

Pre-processing Method for Conversation-Based Fashion Curating Automation Recommendation System

Chulwoong Choi¹⁾, Wolhee Do[†], and Kyungbaek Kim²⁾

¹⁾Dept. of Artificial Intelligence Media Contents, Kwangju Women's University; Gwangju, Korea

²⁾Dept. of Artificial Intelligence Convergence, Chonnam National University; Gwangju, Korea
Dept. of Clothing and Textiles / Research Institute of Human Ecology, Chonnam National University / Healthcare Ware Research and Business Development Center, CNU R&BD Foundation; Gwangju, Korea.

Abstract: Fashion curation is a system that recommends coordination outfits that are tailored to users' situation and environment. Unlike traditional fashion recommendation systems that rely on user profile and survey information to recommend single clothing items, fashion curation assembles multiple items based on the user's context and environment. This approach requires the availability of access to various types and wide range of user information. Personalized fashion curation can be automated by linking item characteristics and fabric attributes of user preferences based on Q&A interactions regarding between the AI coordinator and the user. Over time, the performance of the model can be continuously improved through the accumulation of conversation data, which enables ongoing retraining. Accordingly, a text-embedding pre-processing framework is essential, as it enables the AI coordinator to interpret and integrate both user conversations and fashion item information. This study proposes a conversational data pre-processing method for training an automated fashion curation model using a dataset that includes conversations between the AI coordinator and the user, as well as fashion item information. When using the proposed conversation data pre-processing method, both the Weighted Kendall Tau(WKT) Sum and WKT Avg exhibited significant improvements. The highest performance was achieved when both user conversations and fashion item information were pre-processed.

Key words: fashion curation, recommendation system, data preprocessing, conversation data, chatbot

1. Introduction

As the Fourth Industrial Revolution progresses, the importance of personalized solutions is being emphasized. This trend is impacting the fashion industry, especially evident in the field of online personal styling services. Services like Stitch Fix are leading innovation and experiencing rapid growth in this sector(Kim et al., 2022). These services provide users with personalized fashion styling and product design recommendations based on stylist analysis. However, as the fashion industry shifts online, while users have more choices, it is becoming increasingly challenging to effectively combine various clothing items into coherent outfit sets. Consequently, there is a growing need for a systematic system that can automatically curate outfits considering individual circumstances and environments.

Recently, AI chatbot services like ChatGPT, based on OpenAI's Large Language Model(LLM), are gaining attention in

various fields. These services are being widely adopted in call centers, websites, shopping malls, healthcare, and more. As part of the conversational systems, AI chatbots preprocess and analyze user utterances to understand and generate appropriate responses(Chen et al., 2018; Young et al., 2018). Particularly, integrating such AI chatbots into fashion recommendation systems can effectively collect user preferences and requirements. By combining user utterance data with existing clothing item information, it is possible to automate personalized outfit curation tailored to an individual's circumstances and environment. Additionally, as conversation data accumulates over time, this data can be used to continuously retrain the AI model, enhancing the accuracy of the automated outfit curation system. However, processing utterance and clothing item data requires text embedding, a process of converting text into numerical data. Without proper preprocessing of the data before embedding, the noise in the data can make it difficult to build a high-quality curation system. Therefore, the preprocessing of data before embedding is crucial for AI training.

In this paper, we propose a data preprocessing method for text embedding in an automated fashion curation system using the FASCODE dataset, which includes utterance information between an AI coordinator and users, as well as clothing item information provided by the Electronics and Telecommunica-

†Corresponding author: Wolhee Do

Tel. 

E-mail: whdo@jnu.ac.kr

©2025 The Korean Fashion and Textile Research Journal(KFTRJ). This is an open access journal. Articles are distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

tions Research Institute (ETRI). By applying the proposed pre-processing method to a conversation-based fashion curation recommendation system, it is possible to automate personalized fashion curation and retrain the model using the accumulated data.

2. Related Research

Current fashion recommendation systems primarily use content-based approaches and collaborative filtering. Content-based approaches recommend items by analyzing similarities between users based on user profiles, survey responses, and clothing attributes (Bellini et al., 2013; Bellini et al., 2014; Walek & Spackova, 2018). On the other hand, collaborative filtering employs user interaction data such as ratings, purchase history, site visits, and test responses to make recommendations using neighborhood methods and latent factor models (Omprakash et al., 2019; Koren et al., 2009). This process involves techniques like matrix factorization to understand and characterize the relationships between items and users. However, traditional collaborative filtering based systems mostly rely on static data, which tends to focus recommendations around high-exposure products. To overcome these limitations, techniques using machine learning models like clustering have been developed, but there are still challenges in providing personalized fashion item recommendations tailored to individual user's specific circumstances and environments (Bellini et al., 2023). Specifically, the integration of multimodal information and the development of digital curation techniques have emerged as dominant research areas. Since fashion items are inherently a visual and emotionally driven domain, precisely capturing user preference is challenging using only simple text or quantitative data (Deldjoo et al., 2025). Consequently, multimodal recommendation systems that integrate both visual and textual features have appeared as a crucial solution. A neural architecture combining visual and textual information was proposed by Chen et al. (2019). They demonstrated the utility of the multimodal attention network particularly in sparse interaction scenarios or during the recall stage of the recommendation pipeline. Furthermore, the study by Li et al. (2021) developed an attribute-aware complementary fashion recommendation model based on the BERT framework. This model significantly improved recommendation accuracy and user trust by integrating multimodal features and attention mechanisms to generate attribute-based outfit matching explanations.

Most recently, research is evolving toward utilizing dialogue and generative AI to accurately reflect users' complex and

nuanced preferences. The study by Deldjoo et al. (2025) outlined a research agenda for systems supporting scenarios beyond static queries, aligning with the LLM era. They specifically proposed the Agentic Mixed-Modality Refinement (AMMR) pipeline. Moreover, Deldjoo et al. (2025) emphasized that because the fashion industry is characterized by rapid trend fluctuations, subjective aesthetic preferences, and high return rates, standard recommendation pipelines are inadequate. They stressed that systems capable of responding sensitively to newly emerging styles and fine-grained user intent are essential.

There is also research that proposes an item-set measurement learning framework to calculate the similarity between previously introduced items and newly introduced items by using information obtained from the user's social media to provide more information for fashion recommendations. Although this approach can utilize a slightly larger amount of data compared to existing systems, it has limitations, such as the inability to make recommendations for users who do not use social networks (Haitian et al., 2021). An automated fashion curation system that uses conversational data containing various information is needed for outfit recommendations tailored to specific user needs and situations.

3. Methods

In this study, we study a fashion recommendation system and text embedding. Section 3.1 describes the dataset used in the experiments. Section 3.2 briefly describes the conversation-based fashion system and its operating principles. Section 3.3 describes the data preprocessing method using conversation data and clothing item data. The programming language, libraries, computing environment, and file formats used in this study are shown in Table 1, and the overall data pipeline is shown in Fig. 1.

3.1. Dataset

The FASCODE (FASHion CoOrdination DatasEt) dataset is provided by the ETRI during the '2020 ETRI Autonomous

Table 1. Specifications used in this study

Category	Specification
Programming language	Python 3.8
Morphological analyzer	kkma(from KoNLPy version 0.5.2)
Embedding model	FastText(from Gensim version 3.8.3)
Computing environment	CPU: Intel Core i7-13700K, RAM: 64GB, GPU: NVIDIA GeForce RTX 4070 Ti
Data format	TSV (Tab-Separated Values)

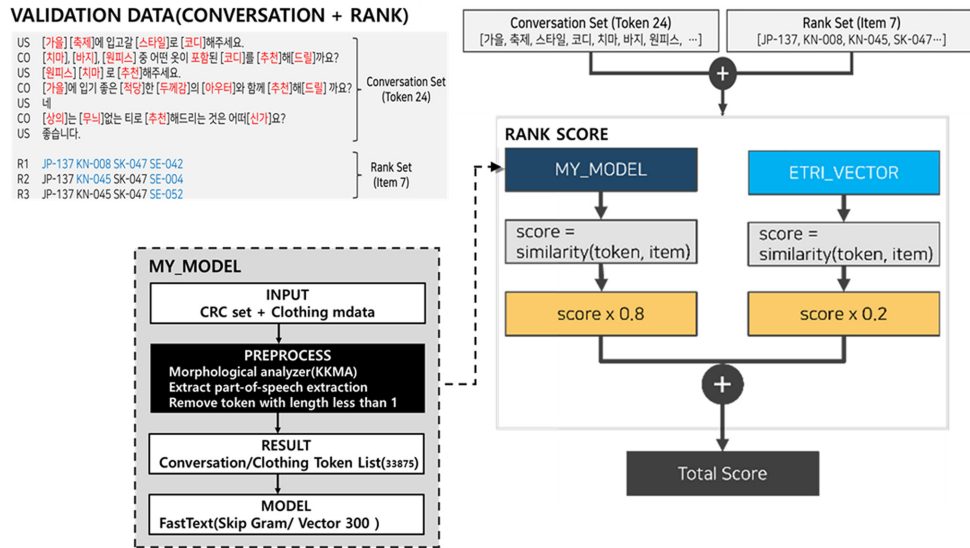


Fig. 1. The overall data pipeline. ETRI Vector: The provided Acbot.wordlist.fasttext.wemb, MY_MODEL:Model developed using FastText, CRC set:Coordination recommended conversation.

Growth Artificial Intelligence Competition(<https://edr.etri.re.kr/data/7b46de7e-42ca-48e4-9018-7dc6f370c2e3>). FASCODE is available in two versions: one with token separation applied and one without token separation. It consists of four components: a training dataset containing conversation information between the AI coordinator and users, clothing metadata containing meta information of clothing items, clothing images, and model validation data. This paper utilizes the version without token separation, employing three components: the training conversation set for outfit recommendations, clothing metadata, and validation data, excluding the clothing images. The training conversation dataset(ddata) from FASCODE contains 7,325 conversation entries between the AI fashion coordinator and

users. Excluding 710 entries where recommendations failed, a total of 6,525 entries are used.

Fig. 2 demonstrates the flow of dialogue between the AI coordinator and the user, culminating in an Outfit Set recommendation based on successive item suggestions. The data begins with fashion coordination recommendation questions that consider specific situations and environments, such as 'Please help me coordinate an outfit for my first day at university.' The AI fashion coordinator recommends items one by one to the user and generates responses based on the user's feedback. Ultimately, a recommended outfit set consisting of 'outerwear', 'top', 'bottoms,' and 'shoes' is created once the user is satisfied, and the conversation concludes.

0	<CO>	어서오세요. 코디봇입니다. 무엇을 도와드릴까요?	INTRO
1	<US>	처음 대학교 들어가는데 입을 옷 코디해주세요.	
2	<CO>	신입생 코디에 어울리게 화사한 스웨터를 추천해드릴게요.	EXP_RES_SITUATION;EXP_RES_DESCRIPTION
3	<AC>	SW-009	
4	<US>	이 옷에 어울리는 치마로 추천해주세요.	USER_SUCCESS
5	<AC>	SK-016	
6	<CO>	고객님의 키 사이즈에 맞추면 이런 옷도 잘 어울릴 것 같은데 어떠신가요?	EXP_RES_BODY;CONFIRM_SATISFACTION
7	<US>	제가 키가 작아서 짧은 치마로 추천해주세요.	USER_FAIL
8	<AC>	SK-052	
9	<CO>	상의 색상과도 매칭이 잘 어울리는 짧은 치마입니다.	EXP_RES_COLOR;EXP_RES_LENGTH
10	<US>	어두운 계열은 없나요?	USER_FAIL
11	<AC>	SK-053	
12	<CO>	연벌런스한 컷팅으로 세련미를 돋보이게 하는 치마인데, 마음에 드시나요? EXP_RES_LENGTH;EXP_RES_DESCRIPTION;CONFIRM_SATISFACTION	
13	<US>	나쁘지 않네요. 외투도 추천해주시겠어요? USER_SUCCESS	
14	<CO>	오늘 계절에는 가디건이나 자켓을 걸치기에 좋는데, 특정 종류로 원하는 게 있으신가요? ASK_TYPE	
15	<US>	트렌치코트 종류로 추천해주세요.	
16	<AC>	CT-019	
17	<CO>	이러 색상과 무난하게 잘 어울릴 트렌치코트입니다.	EXP_RES_TYPE;EXP_RES_COLOR
18	<US>	신발도 추천해주세요.	USER_SUCCESS
19	<CO>	운동화나 구두 중 어떤 걸 선호하시나요? ASK_TYPE	
20	<US>	운동화로 추천해주세요.	
21	<AC>	SE-039	
22	<CO>	어떤 스타일과도 무난하게 잘 어울리는 기본아이템입니다. EXP_RES_ETC	
23	<US>	맘에 드네요. 전체 코디샷 볼 수 있나요? USER_SUCCESS	
24	<AC>	CT-019 SW-009 SK-053 SE-039	
25	<CO>	네, 지금까지 제안해드린 아이템으로 전체 코디샷을 제안해드립니다.	CONFIRM_SHOW
26	<CO>	마음에 드시나요? CONFIRM_SATISFACTION	
27	<US>	네, 마음에 드네요. 감사합니다. USER_SUCCESS	
28	<CO>	마음에 드신다니 다행입니다. SUCCESS	
29	<CO>	이용해 주셔서 감사합니다. CLOSING	

Fig. 2. Example of learning coordination recommended conversation set (ddata).

Table 2. Format of learning coordination recommended conversation set (ddata)

Format
Conversation Number {t}(CO){(US)}{(AC){t}Conversation(t)TAG

Table 3. TAG types of learning coordination recommended conversation set (ddata)

TAG	Describe
INTRO	Conversation Introduction
EXP RES *	Recommended outfit description
USER SUCCESS	Recommended outfit success
USER SUCCESS PART	Some recommended outfits were successful
USER FAIL	Recommended outfit failed
FAIL	Fashion recommendation failed
ASK *	Questions about the clothing type, style, color, etc. that the user wants
CONFIRM *	Confirmation question
SUCCESS	Fashion recommendation success
CLOSING	End conversation
WAIT	Wait request
SUGGEST *	Suggestion utterance
NONE	No clothing
HELP	User support

The format of the training conversation dataset is as shown in (Table 2). The conversation number represents the order 3 of the conversation. "t" represents the tab key and is used as a delimiter in the for mat. "CO" represents the coordinator,"US" represents the user, and "AC" represents the outfit items recommended by the coordinator. Each conversation entry ends with a TAG, which includes the meaning of the conversation(Table 3).

The clothing metadata (mdata) comprises various information such as types and features of outfits including sweaters and shirts, materials, colors, and emotions, organized based on out-

Table 4. Format of clothing metadata(mdata)

Format
Clothing ID {t}O T B S{t}ClothingType{t}F M C E{t}Describe

OTBS stands for Outerwear, Tops, Bottoms, and Shoes. FMCE stands for Features, Materials, Colors, and Emotions, which are the main attributes describing the clothing item.

Table 5. Format type category of clothing metadata(mdata)

Clothing Type(Describe)
CT:Coat, CD:Cardigan, VT:Vest, JK:Jacket, JP:Hood, KN: Jersey, SW:Sweater, SH:Shirt, BL:Blouse, SK:Skirt, PT:Pants, OP:One piece, SE:Shoes

CT:Coat, CD:Cardigan, VT: Vest, JK:Jacket, JP:Hood, KN:Jersey, SW:Sweater, SH:Shirt, BL:Blouse, SK:Skirt, PT:Pants, OP:One-piece, SE:Shoes.

erwear, tops, bottoms, and shoes(Table 4). The clothing ID represents the unique number of the item, and OTBS stands for outerwear, tops, bottoms, and shoes, respectively. The clothing type includes 13 categories ranging from coats and cardigans to shirts and shoes. FMCE represents features, materials, colors, and emotions, respectively(Table 5).

The clothing metadata for each item is composed of various FMCE information(Fig. 3). This shows how a single fashion item is comprehensively defined by multiple FMCE attributes(Features, Materials, Colors, Emotions) in the dataset, which is crucial for contextual matching.

The model validation data consists of 200 validation entries, each paired with the correct answers. Each validation entry is composed of a conversation between the coordinator and the user, along with three recommended outfit sets (outerwear, top, bottoms, shoes). The method of using the model validation data

BL-001	T	BL	F	단추 여밈의 전체 오픈형
BL-001	T	BL	F	스탠드 칼라와 브이넥 네크라인의 결합스타일
BL-001	T	BL	F	손목까지 내려오는 일자형 소매
BL-001	T	BL	F	여유로운 핏
BL-001	T	BL	F	어깨에서 허리까지 세로 절개에 풍성한 러플 장식
BL-001	T	BL	F	와이드 커프스
BL-001	T	BL	M	면 100%
BL-001	T	BL	M	구김이 가기 쉬운
BL-001	T	BL	M	드라이클러닝 권장
BL-001	T	BL	C	시원해 보이는 소라색(SKY BLUE)
BL-001	T	BL	C	단색의 깔끔한 느낌
BL-001	T	BL	E	여성스러운
BL-001	T	BL	E	페미닌한
BL-001	T	BL	E	세련된
BL-001	T	BL	E	사랑스러운
BL-001	T	BL	E	깔끔한
BL-001	T	BL	E	오피스룩
BL-001	T	BL	E	로맨틱한 데이트룩
BL-001	T	BL	E	포멀한 이미지
BL-001	T	BL	E	단정한 오피스룩 룩이미지
BL-002	T	BL	F	넓은 셔츠칼라 네크라인
BL-002	T	BL	F	알중심에 반 오픈 끈여밈있는 스타일
BL-002	T	BL	F	드롭숄더
BL-002	T	BL	F	피프형 소매로 여성스러움을 가미함
BL-002	T	BL	F	세로 주름 조직
BL-002	T	BL	F	풍성한 여유가 있는 몸통
BL-002	T	BL	M	면
BL-002	T	BL	M	세로방향 크릴을 조직
BL-002	T	BL	M	비치는
BL-002	T	BL	M	얇은

Fig. 3. Example of clothing metadata(mdata).

involves the recommendation system analyzing the conversation and ranking the three recommended outfit sets based on user satisfaction, which can then be compared with the correct answer for evaluation. In this paper, the training conversation dataset is defined as 'conversation data' and the clothing metadata as 'clothing data'.

3.2. Conversation-based fashion recommendation system

We introduce a conversation-based fashion recommendation system that combines conversation data and clothing data. Fig. 4 illustrates the architecture of the system using the validation data from 'FASCODE' to recommend outfit sets. This diagram illustrates the system pipeline, from applying morphological analysis to the conversation and clothing data, calculating similarity between tokens and items, and finally ranking the Outfit Sets(R1, R2, R3) using aggregated scores.

After removing duplicates among the item numbers R1, R2, and R3 provided as examples, a list of clothing items is created. After preprocessing the conversation and clothing data, the text embedding process generates tokens and vector values for the items. The similarity between these tokens and items is then calculated. The calculated similarities for R1, R2, and R3 are summed to compute the Score values for R1, R2, and R3, respectively. Finally, the values of R1, R2, and R3 are ranked from highest to lowest, allowing the system to recommend an

outfit set that suits the specific situation and environment.

The core of a conversation-based fashion recommendation system is the calculation of similarity using text-embedded values. The performance of text embedding influences the accuracy of recommendations. To improve the performance of text embeddings, the data preprocessing step is crucial. Conversation data contains more diverse information compared to conventional static information and has a different structure, thus requiring new preprocessing methods to handle it.

3.3. Proposed preprocessing method

We propose preprocessing methods for conversation and clothing data for text embedding in the automated clothing curation system. The conversation data undergoes a five-step process, while the clothing data goes through a three-step process. Fig. 5 flowchart outlines the five steps for conversation data preprocessing and the three steps for clothing data preprocessing, designed to optimize the raw conversational and metadata for text embedding.

Data preprocessing is the process of cleaning data before analysis or the development of artificial intelligence models. Currently used natural language datasets utilize various methods in the preprocessing stage, such as 'morphological analysis', 'Part-of-Speech(POS) tagging', and 'stop word removal'. Since the composition and format of natural language datasets

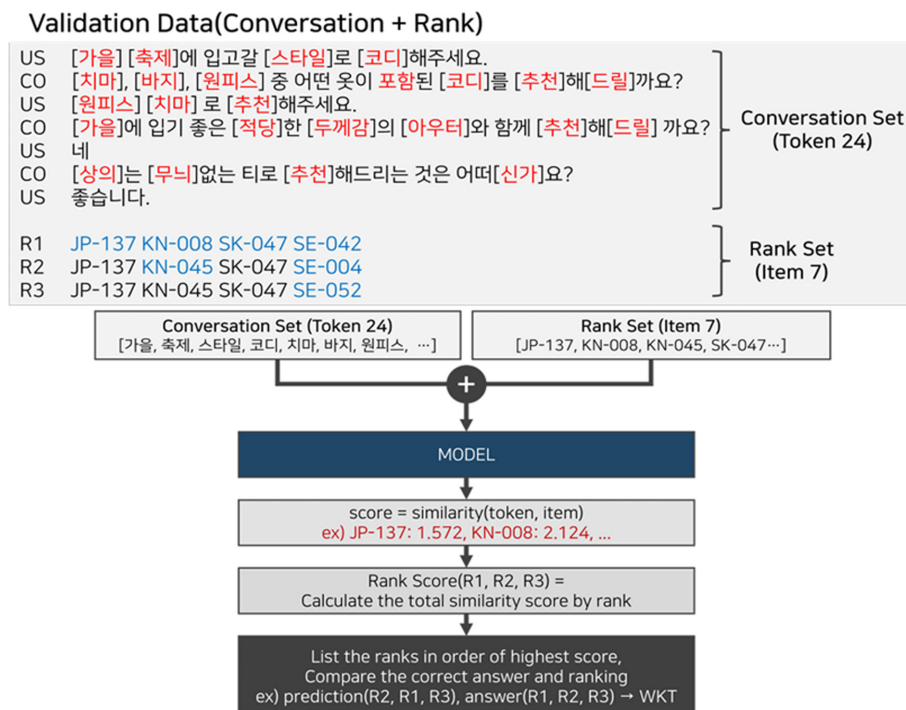


Fig. 4. Architecture of conversation-based fashion recommendation system.

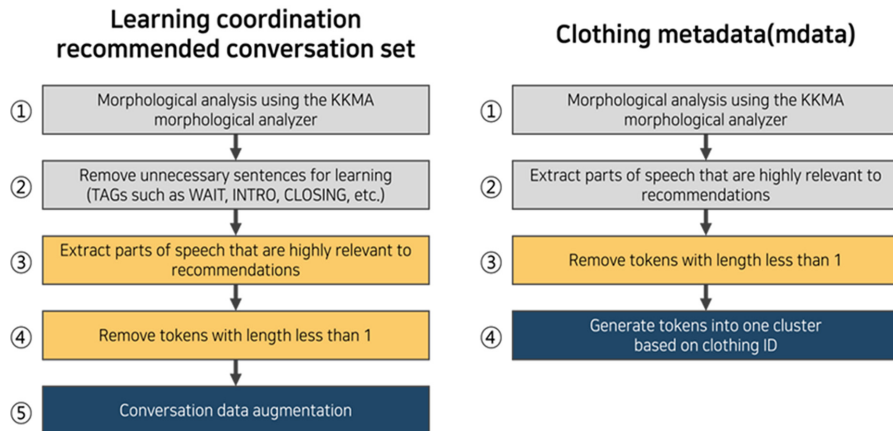


Fig. 5. Proposed data preprocessing stages.

vary depending on the topic, there is no single standard or ‘text-book’ method. Data science efforts are mostly carried out by trying various combinations of the aforementioned ‘morphological analysis’, ‘POS tagging’, etc., to find the optimal approach. Therefore, this study proposes a preprocessing configuration method—5 stages for conversation data and 3 stages for clothing data—which was found to yield the best performance through various experimental attempts. This result has been validated through a recognized challenge.

3.3.1. The preprocessing of conversation data

The preprocessing of conversation data involves five steps: ‘morphological analysis’, ‘removal of unnecessary sentences’, ‘part-of-speech extraction’, ‘removal of tokens with a length of one or less’, and ‘data augmentation’.

The AI coordinator recommends individual items such as ‘outerwear’, ‘tops’, ‘bottoms’, and ‘shoes’ to users through conversation, progressively creating a complete outfit set. If a user is satisfied with the recommended clothing item, the ‘USER SUCCESS’ tag is applied. Conversely, if the user is not satisfied, the ‘USER FAIL’ tag is assigned.

In the first step, conversation data is segmented into paragraphs based on the ‘USER SUCCESS’ and ‘USER FAIL’ tags, and morphological analysis is performed on each segment. A morpheme, the smallest meaningful unit of language, forms the basis of this analysis. Sentences in the conversation data are composed of various morphemes combined, and morphological analysis involves tokenizing speech into morphemes. Unlike English, Korean can be tokenized by performing morphological analysis. In this study, the specialized Korean ‘KKMA Morpheme Analyzer’ is used.

In the second step, unnecessary sentences are removed from

the conversation sentences. conversations in the dataset begin with an ‘INTRO’ tag and end with a ‘CLOSING’ tag. These tags, along with sentences associated with them, are irrelevant to clothing recommendation and are therefore removed. Additionally, tags such as ‘WAIT’ and ‘SUCCESS’ are also removed.

In the third step, parts of speech that are highly relevant to clothing recommendation are extracted. When morphological analysis is conducted, morphemes are segmented, and the parts of speech for these segmented morphemes are also defined. Parts of speech such as NNG (general nouns), NNP (proper nouns), NR (numerals), VA (adjectives), MAG (general adverbs), and XR (roots) that are highly relevant to clothing recommendation are extracted. In the fourth step, tokens with a length of one or less are removed because they are difficult to attribute meaningful interpretation to. Currently, steps 1 to 4 are conducted separately on paragraphs distinguished by ‘USER SUCCESS’ and ‘USER FAIL’ tags.

While these steps allow for understanding individual clothing items within each paragraph, they make it difficult to comprehend a complete set of conversation that starts with an ‘INTRO’ tag and ends with a ‘CLOSING’ tag. To fully understand a conversation set and the recommended outfit sets, all parts of the conversation must be understood. Therefore, in step 5, data augmentation is performed using the existing conversation set data. This involves combining the item numbers associated with ‘USER SUCCESS’ tags, where the user is satisfied, into one outfit set data. Additionally, the various paragraphs from a conversation set and the newly created outfit set data are combined to augment the data. When the five-step preprocessing is applied to a conversation set in Fig. 2, it is divided into six paragraphs by ‘USER SUCCESS’ and ‘USER FAIL’ tags. Steps 1–4 are then carried out to produce six separate data

```

1 [처음,대학교,코디,신입생,코디,화사,스웨터,추천,SW-009,치마,추천]
2 [SK-016,고객,사이즈,제가,치마,추천]
3 [SK-052,상의,색상,매칭,치마,어둡,계열]
4 [SK-053,벨런스,세련미,치마,마음,나쁘,외투,추천]
5 [트렌치,코트,종류,추천,CT-019,이너,색상,무난,트렌치,코트,신발,추천]
6 [운동화,추천,SE-039,스타일,무난,기본,아이템,전체,코디]
7 [CT-019,SW-009,SK-053,SE-039,지금,안해,아이템,전체,코디,제안,마음,마음,감사]
8 [처음,대학교,코디,신입생,코디,화사,스웨터,추천,SW-009,치마,추천,SK-053,벨런스,세련미,치마,마음,나쁘,외투,추천,트렌치,코트,종류,추천,
CT-019, 이너, 색상, 무난, 트렌치, 코트, 신발, 추천, 운동화, 추천, SE-039, 스타일, 무난, 기본, 아이템, 전체, 코디, CT-019, SW-009, SK-053, SE-039,
지금, 안해, 아이템, 전체, 코디, 제안, 마음, 마음, 감사, CT-019, SW-009, SK-053, SE-039]
    
```

Fig. 6. Example of preprocessed conversation data.

```

BL-001,단추,전체,오픈,스탠드,칼라,보이넥,네크라인,결합,스타일,손목,일자,소매,여유,어깨,허리,세로,절개,장식,와이드,컷프스,100,구김,
드라이,클리닝,권장,시원,보이,소라,단색,깔끔,느낌,여성,패미닌,세련,사랑,깔끔,오피스,로맨틱,데이트,이미지,단정,오피스,이미지
    
```

Fig. 7. Example of preprocessed clothing metadata.

pieces. The clothing items associated with ‘USER SUCCESS’ tags are combined to create a coordinated set data.

Fig. 6 shows an example of preprocessing conversation data. Here, we demonstrate how the raw conversation is segmented and refined after five stages of preprocessing, highlighting the process of extracting relevant tokens and augmenting satisfied items into sets. Finally, the preprocessing process combines the six processed data fragments with the adjusted set data, resulting in a total of eight data items.

3.3.2. The preprocessing of Clothing Metadata

Fig. 7 illustrates the preprocessing process for clothing metadata. It shows the structured and refined output after three stages of preprocessing the attributes of a single clothing item. This output is ready for text embedding.

The clothing data(mdata) comprises various attributes from the type of clothing item to color, features, and material. Typically, one clothing item is represented by more than ten attributes(Fig. 3). After preprocessing, it is essential to bundle the attributes that represent a single clothing item together. The preprocessing of clothing data involves three steps: ‘morphological analysis,’ ‘part-of-speech extraction,’ and ‘removal of tokens with a length of one or less.’ The steps are identical to those used for conversation data. After preprocessing a single clothing item as shown in Fig. 3, data as depicted in Fig. 7 can be generated.

4. Experimental

To compare and analyze the impact of preprocessing conversation data and clothing data on text embedding, experiments were conducted using the ‘FASCODE’ dataset at each

stage of preprocessing, with performance compared across stages.

The experiments were set up to progress step by step, interweaving the five stages of conversation data preprocessing with the three stages of clothing data preprocessing, resulting in a total of 15 experimental outcomes. The learning model employed was the FASTTEXT model capable of word embedding, with the word dimension set at 300, a minimum frequency of 1, a training window size of 136, epochs set to 100, and consistently configured for the Skip Gram model. As an evaluation metric, the Weighted Kendall Tau(WKT) was used. The WKT is a type of rank correlation coefficient that calculates the association between two variables by comparing their rankings. The experiments are conducted as follows: The model is trained using preprocessed conversation data and clothing data for each stage.

The conversation set of the model validation data generates conversation set tokens through tokenization. The rank sets (R1, R2, R3), composed of outerwear, tops, bottoms, and shoes, are divided by item unit, and duplicates are removed to create an item list. Using the trained model, the similarity is calculated by using the embedded vector values of the generated tokens and each item in the item list. Once the similarity scores are calculated for each item, the scores are summed for each rank set, and the set with the highest total similarity score is ranked the highest. Evaluation is performed using the correct rank and the WKT evaluation metric.

5. Results

The results of the step-by-step comparative analysis of the impact of preprocessing conversation data and clothing data on

Table 6. Performance comparison using WKT at successive preprocessing stages

Preproces Stage		Evaluation Metrics	
Conversation	Clothing Item	WKKT Sum	WKT Avg.
-	-	23.45	0.11
1	1	65.45	0.32
1	2	66.72	0.33
1	3	62.45	0.31
2	1	63.36	0.31
2	2	59.90	0.29
2	3	67.18	0.33
3	1	55.54	0.27
3	2	56.81	0.28
3	3	55.81	0.27
4	1	49.00	0.24
4	2	55.81	0.27
4	3	55.18	0.27
5	1	73.99	0.36
5	2	75.54	0.37
5	3	82.18	0.41

WKT(Weighted Kendall Tau) is the primary ranking correlation coefficient used to evaluate the model's performance.

text embedding are as follows(Table 6).

Conversation data preprocessing is conducted in five stages: 'morphological analysis', 'removal of unnecessary sentences', 'part-of-speech extraction', 'removal of tokens with a length of one or less', and 'data augmentation'. When only basic tokenization was applied to the conversation data and clothing data, the WKT Sum was 23.45, and the WKT Avg was 0.11, showing the lowest performance. Even compared to using conversation stage 4 and clothing item stage 1, which showed the lowest performance among the cross-experiments of preprocessing stages (WKT Sum 49.00 and WKT Avg 0.24), the basic tokenization method's performance was significantly lower.

The results of the conversation data preprocessing stage experiments show that performance improved as stages were added up to stage 2("Removing Unnecessary Sentences") after stage 1("Morphological Analysis"). This suggests that both stage 1 and stage 2 help model training by removing unnecessary elements from the conversation data. On the other hand, performance declined as stages 3("Part-of-Speech Extraction") and 4 ("Removing Tokens with Length Less Than 1") were added. This decline is likely due to the reduction in the size of the training data as useful content was removed during stages 3 and 4. However, stage 5("Data Augmentation") increased the overall training data size, leading to a synergistic effect with the previous stages and resulting in the highest performance. For

clothing data preprocessing, stage 3 showed the highest performance in most experiments, although there was little difference between stage 1("Morphological Analysis"), stage 2 ("Part-of-SpeechExtraction"), and stage 3("Removing Tokens with Length Less Than 1"). Clothing items are expressed with various details such as features, materials, colors, and sentiments. However, most users, being non-experts, engage in conversations using relatively simple keywords, such as "Recommend clothes for a spring picnic." Therefore, the tokens of clothing items containing many specialized key words did not significantly affect performance.

Consequently, using both the five-stage conversation data preprocessing and the three-stage clothing data preprocessing methods proposed in this paper resulted in a WKT Sum of 82.18 and a WKT Avg of 0.41, the highest performance among all experiments. Additionally, compared to the performance of the basic method without preprocessing (WKT Sum of 23.45 and WKT Avg of 0.11), there was a significant improvement.

6. Discussion and Implications

The core of this research lies in successfully processing complex conversational data to significantly enhance numerical performance(WKT); however, its implications can be expanded beyond mere technical boundaries into the human and aesthetic realms of fashion. This will be discussed further by expanding the analysis across three perspectives: the evolution and conceptual contribution to fashion informatics, the consumer psychological implications regarding preference articulation and personalization perception, and the alignment with the aesthetics of coordination.

First, examining the evolutionary process of data-driven fashion informatics reveals that initial research primarily focused on utilizing static and structured data, based on information such as user purchase history, ratings, basic profile information, and survey results. However, the outcomes of these studies failed to reflect the inherent visual and emotional characteristics of fashion and struggled to grasp contextual preferences based on a user's situation or mood. Subsequently, the introduction of multimodality and attribute enrichment became necessary, as research leveraging the visual features of fashion items alongside deep learning advancements gained importance. Research in this phase is still ongoing, but its limitation lies in the fact that user preferences are often confined to predefined attributes or static images, failing to capture the complex or subtly changing user intent in real-time, which remains a challenge to be addressed. Most recently, the field has evolved towards "Conversation &

Dynamic Curation,” centering on dialogue data to improve user experience and satisfy complex needs. That is, it dynamically grasps the user's context and subtle preferences by integrating conversational data between the user and the AI coordinator with existing clothing metadata. This study corresponds to this latest third stage. It moves beyond simple item recommendation to curate personalized Outfit Sets that reflect the user's situation (weather, location, activity) and psychology, and makes the recommendation process itself interactive via conversational AI. By presenting an empirical preprocessing framework that effectively handles complexly structured dialogue data (a gap in prior research) and integrates it with clothing attributes, this research has established the technical foundation to lead fashion informatics into the era of conversation-based dynamic curation. This evolution demonstrates the fashion industry's shift toward providing personalized consumer experiences through data-driven decision-making.

Next, examining the consumer psychological implications, conversation-based curation brings about a fundamental change in how fashion consumers articulate preferences and perceive personalization. Fashion inherently relies heavily on situational dressing behavior. Dialogue captures user requirements regarding specific situations and environments, and our preprocessing methodology refines this subtle, contextual information to link it with item attributes (FMCE). This achieves a high level of personalization by reflecting the fine-grained needs of customers that traditional systems overlooked. Fashion is a means of self-expression and is intrinsically an emotional domain. This system utilizes the ‘USER SUCCESS / USER FAIL’ information to when users are satisfied or dissatisfied, allowing the model to learn subjective user satisfaction rather than purely algorithmic results. This provides an interactive experience akin to consulting with a real stylist, granting the user a sense of control and emotional attachment, thereby further enhancing the perception of personalization regarding the recommendations.

Finally, this study directly contributes to the aesthetic value of coordination as it ultimately recommends an “Outfit Set,” moving beyond simple item similarity calculations. Although the system technically ranks outcomes via the WKT score, underlying this are aesthetic principles such as Color Harmony, Fabric Compatibility, and Feature (Design Consistency). The system successfully internalizes the aesthetic rules of fashion coordination based on data learning by matching contextual keywords (e.g., ‘neat’, ‘feminine’) extracted from the dialogue with the item attributes (FMCE). This provides significant cross-disciplinary value, linking technical achievement with the research domains of fashion design and styling.

7. Conclusions

This paper proposes and compares the performance of preprocessing methods for training data for text embedding in an automated clothing curation system. Unlike conventional text data, the complexly structured conversation data demonstrated that the proposed preprocessing methods were effective. When conversation data and clothing metadata were combined and used together, good performance was observed.

The proposed preprocessing methods are expected to enable the development of an interactive recommendation system, advancing beyond traditional fashion recommendation systems that recommend individual clothing items. This paper makes significant contributions both academically and industrially. Academically, the study presents a novel and reproducible preprocessing framework tailored for Korean conversational data, merging natural language processing and fashion informatics, thereby providing a foundation for future multimodal or linguistic AI research in the fashion domain. This work contributes to elevating the academic value in the field of fashion data processing by proposing and experimentally validating an effective preprocessing methodology for complexly structured dialogue data, which enhances the performance of automated fashion curation systems and provides an optimized data preparation method for model training. Industrially, we believe this paper can contribute to the successful completion of the digital transformation of the fashion industry by bringing practical improvements—specifically, enhancing the performance and automation of AI fashion curation systems—across relevant industry sectors, including e-commerce, digital styling, and virtual fitting platforms.

The significance of this study lies in successfully internalizing aesthetic rules for fashion coordination based on data learning, matching contextual keywords with item attributes rather than simply extracting simple words from conversations between consumers in online fashion shopping malls. This provides significant interdisciplinary value by connecting the technical achievements of AI learning to the fields of fashion design and styling research.

In the future, the addition of clothing image data is planned to explore preprocessing methods utilizing multimodal data.

Acknowledgement

This study was developed based on the work that received third place in the “Fashion-How” category of the ETRI (Electronics and Telecommunications Research Institute) Auto-

mous Growth Artificial Intelligence Competition, organized by the ETRI.

Reference

- Bellini, P., Bruno, I., Nesi, P., & Paolucci, M. (2013). A static and dynamic recommendations system for best practice networks. In *Human-Computer Interaction: Users and Contexts of Use. Proceedings of the 15th International Conference, HCI International 2013* (pp. 259-268). Springer.
- Bellini, P., Cenni, D., & Nesi, P. (2014). Optimization of information retrieval for cross media contents in a best practice network. *International Journal of Multimedia Information Retrieval*, 3, 147-159. doi:10.1007/s13735-014-0058-8
- Bellini, P., Palesi, L. A. I., Nesi, P., & Pantaleo, G. (2023). Multi clustering recommendation system for fashion retail. *Multimedia Tools and Applications*, 82(7), 9989–10016. doi:10.1007/s11042-021-11837-5
- Chen, W., Zhang, B., Xu, K., & Zhou, X. (2019). MM-Attn: Multimodal attention network for fashion recommendation. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM)*, (pp. 303-312). ACM.
- Chen, Y. N., Celikyilmaz, A., & Hakkani-Tur, D. (2018). Deep learning for dialogue systems. In *Proceedings of the 27th International Conference on Computational Linguistics: Tutorial Abstracts* (pp. 25-31). Association for Computational Linguistics.
- Deldjoo, Y., Rafiee, M., & Ravanbakhsh, S. (2025). Agentic mixed-modality refinement: A research agenda for next-generation fashion recommendation systems. *ACM Computing Surveys*, 57(1), 1-38.
- Electronics and Telecommunications Research Institute(ETRI). (n.d.). *Fascode dataset*. Retrieved October 21, 2025, from <https://fashion-how.org/ETRI/board20.html>
- Zhang, H., Jiang, K., Zhang, W., & Li, J. (2021). Personalized fashion recommendation from personal social media data: An item-to-set metric learning approach. In *2021 IEEE International Conference on Big Data (Big Data)* (pp. 5014–5023). <https://doi.org/10.1109/BigData52589.2021.9671563>
- Kim, J., Kang, S., & Bae, J. (2022). The effects of customer consumption goals on artificial intelligence driven recommendation agents. Evidence from Stitch Fix. *International Journal of Advertising*, 41(6), 997-1016.
- Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42(8), 30-37.
- Li, Z., Chen, J., & Huang, J. (2021). Attribute-aware complementary fashion recommendation based on BERT framework. In *Proceedings of the ACM International Conference on Multimedia (MM)*, 5782–5790. ACM
- Omprakash, S., Muthusamy, C., & Shamik, S. (2019). Personalised fashion recommendation using deep learning. In *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, (pp.368-368).
- Walek, B., & Spackova, P. (2018). Content-based recommender system for online stores using expert system. In *2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*(pp. 164-165). IEEE. doi:10.1109/AIKE.2018.00036
- Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, 13(3), 55-75. doi:10.1109/mci.2018.2840738

(Received September 30, 2025; 1st Revised October 16, 2025; 2nd Revised November 10, 2025; 3rd Revised November 13, 2025;

Accepted December 1, 2025)