

## Personalized Size Recommender System for Online Apparel Shopping: A Collaborative Filtering Approach

Dongwon Lee\*

\*Assistant Professor, Division of Social Sciences, Hansung University, Seoul, Korea

### [Abstract]

This study was conducted to provide a solution to the problem of sizing errors occurring in online purchases due to discrepancies and non-standardization in clothing sizes. This paper discusses an implementation approach for a machine learning-based recommender system capable of providing personalized sizes to online consumers. We trained multiple validated collaborative filtering algorithms including Non-Negative Matrix Factorization (NMF), Singular Value Decomposition (SVD), k-Nearest Neighbors (KNN), and Co-Clustering using purchasing data derived from online commerce and compared their performance. As a result of the study, we were able to confirm that the NMF algorithm showed superior performance compared to other algorithms. Despite the characteristic of purchase data that includes multiple buyers using the same account, the proposed model demonstrated sufficient accuracy. The findings of this study are expected to contribute to reducing the return rate due to sizing errors and improving the customer experience on e-commerce platforms.

▶ **Key words:** Clothing Size, Recommender System, Collaborative Filtering, Personalization, Fashion E-Commerce, Machine Learning

### [요 약]

본 연구는 의류의 디자인 간 치수의 불일치와 비표준화로 인해 온라인 구매 시 발생하는 치수 선택의 오류 문제를 해결할 수 있는 방안을 제시하기 위해 수행되었다. 본 논문은 구매자에게 개인화된 치수를 제시할 수 있는 기계 학습 기반 추천 시스템의 구현 방안을 다루고 있다. 온라인 상거래로부터 발생된 구매 데이터를 사용하여 비음수 행렬 분해(NMF), 특이값 행렬 분해(SVD), k-최근접 이웃(KNN), 공동 클러스터링(Co-Clustering) 등 여러 검증된 협업 필터링 알고리즘을 훈련하였고, 이들 간에 성능을 비교하였다. 연구 결과, 비음수 행렬 분해 (NMF) 알고리즘이 다른 알고리즘들보다 뛰어난 성능을 보임을 확인할 수 있었다. 동일한 계정을 사용하는 여러 구매자가 포함되는 구매 데이터의 특성에도 불구하고, 제안 모형은 충분한 정확도를 보였다. 본 연구의 결과는 치수 선택의 오류로 인한 반품률을 감소하고 전자상거래 플랫폼에서의 고객 경험을 향상시키는 데 기여할 것으로 기대된다.

▶ **주제어:** 의류 치수, 추천 시스템, 협업 필터링, 개인화, 패션 전자 상거래, 기계 학습

- 
- First Author: Dongwon Lee, Corresponding Author: Dongwon Lee
  - \*Dongwon Lee (dongwonlee@hansung.ac.kr), Division of Social Sciences, Hansung University
  - Received: 2023. 07. 19, Revised: 2023. 08. 03, Accepted: 2023. 08. 08.

## I. Introduction

인터넷과 스마트폰의 보급과 이에 따른 전자상거래의 발전으로 인해 온라인 플랫폼을 통해 의류를 구매하는 건수가 급격히 증가하고 있다[1]. 이러한 추세는 특히 최근 코로나19의 대유행과 같이 오프라인 매장 방문이 제한되는 상황에서 더욱 두드러지게 나타난다. 하지만, 온라인 전자상거래 환경에서 구매자는 편안하게 다양한 의류를 탐색하고 주문할 수 있는 반면, 자신에게 적합한 치수를 선택하는 과정에서 겪게 되는 어려움과 잘못된 치수를 선택함으로 인한 부작용도 더욱 심화되고 있다. 이러한 어려움과 부작용은 구매자와 판매자 모두에게 영향을 미친다.

온라인으로 의류를 구매하는 고객들은 제품의 치수를 선택하는 데 어려움을 겪는데[2, 3]. 제품 치수는 제조사나 브랜드마다 차이가 있으며, 개인의 신체 구조, 취향, 제품의 소재나 디자인 등에 따라 적합한 치수가 달라지기 때문이다. 따라서, 판매자가 제공하는 상품 정보나 구매 후기만으로는 완벽하게 적합한 치수를 선택하기 어려울 수 있다.

또한, 온라인 판매자들은 고객이 올바른 제품 치수를 선택할 수 있도록 치수조건표를 제공하는 노력을 기울이지만, 잘못된 치수 선택으로 인한 반품으로 어려움을 겪고 있다. 반품 처리는 비용과 시간이 많이 소요되며, 제품의 상태가 원래대로 유지되지 않았을 때는 재판매가 어렵다. 반품 과정은 재고 관리에 추가적인 노력을 소모하게 하며, 이 과정의 불편함은 고객의 만족도와 장기적인 충성도를 저하할 수 있다. 따라서, 고객이 자신의 체형에 일치하는 치수를 선택하는 것이 반품을 최소화하고 긍정적인 고객 관계를 유지하는 데 중요하며, 이를 위해서는 고객의 신체 치수와 취향을 고려하여 정확한 치수를 예측할 수 있는 치수 추천 시스템이 필요하게 되었다.

그러나, 의류 치수의 비표준화는 온라인 치수 추천 시스템의 개발에 어려움을 초래한다. 브랜드별, 지역별, 심지어는 동일 브랜드 내에서도 디자인에 따른 치수의 차이가 발생하며, 이는 정확한 치수의 추천을 어렵게 함으로써 치수 추천 알고리즘의 성능을 저하하는 원인이 된다. 많은 연구가 키, 몸무게와 같은 신체 치수를 기반으로 가장 적합한 제품 치수를 추천하고자 하는 방법을 시도했으나, 이는 사용자에게 의해 측정된 치수의 부정확성이나 사용자가 치수 제공에서 느끼는 거부감 등으로 인해 정확한 신체 치수의 확보에 어려움을 겪을 수 있을 뿐만 아니라, 이를 확보하더라도 개인의 스타일, 핏(Fit)과 같이 신체 치수를 넘어서는 요소들로 인해 사용자가 실제 선호하는 치수를 찾아내는 데 한계를 겪고 있다.

본 연구에서는 추가적인 노력으로 확보되는 데이터를 활용하지 않고도, 이미 판매자가 보유하고 있는 구매 데이터만으로 적합한 치수를 찾아낼 수 있는 치수 추천 시스템을 제안한다. 이 시스템은 협업 필터링과 앙상블 학습 기법을 사용하며, 신체 치수의 측정 없이도 높은 수준으로 정확한 치수를 예측하는 모형을 개발하고 실제 데이터를 통해 성능을 확인하고자 한다.

이후의 내용은 다음과 같다. II장에서는 협업 필터링 기법 및 전자상거래에서의 치수 추천 현황에 대해 살펴보고, III장에서는 제안 모형을 설명하며, IV장에서 실험 결과를 다룬다. 마지막으로, V장에서는 본 연구의 내용을 간략히 정리하고, 기여점과 한계점에 대해 논의하도록 한다.

## II. Preliminaries

### 1. Related works

#### 1.1 Collaborative Filtering

추천 시스템(Recommender System)은 Netflix, Amazon과 같은 온라인 디지털 플랫폼에서 널리 사용되며 사용자의 선호에 따라 개인화된 맞춤형 추천을 제공함으로써 큰 인기를 얻게 되었다 [3]. 이런 시스템은 사용자가 생성한 대량의 데이터로부터 가치 있는 정보를 추출할 수 있는 알고리즘을 사용함으로써, 비즈니스의 성과를 높이는 데 큰 역할을 하고 있다 [4]. 이들 알고리즘 중, 협업 필터링(Collaborative Filtering, CF)은 추천 시스템을 구축하기 위해 가장 널리 채택되고 있는 방법으로서, 사용자의 행동을 기반으로 추천 목록을 생성한다 [5]. 즉, 협업 필터링 기법은 사용자들이 과거에 선호했던 아이템이 미래에도 그들의 선택에 영향을 미칠 것이라는 가정에 기반한다 [6].

협업 필터링(CF) 기법은 메모리 기반 CF와 모델 기반 CF로 구분된다 [7]. 메모리 기반 알고리즘은 일반적으로 사용자 또는 아이템 간의 유사성을 계산하고 전체 사용자-아이템 데이터 집합을 사용하여 예측 결과를 생성한다. 반면에, 모델 기반 알고리즘은 사용자-아이템 데이터셋을 사용하여 예측 모델을 훈련한 다음 이를 사용하여 추천 목록을 생성한다 [4].

메모리 기반 알고리즘의 예로는 사용자 기반 또는 아이템 기반 협업 필터링을 위한 K-최근접 이웃(K-Nearest Neighbors, KNN) 기법이 있다 [8]. 모델 기반 알고리즘의 예로는 행렬 분해(Matrix Factorization) 방법을 이용하는 특이값 분해(Singular Value Decomposition, SVD)

또는 비음수 행렬 분해(Non-negative Matrix Factorization, NMF)가 있으며, 클러스터링 방법을 이용하는 공동 클러스터링(Co-Clustering)이 있고, 신경망과 같이 훨씬 더 복잡한 머신 러닝 모델이 있다 [9].

K-최근접 이웃(K-Nearest Neighbors, KNN) 기법은 메모리 기반의 협업 필터링 기법의 매우 단순한 알고리즘으로서, 모든 인스턴스를 사전에 저장해두고, 새 인스턴스에 대해 저장된 인스턴스 K개와의 유사도를 측정하는 방식으로 분류를 시행한다 [8]. 추천 시스템의 맥락에서 보면, KNN 알고리즘은 타겟(Target) 사용자 혹은 아이템과 가장 가까운 K개의 사용자 혹은 아이템을 찾고, 그들에 대한 평가를 활용하여 추천 아이템에 대한 타겟(Target) 사용자의 평가를 예측한다 [10].

특이값 분해(Singular Value Decomposition, SVD) 기법은 모형 기반의 행렬 분해 기법으로서, 사용자-아이템 상호작용 행렬을 두 개의 저차원 직사각형 행렬의 곱으로 분해하는 방식으로 작동한다 [9]. 이 행렬 중, 아이템 행렬은 아이템의 잠재적인 특성을 설명하는 것으로, 사용자 행렬은 각 사용자가 이러한 특성을 얼마나 좋아하는지를 설명하는 것으로 볼 수 있다 [4]. 이 두 행렬의 곱(Product)을 통해 얻은 행렬은 각 아이템에 대한 사용자의 평점을 예측한 값을 갖게 된다 [9].

모형 기반의, 또 다른 행렬 인수분해 알고리즘으로서, 비음수 행렬 분해 (Non-negative Matrix Factorization, NMF) 기법은 SVD와 마찬가지로 사용자-아이템 상호작용 행렬을 두 개의 저차원 직사각형 행렬의 곱으로 분해한다. 그러나, SVD와는 달리, NMF기법은 행렬의 값을 비음수(Non-negative)로 제한하는데, 이는 각 아이템의 특성을 엄격하게 양수 또는 0의 합으로 해석할 수 있게 해준다는 점에서 이점을 갖는다 [11].

공동 클러스터링(Co-Clustering) 기법은 모형 기반의 알고리즘으로서, 사용자와 아이템을 동시에 군집화(Clustering)하는 방식으로 작동한다 [12]. 이는 주문이나 평점과 같이 사용자와 아이템 간의 상호작용을 포함하는 행렬의 차원을 줄이는 방식으로 동작하는데, 특히 희소 데이터 세트에서 예측값을 만드는 데 유리하다는 장점을 지닌다.

### 1.2 Size and Fit in Fashion E-Commerce

패션 산업에서 옷의 치수와 착용감은 중요한 요소로, 온라인 쇼핑의 한계로 인해 그 중요성이 더욱 강조된다 [13]. 온라인 쇼핑을 통해 고객들은 직접 옷을 입어보지 않고 치수나 착용감을 판단해야 하기 때문이다. 이로 인해 패션 전자상거래에서는 올바른 치수 예측이 중요하며, 이는 고

객의 만족도를 높이고 반품을 줄이는 데 도움이 된다 [14].

온라인 쇼핑에서 치수 선택의 어려움을 해결하기 위해 다양한 방법들이 제안되었다. 예를 들어, 예측 알고리즘, 가상 피팅, 3D 스캐닝 등이 사용되어왔다. 예측 알고리즘은 고객의 체형 정보, 과거 구매 정보 등을 사용하여 가장 적합한 치수를 예측한다. 가상 피팅은 고객의 디지털 아바타를 사용하여 옷의 착용감을 시뮬레이션한다. 3D 스캐닝은 고객의 신체를 정확하게 측정하여 개인화된 치수를 추천한다.

최근에는, 머신 러닝과 인공지능 기술이 이 영역에서 빠르게 발전하고 있다. 이러한 기술을 사용하면, 고객의 치수와 착용감에 대한 피드백 데이터를 활용하여 더욱 정확한 치수 예측과 개인화된 제품 추천이 가능해진다 [13]. 또한, 이러한 데이터를 활용하여 제품 디자인과 제조 과정을 개선하는 것이 가능하다 [15].

하지만, 이러한 기술들의 활용에는 여전히 도전적인 요소들이 있다. 예를 들어, 정확한 신체 측정 정보를 얻는 것, 개인정보 보호 문제, 고객의 착용감에 대한 주관적인 피드백을 정량화하는 것 등이 있다. 이러한 문제들을 극복하기 위한 연구가 계속 진행되고 있다 [16].

## III. Collaborative Filtering-based Size Recommender System

이 연구에 사용된 방법론은 데이터 수집 및 전처리, 모델 훈련, 모델 평가와 같은 몇 가지 주요 단계를 포함한다. 사용된 기본 모델은 Python의 Surprise 라이브러리에 있는 협업 필터링 모델이다. 본 연구에서의 치수 추천 시스템의 개발 절차는 Fig. 1과 같다.



Fig. 1. Size Recommender System Development Procedure

### 1. Data Collection and Preprocessing

#### 1.1 Data Collection

본 연구에서는 패션 전자상거래 기업으로부터 수집된 데이터를 사용하였다. 이 데이터세트는 온라인을 통해 거래된 40만 건의 주문 정보로서, 사용자 ID, 디자인 ID 및 주문한 옷의 치수를 나타내는 열로 구성된다. 훈련과 검증에 각각 활용되기 위해, 동일 사용자가 최소한 두 번 이상

주문한 경우에만 해당 사용자의 주문을 데이터세트에 유지하였다.

### 1.2 Data Preprocessing

모형의 정확도를 높이기 위한 전처리 과정으로서 이상치를 제거하였다. 본 연구에서는 치수 선택 과정에서 일관성이 떨어지는 주문을 이상치로 간주하였다. 동일 사용자가 여러 벌의 의복을 구매하는 과정에서 각각 서로 다른 치수를 선택한다면 이는 치수 선택 일관성이 결여된 행동으로 보고 이상치로 의심하게 된다. 예를 들어, 동일한 구매자가 의복A는 'Small'을 구매하고, 동일한 복종인 의복B는 'Large'를 구매한다면, 일관성이 낮다고 보는 것이다. 이런 경우는 여러 명이 하나의 계정을 공유하거나, 지인에게 부탁을 받아 주문하는 상황에서 많이 발생하기도 한다. 이렇게 예외적으로 구매된 옷은 사용자 본인의 체형과 무관하므로 해당 구매 내역이 데이터세트에서 제외되는 것이 맞을 것이다.

그러나, 동일한 복종에 속하는 옷 중에서도 제조사에 따라서는 치수의 차이가 큰 경우도 종종 발생한다. 예를 들어, 의류회사A가 만든 'Small' 치수의 셔츠가 잘 맞는 사용자에게, 의류회사B가 만든 'Small' 치수의 셔츠가 맞지 않아 'Medium'을 구매해야 하는 경우가 발생할 수 있다. 따라서, 서로 다른 치수의 의복이 동일 사용자의 주문 데이터에 포함되었다는 이유만으로, 치수 선택의 일관성이 결여된 것으로 보기는 어렵다. 또한, 의류회사마다 서로 다른 체계의 표기법을 사용함으로써, 이들 간의 치수 선택의 일관성은 더욱 확인하기 쉽지 않다. 하지만, 동일 의류회사가 제작한 동일 디자인의 옷에 대해 서로 다른 치수를 중복으로 구매한 경우라면, 이는 치수 선택의 일관성이 결여된 것으로 확신할 수 있을 것이다. 따라서, 본 연구에서는 이런 데이터를 이상치(Outlier)로 간주하고 제거하였다.

### 1.3 Negative Sampling

본 연구의 목적은 특정 디자인이 갖는 여러 치수 중 사용자에게 가장 적합한 것을 예측하는 모형을 구현하는 것이므로, 여러 치수를 선택 가능한 클래스로 간주하고 다중분류(Multi-Class Classification) 모형을 훈련하는 것도 가능할 수 있다. 그러나, 각 디자인에 속한 치수의 개수가 서로 다르기 때문에 클래스의 개수를 한정 지을 수 없을 뿐만 아니라, 표기된 치수가 일치하더라도 서로 다른 옷의 실제 치수 간에 큰 차이가 있을 수 있다는 점을 고려한다면, 'Small', 'Large'와 같이 표기된 클래스를 정확히 예측하도록 훈련되는 분류 모형을 치수 추천 시스템에 적용

하는 것은 무리가 있다. 예를 들어, 고객A가 의복A의 Small 치수를 구매했음을 훈련시킨 모형이 동일 고객이 의복B의 Small 치수를 구매한다고 예측했다고 해서 정확한 결과로 보기 힘들다는 것이다.

이와 같은 상황을 감안하여, 본 연구에서는 각 디자인의 개별 치수에 대한 선호도를 개별적으로 학습하는 모형을 적용한다. 즉, 주어진 디자인에 대해 선호하는 치수를 선택하는 목표가 아니라, 어떤 디자인의 어떤 치수에 대해서라도 구매자가 선호하는 수준을 예측하는 모형을 훈련시킨다. 이후에, 각 디자인에 대해 치수 별로 예측된 평점을 비교하여 이 중에 가장 큰 값을 갖는 치수를 선택하면, 이를 최종 예측 치수로 선정할 수 있다.

이를 위해 추가적인 데이터의 생성이 필요한데, 즉, 어떤 고객이 Small 치수를 구매한 기록은 해당 치수를 선호하는 것으로 간주하여 훈련 데이터로 사용할 수 있지만, 그 외의 치수(Medium, Large)에 대한 선호도 값이 존재하지 않아 모형의 훈련이 불가능하기 때문이다. 이 때, 구매되지 않은 치수는 해당 고객이 선호하지 않는 것으로 간주함으로써 추가 데이터를 생성하는 네거티브 샘플링(Negative Sampling)이 필요하다. 이렇게 데이터세트를 생성하면 선호여부('1' 또는 '0')를 활용하여 협업 필터링 모형(Collaborative Filtering)을 훈련시키는 것이 가능하다.

## 2. Model Training

### 2.1 Collaborative Filtering Algorithms

본 연구의 모형 훈련 프로세스에는 네 가지 고급 협업 필터링(Collaborative Filtering) 알고리즘을 사용하였다. 이들은 각각 특이값 분해(SVD), 비음수 행렬 분해(NMF), K-최근접 이웃(K-Nearest Neighbors, KNN), 공동 클러스터링(Co-Clustering)이다. 각 알고리즘은 데이터세트의 특정 구조에 맞게 조정되었으며, 사용자와 아이템 간에 존재하는 상관관계를 통해 사용자가 특정 아이템에 대해 얼마나 선호하는지를 학습하도록 훈련된다. 구매 데이터를 훈련에 사용하였으므로, 훈련된 모형이 예측하는 선호도는 구매 가능성으로도 해석될 수 있다.

그 외에도, 치수 추천의 정확도를 높이기 위해, 생성한 앙상블 모형(Ensemble Model)과 성능의 비교를 위한 랜덤 모형(Random Model) 및 베이스라인 모형(Baseline model)이 있다. 이 중, 앙상블 모형은 선형회귀(Linear Regression) 모형을 기반으로 하며, 개별로 훈련된 협업 필터링 기반 알고리즘을 가중 평균 접근 방식으로 결합한다. 랜덤 모형은 무작위로 평점을 예측하는 알고리즘으로서, 훈련된 협업 필터링 모형의 성능이 모형을 사용하지

않을 때보다 얼마나 향상되는지 살펴보기 위한 목적으로 훈련된다. Baseline 모형은 전체 평균과 특정 사용자 및 아이템 편향을 기반으로 사용자-아이템 쌍에 대한 기준 추정치를 계산하여 사용자의 평점을 예측하는 방식으로 작동한다.

### 2.2 Model Fitting through Cross-Validation

각 알고리즘별로 가장 효과적인 하이퍼파라미터 (Hyperparameter) 집합을 결정하기 위해 모든 알고리즘에 대해 그리드 검색(Grid Search)을 통한 교차 검증(Cross-Validation)을 수행한다. 하이퍼파라미터의 선택은 모형의 성능을 결정짓는 요소이므로 매우 중요하다. 각 모형은 서로 다른 하이퍼파라미터에 대해 서로 다른 민감도를 보이기 때문에, 각 알고리즘에 따라 설정 가능한 하이퍼파라미터 값의 다양한 조합을 체계적으로 탐색하도록 한다.

각 알고리즘에 가장 적합한 하이퍼파라미터를 식별한 후, 훈련 데이터세트에 모델을 적합(Fitting)시킨다. 이를 통해 맞춤형 모델이 개발되어 사용자-아이템 상호작용을 예측할 수 있게 되고, 추천 시스템의 효율성이 향상된다.

## 3. Model Evaluation

### 3.1 Root Mean Squared Error (RMSE)

모델 훈련이 완료된 후에는 또 하나의 중요한 단계인 모델 평가(Model Evaluation)가 진행된다. 평가를 위해 선택된 지표는 평균제곱근오차(RMSE)로서, 예측 등급과 실제 등급 간의 평균 제곱 차이를 정량화하는 방식으로 모형의 성능 평가에 적용된다.

치수 추천은 주어진 디자인에 대해 여러 치수 중 하나를 선택하는 다중 분류(Classification) 문제로 볼 수 있으나, 각 디자인이 갖는 치수의 개수가 일정하지 않다는 점 때문에 분류 모형을 적용하는 것은 쉽지 않다. 이런 이유로 주어진 디자인에 대해 치수를 예측하는 방식으로 모형을 훈련하지 않고, 디자인과 치수가 조합된 개별 아이템에 대한 평점을 예측하는 방식으로 모형을 훈련한 후, 이 예측값이 가장 높은 치수를 선택하는 방식으로 작동하도록 시스템을 설계하였다. 즉, 개별 알고리즘에 대해 평균제곱근오차(RMSE)로 측정되는 성능을 높이는 방향으로 훈련을 진행한 후, 훈련된 각각의 추천 알고리즘이 주어진 디자인의 각 치수에 대한 선호도, 즉 평점을 예측하면, 주어진 사용자와 디자인에 대해 선택 가능한 모든 치수의 예측 평점을 조합하여, 이들 중 최대값을 갖는 치수를 추천하는 방식으로 추천 시스템이 작동하게 된다(Fig. 2).

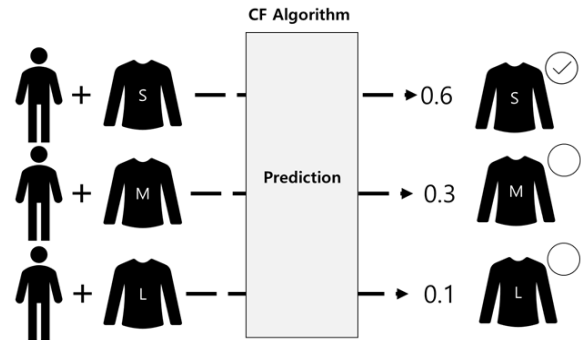


Fig. 2. CF Algorithm-based Size Recommendation Process

### 3.2 Hit Rate

본 연구에서 제안하는 치수 추천 시스템은, 개별 추천 알고리즘이 예측한 선호도를 결합하여, 최대값을 갖는 치수를 선정하는 방식으로 동작하므로, 개별 추천 모형의 성능과는 별개로 추천 시스템 자체에 대한 성능 평가가 필요하다.

추천 시스템의 성능을 측정하는 데 보편적으로 사용되고 있는 적중률(Hit Rate)은 아래의 수식과 같이 계산된다. 여기서, “적중 건수”는 사용자에게 관련성이 있거나 유용한 것으로 판단된 추천 아이템의 수, “총 추천 건수”는 총 추천 아이템 수를 의미한다.

$$\text{적중률(Hit Rate)} = \frac{\text{적중 건수(Number of hits)}}{\text{총 추천 건수(Total number of recommendations)}}$$

## IV. Results

훈련 결과 얻어진 기준 모형(Base Model)의 성능은 앞서 언급한 두 가지 지표(RMSE, Hit Rate)에 의해 평가되었고, 그 결과는 Table 1과 같다. Fig. 3에서도 볼 수 있듯이, 이들 중 NMF가 가장 낮은 RMSE(0.4095)와 가장 높은 적중률(0.6495)을 보였으며, KNN이 가장 높은 RMSE(0.4993)와 가장 낮은 적중률(0.4847)을 보였다. 이들 기준 모형(Base Model)을 결합한 앙상블 모형의 경우, 미세한 차이기는 하지만 NMF보다 약간 높은 RMSE(0.4414)와 약간 낮은 적중률(0.6428)을 보였다. 이는 대체로 앙상블 모형이 개별 모형보다는 높은 성능을 보이는 것과는 다른 결과를 보인 것이다.

일반적으로, 앙상블 모형을 구성하는 기준 모형이 다양하지 못한 경우나 성능이 떨어지는 경우, 또는 과적합(Overfitting)이 발생한 경우에도 개별 알고리즘에 비해 낮은 성능을 보일 수 있다. 본 실험에서는, 서로 다른 알고리

즘을 충분히 포함하고 교차 검증을 실시하여 과적합의 가능성을 낮췄으나, KNN 모형의 낮은 성능이 앙상블 모형의 성능을 저하시킨 원인으로 추측된다.

Table 1. Performance Comparison of Collaborative Filtering Algorithms

Category	Algorithm	RMSE	Hit Rate
Base Models	SVD	0.4409	0.5054
	NMF	0.4095	0.6496
	KNN	0.4993	0.4847
	Co-Clustering	0.4414	0.5860
Ensemble Model	Linear Regression	0.4338	0.6428
Comparative Models	Random	0.5894	0.4097
	Baseline	0.4379	0.5033

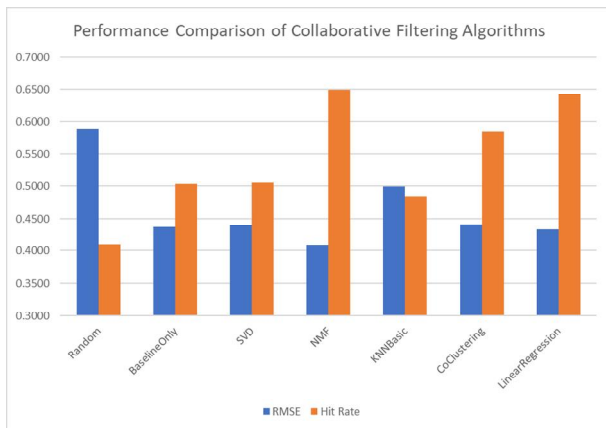


Fig. 3. Performance Comparison of Collaborative Filtering Algorithms

둘 다 행렬 분해(Matrix Factorization) 방식을 기반으로 하는 알고리즘이지만, SVD보다 NMF가 더 높은 성능을 보인 것은, 주어진 상황에 따라 머신러닝 모형의 성능이 크게 달라질 수 있음을 보여준다. 즉, 모형의 특성과 데이터가 얼마나 서로 잘 맞는지가 성능에 큰 영향을 미친다는 것이다. SVD의 경우 양의 평점(Rating)과 음의 평점(Rating)을 모두 다룰 수 있지만, NMF는 양의 평점에 대해서만 작동하는 특징을 가지며, 이런 특징이 더욱 해석하기 쉬운 모형을 만든 것으로 볼 수 있다. 즉, 구매 데이터에 포함된 데이터는 구매 혹은 비구매이므로 '1' 또는 '0'으로 평점을 부여하는 상황에서, 음수값을 고려하지 않는 NMF모형이 상대적으로 더 잘 훈련된 것으로 보인다.

또한, NMF는 희소성이 강한 데이터에 더욱 강한 특성을 갖는다. 데이터의 희소성은 많은 추천 시스템이 공통적으로 겪는 어려움이기도 하지만, 계절성의 영향을 크게 받는 의류의 특성, 즉 한정된 기간에만 발생하는 의류 구매

데이터의 희소성은 모형의 성능에 크게 영향을 미쳤을 것으로 짐작할 수 있다. 데이터의 희소성 문제는 KNN의 성능이 매우 낮게 나온 것에도 영향을 미쳤을 것으로 보인다. 또한, 의류가 출시된 초기 시점이나, 신규 고객이 가입한 시점에는 충분한 구매 데이터를 확보하기 힘들기 때문에 발생하게 되는 콜드 스타트(Cold Start) 문제에도 KNN 모형이 상대적으로 취약했을 것으로 추측된다.

Baseline 모형은, 협업 필터링 알고리즘처럼 사용자-아이템 간 상호작용을 반영하지는 않지만, 각 사용자와 각 항목의 평균 평점을 고려하고 이를 기반으로 예측을 수행한다. 이 모형은, 일부 사용자는 대부분의 아이템에 대해 항상 높은 평점을 주거나 혹은 반대로 낮은 평점을 주는 경향이 있고, 일부 아이템은 많은 사용자가 보편적으로 좋아하거나 반대로 싫어하는 경향이 있다는 사실을 고려하기 때문에, 매우 훌륭한 예측 결과를 제공하기도 한다. 그 결과 KNN보다 전반적인 성능이 우수했을 것으로 추측해 볼 수 있다.

본 연구는 실무 환경에서 적용 가능한 수준으로 치수 추천 시스템을 구현하는 것을 목표로 삼고 있다. 따라서, 어떤 상황에서 추천 시스템이 잘 작동하는지를 확인하기 위한 분석을 시행하였다(Table 2, Fig. 4).

구매 데이터로부터 얻을 수 있는 정보 중 추천 성능에 영향을 미칠 수 있는 두 가지 요소를 각각 아이템(옷)과 사용자(구매자)로부터 선정하였으며, 이들 중 하나는 각 디자인별 치수의 개수이고, 다른 하나는 구매자의 구매 건수이다.

각 의류 디자인은 다양한 치수를 갖는다. 본 연구에서 사용된 데이터셋에는 한 개의 치수를 갖는 디자인으로부터 최대 12개의 치수를 갖는 디자인까지 존재한다. 치수가 다양하다는 것은 사용자의 다양한 체형에 맞도록 제작되었음을 의미하므로, 이들 중 하나를 예측하는 것은 더욱 힘든 과제가 될 수밖에 없다. Table 2와 Fig. 4로부터, 적중률이 치수의 개수에 크게 영향을 받는 것을 알 수 있다. 즉, 1개의 치수를 갖는 의류에 대해서는 모두 1.0000의 적중률을 보이다가 치수의 개수가 증가할수록 점점 낮은 적중률을 보이게 된다. 여러 기준 모형(Base Model) 중에서도 KNN이 치수의 개수 증가에 따라 가장 큰 폭으로 적중률이 낮아지는 것을 볼 수 있고, NMF와 앙상블 모형은 적중률의 저하가 가장 적은 것을 볼 수 있다.

Table 2. Size Count vs. Cumulative Hit Rate

#Sizes	Random	Baseline	SVD	NMF	KNN	Co-Clustering	Ensemble
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0.7520	0.7912	0.7930	0.8490	0.7619	0.8354	0.8457
3	0.6106	0.6770	0.6789	0.7737	0.6555	0.7345	0.7679
4	0.5169	0.5956	0.5952	0.7017	0.5656	0.6582	0.6913
5	0.4519	0.5407	0.5400	0.6659	0.5256	0.6065	0.6575
6	0.4050	0.4946	0.4926	0.6274	0.4852	0.5637	0.6206
7	0.3687	0.4600	0.4588	0.6130	0.4733	0.5308	0.6074
9	0.3380	0.4497	0.4475	0.6208	0.4789	0.5291	0.6159
10	0.3087	0.4230	0.4266	0.5875	0.4517	0.4964	0.5859
12	0.2778	0.3807	0.3839	0.5287	0.4065	0.4467	0.5273

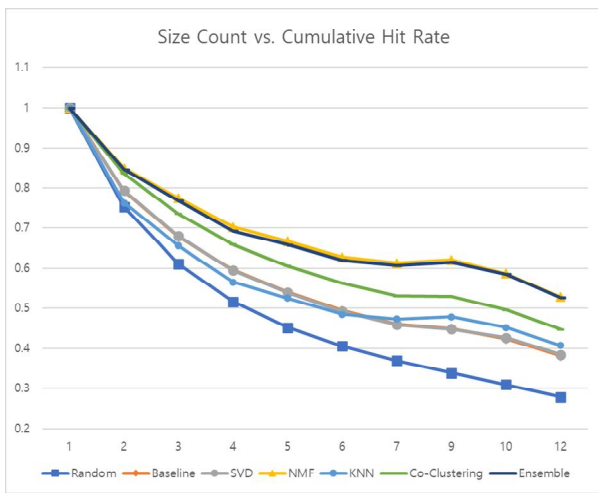


Fig. 4. Size Count vs. Cumulative Hit Rate

개별 구매자의 구매 건수가 증가하면, 각 구매자의 선호도를 파악하기 위한 정보가 더 많이 확보됨을 의미하는 것으로 볼 수 있고, 이는 치수 추천에 긍정적인 효과를 미칠 것으로 기대해볼 수 있다. Table 3과 Fig. 5는 구매 건수가 증가함에 따라, 본 연구에서 실험한 각 모형의 적중률이 어떻게 달라지는지를 보여주고 있다. Fig. 5를 살펴보면, 기준 모형으로 사용된 Random 모형과 Baseline 모형에서는 건수의 증가에 따른 성능 변화가 눈에 잘 띄지 않는 것을 확인할 수 있다. 즉, 구매 건수를 두 개의 구간으로 나눠, '16건 이하'와 '17건 이상'으로 구분했을 때, 두 구간의 적중률 중 어느 쪽이 다른 쪽에 비해 높거나 낮은 양상이 보이지 않는다. 반면, 협업 필터링 알고리즘인 SVD, NMF, KNN, Co-Clustering 기법과 이들 모형을 결합한 앙상블 모형에서는, 두 구간에서의 성능 차이가 뚜렷하게 나타난다. 즉, 구매 건수가 낮은 구간(16건 이하)보다 구매 건수가 높은 구간(17건 이상)에서 적중률이 더 높게 나타나며, 특히, 구매 건수가 낮은 구간에서는 구매 건수의 증가에 따른 적중률 향상이 좀 더 확연하게 드러나는 것을 볼 수 있다.

Table 3. Purchase Count vs. Cumulative Hit Rate

#Purchases	Random	Baseline	SVD	NMF	KNN	Co-Clustering	Ensemble
3	0.4100	0.5074	0.5096	0.6354	0.4638	0.5782	0.6271
4	0.4102	0.5033	0.5050	0.6499	0.4784	0.5802	0.6415
5	0.4123	0.4982	0.5023	0.6543	0.4842	0.5803	0.6481
6	0.4138	0.5041	0.5073	0.6570	0.4849	0.5817	0.6490
7	0.4151	0.5069	0.5079	0.6579	0.4854	0.5883	0.6509
8	0.4166	0.5059	0.5085	0.6610	0.4890	0.5941	0.6533
9	0.4150	0.5089	0.5111	0.6677	0.4949	0.5970	0.6584
10	0.4144	0.5091	0.5110	0.6736	0.4997	0.6031	0.6650
11	0.4135	0.5077	0.5084	0.6758	0.5028	0.6021	0.6685
12	0.4143	0.5074	0.5085	0.6745	0.5018	0.6044	0.6689
13	0.4138	0.5108	0.5101	0.6770	0.5055	0.6073	0.6721
14	0.4096	0.5101	0.5099	0.6805	0.5097	0.6087	0.6762
15	0.4052	0.5056	0.5052	0.6794	0.5129	0.6080	0.6757
16	0.4028	0.5102	0.5097	0.6824	0.5113	0.6108	0.6790
17	0.4049	0.5098	0.5091	0.6805	0.5103	0.6124	0.6755
18	0.4077	0.5075	0.5083	0.6814	0.5117	0.6146	0.6778
19	0.4051	0.5114	0.5137	0.6787	0.5114	0.6115	0.6735
20	0.4067	0.5144	0.5162	0.6786	0.5099	0.6122	0.6733
21	0.4064	0.5093	0.5122	0.6731	0.5096	0.6123	0.6673
22	0.4084	0.5107	0.5125	0.6708	0.5082	0.6099	0.6671
23	0.4067	0.5105	0.5111	0.6744	0.5104	0.6084	0.6715
24	0.4067	0.5057	0.5102	0.6780	0.5122	0.6091	0.6772
25	0.3977	0.5011	0.5098	0.6799	0.5091	0.6069	0.6790
26	0.3992	0.4992	0.5085	0.6814	0.5087	0.6043	0.6788
27	0.4000	0.5035	0.5105	0.6811	0.5144	0.6052	0.6787
28	0.3991	0.5033	0.5076	0.6874	0.5199	0.6096	0.6814
29	0.3987	0.5100	0.5123	0.6926	0.5232	0.6069	0.6878
30	0.3945	0.5119	0.5152	0.6980	0.5235	0.6087	0.6911

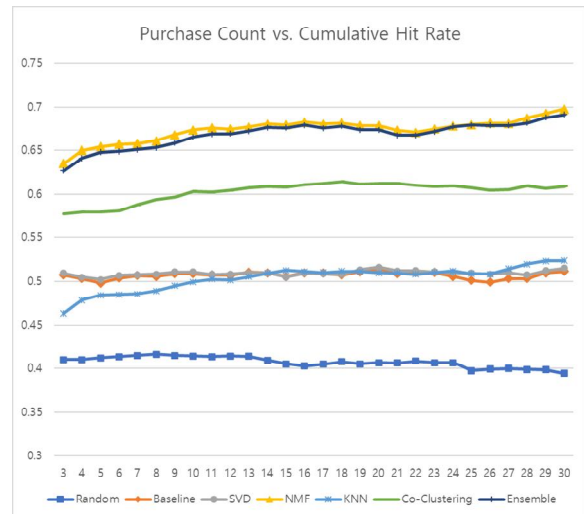


Fig. 5. Purchase Count vs. Cumulative Hit Rate

앞서 언급한 바와 같이, 구매 건수가 낮은 구간과 구매 건수가 높은 구간에서 기준 모형의 성능은 거의 일정한 수준을 유지하는 것을 볼 수 있는데(Fig. 5), 이 중 Baseline 모형은 전체 평균과 특정 사용자 및 아이템 편향을 기반으로 사용자-아이템 쌍(Pair)에 대한 기준 추정치(평균)를 계산하며, 구매 건수가 적은 사용자의 행동은 평균에 가까울

가능성이 높기 때문에, 일정 수준 이상으로 정확하게 예측할 수 있다는 점을 원인으로 생각할 수 있다. 즉, 사용자 또는 아이템 유사성에 의존하지 않기 때문에 구매 횟수가 적은 사용자에게 흔히 발생하는 사용자-아이템 상호 작용 데이터 부족으로 인해 성능에 부정적인 영향을 받지 않은 것이다. 반면, KNN은 오히려 사용자 또는 아이템 간의 유사성을 사용하여 예측을 수행하므로, 사용자와 아이템 간의 상호 작용이 적은 경우(즉, 사용자의 구매 건수가 적은 경우) KNN 알고리즘이 유사한 사용자나 아이템을 찾아 정확한 예측을 수행하기 어려워 성능이 저하될 수 있다. 그러나, 구매 건수가 증가하면 KNN 알고리즘이 더 많은 유사한 사용자 또는 아이템을 찾을 수 있으므로, 성능이 향상될 수 있다(Fig. 5).

## V. Conclusions

최근 많은 전자상거래 사이트들이 상품 추천에 효과적인 머신러닝 기반의 추천 시스템을 적극적으로 도입하고 있다. 본 연구에서는 이러한 추천 시스템을 활용하여, 온라인 환경에서 옷을 구매하려는 사용자가 자신에게 적합한 치수를 선택할 수 있게 돕는 방안을 제시하였다. 또한, 추천 시스템에서 매우 빈번하게 활용되는 협업 필터링(Collaborative Filtering) 기법 중 대표적인 기법들을 적용하여 치수 추천 모형을 구현하였으며, 실제 데이터를 통해 각 모형의 성능을 측정 비교하였다. 그리고, 모형의 예측력을 좀 더 높이기 위해 앞서 구현된 알고리즘을 하나로 결합한 앙상블 학습(Ensemble Learning) 모형을 구현하였다.

온라인 쇼핑이 성장할수록 패션 업계 고객들은 의류를 구매할 때 적절한 치수를 선택하는 데 더 큰 어려움을 겪고 있다. 이는 종종 만족스럽지 않은 제품 선택과 높은 반품률로 이어진다. 이는 고객의 불편으로 인한 불만족과 충성도 저하로 이어질 수 있으며, 판매자에게는 반품 처리 과정에서 큰 비용과 노력이 소요되는 문제를 야기한다. 판매자들은 치수조건표를 제공하고, 구매자는 구매 후기를 활용하는 등 올바른 치수의 선택을 위한 노력이 기울여지고 있으나, 제조사나 브랜드별 치수 체계의 불일치, 구매자 간 체형의 차이, 개인적 취향의 차이, 제품의 소재나 디자인 등에 따라 구매자가 선호하는 치수는 달라질 수 있다는 점에서 한계점을 갖는다.

이런 문제를 해결하기 위하여 수행된 많은 연구들은, 실측된 체형을 기반으로 하거나, 이미지 등으로부터 체형의

실측치를 예측함으로써, 적합한 치수를 선택할 수 있는 기법을 제안해왔다. 적합한 치수를 찾으려 하는 것은 판매자의 역할이나, 이에 필요한 정보는 구매자의 것이므로, 구매자가 자발적으로 제공하지 않는 이상, 판매자는 이런 정보를 활용할 수 없다는 점에서 실용성이 높지 않을 수 있다. 추가로 정보를 확보하는 것보다는, 이미 판매자가 충분히 확보한 구매 데이터를 활용하는 것이 현실적으로는 가장 좋은 대안이라고 할 수 있다. 이런 이유로, 본 연구는 구매 데이터에 포함된 구매자와 의류 간의 상호관계를 이용하여 구매자가 구매하려는 옷의 적합한 치수를 예측할 수 있는 모형을 개발하였다.

다양한 협업 필터링 기법에 대해 예측 모형을 구현한 후, 온라인 전자상거래에서 수집된 실제 구매 데이터를 사용한 실험을 통해 각 모형의 성능을 평가할 수 있었다. 데이터 전처리 단계에서는 동일 의류에 대해 여러 치수를 구매하는 데이터를 제외함으로써 모형의 성능이 향상될 수 있었다.

치수 추천 시스템을 구축하기 위해 다양한 협업 필터링 알고리즘을 훈련하였다. 행렬 분해 방식을 기반으로 한다는 공통점을 갖는 SVD와 NMF 알고리즘 중, NMF가 더 높은 성능을 보였는데, 이는 모형과 데이터의 적합성이 성능에 큰 영향을 미쳤기 때문으로 해석된다. 특히, 양의 평점만 다루는 NMF는 구매/비구매 데이터를 잘 처리한 것으로 판단되었다. 희소성과 계절성을 갖는 의류 데이터의 특성으로 인해 KNN의 성능이 낮게 나온 반면, 비교를 위해 훈련된 Baseline 모형은 KNN보다 전반적으로 높은 성능을 보였다. 그러나, 사용자와 아이템 간 상호작용을 고려하는 알고리즘의 특성상, KNN의 성능은 구매 건수가 많아질수록 향상되었다.

본 연구의 결과는 옷의 치수를 예측하는 것이 전통적인 추천 시스템보다 어려울 수 있음을 보여준다. 이는 앞서 언급한 바와 같이, 동일 계정으로 여러 명의 구매자가 의류를 구매하는 것이 가장 큰 이유라고 생각된다. 따라서, 동일 계정이라고 하더라도 다양한 구매자를 구분할 수 있는 방안을 고안할 수 있다면 성능의 개선을 크게 기대할 수 있을 것이다. 이런 한계에도 불구하고, 본 연구는 구매 데이터만을 활용하여 높은 수준으로 치수를 예측할 수 있다는 것을 보였다는 측면에서 평가받을 만하다. 향후에는 신경망과 같이 본 연구에서 다루지 않은 알고리즘에 대한 연구를 통해 더 높은 성능의 치수 추천 시스템을 개발할 수 있을 것으로 기대된다.

본 연구를 통해, 추천 시스템은 단순히 상품을 추천하는 것뿐만 아니라, 사용자의 개별적인 필요에 맞춰 세부적인

정보(예: 옷의 치수)를 제공하는 데에도 효과적일 수 있음을 확인할 수 있었다. 이러한 접근 방식은 개인화된 쇼핑 경험을 제공하는데 기여할 수 있으며, 전자상거래 사이트들이 더욱 세밀하게 사용자의 요구를 이해하고 충족시키는 데 도움이 될 것이다. 이 연구의 결과는 향후 더 복잡하고 세밀한 추천 시스템의 개발에 대한 진전을 이끌어낼 수 있을 것으로 기대된다.

## ACKNOWLEDGEMENT

This research was financially supported by Hansung University.

## REFERENCES

- [1] Y. Chen, and Z. Yang, "The behavioral analysis of choice difficulty states during clothing online shopping", *International Journal of Clothing Science and Technology*, Vol. 33, No. 4, pp. 577-589, July 2021. DOI: 10.1108/IJCST-12-2019-0189
- [2] A. Dimoka, Y. Hong, and P. A. Pavlou, "On Product Uncertainty in Online Markets: Theory and Evidence," *MIS Quarterly*, Vol. 36, No. 2, June 2012. pp. 395-426. DOI: 10.2307/41703461
- [3] F. Ricci, L. Rokach, and B. Shapira, "Introduction to Recommender Systems," *Recommender Systems Handbook*, Boston, MA: Springer US., pp. 1-35, 2010. DOI: 10.1007/978-0-387-85820-3\_1
- [4] Y. Koren, "Factorization Meets The Neighborhood: A Multifaceted Collaborative Filtering Model," *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 426-434, Las Vegas Nevada USA, August 2008. DOI: 10.1145/1401890.1401944
- [5] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, "Collaborative Filtering Recommender Systems," *The Adaptive Web*, Springer, pp. 291-324, 2007. DOI: 10.1007/978-3-540-72079-9\_9
- [6] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An Algorithmic Framework for Performing Collaborative Filtering," *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Vol. 51, No. 2, pp. 227-234, New York NY USA, August 2017. DOI: 10.1145/3130348.3130372
- [7] X. Su, and T. M. Khoshgoftaar, "A Survey Of Collaborative Filtering Techniques," *Advances in Artificial Intelligence*, Vol. 2009, pp.1-19, 2009. DOI: 10.1155/2009/421425
- [8] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based Collaborative Filtering Recommendation Algorithms," *Proceedings of The 10th International Conference on World Wide Web*, pp. 285-295, Hong Kong Hong Kong, May 2001. DOI: 10.1145/371920.372071
- [9] Y. Koren, R. Bell, and C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," *Computer*, Vol. 42, No. 8, pp.30-37. August 2009. DOI: 10.1109/MC.2009.263
- [10] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: An Open Architecture for Collaborative Filtering of Netnews," *Proceedings of the 1994 ACM Conference on Computer-Supported Cooperative Work*, pp. 175-186, Chapel Hill North Carolina USA, October 1994. DOI: 10.1145/192844.192905
- [11] D. Lee, and H. S. Seung, "Algorithms for Non-negative Matrix Factorization," *Advances in Neural Information Processing Systems 13 - Proceedings of the 13th International Conference on Neural Information Processing Systems (NIPS'00)*, pp. 535-541, Denver Colorado USA, January 2000.
- [12] T. George, and S. Merugu, "A Scalable Collaborative Filtering Framework based on Co-Clustering," *Fifth IEEE International Conference on Data Mining (ICDM'05)*, pp. 625-628, Houston Texas USA, November 2005, DOI: 10.1109/ICDM.2005.14.
- [13] A. S. Sheikh, R. Guigourès, E. Koriagin, Y. K. Ho, R. Shirvany, R. Vollgraf, and U. Bergmann, "A Deep Learning System for Predicting Size and Fit in Fashion E-Commerce," *Proceedings of the 13th ACM Conference on Recommender Systems*, pp. 110-118, Copenhagen Denmark, September 2019. DOI: 10.1145/3298689.3347006
- [14] A. Nestler, N. Karessli, K. Hajjar, R. Weffer, and R. Shirvany, "Sizeflags: Reducing Size and Fit Related Returns in Fashion E-Commerce," *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3432-3440, Virtual Event Singapore, August 2021. DOI: 10.1145/3447548.3467160
- [15] G. Pisut, and L. Jo Connell, "Fit Preferences of Female Consumers in the USA," *Journal of Fashion Marketing and Management*, Vol. 11, No. 3, pp. 366-379, July 2007. DOI: 10.1108/13612020710763119
- [16] K. Liu, H. Wu, C. Zhu, J. Wang, X. Zeng, X. Tao, and P. Bruniaux, "An Evaluation of Garment Fit to Improve Customer Body Fit of Fashion Design Clothing," *The International Journal of Advanced Manufacturing Technology*, Vol. 120, No.3-4, pp. 2685-2699, May 2022. DOI: 10.1007/s00170-022-08965-z

## Authors



Dongwon Lee earned his B.S. degree in Materials Science and Engineering from Hanyang University in 1997, M.S. degree in Management Information Science and Ph.D. degree in Management Engineering from

Korea Advanced Institute of Science and Technology (KAIST) in 2005 and 2014. He worked as a system engineer for LG CNS from 1997 to 2002. Dr. Lee joined the faculty of the Division of Social Sciences at Hansung University, Seoul, Korea, in 2015. He is currently an Assistant Professor in the Division of Social Sciences, Hansung University. He is interested in business analytics, data mining, deep learning, consumer behavior, and recommender systems.