Research Statement

Yunshan Ma

National University of Singapore

yunshan.ma@u.nus.edu

July 2021

Abstract

My research mainly focuses on computational fashion analysis, that is developing computational methods to solve the problems in the fashion area. Fashion is an indispensible part of human life, and the matured fashion industry enables people to conveniently purchase their preferred fashion items. Meanwhile, the past decade has witnessed the fast development of Internet, Big Data, and Artificial Intelligence, which raise new challenges as well opportunities to the entire fashion industry. Therefore, I aim to target on the new problems in fashion domain and seek to tackle them with the computational methodology of Multimedia, Information Retrieval, Natural Language Processing, etc. In this research statement, several of my research topics will be introduced: 1) fashion knowledge extraction; 2) fashion trend forecasting; 3) fashion recommendation; 4) fashion chatbot.

1 Fashion Knowledge Extraction

We aim to extract fashion knowledge from large-scale unstructured data, such as images, texts, and multimodal data. Different from most of current works, which only focus on recognizing fashion elements from visual inputs, we aim to extract more comprehensive user-centric fashion knowledge from multimodal data. In [1, 2], we specifically define fashion knowledge in a triplet form of ¡occasion, person, clothing¿, each instance of which describes a tip about what to wear for certain type of people in specific of occasion.

For person detection and analysis, we adopt the off-the-shelf tools to detect the body and face within each image and estimate the age and gender of the person. For occasion and clothing, we design a contextualized fashion concept learning module to take advantage of the correlations between all the appeared fashion elements. To reduce the cost of labelling data manually, we leverage a weak label modeling module to achieve good performance largely based on machine-labelled data, only with a small portion of human-labelled data. We also build a benchmark dataset and a demo system to demonstrate the acquired fashion knowledge.

2 Fashion Trend Forecasting

The essence of fashion is changeability. People's fashion preference as well as the fashion knowledge are evolving all the time, which forms the fashion trend. How to capture the evolution and dynamics and further give forecasting of fashion trend have always been the core research problems in fashion domain. Thanks to the works of fashion knowledge extraction, we are well-prepared to study fashion trend based on extracted large-scale structured data. We formulate the task as a time series forecasting problem and aim to develop models to precisely predict the future trends based on the historical observations.

In [3, 4], we propose to incorporate the relations among different fashion elements and user groups to enhance the neural network model for better forecasting performance. Particularly, we first introduce an LSTM encoder decoder framework as the backbone, which is a deep learning-based sequence to sequence model for multi-horizon time series forecasting. In addition, since the trends of different fashion elements usually affect one another, we adopt a message passing module to propagate information from the influencer to the target within a pair. We also create a fashion trend dataset FIT, which covers more fine-grained fashion elements, more user groups, and longer time span compared with current fashion trend dataset.

3 Fashion Recommendation

As information explosion posits great challenges for people, recommender systems have become a stakeholder in most of the online services. When it comes fashion domain, recommender systems are also essential for consumers especially when people shop online facing with millions of fashion items. In terms of fashion recommendation, we have explored two strands of work: 1) item to item recommendation, i.e, fashion matching; 2) sequential fashion recommendation.

3.1 Fashion Matching

Given a fashion item as query, fashion matching model is required to recommend an ordered list of items that are probably matched with the query. It will greatly benefit consumers' online shopping experience and improve the online retailers' profit. The key to fashion matching model lies in modeling the mix-and-match relationship between a pair of fashion items or a set of fashion items (outfit).

In [5], we build a Translation-based Neural Fashion Compatibility Modeling (TransNFCM) model. It can jointly learn fashion item embeddings and category-specific complementary relations in the same space. Both visual and textual information are utilized though deep encoder networks. Moreover, besides recommendation precision, explanations of recommendation results are also required to convince the users. Motivated by this, in [6], we develop a method named Attribute-based Interpretable Compatibility (AIC) for interpretable fashion matching. Specifically, we borrow the explainable model of decision tree to generate explainable decision rules, and then learn embeddings based on the extracted rules for final recommendation. Those decision rules can serve as explanations along with the recommendation results.

3.2 Sequential Fashion Recommendation

Sequential fashion recommendation aims to recommend users with fashion items, which should not only consider users' long-term preference but also the short-term preference. Given a chronologically ordered user-item interaction history, sequential fashion recommendation model aims to predict the next item that the user will interact with. The key to this task lies in two aspects: the user-item interaction modeling and item-item transition modeling. However, the large set of user and items as well the sparse interaction history posit challenges for this problem.

To tackle these challenges, in [7], we propose a model named Dual-Graph Sequential Recommender, which leverages two types of global graphs and takes advantage of great modeling power of Graph Neural Network (GNN). As a result, the representations of user and item will be greatly enhanced by their neighbors in the two types of global graph, i.e., the global user-item interaction graph and item-item transition graph.

From another independent direction, we aim to uncover the users' instant intent and capture the content-level transition when predicting the next interaction. In [8], we split the next items into three sub-groups according to user's instant intent, i.e. mix-and-match, substitution, and others. The instant intent is derived by the category composition of two adjacent items in the interaction sequence. In addition, since the content-level attributes are crucial for user's decision for next pick, we develop an attention module to consider the content-level transition for next-item recommendation.

4 Fashion Chatbot

Serving as an online assistant, fashion chatbot can automatically response to consumers' inquiries and efficiently route users to their preferred items. Compared with other types of chatbot, which are usually purely text-based, fashion chatbot must be able to understand the multimodal information when interacting with the real users. Therefore, multimodal dialogue system is in demand for a successful online fashion chatbot. In [9], we develop a Knowledge-aware Multimodal Dialogue (KMD) model to address the problem of building multimodal fashion chatbot. Specifically, it is characterized with a taxonomy-based learning module, an end-to-end neural conversational model, and a deep reinforcement learning-enhanced optimization method. Later on, we also build an online demo [10], which not only realize the functions mentioned in [9] but also integrate a fashion matching module [5] into the system.

References

[1] Y. Ma, X. Yang, L. Liao, Y. Cao, and T.-S. Chua, "Who, where, and what to wear? extracting fashion knowledge from social media," in *Proceedings of the 27th ACM International Conference on Multimedia*, 2019, pp. 257–265.

[2] Y. Ma, L. Liao, and T.-S. Chua, "Automatic fashion knowledge extraction from social media," in *Proceedings of the 27th ACM International Conference on Multimedia*, 2019, pp. 2223–2224.

[3] Y. Ma, Y. Ding, X. Yang, L. Liao, W. K. Wong, and T.-S. Chua, "Knowledge enhanced neural fashion trend forecasting," in *Proceedings of the International Conference on Multimedia Retrieval*, ACM, 2020, pp. 82–90.

[4] Y. Ding, Y. Ma, L. Liao, W. Wong, and T.-S. Chua, "Leveraging multiple relations for fashion trend forecasting based on social media," *IEEE Transactions on Multimedia*, 2021.

[5] X. Yang, Y. Ma, L. Liao, M. Wang, and T.-S. Chua, "Transnfcm: Translation-based neural fashion compatibility modeling," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 403–410.

[6] X. Yang, X. He, X. Wang, Y. Ma, F. Feng, M. Wang, and T.-S. Chua, "Interpretable fashion matching with rich attributes," in *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2019, pp. 775–784.

[7] Y. Ding, Y. Ma, W. K. Wong, and T.-S. Chua, "Leveraging two types of global graph for sequential fashion recommendation," *arXiv preprint arXiv:2105.07585*, 2021.

[8] Y. Ding, Y. Ma, W. Wong, and T.-S. Chua, "Modeling instant user intent and content-level transition for sequential fashion recommendation," *IEEE Transactions on Multimedia*, 2021.

[9] L. Liao, Y. Ma, X. He, R. Hong, and T.-s. Chua, "Knowledge-aware multimodal dialogue systems," in *2018 ACM Multimedia Conference on Multimedia Conference*, ACM, 2018, pp. 801–809.

[10] L. Liao, Y. Zhou, Y. Ma, R. Hong, and T.-s. Chua, "Knowledge-aware multimodal fashion chatbot," in *Proceedings of the 26th ACM international conference on Multimedia*, 2018, pp. 1265–1266.