A Dataset for Adapting Recommender Systems to the Fashion Rental Economy

Karl Audun Kagnes Borgersen
University of Agder
Grimstad, Norway
karl.audun.borgersen@uia.no

Morten Goodwin
University of Agder
Grimstad, Norway
morten.goodwin@uia.no

Morten Grundetjern
University of Agder
Grimstad, Norway
morten.grundetjern@uia.no

Jivitesh Sharma
University of Agder
Grimstad, Norway
jivitest.sharma@uia.no

ABSTRACT

In response to the escalating ecological challenges that threaten global sustainability, there's a deep need to investigate alternative methods of commerce, such as rental economies. Like most online commerce, rental or otherwise, a functioning recommender system is crucial for their success. Yet the domain has, until this point, been largely neglected by the recommender system research community.

Our dataset, derived from our collaboration with the leading Norwegian fashion rental company Vibrent, encompasses $64k$ transactions, rental histories from $2.2k$ anonymized users, and $15.8k$ unique outfits in which each physical item's attributes and rental history is meticulously tracked. All outfits are listed as individual items or their corresponding item groups, referring to shared designs between the individual items. This approach underlines the novel challenges of rental as compared to more traditional recommender system problems where items are generally interchangeable. Each outfit is accompanied by a set of tags describing some of their attributes. We also provide a total of $50k$ images displaying across all items along with a set of precomputed zero-shot embeddings.

We apply a myriad of common recommender system methods to the dataset to provide a performance baseline. This baseline is calculated for the traditional fashion recommender system problem of recommending outfit groups and the novel problem of predicting individual item rental. To our knowledge, this is the first published article to directly discuss fashion rental recommender systems, as well as the first published dataset intended for this purpose. It is our hope that the publication of this dataset will serve as a catalyst for a new branch of research for specialized fashion rental recommender systems.

The dataset is temporarily hosted at https://anonymous.4open.science/r/clothes_rental_dataset-B1E8/ and Google Cloud at https://console.cloud.google.com/storage/browser/clothes-rental-dataset.

The code for collecting, processing, and evaluating the dataset is hosted at https://anonymous.4open.science/r/Vibrent_Dataset_Collection-03DC/

CCS CONCEPTS

• **Information systems → Recommender systems**; **Personalization**.

KEYWORDS

Recommender Systems, Fashion, Rental, Dataset

ACM Reference Format:
Karl Audun Kagnes Borgersen, Morten Goodwin, Morten Grundetjern, and Jivitesh Sharma. 2018. A Dataset for Adapting Recommender Systems to the Fashion Rental Economy. *Proc. ACM Meas. Anal. Comput. Syst.* 37, 4, Article 111 (August 2018), 5 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

In the estimation of the EU, the fashion industry is responsible for between 2% and 10% of the environmental impact of consumption in its member states[13]. Meanwhile, the average number of times an outfit has been worn before being disposed of has decreased by 36% between 2000 and 2015[5]. Though an increased focus on the recycling of clothes and materials may be a component of a more sustainable economy, reuse is substantially more beneficial than recycling[5]. Even though companies built around Circular Business Models (CBMs) and sustainability are generally perceived positively by the general public[11], This positive perception is not enough to gain traction for a functioning rental economy. Rather, the move to a rental economy needs to be made as frictionless as possible. To minimize the perceived inconvenience of using digital platforms, users should be presented with clothing options that are immediately appealing[2]. One of the most pertinent ways to accomplish this is through effective Recommender Systems (RS) tailored to this domain, a topic that, until this point, has seen remarkably little written about it.

Vibrent is a leading Norwegian fashion rental company established in 2020 originally called Fjong. The company is particularly

ACM 2476-1249/2018/8-ART111
https://doi.org/XXXXXXX.XXXXXXX

focused on providing women's everyday and party wear. During the period since its founding, it has collected troves of useful, implicit RS data of its customers and product line. Though it does offer one-time rentals, its primary focus is on its subscription service with which its users may rent a set number of clothes per month.

Recommendations for clothes rental services offer several unique challenges not present in other domains, such as managing temporal item availability and modeling item condition. Though most fashion rental companies will likely find the use of more traditional RS adequate for the job of providing recommendations, the creation and application of tailored RS will improve performance in the domain.

The key contribution in this paper is the Vibrent Clothes Rental dataset. The lack of existing literature on clothes rental RS is remarkable, but could be explained by the lack of datasets dedicated to this purpose with the exception of the dataset discussed in Section 2.1. With the introduction

2 RELATED WORK

In contrast to fashion rental, RSs for the domain of fashion retail are largely well-explored. Deldjoo et al.[4] has published a survey paper detailing the different aspects of RS research regarding fashion. Two examples of well-established datasets for this domain are the H&M dataset[6] and the Amazon Fashion dataset[12]. These document the transactions of two of the world's largest fashion retailers and will, therefore, naturally be far more extensive than our own dataset. These are still retail companies, however, and won't be suitable for capturing the signals inherent in rental recommendations. Further, the level of detail concerning the outfits is lacking as compared to our dataset with the documentation of both attributes (tags) and product images being less comprehensive. While fashion CBMs and rental enterprises have received some previous academic attention, this research has predominantly emerged from less quantitative disciplines [2, 5, 8, 11]. These generally cover the aspects of clothing rental relating to its sustainability and customer perception.

There is a previous precedent for the publication of papers announcing datasets for relevant tasks. This is particularly common in the context of Natural Language Processing and Large Language Models as a method for acquiring quantitative metrics for performance on specific tasks. QASPER[3], Hellaswag[18], are two prominent examples of this. The paper introducing LiveRec by Rappaz et al.[14] also released their accompanying dataset. There is some existing literature relating to the application of RSs to rental in other areas aside from fashion, for example, Iqbal et al discuss applying machine learning to optimize the rental activities of a Korean library[7].

2.1 Existing datasets

To our knowledge the only existing dataset relating to fashion rental is the RentTheRunway dataset collected by Misra et al. for their article discussing adapting RSs to account for clothing fit[10]. It appears to have been collected through scraping Rent the Runway's website. The dataset contains 192, 544 transactions and 5, 850 items in total, but little supplemental information about the relevant items. Remarkably few of the transactions discussed are from repeat customers, only 32% of the customerbase has more than a single transaction associated with them. The dataset has been made available at https://cseweb.ucsd.edu// jmcauley/datasets.html.

3 METHOD

This paper introduces the Vibrent Clothes Rentaldataset, which consists of four separate components: outfits, users, pictures, and transactions. Of these, there are $15.8k$ unique outfits and $9.7k$ outfit groups, $2.2k$ anonymized users with accompanying rental histories, $50k$ pictures, and a total of $64k$ transactions. These users represent all users that have rented at least one item via Vibrent's subscription service. As Vibrent worked off a subscription model, the average user in our dataset has rented a significantly higher number of outfits than the RentTheRunway dataset discussed in Section 2.1. Overall less than 3% of users have rented only a single item.

In the context of this RS rental problem, there are two distinct approaches for evaluating recommendations, either for individual items or groups of items. As Vibrent's initial priority was to gather a catalog with the highest possible diversity of clothing items, a significant portion of them (49.7%) are completely unique compared to the rest of the dataset. The remaining half has two or more copies of the same item available, though not always in the same size. All in all, of the 9791 distinct outfit groups available, 76.8% have a single item tied to them while the rest have two or more items. Predicting a user's interest in an outfit group will naturally be easier than predicting interest in individual items. Later benchmarks will refer to these as either "Groups" or "Individuals" and will evaluate these separately. None of the data contains explicit product feedback from the user, any benchmarks applied will therefore be compatible with implicit training methods.

As the dataset consists of relatively few users with comparatively sizable rental histories, we evaluate performance by hiding a portion of the most recent rentals of a given user and quantify performance using the Hit Rate at 10 and 100 (HR@10, HR@100). The size of the test data is 30% of the user's entire rental history (rounded down). To measure a method's reliance on previous rentals to improve performance, we also record HR@10-new and HR@100-new, which prevents the hidden test outfit from the previous rental history from being one the user previously rented. Hit rate is defined as HR@n = $\frac{H_n}{U}$ in which H_n is the number of users for whom the correct item is in the top n recommendations and U be the total number of users.

3.1 Picture Embeddings

Although images are vital for ensuring humans are interested in renting a clothing item, they provide little useful information to machines and RSs in their raw form. As these images have been made available alongside the rest of the dataset, any parties interested in applying their own methods of embedding these images are free to do so. This process can be compute and time intensive, however. With this in mind, we provide a set of pre-computed embeddings for each image in our dataset. These embeddings were arrived at via retrieving the embeddings from the second to last layer of the torchvision[9] pre-trained EfficientNet_V2_L[16] model. We chose this model because it is the top-performing model for zero-shot fashion image similarity[1].

3.2 Heuristic Baselines

We evaluate the following heuristic baselines:

- **Popular Outfits (Pop)**: Recommends the n globally most commonly rented outfits to every single user.
- **Repeat Outfits (Rep)**: Recommends the n outfits the user has previously rented.
- **Popular and Repeated Outfits (Pop + Rep)**: A combination of Popular Outfits and Repeat Outfits. Recommends repeated items if the user has less than n rentals. Otherwise, the recommendations are padded with the most popular global items.

3.3 Collaborative Filtering Baselines

We evaluate two collaborative filtering techniques:

- **Alternating Least Squares (ALS)**: ALS is a matrix factorization technique particularly effective for handling large-scale and sparse datasets and can incorporate both explicit and implicit feedback [19]. ALS has been applied to our data with 32 factors and regularization of 0.01.
- **Bayesian Personalized Ranking (BPR)**: BPR is a pairwise learning method designed to optimize for personalized ranking. It utilizes a matrix factorization model with the objective of maximizing the difference between the observed positive interactions and unobserved negative interactions [15]. BPR has been applied to our dataset with 128 factors, a regularization of 0.01, and a learning rate of 0.01.

3.4 Content-based Baselines

We evaluate two content-based approaches:

- **Image Embeddings (Img Embed)**: A content-based approach built on the Vibrent Clothes Rentaldataset image embeddings as discussed in Section 3.1. Recommendations are made by calculating the n nearest neighbors of a mean representation of all images in a user's rental history.
- **Tag Embeddings (Tag Embed)**: A content-based approach built on each outfit's tags. Each outfit's set of tags is one-hot encoded, and a user is represented as the mean embeddings across its rental history. Recommendations are made by calculating the n nearest neighbor outfit embeddings of these user embeddings.

4 RESULTS AND DISCUSSION

The results from the baseline experiments are presented in Table 1. Overall, the simple heuristics of recommending popular and repeated items fared remarkably well, outperforming the other methods when evaluation allows for repeated items. Despite this domain being related to rentals, users are seemingly still most inclined to interact with items they have previously engaged with. This phenomenon is well established in other domains too, however, a particularly relevant example of this is within live streaming recommendation[14]. The collaborative filtering models outcompete the other methods when recommending novel items.

Recommendations within the domain of fashion rental have the added advantage of far denser interaction matrixes than traditional retail fashion RSs. Quantifying a sparsity score as $Sparsity = 1 - \frac{T}{UxI}$ where T is the total number of unique transactions, U is the number of users, and I is the number of items. Figure ?? displays a direct comparison of density between the Vibrent Clothes Rental-dataset and some prominent Collaborative Filtering (CF) datasets. While still more sparse than the Netflix Prize Competition dataset, both versions are still significantly denser than most other RS datasets, which speaks to the applicability of CF-based methods compared to most other RS problems.

The top-performing CF models differ when recommending item groups or individual products. Indicating that algorithms can be more suited to one of the tasks and not the other. The content-based methods recommending items based on tag and image attributes also perform far better on individual product recommendations as compared to group recommendations. In the case of the tag-based method, this could to some extent be chalked up to outfit size being an explicit factor and brand loyalty. Minor adjustments to the scheme for the domain, such as applying a simple machine learning approach to weight tag categories, would likely make it the most performant on this topic.

4.1 Future work

Although traditional fashion RSs perform adequately in the context of fashion rental as well, as we've already demonstrated, there are some unique challenges to this domain.

4.1.1 Clothing wear and decay. As clothing items are worn over time, their condition will naturally deteriorate as they accumulate wear from multiple uses. The Vibrent Clothes Rentalcould be used to model this gradual decay by observing when clothes are either deleted or retired and re-sold. Any CBM company interested in developing an RS tailored for rental would likely benefit from modeling the likelihood of a clothing piece being retired after any number of successive uses and weighting tentative recommendations accordingly.

4.1.2 Clothing availability. Similarly to the live-streaming RS by Rappaz et al.[14] a clothing rental RS would benefit from consciously factoring in whether a clothing item was available when factoring in positive and negative signals.

As the same individual items cannot be rented to different people for any given period, a tailored clothes rental RS could also benefit from consciously balancing who gets recommended items in high demand. As we've demonstrated in Table 1, recommending the globally most popular items to all users is a reasonably effective strategy. However, suppose user x and user y both have a popular item among their highest-scoring recommendations. Still,

	HR@10	HR@100	HR@10-new	HR@100-new
Pop Ind	0.0221	0.1372	0.0271	0.1711
Pop Groups	0.0785	0.3439	0.1043	0.3467
Rep Ind	0.1340	0.2671	0.0000	0.0000
Rep Groups	0.1593	0.2974	0.0000	0.0000
Rep + Pop Ind	**0.1354**	**0.3145**	0.0018	0.1530
Rep + Pop Groups	**0.1638**	**0.4499**	0.0086	0.3219
BPR Ind	0.0582	0.2685	**0.0650**	**0.2912**
BPR Groups	0.0609	0.3051	0.0835	0.3345
ALS Ind	0.0469	0.2324	0.0619	0.2569
ALS Groups	0.0857	0.3556	**0.1084**	**0.3589**
Img Embed Ind	0.0338	0.1503	0.0312	0.1670
Img Embed Groups	0.0338	0.1620	0.0361	0.1783
Tag Embed Ind	0.0740	0.3096	0.0623	0.2880
Tag Embed Groups	0.0979	0.3281	0.0659	0.2849

Table 1: Hit rates evaluations for our baseline methods. For a deeper explanation of the different approaches see Sections 3.2, 3.3, 3.4

Dataset	**Density**
Netflix Prize	1.18e-02
Clothing Rental Groups	3.71e-03
Clothing Rental Individual	2.32e-03
Rent The Runway	3.12e-04
H&M Fashion	1.92e-04
Book Rental	3.41e-05
Goodreads	1.63e-05
Amazon Fashion	6.30e-06

Table 2: Comparison between the density of various datasets.

user x has several similar scoring other recommendations, and user y has a larger gap between the top-scoring popular item and the less popular items. In that case, user y should be the only user recommended the popular item. Thereby optimizing global user satisfaction. Wang et al.'s survey on fairness in RSs[17] mentions some relevant existing approaches to this topic when discussing fairness from an item perspective and a system perspective.

5 CONCLUSION

This article has introduced the Vibrent Clothes Rentaldataset and presented some of the unique and novel problems inherent in the clothes rental domain of fashion, such as the recommendation of individual items rather than product lines, the gradual wear and decay of used clothes, and the varying availability of clothing items. We have also made and evaluated a few initial implementations to provide recommendations in this environment to serve as a baseline for later work. The overall best-performing methods among our models are popularity heuristics for repeated recommendations and ALS and BPR for novel product recommendations. The collaborative filtering models differ in performance based on whether they recommend individual outfits or outfit groups, indicating that proficiency in one of the tasks isn't necessarily directly correlated with proficiency in the other. We hope this publication will serve as a springboard to engage more researchers on this issue, aiming to lower the entry barrier for future commercial actors within this field.

6 ACKNOWLEDGEMENTS

We sincerly thank the employees at Vibrent for their diligent work in collecting the data presented in the dataset, as well as for their willingness to share it with the rest of the world.

REFERENCES

[1] Karl Audun Kagnes Borgersen, Morten Goodwin, Jivitesh Sharma, Tobias Aasmoe, Mari Leonhardsen, and Gro Herredsvela Rørvik. 2023. CorrEmbed: Evaluating Pre-trained Model Image Similarity Efficacy with a Novel Metric. In *Artificial Intelligence XL*, Max Bramer and Frederic Stahl (Eds.). Vol. 14381. Springer Nature Switzerland, Cham, 89–102. https://doi.org/10.1007/978-3-031-47994-6_7 Series Title: Lecture Notes in Computer Science.

[2] Fiona Charnley, Fabienne Knecht, Helge Muenkel, Diana Pletosu, Victoria Rickard, Chiara Sambonet, Martina Schneider, and Chunli Zhang. 2022. Can Digital Technologies Increase Consumer Acceptance of Circular Business Models? The Case of Second Hand Fashion. *Sustainability* 14, 8 (April 2022), 4589. https://doi.org/10.3390/su14084589

[3] Pradeep Dasigi, Kyle Lo, Iz Beltagy, Arman Cohan, Noah A Smith, and Matt Gardner. 2021. A dataset of information-seeking questions and answers anchored in research papers. *arXiv preprint arXiv:2105.03011* (2021).

[4] Yashar Deldjoo, Fatemeh Nazary, Arnau Ramisa, Julian Mcauley, Giovanni Pellegrini, Alejandro Bellogin, and Tommaso Di Noia. 2023. A review of modern fashion recommender systems. *Comput. Surveys* 56, 4 (2023), 1–37.

[5] Ellen MacArthur Foundation. 2017. *A new textiles economy: Redesigning fashion's future*. Technical Report. Ellen MacArthur Foundation. https://www.ellenmacarthurfoundation.org/a-new-textiles-economy

[6] Carlos García Ling, ElizabethHMGroup, Frida Rim, inversion, Jaime Ferrando, Maggie, neuraloverflow, and xlsrln. 2022. H&M Personalized Fashion Recommendations. https://kaggle.com/competitions/h-and-m-personalized-fashion-recommendations

[7] Naeem Iqbal, Faisal Jamil, Shabir Ahmad, and Dohyeun Kim. 2020. Toward effective planning and management using predictive analytics based on rental book data of academic libraries. *Ieee Access* 8 (2020), 81978–81996.

[8] Chunmin Lang, Muzhen Li, and Li Zhao. 2020. Understanding consumers' online fashion renting experiences: A text-mining approach. *Sustainable Production and Consumption* 21 (Jan. 2020), 132–144. https://doi.org/10.1016/j.spc.2019.12.003

[9] TorchVision maintainers and contributors. 2016. *TorchVision: PyTorch's Computer Vision library*. https://github.com/pytorch/vision

[10] Rishabh Misra, Mengting Wan, and Julian McAuley. 2018. Decomposing fit semantics for product size recommendation in metric spaces. In *Proceedings of the 12th ACM Conference on Recommender Systems*. ACM, Vancouver British Columbia Canada, 422–426. https://doi.org/10.1145/3240323.3240398

[11] Amira Mukendi and Claudia Elisabeth Henninger. 2020. Exploring the spectrum of fashion rental. *Journal of Fashion Marketing and Management: An International Journal* 24, 3 (April 2020), 455–469. https://doi.org/10.1108/JFMM-08-2019-0178

[12] Jianmo Ni, Jiacheng Li, and Julian McAuley. 2019. Justifying recommendations using distantly-labeled reviews and fine-grained aspects. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*. 188–197.

[13] SAJN Nikolina. 2019. Environmental impact of the textile and clothing industry: What consumers need to know. (2019).

[14] Jérémie Rappaz, Julian McAuley, and Karl Aberer. 2021. Recommendation on Live-Streaming Platforms: Dynamic Availability and Repeat Consumption. In *Fifteenth ACM Conference on Recommender Systems*. ACM, Amsterdam Netherlands, 390–399. https://doi.org/10.1145/3460231.3474267

[15] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2012. BPR: Bayesian Personalized Ranking from Implicit Feedback. (2012). https://doi.org/10.48550/ARXIV.1205.2618 Publisher: [object Object] Version Number: 1.

[16] Mingxing Tan and Quoc Le. 2021. EfficientNetV2: Smaller Models and Faster Training. In *Proceedings of the 38th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 139)*, Marina Meila and Tong Zhang (Eds.). PMLR, 10096–10106. https://proceedings.mlr.press/v139/tan21a.html

[17] Yifan Wang, Weizhi Ma, Min Zhang, Yiqun Liu, and Shaoping Ma. 2023. A Survey on the Fairness of Recommender Systems. *ACM Transactions on Information Systems* 41, 3 (July 2023), 1–43. https://doi.org/10.1145/3547333

[18] Rowan Zellers, Ari Holtzman, Yonatan Bisk, Ali Farhadi, and Yejin Choi. 2019. HellaSwag: Can a Machine Really Finish Your Sentence? (2019). https://doi.org/10.48550/ARXIV.1905.07830 Publisher: [object Object] Version Number: 1.

[19] Yunhong Zhou, Dennis Wilkinson, Robert Schreiber, and Rong Pan. 2008. Large-Scale Parallel Collaborative Filtering for the Netflix Prize. In *Algorithmic Aspects in Information and Management*, Rudolf Fleischer and Jinhui Xu (Eds.). Vol. 5034. Springer Berlin Heidelberg, Berlin, Heidelberg, 337–348. https://doi.org/10.1007/978-3-540-68880-8_32 ISSN: 0302-9743, 1611-3349 Series Title: Lecture Notes in Computer Science.