Pointwise Mutual Information) 행렬을 만든다.

[0066] 마코위츠 최적 포트폴리오에 대한 상세한 내용은 [Markowitz, H. M. (1968). Portfolio selection. In *Portfolio selection*. Yale university press.]에 개시되어 있다.

[0067] 아이템간 PMI(점별 상호정보량) 행렬을 만드는 식은 수학적식 5와 같다.

[0068] [수학적식 5]

[0069]
$$PMI(i,j) = \frac{P(i,j)}{P(i)P(j)}$$

[0070] 여기서, $P(i)$, $P(j)$ 는 마코위츠 최적 포트폴리오에 아이템 i, j 가 포함될 확률이고, $P(i,j)$ 는 아이템 i, j 가 동시에 포함될 확률이다.

[0071] [Levy, O., & Goldberg, Y. (2014). Neural word embedding as implicit matrix factorization. *Advances in neural information processing systems*, 27.]에서는 PMI 행렬에 어떠한 상수를 빼고 양수 부분만을 취한 SPPMI(Shifted Positive PMI) 행렬을 분해함으로써 주식 아이템에 대한 잠재요인을 구할 수 있다는 것을 보여주고 있다. 이렇게 얻어낸 잠재요인은 하나의 아이템에 대해 포트폴리오에 공존하는 아이템을 찾을 확률을 높이는 예측에 사용될 수 있다.

[0072] 투자 성과를 높이는 추천을 위해 고객-아이템 선호행렬을 분해하는 아이템 잠재요인 Q 가 SPPMI 행렬 또한 같이 분해하도록 하는 모델 학습 방법은 수학적식 6과 같다.

[0073] [수학적식 6]

[0074]
$$\min_{p,q,t} \sum_{u,i} c_{ui} (y_{ui} - p_u q_i^T)^2 + \sum_{i,j} c_{cof} (m_{ij} - q_i^T t_j - d_i - e_j)^2 + \lambda \left(\sum_u \|p_u\|^2 + \sum_i \|q_i\|^2 + \sum_j \|t_j\|^2 \right)$$

[0075] 여기서, m_{ij} 는 SPPMI 행렬의 i, j 번째 항이며, 이를 분해한 행렬이 Q 와 T (아이템 잠재요인 행렬)가 된다. t_j 는 T

의 행 벡터이며 d_i, e_j 는 상수 값이다. c_{cof} 는 Q 가 마코위츠(Markowitz) 최적 포트폴리오의 잠재요인에 대한 학습과 고객-아이템 선호행렬에 대한 학습의 비중을 결정하는 변수이다.

[0076] Q 와 T 가 모두 최적 포트폴리오의 SPPMI 행렬을 분해하는 아이템 잠재요인을 의미하고, 이때 Q 가 고객의 보유벡터를 추정하는 아이템 잠재요인과 같은 것을 사용함으로써 고객의 보유벡터가 최적 포트폴리오에서의 공존하는 아이템의 특성을 반영할 수 있도록 하는 것이다. T 는 고객 보유벡터를 추정하는 데에 사용되지는 않는다.

[0077] 단계(S240)에서 프로세서(120)는 투자 성과 기반의 모델 학습과 선호도 기반의 모델 학습을 적용하여 주식 추천을 위한 협업 필터링 추천 모델의 학습을 진행할 수 있다. 프로세서(120)는 고객의 주식 아이템에 대한 선호도, 고객의 보유 포트폴리오에 대한 위험도와 기대수익, 그리고 마코위츠 최적 포트폴리오의 아이템 잠재요인을 기초로 주식 추천을 위한 협업 필터링 모델을 학습할 수 있다.

[0078] 고객의 보유 포트폴리오 위험을 감소시키고 포트폴리오 기대수익을 증가시키는 항을 학습에 추가하는 방법과 아이템 잠재요인들이 마코위츠 최적 포트폴리오들이 만드는 아이템 잠재요인이 되도록 강제하는 방법을 같이 적용한 협업 필터링 추천 모델의 학습 방법은 수학적식 7과 같다.

[0079] [수학적식 7]

[0080]
$$\min_{p,q,t} \sum_{u,i} c_{ui} (y_{ui} - p_u q_i^T)^2 + \sum_{i,j} c_{cof} (m_{ij} - q_i^T t_j - d_i - e_j)^2 + \sum_u \lambda_{MV} \left(\frac{\gamma}{2} \sigma_u^2 - \mu_u \right) + \lambda (\|p_u\|^2 + \|q_i\|^2 + \|t_j\|^2)$$

[0081] 수학식 2, 수학식 6, 및 수학식 7에 추가한 항들은 보다 나은 포트폴리오 위험대비 수익을 추정하기 위한 항이다.

[0082] 수학식 6의 $\sum_{i,j} m_{ij} \neq 0 c_{cof} (m_{ij} - q_i^T t_j - d_i - e_j)^2$, 수학식 7의 $\sum_{i,j} m_{ij} \neq 0 c_{cof} (m_{ij} - q_i^T t_j - d_i - e_j)^2$ 은 먼저 최적 포트폴리오를 이용하여 아이템간 PMI 행렬 M을 만들고 고객의 아이템 선호도를 추정할 때 사용하는 아이템 잠재요인 Q가 M 또한 분해하게 만들어 최적 포트폴리오의 공존 아이템의 특징을 가지게 만드는, 결과적으로 추천 아이템으로 인한 포트폴리오 성능을 향상시키는 역할을 할 수 있다.

[0083] 수학식 2의 $\sum_u \lambda_{MV} (\frac{Y}{2} \sigma_u^2 - \mu_u)$, 수학식 7의 $\sum_u \lambda_{MV} (\frac{Y}{2} \sigma_u^2 - \mu_u)$ 는 고객의 추정된 보유벡터의 위험을 낮추고 기대수익을 높이는 역할을 한다.

[0084] 따라서, 목적식의 각 항인 1) $\sum_{u,i} c_{ui} |y_{ui} - p_u q_i^T|^2$, 2) $\sum_u \lambda_{MV} (\frac{Y}{2} \sigma_u^2 - \mu_u)$, 3)

$\sum_{i,j} m_{ij} \neq 0 c_{cof} (m_{ij} - q_i^T t_j - d_i - e_j)^2$ 을 각 단계로 정리한다면, 1) 항을 고객의 주식 아이템에 대한 선호도를 추정하는 과정(S210), 2) 항을 고객의 포트폴리오에 대한 위험도와 기대수익을 추정하는 과정(S220), 3) 항을 마코위츠 최적 포트폴리오의 아이템 잠재요인을 구하는 과정(S230)으로 구분할 수 있다.

[0085] 도 3은 주식 추천에 의한 샤프지수를 나타내고 있고, 도 4는 추천 성능 재현율을 나타내고 있다.

[0086] 도 3을 참조하면, 각 모델의 결정변수를 증가시킴에 따라 추천에 의한 샤프지수(수익÷변동성)가 증가함을 알 수 있다. 도 4를 참조하면, 각 모델의 결정변수를 증가시킴에 따라 추천 성능 재현율(실제 보유한 것을 추천한 수÷실제 보유한 수)이 감소함을 알 수 있다.

[0087] 이처럼 본 발명의 실시예들에 따르면, 고객이 원하는 특성을 가진 아이템을 추천함과 동시에 고객의 포트폴리오 관점에서 위험대비 기대수익률을 증가시킬 수 있는 방향으로 추천을 제공함으로써 기존 추천보다 높은 투자 성과를 달성할 수 있고, 각 모델의 결정변수를 통하여 고객의 선호도를 맞추기 위한 추천과 투자 성과를 위한 추천의 비중을 결정할 수 있다.

[0088] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate 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 어플리케이션을 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 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0089] 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, 컴퓨터 저장 매체 또는 장치에 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.

[0090] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 이때, 매체는 컴퓨터로 실행 가능한 프로그램을 계속 저장하거나, 실행 또는

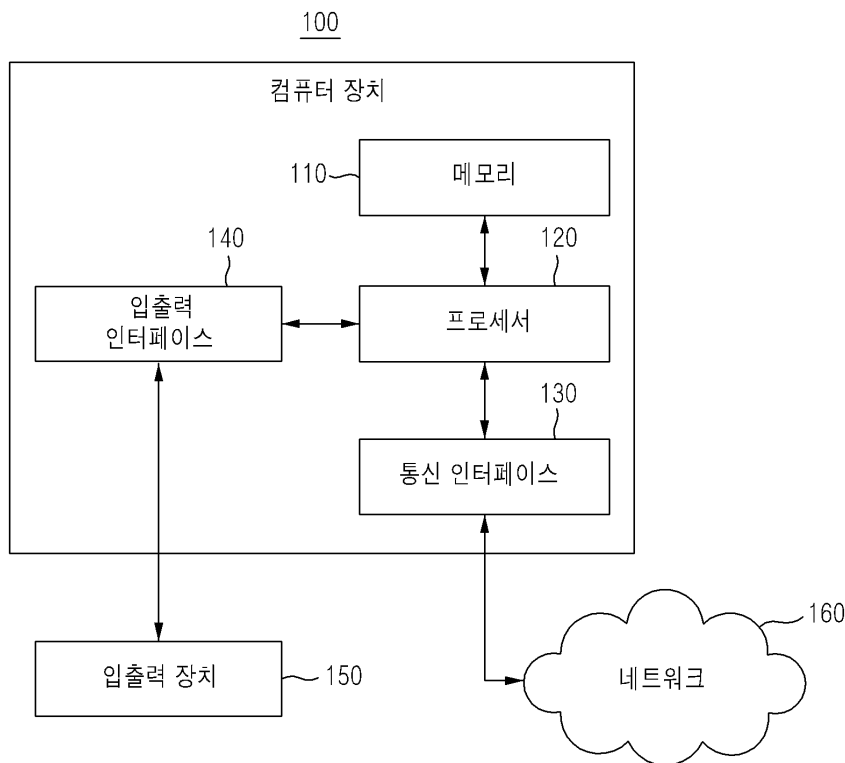
다운로드를 위해 임시 저장하는 것일 수도 있다. 또한, 매체는 단일 또는 수 개의 하드웨어가 결합된 형태의 다양한 기록수단 또는 저장수단일 수 있는데, 어떤 컴퓨터 시스템에 직접 접속되는 매체에 한정되지 않고, 네트워크 상에 분산 존재하는 것일 수도 있다. 매체의 예시로는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등을 포함하여 프로그램 명령어가 저장되도록 구성된 것이 있을 수 있다. 또한, 다른 매체의 예시로, 어플리케이션을 유통하는 앱 스토어나 기타 다양한 소프트웨어를 공급 내지 유통하는 사이트, 서버 등에서 관리하는 기록매체 내지 저장매체도 들 수 있다.

[0091] 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.

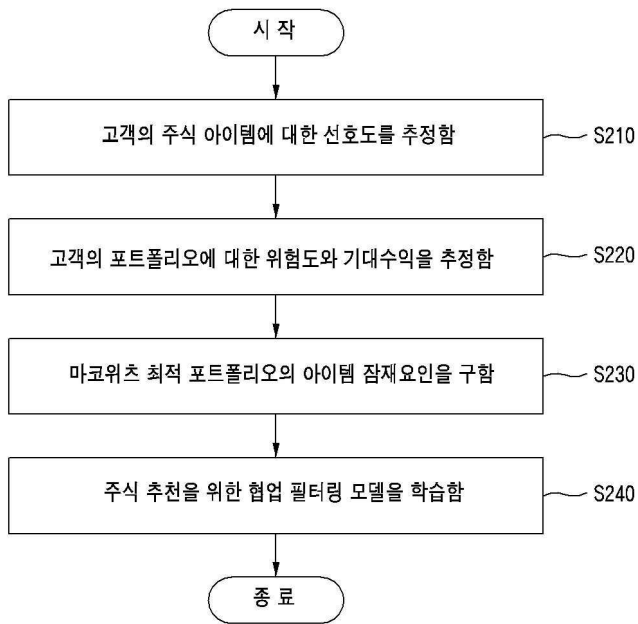
[0092] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

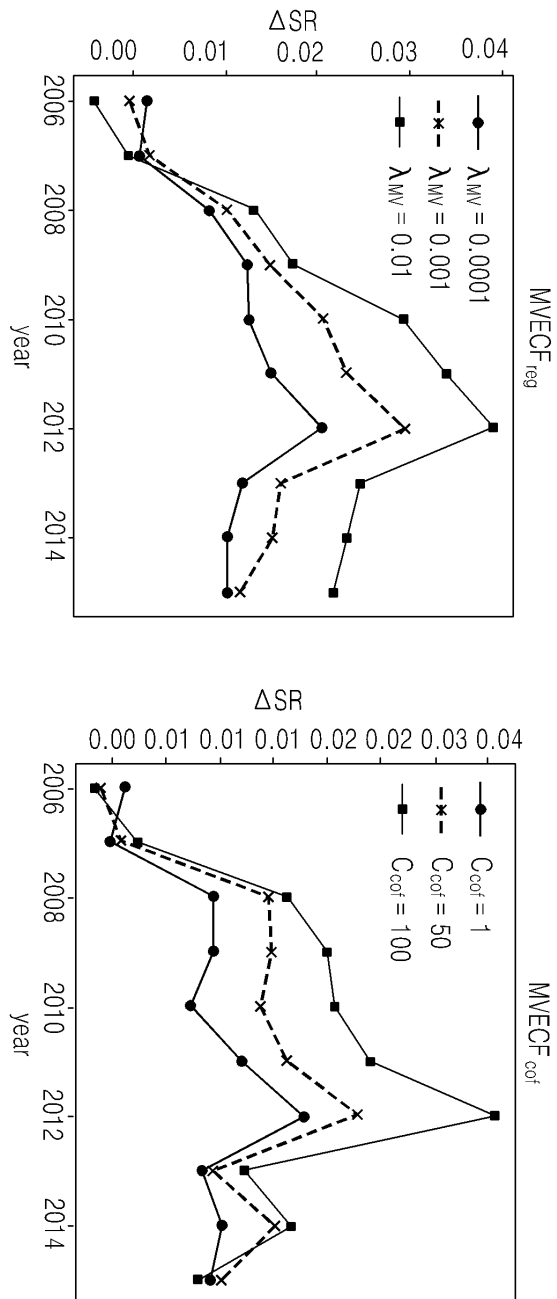
도면1



도면2



도면3



도면4

