Information Systems Review
Vol. 25, No. 1 February 2023 http://dx.doi.org/10.14329/isr.2023.25.1.111

고객의 투자상품 선호도를 활용한 금융상품 추천시스템 개발

Financial Products Recommendation System Using Customer Behavior Information

김 효 중 (Hyojoong Kim) 연세대학교 정보시스템학 석사
김 성 범 (SeongBeom Kim) 연세대학교 정보대학원 석박통합과정
김 희 웅 (Hee-Woong Kim) 연세대학교 정보대학원 교수, 교신저자

요 약

인공지능(AI) 기술이 발전함에 따라 빅데이터 기반의 상품 선호도 추정 개인화 추천시스템에 관심이 증가하고 있는 추세이다. 하지만 개인화 추천이 적합하지 않은 경우 고객의 구매 의사를 감소시키고 심지어 금융상품의 특성상 막대한 재무적 손실로 확대될 수 있는 위험을 가지고 있다. 따라서 고객의 특성과 상품 선호도를 포괄적으로 반영한 추천시스템을 개발하는 것이 비즈니스 성과 창출과 컴플라이언스 이슈 대응에 매우 중요하다. 특히 금융상품의 경우 개인의 투자성향과 리스크 회피도에 따라 고객의 상품 선호도가 구분되므로 축적된 고객 행동 데이터를 활용하여 맞춤형 추천서비스를 제안하는 것이 필요하다. 이러한 고객의 행동 특성과 거래 내역 데이터를 사용하는 것뿐만 아니라, 고객의 인구통계정보, 자산정보, 종목 보유 정보를 포함하여 추천 시스템의 콜드 스타트 문제를 해결하고자 한다. 따라서, 본 연구는 고객의 거래 로그 기록을 바탕으로 고객의 투자성향과 같은 특성 정보와 거래 내역 및 금융상품 정보를 통해 고객별 금융상품 잠재 선호도를 도출하여 딥러닝 기반의 협업 필터링을 제안한 모형이 가장 성능 우수한 것을 확인하였다. 본 연구는 고객의 금융 투자 메커니즘을 기반으로 금융상품 거래 데이터를 통해 미거래 금융상품에 대한 예상 선호를 도출하는 추천 모델을 구축하여, 선호가 높을 것으로 예상되는 상위 상품군을 추천하는 서비스를 개발하는 것에 의의가 있다.

키워드 : 개인화 추천시스템, 신경망 모델, 잠재 선호도, 고객 특성 정보, 금융상품

Ⅰ. 서 론

인터넷 보급과 정보통신기술(ICT, Information and Communication Technology)의 발달로 개인 고객 주식 거래 대금이 2019년 기준 일평균 13조 원에서 2021년 28조 원으로 증가하는 추세이다(한국거래소, 2021). 이러한 배경 속에서 앱을 통한 금융상품 거래가 보편화 되었으며, 앱 시장에 유통되는 금융상품의 다양성도 증가하고 있다. 모바일 증권거래시스템인 MTS(Mobile Trading Systems)

에서 금융상품 종류의 증가는 서로 다른 투자자들의 기호를 충족시켜주어 소비자들의 편의성을 높일 수 있다. 하지만, 특정 선호가 형성되지 않은 소비자들에게 부적합한 금융상품을 권유하거나 정보 공시 위반 사례가 높은 비중을 차지하는 등 금융상품의 다양성 속에서 투자결정을 결정하는데 비용을 증가시키고 있다(KDB미래전략연구소, 2019). 즉, 다양한 금융상품 정보 속 복잡성에 직면한 투자자들이 많은 정보 속에서 자신에게 적합한 금융상품을 선택하지 못한 채 오로지 금융상품 정보 수집에 몰두하여 투자로 연결되지 않는 문제가 제기되고 있다(Jannach *et al.*, 2010; Persson, 2018). 여러 금융상품을 검색하는 것이 쉬운 MTS(Mobile Trading Systems)와 HTS(Home Trading Systems)의 특성상 정보 과부하로 인한 문제는 일선 영업점을 이용하는 것보다 더 자주 발생하고 이는 투자 방해 요소로써 영향을 미친다. 아울러, 증권투자권유자문인력은 고객 별 다양한 투자 성향과 수 천 개의 금융상품의 종류로 인해 고객에게 맞춤형 금융상품 판매가 사실상 어려운 실정이며, 증권 계좌 수 증가에 비해 판매 전문 인력의 증가 속도의 한계도 존재한다(KDB미래전략연구소, 2019). 이러한 점을 고려할 때, MTS 플랫폼에게 있어 고객들의 금융상품 정보 탐색 비용을 낮추고 투자 의사결정을 돕는 선호도 기반 맞춤형 추천서비스 제공에 대한 수요가 급증하고 있다(송희석, 2019). 특히, 방대한 양의 축적된 고객 행동 데이터와 거래 로그 데이터를 통해 고객이 선호할 만한 금융상품을 추천하여 고객들의 투자 탐색 비용을 줄이고 투자 의사결정을 촉진시키는 개인화 맞춤 추천 시스템이 주목을 받고 있다. 특히, 금융 고객들은 금융상품 특성 상 처분 효과, 성과 추종 등 심리학적인 메커니즘에 따라 유전적 차이가 관찰 가능한 개인 특성을 통제한 후 개인 투자자 간의 투자 행동 변동이 최대 45%임을 감안하면, 고객의 금융 관련 경험과 투자 편향과 사건으로 인해 투자 행동이 상이하다는 것을 알 수 있다(Cronqvist and Siegel, 2014; 원종윤, 이건창, 2021). 이러한 행동의 투자 여정 환경을 고려하여 고객의 투자 선호와 리스크 회피도와 같은 금융상품 경험에 따른 행동을 기반으로 적합한 금융상품을 추천할 필요성이 존재한다고 볼 수 있다.

추천시스템의 경우 개인에게 적합하지 않은 상품을 추천하면 오히려 구매 의사를 감소시키고 고객에게 불만 요인으로 확대될 수 있는 위험을 가지고 있다(Kim and Kim, 2016). 특히, 금융상품의 경우 개인의 투자성향과 더불어 상품 민감도에 따라 특정 리스크가 포함된다. 즉, 상품에 대한 명확한 선호도가 구분되므로 축적된 고객 행동 데이터와 금융상품 정보를 포함한 거래 로그 데이터를 융합하여 맞춤형 금융상품 추천을 하는 것이 시스템의 성능과 만족도에 영향을 미칠 수 있다. 하지만 현재까지 금융상품 관련 추천시스템 연구들은 고객의 투자성향과 리스크 회피도를 반영하지 않은 채 상품 및 포트폴리오 기반의 알고리즘이나 금융 중개인의 소셜 네트워크 분석을 토대로 금융상품 추천을 제안하였다(Gholamian *et al.*, 2011; Gigli *et al.*, 2017; Zatevakhina *et al.*, 2019). 또 다른 연구에서는 고객의 자산 및 거래 정보를 반영하지 않고 단순히 수익창출 수준에 따라 고객을 구분하여 고객이 속한 투자 그룹에 적합한 상품을 연결하는 것에 그치는 한계가 존재한다(Patty *et al.*, 2018; Yim *et al.*, 2016). 즉, 현재까지 고객의 구체적인 인적 정보와 자산 및 거래 정보를 포괄적으로 반영한 연구는 찾기 어렵다고 볼 수 있다. 아울러, 기존 연구들은 방법론 측면에서 사용자와 아이템 간의 유사도를 산출하여 상품을 제안하는 협업 필터링(Collaborative Filtering, CF)을 추천 모형으로 활용하였으나, 이는 신규 고객 또는 새로운 아이템이 존재할 때 추천 성능에 제약적인 콜드 스타트(Cold Start) 이슈가 발생하는 한계가 존재한다(Kim and Kim, 2016). 이처럼 기존 금융상품 분야에 대한 추천시스템은 아이템 기반 협업 필터링과 클러스터링 기법 등 머신러닝 중심의 방법론을 주로 활용하였으나, 고객의 비선형적인 선호도를 파악하는데 한계를 확인하였다(Gholamian *et*

al., 2011; Matsunami *et al.*, 2017; Patty *et al.*, 2018). 아울러, 고객의 평점, 설문조사 등 소수 고객의 주관적 선호도 지표를 활용하는 등 주관적인 데이터를 사용하여, 단기간 동안의 소수 데이터 활용은 모델의 편향성을 가중시키는 문제를 확인하였다 (Gholamian *et al.*, 2011).

따라서, 본 연구에서는 딥러닝을 활용한 협업 필터링 모형을 기반으로 고객의 금융상품 거래 기록과 더불어 고객의 자산정보, 종목 포트폴리오 정보 등 고객의 잠재적인 행동 특성을 포괄적으로 반영한 심층 신경망 기반의 추천시스템 모형을 제시하고자 한다. 고객의 특성 정보를 활용하여 새로운 고객에 대해서도 기존 고객과의 특성 유사도를 산출하여 실용적인 추천 모델을 제시한다. 이를 통해 기존 협업 필터링의 콜드 스타트 이슈를 완화하고자 한다. 본 연구가 제안하는 추천 모형을 검증하기 위해 특성 정보를 제외한 딥러닝 협업 필터링 모형과 몇 가지의 머신러닝 모델의 비교를 시행하였다. 이러한 금융상품 추천시스템을 개발함으로써, 고객에게 투자 선호와 리스크 회피도에 적합한 상품을 제안할 수 있으며, 전문 판매인에게는 고객의 니즈에 적합한 금융상품을 효율적으로 추천하고 기존 금융 규제 이슈에 대응이 가능하다는 이점이 존재한다. 나아가, 금융상품을 운영하는 회사의 경우 ESG 경영 일환으로 금융분쟁 이슈에 따른 리스크를 감소시킬 수 있으며, 각종 소송 비용과 같은 제비용 경감효과를 유발할 수 있을 것이다.

Ⅱ. 문헌 연구

2.1 금융상품 관련 추천시스템

고객 정보를 적극적으로 활용한 상품 추천시스템 관련 연구는 활발하게 진행되었다. 고객 정보를 활용하거나 상품의 카테고리 특성 정보에 대해 TF-IDF를 통해 상품의 속성 정보를 추출한 추천시스템에 적용한 연구가 진행되었다. Matsunami *et al.*(2017)은 화장품 상품에 대한 사용자 후기를 기반하여 사용자 간 거리 유사도 측정 방법을 제시하였다. 고객의 리뷰 데이터를 기반으로 동시발생 키워드를 활용한 자동 점수 추출법을 제안하였으며, 산출된 평점으로 상품을 군집화하고 고객 간 유사도를 계산하였다. 이는 고객 리뷰 데이터를 기반으로 추천 모델 내 희소성 문제를 해결하며 근접한 사용자에 따라 선호할 만한 제품을 추천하는 방식의 시스템을 제안하였다. Patty *et al.* (2018)은 30개의 상품 카테고리를 기반으로 TF-IDF을 활용한 상품 정보 기반 추천시스템을 제안하였다. TF-IDF 분석을 위해 상품 정보와 고객 연령대, 고객 타입, 상품 사용시기 등을 고객 정보로 구성하였다. 그리고 앞서 분류한 40개의 상품과의 코사인 유사도(Cosine Similarity)를 계산하여 상품을 추천하는 방식의 시스템을 제안하였다. Yoon and Joung(2020)은 상품의 속성과 평가 데이터를 수집하고자 웹크롤링을 통하여 비정형 데이터인 텍스트를 수집하였으며, TF-IDF 방식을 활용하여 상품과 고객 간 유사도를 산출하는 시스템을 구축하였다.

이처럼 상품 정보와 고객 정보를 반영한 추천시스템 연구는 다양한 고객 행동을 통해 선호도를 파악할 수 있으며, 금융상품에 고객 정보를 반영하는 추천시스템 연구가 진행되어왔다. 김석수(2014)는 금융상품에 가입된 고객의 여러 특성을 파악하여 반영하였고, 각 고객의 특성에 적합한 금융상품을 적시에 추천을 하고자 하였다. K-means 알고리즘 기법을 적용하여 고객의 거래특성, 보유상품 포트폴리오, 인구통계정보 등의 고객 정보를 활용하여 Hadoop과 Mahout을 통해 금융상품 추천 서비스를 제시하였다. 또한, 이재웅 등(2016)은 머신러닝 기술을 통해 고객의 비정형 데이터인 텍스트 VOC를 수집하고자 하였다. 온라인상 고객의 비정형 텍스트를 통해 고객 특성을 추출하고 이에 적합한 금융상품 정보를 자동 추천하는 알고리즘을 구축하여 금융상품 추천을 위한 정보의 범위를 확대하고자 하였다. 나아가 금융상품을 선택하고

자 하는 고객의 패턴을 분석하여 고객의 정보를 반영하고자 하였다. 즉, 고객의 금융 소비 유형을 중시하는 요인을 중심으로 텍스트 내 단어 집합을 도출하고, 온라인 상 고객 로그와 비교하여 사용자의 투자 성향을 예측하였다.

금융상품의 특성 상 고객은 과도한 거래, 처분 효과, 성과 추종 등 심리학적인 메커니즘에 따라 고객의 투자 행동과 성향이 정해진다(Cronqvist and Siegel, 2014). 즉, 금융상품에 대한 고객은 과도한 거래를 통해 자기과신(Overconfidence)을 하거나, 처분 효과를 통한 손실 회피(Loss Aversion)를 지향하며 성과 추종(Performance Chasing)을 하거나 현상을 유지(Status quo)하는 투자 행동을 보인다는 것이다(Cronqvist and Siegel, 2014). 따라서, 고객마다 투자 성향의 차이가 존재하며 리스크도 상이하기 때문에 상품에 대한 선호도를 비롯한 고객의 다양한 특성을 반영해야 할 필요성이 존재한다(이은주 등, 2018). 이러한 배경에서 선행 연구들은 리뷰 데이터, 상품 정보, 거래 정보 등을 활용하여 추천 모델 내 데이터 희소성 문제를 해결하며 다양한 특성 데이터를 활용한 추천시스템 연구를 시도하였다. 하지만 선행연구 중 일부는, 상품의 특성을 고객이 선호해왔던 상품과 유사한 상품을 연결하는 컨텐츠 혹은 K-means기반의 추천시스템 방식을 제안하였다. 특히, 금융상품의 추천 시스템과 관련된 선행 연구는 고객의 투자 성향, 리스크 회피도, 종목 선호 등의 고객 특성 정보를 실질적으로 반영하지 못하고, 고객이 선호하는 특성을 가진 상품을 반복적으로 추천하는 약점이 존재한다. 또한, 거래정보, 종목 정보 같은 고객 특성 정보를 반영하기 위해 다양한 노력도 있었지만, 주로 활용된 머신러닝의 특성 상 상품에 대한 선호도와 고객의 특성 정보를 함께 반영한 추천 모델을 개발하기에는 한계가 있었다. 고객 특성을 반영한 상품과 금융상품 관련 추천시스템 선행 연구는 <표 1>을 통해 정리된 내용을 확인할 수 있다. 따라서, 본 연구는 고객과 상품 선호도 사이의 비선형 관계를 중심으로 모델링을 진행하고, 나아가 고객의 특성 정보를 투입 변수의 차원 확장 형태로 활용하는 추천 모델을 제안하고자 한다.

2.2 추천시스템 알고리즘 연구

추천시스템은 고객의 선호도와 특성 정보를 반영하기 위해 내용 중심의 콘텐츠 기반 필터링(Content-based Filtering, CBF)과 상호작용 정보를 반영하는 협업 필터링(Collaborative Filtering, CF)으로 구분할 수 있다. 내용 기반 추천은 추천 대상의 컨텐츠가 비슷한 아이템을 추천하는 방식으로, 고객이 과거 선택했던 아이템 중 유사한 속성의 아이

〈표 1〉 고객 특성을 반영한 추천시스템 선행 연구

관련 연구	연구 방법	내용 및 시사점
Matsunami *et al.*(2017)	Correlation coefficient (Similarity)	데이터의 희소성 문제를 해소하며 사용자 간 유사도 기반 추천시스템 개발
Patty *et al.*(2018)	Content-based Filtering with TF-IDF	제품의 카테고리에 따른 코사인 유사도(Cosine Similarity)를 계산하여 유사도 기반 제품 군을 추천시스템 제안
Yoon and Joung(2020)	Content-based Filtering with TF-IDF	상품 텍스트 데이터를 수집하기 위해 웹크롤링을 활용하고, TF-IDF 방식을 통해 상품 간의 유사도 판별 추천시스템 구축
김석수(2014)	K-means	고객의 거래 특성, 보유 상품의 포트폴리오, 인구통계정보 등 고객의 정보를 반영하여 K-means를 통해 비슷한 유형 고객에게 금융상품 추천 서비스 제시
이재웅 등(2016)	Machine Learning(SVM, DT, LR, K-NN)	고객 VOC 정보를 인식하기 위해 딥러닝 기술을 활용하여 고객 투자 특성을 추출하고 적합한 상품 추천 알고리즘 구축

템을 분석하여 추천 아이템으로 제시한다. 특히, 비정형 데이터인 텍스트를 중심으로 TF-IDF(Term Frequency - Inverse Document Frequency)를 활용하거나, 특성 기반으로 클러스터링하여 아이템을 연결하는 방식으로 활용되었다(Bobadilla *et al.*, 2013; Schafer *et al.*, 2007). 하지만, 과거 선택 기록이 없는 신규 유저에 대한 아이템 추천이 어렵고, 선호하는 아이템을 반복적으로 추천하는 한계가 존재하기 때문에 협업 필터링 기반 추천시스템이 활발하게 연구되었다(Bobadilla *et al.*, 2013). 협업 필터링은 사용자-아이템 행렬을 이용하여 고객 및 제품 간의 유사도를 계산하는 방식이다. 이처럼 고객 데이터를 기반으로 특정 고객과 비슷한 선호 패턴을 가지는 다른 고객 정보를 활용하여 신규 유저와 새로운 아이템을 추천하기 어려운 내용 기반 추천의 한계를 보완한다는 점에서 산업계에서 활발하게 활용되고 있다(Cremonesi *et al.*, 2012).

하지만 전통적인 메모리 기반의 협업 필터링(Memory-based Collaborative Filtering)은 아이템 종류와 사용자 수가 증가함에 따라 유사도 계산을 위한 많은 계산량이 요구된다는 한계점이 존재한다(Bobadilla *et al.*, 2013). 또한, 신규 사용자나 새로운 아이템에 대한 과거 이력이 존재하지 않으면, 고객의 선호도 예측과 적절한 제품을 추천하지 못해 성능적인 측면에서 어려운 콜드 스타트 이슈가 제기되었다(Su and Khoshgoftaar, 2009). 콜드 스타트는 데이터의 희소성과 관련이 있으며, 이를 해결하여 사용자가 만족할 수 있는 제품을 추천하도록 하는 것이 추천시스템의 핵심으로 떠올랐다(Park and Chu, 2009; Zhang *et al.*, 2020).

이러한 콜드 스타트 문제를 해결하고자 제안된 방식은 모델 기반의 협업 필터링(Model-based Collaborative Filtering)으로, 사용자와 아이템의 관계를 행렬분해 모형(Matrix Factorization, MF)과 심층 신경망을 통한 딥러닝 기반 협업 필터링을 통해 추천시스템 내 높은 성능을 보이고 있다. 메모리 기반의 협업 필터링에서 활용하는 행렬분해 모형은 사용자와 아이템의 관계 행렬을 사용자의 잠재 벡터(User Latent Vector)와 아이템의 잠재 벡터(Item Latent Vector)로 분해하며, 이를 손실 함수 최적화를 통해 새롭게 계산된 행렬을 기반으로 아이템을 추천하는 방식이다(Koren *et al.*, 2009). 나아가 행렬분해 모형의 잠재 벡터 기반 행렬을 계산한다는 장점을 딥러닝 기반 협업 필터링 모형인 심층 신경망 협업 필터링 모형(Neural Collaborative Filtering, NCF)이 새로운 방식이 제안되었다(Gao *et al.*, 2019; He *et al.*, 2018). 심층 신경망 협업 필터링은 원-핫 인코딩(One-hot Encoding) 방식으로 구성된 사용자 정보와 아이템 정보를 임베딩 과정을 진행하여 사용자 잠재 벡터와 아이템 잠재 벡터로 표현한다. 이를 다층 퍼셉트론(Multi-Layer Perceptron)에 통과하여 사용자와 아이템 간의 비선형적 관계를 학습하여 기존 메모리 기반의 협업 필터링의 한계를 해결하는데 기여하였다(Gao *et al.*, 2019).

2.3 딥러닝 기반의 협업 필터링

본 연구에서 제안하는 추천시스템은 모델 기반의 협업 필터링 중에서 딥러닝 기반의 협업 필터링 모델을 기반으로 진행되었다. He *et al.*(2018)은 일반화 행렬분해 모형(Generalized Matrix Factorization, GMF)의 비선형관계를 포착할 수 없다는 한계를 지적하고 딥러닝 기반의 협업 필터링을 통해 해결하고자 하였다. 즉, 일반화 행렬분해 모형에서 다층 퍼셉트론 모형(Multi-Layer Perceptron, MLP)을 결합하여 심층 신경망 기반 협업 필터링을 개발한 것이다(He *et al.*, 2017). 이는 고객의 잠재 요인을 상호작용 유무에 따른 이항 출력 결과(Binary output)를 기반 모형을 제시하여 구매 가능성이 높거나 예상 선호도가 높게 계산된 항목을 추천하고자 하는 목적으로 개발된 방법이다.

심층 신경망 협업 필터링의 입력 층(Input Layer)은 고객 정보와 상품 정보의 잠재 요인을 포함하고 있는 임베딩 층(Embedding Layer)을 포함한다. 다음으로, 일반화 행렬분해와 다층 퍼셉트론 모형

을 포괄하는 은닉 층(Hidden Layer)으로 구성되며, 고객의 예상 상품 선호도를 도출하는 출력 층(Output Layer)의 구조로 구성되어 있다. 은닉 층에는 기존의 일반화 행렬분해 모형의 선형성 한계를 극복하기 위해 시그모이드(Sigmoid)와 같은 활성화 함수(Activation Function)를 사용하는 것이 특징이다(He *et al.*, 2017). 마지막으로, 일반화 행렬분해의 출력 값과 고객 특성 및 제품의 잠재 요인을 임베딩 층의 출력 값을 결합하고 다수의 완전연결 계층(Fully Connected Layer)을 통해 출력 층으로 이어지는 구조로 설계되었다(He *et al.*, 2017). 심층 신경망 협업 필터링 모형은 확률적 경사 하강법(Stochastic Gradient Descent, SGD)을 통해 학습이 이루어지며, 출력과 실제의 값 비교를 통해 로그 손실(Log Loss)의 목적 함수를 최소화하도록 진행된다(배은영, 신오순, 2020). 심층 신경망 협업 필터링 모형은 다른 비교 추천 알고리즘 모델에 비해 성능이 우수하였다. 이는 고객의 제품 선호에 대해 잠재 요인을 비선형 신경망 구조를 통해 기존 일반화 행렬분해의 한계를 해결하였다는 점에 기인한다고 볼 수 있다(He *et al.*, 2017).

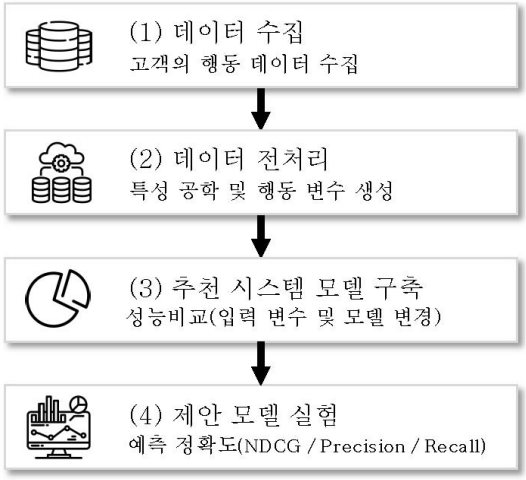
2.4 고객 특성 정보를 반영한 협업 필터링

Ricci *et al.*(2011)은 단순히 사용자와 아이템의 상호관계를 파악하는 것뿐만 아니라 고객의 특성 정보도 반영하는 포괄적인 추천 모델을 제안하였다. 고객의 특성 정보는 고객의 나이와 성별 등 정적인 요소부터 제품의 구매 주기, 앱 활동 및 접속 빈도 등 다양한 활동 요소까지 포함하는 개념이다. Ricci *et al.*(2011)가 제시한 모델은 협업 필터링 구조에서 K개의 특성 정보(Contextual Information)를 반영한 만큼 새로운 차원을 확장한 모형으로 볼 수 있다. 즉, 사용자와 아이템의 관계인 2차원 행렬에서 특성 정보의 수만큼 차원을 확장하며 모형의 입력 층에 투입하는 것이다. 이러한 고객의 특성 정보 반영은 고객의 제품 선호도를 보다 정교하게 반영할 수 있으며, 추천시스템 모델의 성능을 향상시킬 수 있다(김효중 등, 2021). Covington *et al.*(2016)은 유튜브 클립을 고객에게 제안하는 추천 모델을 개발하는 과정 중 고객의 특성 정보를 포함하는 것이 추천시스템의 성능 측면에서 상당한 기여를 한다는 것을 실증적으로 입증하였다. 이러한 협업 필터링에 고객 특성을 반영한다는 기존 연구를 통해, 본 연구에서 제안하는 모델은 고객의 금융상품 선호도에 유의미한 영향을 주는 고객 특성 정보가 입력 층 단계에서 받을 수 있도록 진행되었다.

Ⅲ. 추천시스템 모형

3.1 연구 절차

본 연구는 금융상품을 MTS를 통해 거래한 기록 데이터와 고객의 다양한 특성 정보 데이터를 활용한다. 거래 데이터에는 고객이 거래한 금융상품 대한 잠재적 선호(Implicit feedback)가 포함된다고 볼 수 있으며, 고객 특성 정보에는 고객의 성별, 자산정보, 종목 포트폴리오 등이 있다. 데이터 전처리 및 학습, 검증, 테스트 단계를 거쳐 실제 딥러닝 모델을 통해 고객 선호를 산출할 수 있도



〈그림 1〉 연구 절차

록 한다. 마지막으로 모델 예측력을 다양한 평가 지표를 통해 검증을 진행하고자 한다. 이때 실험 결과는 평가지표를 토대로 기술하고자 한다. 아래 <그림 1>은 본 연구의 절차를 나타내는 그림 자료이다. 연구 단계는 고객의 거래 정보 및 특성 데이터 수집하는 것에서 시작하여, 고객의 금융 데이터 전처리 작업을 진행한다. 이후 학습과 검증 그리고 평가 데이터셋으로 구분한 뒤 추천 모델링 구축과 평가를 수행한다.

3.2 데이터 수집

본 연구는 고객의 투자 행동을 고려한 변수를 활용하여 금융상품 추천시스템 모형을 개발하고자 한다. 본 연구는 글로벌 투자회사인 M사의 실제 고객 특성 및 금융상품 거래 데이터를 기반으로 진행되었으며, 고객의 거래 로그 기록을 바탕으로 고객별 금융상품 선호를 도출할 수 있다. 이를 바탕으로 금융투자업과 금융상품 판매 등에 상품 개발 및 마케팅 전략을 제시함으로써 수익 창출에 기여할 수 있다. 고객의 과거 거래 기록을 토대로 고객이 향후 높은 관심을 나타낼 것으로 예상되는 금융상품을 선제적으로 추천해 주는 모델을 개발하여, 고객의 상품 투자 편의를 제공하고 동시에 금융 투자 내 추천 서비스의 활성화를 기여할 수 있을 것으로 생각한다. 본 연구는 고객의 금융상품 거래 데이터를 기반으로 미거래 금융상품에 대한 예상 선호를 도출하는 추천 모델을 구축하여, 선호가 높을 것으로 예상되는 상위 상품군을 추천하는 서비스를 개발하는 데 실무적 의의가 있다. 고객 특성 및 거래 데이터는 2020년 6월 1일부터 2021년 5월 31일까지 수집된 것을 활용하였다. 해당 데이터에 포함된 1,499,992명의 데이터가 포함되어 있으며, 900개의 금융상품의 거래 내역으로 구성되어 있다. 고객은 평균적으로 3.5개의 계좌를 보유하고 있으며, 주 거래 계좌 및 자산 데이터에 높은 가중치를 부여하여 추천 시스템 모형을 구성하고자 한다. 해당 데이터는 금융보안원의 가명처리 규정에 따라 고객 비식별화 작업을 진행한 데이터이다. 활용 데이터의 경우 개인 레벨의 일별 데이터를 추출하고, 고객별 특성이 주 계좌 거래 및 자산 데이터를 중심으로 활용하여 추천시스템 모형에 대한 금융 유동성 시장의 일반화를 제고하고자 한다. 하기 <표 2>는 입력층에서 투입된 고객 특성과 거래 내역 데이터 항목에 관해 정리한 것이며, 고객의 성별, 자산정보, 종목 보유정보 및 거래 내역에 관해 포함되어 있다.

3.3 모형 개요

본 연구는 고객의 투자상품 선호도 및 리스크 회피도를 반영한 금융상품 추천시스템 모형을 제안한다. 나아가 기존 연구 모델인 행렬분해 모형과 딥러닝 협업 필터링 모형과의 비교를 진행하고

〈표 2〉 활용 데이터 소스 및 활용 변수

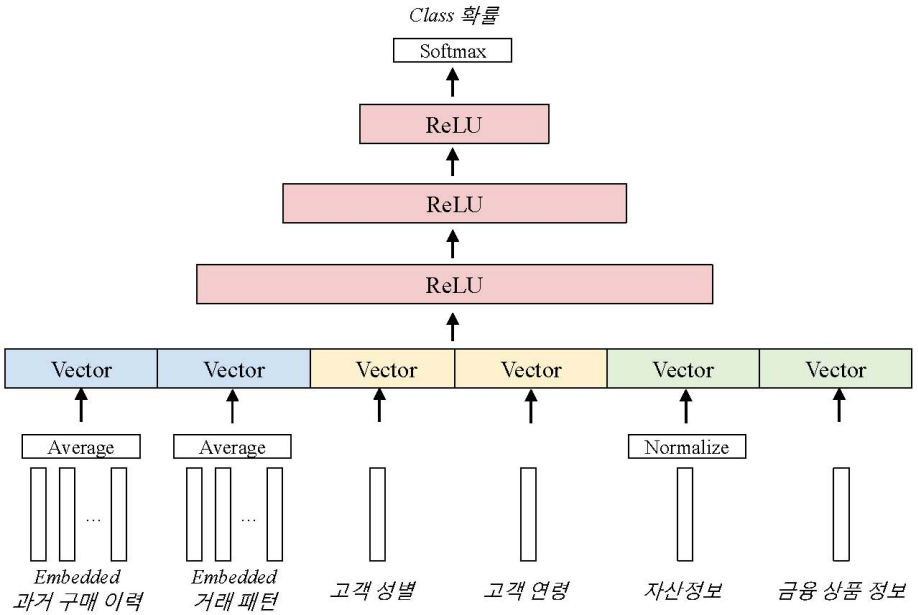
데이터 소스	활용 변수
고객정보	계좌개설일, 연령, 성별, 고객등급, 주거래상품, 주거래매체
주식정보	주식자산, 매수금액, 매도금액, 보유종목수, 거래종목수 등
매체이용 성향	MTS-국내/해외, 자산관리앱 접속일수
퇴직연금/개인연금	퇴직연금잔고, 개인연금잔고
거래패턴	코스피 거래비중, 선호시장, 시총 1천 억 미만선호도 등
신용	신용잔고
종목정보	보유, 거래 Top 3 종목정보
AI 서비스 사용여부	파워랭킹, AI투자진단

〈표 3〉 모델의 핵심 이슈 대응

고려 사항	관련 이슈
고객 특성 정보	입력 변수의 차원 확장(Input Dimension Extension)
제품 선호도	잠재 요인 임베딩 레이어(Latent Factor Layer)
콜드 스타트	모델 기반 협업 필터링(Model based Collaborative Filtering)

자 한다. 우선 본 연구에서 제안하는 모형은 고객의 특성 정보와 고객의 거래내역 정보를 반영하고, 메모리 기반 협업 필터링의 한계점 중 콜드 스타트 이슈를 완화할 수 있는 모델 기반 협업 필터링으로 제안한 점이다. 고객 특성 정보는 입력 변수 차원 확장으로, 고객의 금융상품 선호도는 잠재 요인 임베딩 층(Latent Factor Layer), 행렬분해와 내적 연산, 그리고 딥러닝 방법을 활용한다. 아래 <표 3>은 본 연구에서 제안하는 추천 모델의 특성을 정리한 도식 자료이다.

한편, 제안 모형은 고객의 금융상품 선호에 포함된 잠재 요인을 중심으로 개발한 협업 필터링 모델이며, 이는 기존의 머신러닝 모형의 선형성 한계를 완화한 딥러닝 협업 필터링 모형을 응용하여 개발하였다. 딥러닝을 기반으로 추론하는 고객 금융상품 선호 예측 모델이라는 틀 속에 고객 고객별 투자 성향 및 리스크 회피도를 입력 변수로 활용하도록 구성하였다. 이러한 특성 정보를 딥러닝 모델의 변수로 활용하는 이점은 콜드 스타트(Cold-start) 이슈를 해결할 수 있으며, 데이터 수집이 미약할 때 콘텍스트 정보를 보완할 수 있다는 점에 있다(Covington *et al.*, 2016). 또한 고객의 특성 정보를 파악할 수 있는 다양한 콘텍스트를 통해 초기 시그널이 약한 모델과 신규 데이터의 유입 등을 보완하는 역할로 활용하고 있다(Covington *et al.*, 2016). 해당 데이터를 통해 고객의 금융상품 선호도에 미치는 요인에 대한 모델 설명 가능성을 제고할 수 있을 것으로 기대하며, 모형은 아래 <그림 2>와 같다.



〈그림 2〉 제안 모델 구조

3.4 제안 모형

본 연구에서의 모형은 고객의 거래 로그 데이터 및 특성 정보를 기반으로 고객에게 맞춤형 금융상품을 추천하는 모델로 기획하였다. 이 같은 목적을 실현하고자 상기 <그림 2>의 제안 모형에는 단계별 모형 구축 논리를 포함하고 있다. 먼저 고객과 금융상품을 각각 원-핫 인코딩(One-Hot Encoding)을 활용하여 입력한다. 또한 고객의 특성 정보는 고객마다 라벨인코딩(Label Encoding)으로 입력하게 된다. 본 연구는 금융상품 거래 로그 데이터 및 특성 정보를 활용한 금융상품 추천을 목적으로 진행하는 연구로써, 고객이 금융상품 거래 시 남긴 다양한 거래 정보, 종목 보유 정보, 자산 정보 등 특성 정보들을 적극적으로 활용하였다.

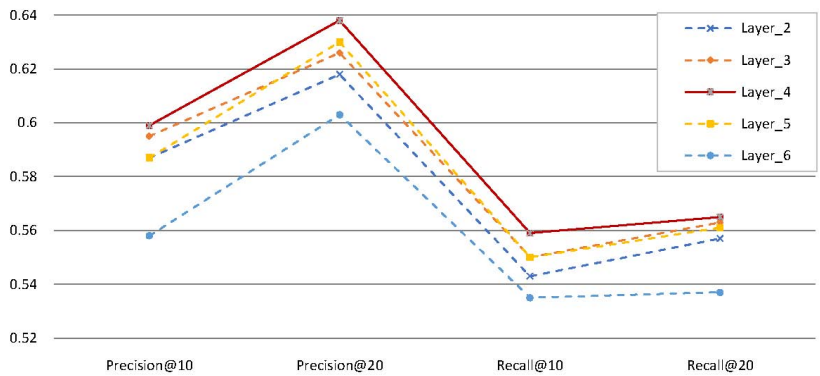
인풋 단계를 넘어선 후로는 잠재 요인의 임베딩 층을 구성하는 단계이며, 고객과 금융상품은 각각 잠재 요인 K개의 의미를 포함하는 임베딩 층으로 구성한다. 본 연구에서는 다양한 K값(K: 200 ~ 400 범위)의 성능 실험을 통해 최적의 K는 300으로 활용하였다. 한편, 고객 특성 정보는 특성 분류 개수와 관련이 있도록 잠재요인 K값을 지정하였다.

임베딩 층을 거친 후에는 일반화 행렬분해 모형과 고객의 거래 내역 및 특성 정보의 출력 값을 각각 계산하는 단계이다. 고객과 금융상품의 잠재요인 임베딩 층은 행렬분해 모형으로 출력 값을 계산한다. 이는 행렬분해가 잠재요인 기반의 고객 선호 계산 및 예측력이 우수하여, 잠재요인 임베딩 층의 계산 과정에 유의미한 역할을 의미하기 때문이다(He *et al.*, 2017). 한편, 고객의 성별과 같은 특성 정보는 고객과 금융상품의 잠재요인 임베딩 층과 달리 행렬분해 모형을 거치지 않고 산출된다. 이는 고객과 금융상품의 잠재요인 임베딩 층에 대비하여 비중이 크지 않고, 계산 비용 측면에서 비효율적이기 때문이다.

행렬분해 모형과 고객의 다양한 특성 정보의 임베딩 층의 출력 값을 연결하고 딥러닝 모형으로 확장하는 단계이다. 행렬분해 출력 값과 고객의 다양한 특성 정보의 출력 값을 단순 결합(Concatenate)하여 DNN모형을 구성한다. 아래 식 (1)을 통해 행렬분해 출력 값과 고객의 다양한 특성 정보인 임베딩 층의 출력 값을 연결한 행렬을 입력 값으로 처리한 후 각 은닉 층과 노드를 거쳐 출력값($\varnothing^{MLP}$)을 계산하는 함수를 적용하였다. 이때 α_L은 각 계층의 활성화 함수를 뜻하며, W는 가중치, 그리고 b_L은 편향(Bias)을 의미한다.

$$\varnothing^{MLP} = \alpha_L\left(W_L^T\left(a_{(L-1)}\left(\cdots a_1\left(W_1^T\begin{bmatrix}p_u^{MLP}\\ q_i^{MLP}\end{bmatrix}+b_2\right)\cdots\right)\right)+b_L\right)+b_L \quad (1)$$

아래 식 (2)를 참조하면, a_{out}에 비선형 모델링을



〈그림 3〉 모델 은닉층 개수 비교

위한 활성화 함수를 적용한 것을 볼 수 있다. 본 모형에서의 최종 활성화 함수는 ReLU 등 다른 활성화 함수를 사용한 실험 결과와 비교해 가장 우수한 성능을 보이는 소프트맥스(Softmax)를 활용하였다.

$$\hat{y}_{ui} = a_{out}(h^T(p_u \circ q_i)) \quad (2)$$

이는 각 출력 계층에서 비선형 활성화 함수 활용을 통해 선형성의 한계를 보완한 기존모형의 연구를 참조하였다(He *et al.*, 2017). 모델의 학습 과정은 확률 경사 하강법(Stochastic Gradient Descent, SGD) 방법을 활용하였으며, 목적 함수를 평균 제곱 오차(Mean Squared Error, MSE)로 정의하여 이를 최소화하도록 모델은 학습하였다. 한편 모델 구성은 5개로 구분하여, 은닉 층 구성을 달리하는 모델을 구축한 이후 성능이 가장 좋은 모델 하나를 선택하여 개발하였다. 은닉 층과 노드 개수에 따른 성능 지표의 차이는 위 <그림 3>을 통해 확인할 수 있으며, 성능이 가장 우수한 Layer_4를 선택하여 개발하였다.

Ⅳ. 실 험

4.1 분석 결과

본 연구에서 제안하는 금융상품 추천 모형은 고객의 거래 및 자산, 종목 보유 정보 등을 기반으로 구축한 추천 모델로, 전통적인 추천 모델과 성능을 비교하여 평가하고자 한다. 이때 평가 기준은 상위 N개 상품을 추천한 후 평가지표의 차이를 확인하려고 한다. 비교하고자 하는 모델은 머신러닝 모델과 다양한 고객 특성 정보를 반영하지 않은 딥러닝 협업 필터링 모형이다. 상위 N개 금융상품 중 추천 모델 성과 평가를 하기 앞서, 본 모형을 통해 계산된 예측 값과 실제 선호 지표와의 MAE를 계산하여 전반적인 모델 예측력을 확인하고자 한다. 이때 MAE는 선호도 평가 수치(실제)를 나타내는 p와 예측 수치를 나타내는 q 사이에 오차 절댓값의 평균을 뜻한다. 아래 식 (3)을 확인하면, MAE를 산출하는 식이며, N은 평가 대상의 개수를 뜻한다.

$$MAE = \frac{\sum_{i=1}^{n} |p - q|}{N} \quad (3)$$

본 연구의 제안 모델인 다양한 고객 정보를 포함한 신경망 모형과 모델 기반의 협업 필터링인 행렬분해 모형 그리고 다양한 고객 특성 정보가 없는 심층 신경망 협업 필터링 추천 모형의 MAE 및 MSE 수치는 다음 <표 4>와 같다. 이때, 평가의 신뢰성 검증을 위해 K(K=5) 교차 확인 방식을 활용하였으며, 각각의 MAE와 MSE 값을 산출하였다. 모델 기반 협업 필터링인 행렬분해 모형(MF)은 MAE 3.012에 그쳤으며, 고객 특성 정보가 없는 딥러닝 협업 필터링 모델(NCF)은 2.783, 그리고 본 연구에서 개발한 다양한 특성 정보를 포함한 딥러닝 협업 필터링 모델(YTB(o))은 2.051을 기록하면서, 본 연구의 예측 모델 성능이 우수하다고 볼 수 있다.

본 연구는 MAE 성능 비교에서 멈추지 않고, TOP N 성능 평가를 진행하였다. TOP N 성능 평가

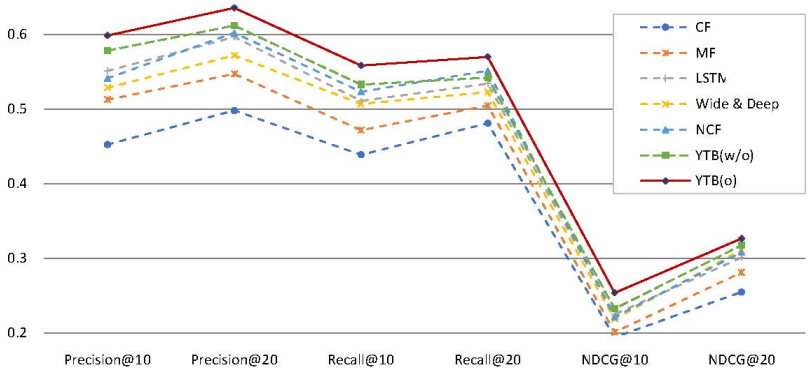
〈표 4〉 MAE와 MSE(CV=5)

Model	MF		NCF		YTB(o)	
	Mean	Std	Mean	Std	Mean	Std
MAE	3.012	0.05029	2.783	0.034694	**2.051**	**0.021187**
MSE	0.8226	0.005161	0.5281	0.006529	**0.4469**	**0.0044872**

방식에서 N은 추천 제품 개수를 뜻하며, 테스트 데이터를 기반으로 내림차순 정렬하여 추천 목록을 생성한 뒤, N개 내 고객의 제품 상호작용 여부에 따른 정밀도(Precision), 재현율(Recall), NDCG (Normalized Discounted Cumulative Gain) 값을 측정하여 비교하는 방식이다. 아래 <그림 4>는 N의 개수 변화를 달리 적용하였을 때의 각 모델별 정밀도, 재현율, NDCG값을 나타낸 자료이다. 이때, NDCG는 추천의 순서에 따라 가중치를 다르게 두어 성능을 평가하는 지표를 뜻한다. 추천된 아이템 순서 목록에서 실제 고객이 선택했는지 여부를 파악하는데 유용하다.

<그림 4>에서 협업 필터링 모형(CF, MF)보다는 심층 신경망 협업 필터링 모형(LSTM, Wide&Deep, NCF)이, 심층 신경망 협업 필터링 추천 모형보다는 본 연구에서 개발한 YTB 모형이 정밀도, 재현율, NDCG 측면에서 탁월한 성과를 나타내는 것을 볼 수 있다. YTB 모형 중에서도 고객 특성 정보가 더 정교하게 반영된 YTB(o) 모형이 그렇지 않은 YTB(w/o) 모형보다 더 우수한 성능을 확인할 수 있다. 결론적으로, 다수의 머신러닝과 딥러닝 모델에 비해 본 연구에서 개발한 YTB 모델이 금융상품 관련 고객 선호 및 리스크 회피도를 예측하여 맞춤형 금융상품을 추천하는 목적 측면에서 탁월하다고 볼 수 있다. 이는 기존의 추천시스템 모형의 한계였던 콜드 스타트 문제를 해결하였다고 볼 수 있다. 즉, 고객의 특성 정보를 반영하지 않은 협업 필터링 기반의 추천시스템 모형보다 리스크 회피도와 투자 선호도 등 고객의 특성을 반영한 협업 필터링 기반의 추천시스템 모형이 우수한 성능을 보일 뿐만 아니라, 금융상품 정보를 반영하였기 때문에 기존 대비 성능이 우수하다고 볼 수 있다. 한편, N값이 20개로 증가함에 따라 우수한 성능을 보이는 점도 주목할 필요가 있다. 이는 고객의 예상 선호도 산출 값이 고객 선호를 충분히 반영하지 못하고 있는 한계점을 보인 것으로 볼 수 있다. 즉, 산출 값이 증가할수록 고객의 다음 행동이 위치해 있다는 뜻이기 때문이다. 하지만, 모델 기반의 협업 필터링과 심층 신경망 협업 필터링 모형에 비해 비교적으로 우수한 성능을 나타냈다는 점에서 의미가 있다고 볼 수 있다. 구체적인 실험결과 값은 아래 <표 5>를 통해 확인할 수 있다.

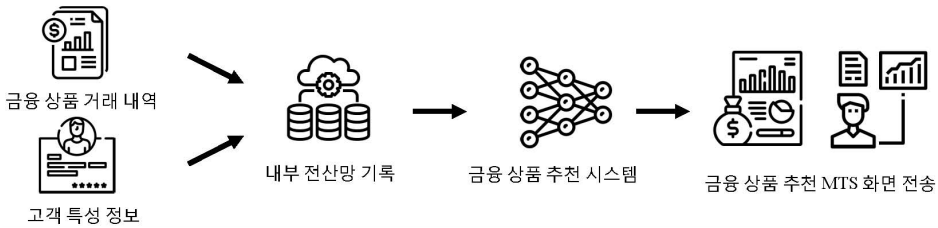
금융상품 추천 모델의 서빙 방식은 <그림 5>과 같이 제안한다. 먼저, 고객의 금융상품 거래 내역 및 특성 정보가 내부 전산망에 입력되면, 추천 금융상품이 상품 번호 형태로 도출되어 해당 품목과 연결된 실제 상품명이 고객 MTS 화면에 전송되는 형태로 이루어질 있다. 이러한 형태의 금융상품 추천시스템을 구축하여 증권 기업의 고객들의 행동 특성을 파악하고 고객 군별 프로파일을 생성하거나 고객 행동을 중심으로 적절한 프로모션과 마케팅 방향을 수립할 수 있다.



〈그림 4〉 TOP-N 변화 실험 결과

〈표 5〉 실험 결과

Model	Precision@10	Precision@20	Recall@10	Recall@20	NDCG@10	NDCG@20
CF	0.4524	0.4981	0.4388	0.4810	0.1942	0.2549
MF	0.5125	0.5471	0.4717	0.5046	0.2012	0.2812
LSTM	0.5512	0.5967	0.5109	0.5341	0.2251	0.3011
Wide & Deep	0.5285	0.5721	0.5068	0.5228	0.2186	0.3114
NCF	0.5411	0.6021	0.5232	0.5513	0.2214	0.3083
YTB(w/o)	0.5782	0.6119	0.5328	0.5422	0.2325	0.3174
YTB(o)	**0.5983**	**0.6355**	**0.5582**	**0.5699**	**0.2538**	**0.3265**



〈그림 5〉 모델 활용 방안 예시

Ⅴ. 토의 및 시사점

5.1 연구 토의 및 시사점

본 연구를 통해 고객의 금융상품 선호도 예측을 통한 금융 투자 상품을 추천하는 서비스 개발을 위한 것으로, 고객의 선호도를 복합적이고 정교하게 반영할 수 있는 방안에 관해 고민하였다. 개발한 모델에서는 크게 세 가지 특장점을 가지고 있다. 첫째, 기존 금융 고객들은 은행 및 기관 투자자와 같은 기존 금융 기관에 대해 기술에 대한 생소함, 규제 위험, 짧은 투자 기간, 높은 거래 비용 및 적절한 금융 메커니즘의 부족으로 인해 투자에 대해 주저한다(Bolton and Foxon, 2015; Hall *et al.*, 2017). 이러한 현상을 해결하기 위해 본 연구는 투자 선호와 리스크 회피도와 같은 고객의 선호도를 반영한 추천시스템이 개발되었다. 둘째, 고객과 금융상품 예상 선호도 간 비선형 관계를 모델에 반영한 것이다. 기존의 행렬분해 머신러닝 모델은 고객과 금융상품 선호 간 평면 형태로 모델링하여 고객의 잠재 선호 요인(Latent Factor)을 오롯이 반영하지 못한 부분이 있다. 하지만 본 모델은 딥러닝 기반의 비선형 모델링을 활용하여 고객의 금융상품 잠재 선호 요인을 충분히 반영하고자 하였다. 세 번째는 고객마다 다양한 특성 요소를 추천 모델에 포함한 것이다. 고객의 금융 잔고 현황, 거래 내역, 종목 보유, 인적 정보 등 개인별 특성이 상이하기 때문에 추천하는 금융상품 또한 이에 맞춤형으로 제공될 필요가 있다. 그런 측면에서 본 연구에서의 모델은 고객의 복합적이고 정교한 특성을 입력 변수에 반영하여 고객 맞춤형 금융상품 추천 서비스가 실현되도록 설계되었다. 한편, 실험 결과에 따른 추천 모델의 성능에 주목할 필요가 있다. 개발 모델의 성능 평가는 행렬분해 머신러닝 모형과 딥러닝 협업 필터링 모형을 중심으로 이루어졌다. 우선, 행렬분해 모형을 사용한 이유는 평면 형태의 일반화 행렬분해 머신러닝 모델에 비해 비선형 관계성을 모델링한 본 연구의 의의를 확인하기 위함이다. 딥러닝 기반 협업 필터링 모형을 활용한 것은 고객의 다양하고 복합적인 특성 정보를 반영한 부분의 효과성을 확

인하고자 한 것이다. 복수의 머신러닝과 딥러닝 모델들의 성능 비교를 보아, 본 연구에서 고객 및 금융상품 선호도 간 비선형 관계를 모델링 한 것과 고객의 다양한 금융 거래 특성 요소를 추천 모델에 직접 반영한 것은 상당히 성능측면에서 효과가 있는 것으로 평가할 수 있다.

본 연구에서 지닌 학술적 시사점은 다음과 같다. 첫째, 기존의 금융상품 추천 연구는 고객의 정교한 금융 거래 및 특성 정보를 반영하지 않은 채 금융상품에 대한 고객 잠재 선호도나 제품 관련 컨텐츠를 중심으로 한 추천시스템 연구가 진행되었다(김석수, 2014). 즉, 기존 금융상품 추천시스템 관련 연구는 다양한 금융 정보와 고객 행동 정보를 반영하지 못하여 고객의 선호도와 리스크 회피도와 같은 컨텍스트 정보를 반영하지 못한 한계점이 존재하였다. 따라서, 본 연구는 금융 거래 정보를 포함하는 고객의 다양한 요소를 반영한 추천시스템 모델을 제안한 것이다. 금융상품 거래 정보부터 고객의 성별과 나이, 자산 정보까지 고객의 다양한 특성과 거래 정보를 동시에 반영하여 고객의 선호도와 리스크 회피도를 산출한 추천시스템을 제안하여 기존 금융상품 추천시스템을 확장한 것에 의의가 있다.

둘째, 기존 금융상품 분야의 추천시스템 연구는 고객에게 평점, 설문조사 등 소수 고객고객부터 주관적인 선호도 지표를 활용한 직접 평가 데이터를 제공받아 진행하였다(이재웅 등, 2016). 이는 주관적이고 단기간의 데이터를 사용하였으며, 모델의 편향성을 가중시킨다는 한계가 존재한다. 본 연구는 고객의 실제 거래 행위 등 잠재적 선호(Implicit feedback)를 활용하여 고객 행동을 직접 평가하였다. 이를 통해 기존 추천시스템에서 갖는 데이터 희소성, 콜드 스타트, 주관성 편향 등의 한계점을 보완하였으며, 고객의 다양한 특성을 파악하여 적시에 금융상품 추천을 제안할 수 있다. 이러한 점을 고려하면, 온라인 고객의 잠재적 선호도 행동을 고려하고 고객의 의도 정보를 반영한 추천시스템 연구로 확장할 수 있다(이동원, 2020; 이동훈, 김남규, 2021) 아울러, 방대한 규모의 금융 고객 서비스에 활용하기 위해 대용량의 데이터에서 모델 효율화 연구에 대한 중요성을 강조할 수 있다(김석수, 2014).

셋째, 최근접 이웃 또는 컨텐츠 유사도 산출을 위해 막대한 컴퓨팅 능력이 요구되는 모델 연구(Matsunami *et al.*, 2017; Patty *et al.*, 2018; Yim *et al.*, 2016)와 달리 본 연구는 모델 기반 협업 필터링 모델을 제안함으로써, 비용-효익(Cost-benefit) 관점에서 효과적인 모델을 개발한 부분이다. 이는 고객의 상호작용이 지속적으로 증가하여 실시간 금융상품에 대한 추천 서비스를 위해 데이터 처리를 시도할 때, 본 연구에서 제시하는 고객 특성을 반영한 협업 필터링 방법의 활용성을 제시할 수 있다. 즉, 다양한 고객 행동과 특성 정보를 반영한다는 점에서 추천 서비스의 성능이나 고객의 만족도 측면에서 효율성이 높아질 것으로 볼 수 있다(장예화 등, 2021).

본 연구가 가진 실무적 시사점은 아래와 같다. 첫째, 쏟아지는 금융상품 정보 속에서 선택의 복잡성과 금융상품의 위험성에 직면한 고객에게 적합한 상품을 추천하여 올바른 투자 문화를 이끌 수 있다는 것이다. 즉, 리스크 회피도, 거래 내역, 성별, 나이 등 고객의 행동과 특성을 고려하며 각각의 적합한 상품을 제공함으로써 금융상품 추천 서비스에 대한 긍정적인 반응을 유도하고, 투자 서비스에 대한 고객 충성도와 만족도를 높이는 등 금융 산업의 패러다임을 변화할 수 있을 것이다(Komiak and Benbasat, 2006).

둘째, 금융투자회사는 고객의 선호를 조기에 파악하여, 맞춤형 금융상품 개발과 타겟 프로모션을 통해 불필요한 마케팅 비용을 낮출 수 있다. 이러한 금융상품 추천시스템을 통해 금융투자회사의 인건 비용 절감, 고객 금융상품의 가입 증가를 통해 이익 확대 등을 기대할 수 있다. 또한, 고객에게 적절한 상품과 혜택을 제안한다는 점에서 기업의 경쟁 전략 중 중대한 활용방안으로 여겨질 것이다(방영석 등, 2011).

셋째, 고객의 복합적인 특성 정보를 바탕으로 적합한 금융상품을 추천함으로써, 고객의 투자 경험을 개선할 수 있을 것이다. 고객들은 스팸 메시지와 무의미한 상품 추천보다 자신의 관심사와 취향을 고려한 추천 서비스에 대해 우호적인 것으로 알려져 있다(최용석, 2010). 이러한 점을 고려할 때, 고객의 행동과 특성을 고려하지 않은 추천시스템보다 개개인의 상황에 적절한 금융상품을 추천하여 고객이 만족할 것으로 기대한다. 이를 통해 고객 투자 여정 및 성향을 고려한 맞춤형 금융상품 추천을 진행하여 금융 분쟁 사유인 '부적합 금융상품권유에 따른 금전적 손해 발생'에 대한 법률적 비용(Legal Cost) 및 고객 서비스(CS) 비용을 감축할 수 있으며, ESG 경영을 위한 초석을 마련할 수 있다.

5.2 한계점 및 향후 연구 방향

본 연구의 한계점은 세 가지가 존재한다. 첫째, TOP-N 추천 성능 지표 검증에서 10개와 20개의 상품을 기준으로만 검증한 것이다. 그 사이 추천 상품 개수에 따른 성능 지표 변화 모두 실험하지 않았기 때문에, 최적화된 상품 추천 개수를 제시하지 못하였다. 두 번째는 고객의 다양한 특성 요소를 제한적으로 활용하였다. 본 연구에서 활용한 기록과 데이터는 실제 금융 서비스를 진행하고 있는 기업 데이터이기 때문에 금융 당국의 규제에 저촉되지 않는 수준에서 정보를 수집할 수밖에 없었다. 따라서, 고객의 다양한 특성 데이터를 대상으로 복합적이고 확장성 있는 실험에 제약하여 진행할 수 없었던 것이 한계로 볼 수 있다. 마지막으로, 콜드 스타트 문제를 고객의 특성과 금융상품 정보를 통해 해결하고자 하였으나, 이에 대한 고객 설문과 같은 실질적인 검증이 필요하다.

향후 연구 방향은 다음과 같다. 첫째, TOP-N 추천 방식을 넘어서서 최적의 N 값을 도출하는 부분에 관한 연구가 필요하다. 본 연구에서 금융상품에 대한 예상 선호도 상위 10개와 20개를 추천하였을 때 고객이 해당 금융상품을 선택한 비율을 토대로 실험을 수행하였지만, 고객 경험 개선과 추천의 효과적인 측면에서 적합한 N 값을 찾는 것은 상당한 의미가 있기 때문이다. 둘째, 고객이 남긴 VOC와 STT와 같은 비정형 데이터이자 텍스트 데이터는 고객의 선호와 리스크 회피도를 예측하는 데 도움이 된다. 따라서 텍스트 자료를 포함한 다양한 데이터 기반의 금융상품 추천시스템 모델 연구를 진행한다면 학술적인 측면과 실무적 관점에서 크게 역할을 다할 수 있을 것으로 기대한다. 마지막으로, 기존 협업 필터링의 추천시스템 모형 대비 고객의 특성과 금융상품 정보를 반영한 추천시스템 모형이 초기 유저에 대한 상품의 다양성이나 고객 만족에 영향을 미쳤는지 사후 분석이나 설문이 필요하다. 이를 통해 실질적으로 콜드 스타트 문제에 대한 본 모형의 영향력을 확인할 수 있을 것이다.

참 고 문 헌

[1] 김석수, "금융 상품 추천에 관련된 빅 데이터 활용을 위한 개발 방법", *한국컴퓨터정보학회 논문지*, 제19권, 제8호, 2014, pp. 73-81.

[2] 김효중, 신우식, 신동훈, 김희웅, 김화경, "고객의 특성 정보를 활용한 화장품 추천시스템 개발", *Information Systems Review*, Vol.23, No.4, 2021, pp. 69-84.

[3] 방영석, 이동주, 배윤수, "개인화 서비스의 수용에 있어서 인지된 개인화와 이해의 역할", *경영학연구*, 제40권, 제2호, 2011, pp. 355-382.

[4] 배은영, 신오순, "Neural Collaborative Filtering 기반 개인 맞춤형 운동 추천 알고리즘", *한국통신학회 학술대회논문집*, 2020, pp. 313-314.

[5] 송희석, "심층신경망 기반의 뷰티제품 추천시스템", *Journal of Information Technology Applications & Management*, 제26권, 제6호, 2019, pp. 89-101.

[6] 원종윤, 이건창, "금융상품의 재무정보 형식에

따른 사용자의 주의할당과 선호도에 관한 실증연구: 아이트래킹 실험을 중심으로", *Information Systems Review*, 제23권, 제2호, 2021, pp. 1-19.

[7] 이은주, 송재오, 김이나, 유재수, "화장품 추천을 위한 개인의 피부 유형 및 유전자를 이용한 빅데이터 분석 기반 모바일 서비스", *한국콘텐츠학회 종합학술대회 논문집*, 2018, pp. 495-496.

[8] 이재웅, 김영식, 권오병, "비정형 데이터 분석을 통한 금융소비자 유형화 및 그에 따른 금융상품 추천 방법", *한국 IT 서비스학회지*, 제15권, 제4호, 2016, pp. 1-24.

[9] 장예화, 이청용, 최일영, 김재경, "리뷰 데이터 마이닝을 이용한 하이브리드 추천시스템 개발: Amazon Kindle Store 데이터 분석사례", *Information Systems Review*, 제23권, 제1호, 2021, pp. 155-172.

[10] 최용석, *아이폰과 아이패드 애플의 전략*, Initial Communications Corp., 2010.

[11] Bobadilla, J., F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey", *Knowledge-Based Systems*, Vol.46, 2013, pp. 109-132.

[12] Bolton, R. and T. J. Foxon, "A socio-technical perspective on low carbon investment challenges: Insights for UK energy policy", *Environmental Innovation and Societal Transitions*, Vol.14, 2015, pp. 165-181, Available at https://doi.org/https://doi.org/10.1016/j.eist.2014.07.005.

[13] Covington, P., J. Adams, and E. Sargin, "Deep neural networks for youtube recommendations", *Proceedings of the 10th ACM Conference on Recommender Systems*, 2016.

[14] Cremonesi, P., F. Garzotto, and R. Turrin, "Investigating the persuasion potential of recommender systems from a quality perspective: An empirical study", *ACM Transactions on Interactive Intelligent Systems (TiiS)*, Vol.2, No.2, 2012, pp. 1-41.

[15] Cronqvist, H. and S. Siegel, "The genetics of investment biases", *Journal of Financial Economics*, Vol.113, No.2, 2014, pp. 215-234, Available at https://doi.org/https://doi.org/10.1016/j.jfineco.2014.04.004.

[16] Gao, C., X. He, D. Gan, X. Chen, F. Feng, Y. Li, T.-S. Chua, and D. Jin, "Neural multi-task recommendation from multi-behavior data", *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, 2019.

[17] Gholamian, M., M. Fathian, M. Julashokri, and A. Mehrbod, "Improving electronic customers' profile in recommender systems using data mining techniques", *Management Science Letters*, Vol.1, No.4, 2011, pp. 449-456.

[18] Gigli, A., F. Lillo, and D. Regoli, *Recommender Systems for Banking and Financial Services*. RecSys Posters, 2017.

[19] Hall, S., T. J. Foxon, and R. Bolton, "Investing in low-carbon transitions: Energy finance as an adaptive market", Climate Policy, Vol.17, No.3, 2017, pp. 280-298.

[20] He, X., L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering", *Proceedings of the 26th International Conference on World Wide Web*, 2017.

[21] He, X., X. Du, X. Wang, F. Tian, J. Tang, and T.-S. Chua, "Outer product-based neural collaborative filtering", 2018, arXiv preprint arXiv:1808.03912.

[22] Jannach, D., M. Zanker, A. Felfernig, and G. Friedrich, *Recommender Systems: An Introduction*, Cambridge University Press, 2010.

[23] Kim, M.-G. and K.-J. Kim, "Recommender systems using SVD with social network information", *Journal of Intelligence and Information*

Systems, Vol.22, No.4, 2016, pp. 1-18.

[24] Komiak, S. Y. and I. Benbasat, "The effects of personalization and familiarity on trust and adoption of recommendation agents", *MIS Quarterly*, 2006, pp. 941-960.

[25] Koren, Y., R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems", *Computer*, Vol.42, No.8, 2009, pp. 30-37.

[26] Matsunami, Y., M. Ueda, and S. Nakajima, "How to find similar users in order to develop a cosmetics recommender system", *International MultiConference of Engineers and Computer Scientists*, 2017.

[27] Park, S.-T. and W. Chu, "Pairwise preference regression for cold-start recommendation", *Proceedings of the third ACM conference on Recommender systems*, 2009.

[28] Patty, J. C., E. T. Kirana, and M. S. D. K. Giri, "Recommendations system for purchase of cosmetics using content-based filtering", *International Journal of Computer Engineering and Information Technology*, Vol.10, No.1, 2018, pp. 1-5.

[29] Persson, P., "Attention manipulation and information overload", *Behavioural Public Policy*, Vol.2, No.1, 2018, pp. 78-106.

[30] Schafer, J. B., D. Frankowski, J. Herlocker, and S. Sen, "Collaborative filtering recommender systems", In The adaptive web (pp. 291-324), Springer, 2007.

[31] Su, X. and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques", *Advances in Artificial Intelligence*, 2009.

[32] Yim, Y.-J., H.-S. Bae, Y.-J. Jeong, M.-Y. Kim, A. Nasridinov, K. H. Yoo, and J.-E. Hong, "A user driven cosmetic item recommendation system by character recognition", *Proceedings of the Korea Information Processing Society Conference*, 2016.

[33] Yoon, J. and S. Joung, "A big data based cosmetic recommendation algorithm", *Journal of System and Management Sciences*, Vol.10, No.2, 2020, pp. 40-52.

[34] Zatevakhina, A., N. Dedyukhina, and O. Klioutchnikov, "Recommender systems - The foundation of an intelligent financial platform: Prospects of development", *2019 International Conference on Artificial Intelligence: Applications and Innovations (IC-AIAI)*, 2019.

[35] Zhang, S. X., Y. Wang, A. Rauch, and F. Wei, "Unprecedented disruption of lives and work: Health, distress and life satisfaction of working adults in China one month into the COVID-19 outbreak", *Psychiatry Research*, Vol.288, 2020, pp. 112958, Available at https://doi.org/https://doi.org/10.1016/j.psychres.2020.112958.

Information Systems Review
Volume 25 Number 1
February 2023

Financial Products Recommendation System Using Customer Behavior Information

Hyojoong Kim* · SeongBeom Kim** · Hee-Woong Kim***

Abstract

With the development of artificial intelligence technology, interest in data-based product preference estimation and personalized recommender systems is increasing. However, if the recommendation is not suitable, there is a risk that it may reduce the purchase intention of the customer and even extend to a huge financial loss due to the characteristics of the financial product. Therefore, developing a recommender system that comprehensively reflects customer characteristics and product preferences is very important for business performance creation and response to compliance issues. In the case of financial products, product preference is clearly divided according to individual investment propensity and risk aversion, so it is necessary to provide customized recommendation service by utilizing accumulated customer data. In addition to using these customer behavioral characteristics and transaction history data, we intend to solve the cold-start problem of the recommender system, including customer demographic information, asset information, and stock holding information. Therefore, this study found that the model proposed deep learning-based collaborative filtering by deriving customer latent preferences through characteristic information such as customer investment propensity, transaction history, and financial product information based on customer transaction log records was the best. Based on the customer's financial investment mechanism, this study is meaningful in developing a service that recommends a high-priority group by establishing a recommendation model that derives expected preferences for untraded financial products through financial product transaction data.

Keywords: Recommender System, Deep Neural Network, Latent Factor, Customer Attributes Information, Financial Product

* Graduate School of Information, Yonsei University
** Integrated PhD program, Graduate School of Information, Yonsei University
*** Corresponding Author, Professor, Graduate School of Information, Yonsei University

◐ 저 자 소 개 ◑

김 효 중 (libertygood9@naver.com)

연세대학교 정보대학원에서 비즈니스 빅데이터 분석을 전공하고, 정보시스템학 석사 학위를 취득하였다. 디지털 서비스 연구실에서 데이터분석 관련 연구들을 진행하였으며, 주요 연구 분야는 빅데이터 분석, 딥러닝, 추천시스템 등이다.

김 성 범 (sb_kim@yonsei.ac.kr)

연세대학교 정보대학원에서 비즈니스 빅데이터 분석 트랙 석박사통합과정이다. 현재 디지털 서비스 연구실에서 빅데이터 분석 관련 연구를 진행 중이다. 주요 연구분야는 Customer Behavior, Personalized Service, Deep Learning, Recommender System 등이다.

김 희 웅 (kimhw@yonsei.ac.kr)

National University of Singapore 정보시스템학과에서 근무한 후, 현재 연세대학교 정보대학원 교수로 재직 중이다. 주요 연구분야는 디지털 비즈니스, 정보시스템 관리 및 활용 등이다. 관련 연구들은 MIS Quarterly, Information Systems Research, Journal of Management Information Systems 등에 60여 편의 논문이 게재되었다. MIS Quarterly, JAIS, IEEE TEM의 편집위원으로 활동했고, KrAIS 회장을 역임했다.

논문접수일 : 2022년 07월 30일 **게재확정일 : 2023년 01월 10일**
1차 수정일 : 2022년 10월 10일 **2차 수정일 : 2022년 11월 30일**